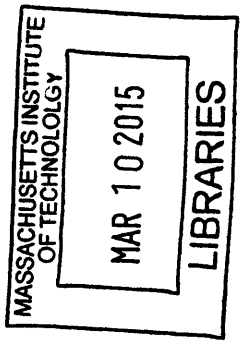

Finding Order in a Contentious Internet

by

Jesse Horton Sowell, II
B.S., Clemson University (2001)
M.S., Michigan State University (2005, 2007)
S.M. Massachusetts Institute of Technology (2010)

Submitted to the Engineering Systems Division
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Engineering Systems
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
February 2015

Signature redacted

Author..... .....
Engineering Systems Division
January 31, 2015

Signature redacted

Certified by..... .....
Ken Oye
Associate Professor, Political Science and Engineering Systems Division
Committee Chair

Signature redacted

Certified by..... .....
David D. Clark
Senior Research Scientist, Computer Science and Artificial Intelligence Laboratory
Thesis Supervisor

Signature redacted

Certified by..... .....
Charles Fine
Professor, Management and Engineering Systems Division
Committee Member

Signature redacted

Certified by..... .....
Nazli Choucri
Professor, Political Science
Committee Member

Signature redacted

Certified by..... .....
Frank Field
Senior Research Scientist, Engineering Systems Division
Committee Member

Signature redacted

Accepted by..... .....
Munther A. Dahleh
William A. Coolidge Professor of Electrical Engineering and Computer Science
Acting Director, Engineering Systems Division

Finding Order in a Contentious Internet

by

Jesse Horton Sowell, II

Submitted to the Engineering Systems Division
on January 31, 2015, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Engineering Systems

Abstract

This inquiry started with the simple question, "Who manages the Internet infrastructure and how?" Since, this question evolved into an evaluation of the routing system and the institutions that manage it. This institutional complex is referred to as the number resource system (NRS). NRS authority is contingent, rooted in consensus-based knowledge assessment necessary to adapt apace with Internet growth. The efficiency with which observable negative externalities are remediated is a compelling entry point to this work. The Pakistan-YouTube story is a halcyon parable of "self-repair." Network operators recognized a *global* negative externality, traced it to the origin, and remediated the complicit networks in approximately three hours. To the casual observer, organic cooperation surfaced to remediate damages, then dissolved into the background noise of "normal operations."

Remediation is far from organic; rather, it is a consequence of distinct rights and obligations amongst, and enforced by, NRS participants. Explaining the rationale and mechanics of "ad hoc" crisis management is the first contribution of this work. The early NRS comprised "close-knit yet loosely organized" communities created to *1*) share operational knowledge (network operator groups, NOGs); *2*) delegate unique network identifiers (Regional Internet Registries, RIRs); *3*) create neutral markets for exchanging routes and traffic (Internet eXchanges, IXes); and *4*) limit abusive messaging (anti-spam, later anti-abuse). Alongside Internet growth, NRS norms evolved into distinct institutions, replete with function-specific constitutional, collective choice, and operational rules for managing the knowledge commons and facilities supporting routing system function. The NRS institutions form a contingent social order, rooted in shared, authoritative images of system function and externalities management. NRS institutions collectively ensure participants *common interests* in the jointly provisioned routing system stability.

The second contribution of this work explains NRS institutional structures and how the attendant rules keep pace with a high clockspeed Internet infrastructure. NRS institutions are characteristically, and *necessarily*, adaptive: each comprises a unique consensus process, animated by a diverse set of nominal competitors, that creates and adapts function-specific rules and processes contributing to routing system integrity. Consensus processes evaluate the performance of common resource management rules and, when—not if—necessary, adapt these rules to satisfy changing resource demands and patterns of use in the broader Internet infrastruc-

ture industry. Anticipation and evaluation in the consensus process are essential to adaptive capability, framed as a form of joint knowledge assessment. Moreover, diverse representation, comprising experts across industry sub-sectors, animated by constructive conflict amongst these experts, mediated by consensus processes, makes for a *durable* family of *credible* knowledge assessment processes that are rare amongst conventional regulatory arrangements.

Processes described thus far are largely endogenous to the NRS and its constituencies. Historically these institutions have operated quietly underneath the hood. Adaptation and the resulting policy is scoped to *common* interests, explicitly avoiding impinging on public policy. In contrast to conventional international regimes, the NRS self-limits to the *scope of its authority*, namely supporting and enhancing routing system function. Thus far, the NRS's common interests have not run counter to the public interest. Nonetheless, a path-dependent history of harmonious alignment between a common and the public does not carry the assurances of alignment resulting from explicit coordination and cooperation. Some states and state-sanctioned international governance organizations see control of NRS facilities as critical to preserving their own authority. Predatory claims to stewardship of routing system resources further complicates the alignment problem.

To better frame and understand this alignment problem, the concluding chapters of the dissertation explore the question: "Are the incentives and resources of NRS institutions commensurate with the aggregate social loss due to a partial, or worse yet, systemic, failure?" Simply put, absent the progress on the explicit assurances above, the answer is no. Would-be state principals also fall short. State-based authorities are severely deficient in basic operational capacity that form the foundation of knowledge assessment capabilities and subsequent adaptive capabilities in the NRS. States' deficiencies correspond to those capabilities engendered by the NRS. Adding NRS stewardship to a state's portfolio of domestic regulatory interests will expose management processes to powerful short-term interests that will inevitably weaken, if not eliminate, extant credible knowledge assessment and adaptive capabilities. In effect, aggressive predatory rule would likely eliminate precisely the characteristics that make the NRS a valuable steward of a high clockspeed infrastructure.

This initial conclusion is not a prediction of adaptive management doomed to failure. Although neither the NRS, nor state authorities, have sufficient capabilities and modes of authority to manage an Internet underpinning an ever-increasing array of public, private, and social goods on their own, a mix of their capabilities *is sufficient*. Rather, the conclusion frames a discussion of what explicit assurances will look like and the barriers to developing those assurances. The last part of this dissertation lays out the challenges for establishing such a comity, a mutual recognition of the norms and authority between the NRS and state authorities.

In the global political arena, the NRS's political capital is credible knowledge assessment and adaptive capacity as the roots of authoritative policy advice. Barriers to explicit assurances draw lessons from the deconstruction and reconstruction of scientific knowledge in political environments, instances of international epistemic consensus, and characteristics of elusive, but effective, adaptation that has

survived in conventional regulatory environments. Analytically, the dissertation argues the NRS and state authority need not be competitors—the two can be quite complementary. If these two sets of institutions can avoid the pitfalls of previous efforts, in particular short-term usurpation of the others' authority, the global, non-discriminatory character of the Internet may be sustainable.

Thesis Supervisor: David D. Clark
Title: Senior Research Scientist, Computer Science and Artificial Intelligence Laboratory

Acknowledgments

First, I would like to thank my advisors, David Clark and Ken Oye, who played two distinctly complementary roles in my academic career at MIT. Dave has been generous with his time and extraordinarily deep knowledge of the Internet, how it has evolved, and the constructs that hold it together. I am indebted to Dave for giving me the latitude to explore this space to find my own topic and approach to the problem. I believe I will always think of one of Dave's favorite phrases, "Writing is God's way of telling you that you don't know what you're talking about," when I sit down to compose. Ken has been the consummate constructive skeptic, helping me understand where I fit in the political economy and broader academic communities, both through advising and my teaching experience in his course Science, Technology, and Public Policy. I am also indebted to Ken for being a champion of my work among his colleagues in these communities. Ken helped me (after just a bit of resistance) to cut to the chase and offer a rich, yet pragmatically compelling narrative.

My committee also comprised Frank Field, Nazli Choucri, and Charles Fine. I am thankful to Frank for helping me navigate MIT, TPP, and ESD—not simply the mechanics of the programs—but also the intellectual landscape of how to think about science, technology, engineering, and society. Our conversations on the nature and character of TPP and ESD have given me insight into how to direct my own work, for which I will always be grateful. Nazli has a fantastic ability to pull structure from a seeming chaos of ideas, helping me understand how to structure writing that was at times more than a bit chaotic in and of itself. In this process she continually helped me explore, develop, and reason about an array theoretical constructs, reminding me not to hew too closely to any of these ideal forms. In one of my first research meetings with Charlie, he suggested I "be a good cartographer," which I came back to frequently as I tried to make sense of the institutional complex described here. I am grateful to Charlie for being able to ask me the profoundly fundamental questions, making me think about the assumptions and offering fantastic advise on how to proceed.

The topic of this work is the network operator groups and anti-abuse communities. Both communities have been phenomenally generous with their time and knowledge of how the Internet's infrastructure operates, how they reason about it, and how their institutions work. I cannot begin to list the diverse set of individuals that shared their experiences and perspectives with me. I am immensely grateful that, from the outset, these communities have welcomed me and encouraged my work, especially when I was, in the community vernacular, "less than clueful." A number of magnanimous individuals went out of their way to introduce me to the rest of the community, providing me with more information, experiences, and exposure to the community than I imagined possible at the outset. I consider many of these individuals colleagues and friends and I am looking forward to those continued friendships.

Last, and certainly not least, I would like to thank my friends, my colleagues

in the Advanced Network Architecture group and ESD, my mother, and my partner Katie for their support during my time in the PhD program. At MIT I have met more amazing people than I can count and am privileged to consider them friends. I am thankful for their support during the program, over numerous coffee breaks, meals, and happy hours, as well as in our many conversations on research and academia. My mother has been a continuous support, always encouraging me to pursue my interests since I was small, and equally interested in everything. Most especially, I owe my incalculable gratitude to my partner Katie. She came along with me on this adventure, forgiving my long work nights, supporting my effort, and ensuring I kept in touch with the world outside MIT. Mason (our dog) also helped—I'm pretty sure our long walks around the neighborhood were more cathartic for me than him.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In December 2008 Pakistan censored a YouTube video that, mocking the Prophet Muhammad, is illegal in Pakistan. Pakistan's censorship strategy manipulated how users originating in Pakistan were routed to YouTube. The Pakistan Telecommunications Authority (PTA) injected a route into the Internet's routing system that was *intended* to (re)direct *Pakistani* users destined for YouTube to a local Pakistani system hosting a censorship message. Unfortunately, the PTA's route "leaked" to the rest of the world. Within a few minutes, YouTube users *around the world* were met with a censorship message from the PTA.

From Pakistan's perspective, local route manipulation was a legitimate, low-cost solution to a local problem within its national jurisdiction. The Internet routing system does not neatly fit into Westphalian jurisdictions. Rather, it is a *global* resource system *jointly* provisioned by a collective of largely *private* actors. Historically, state preferences have held little direct sway over the norms and operational rules that order the management of this system. From the perspective of routing system participants, the global impact of Pakistan's censorship strategy, regardless of its intended scope, damaged the integrity of the global routing system. Pakistan's censorship strategy did not last long in this Internet.

Every network participating in the Internet infrastructure relies on the routing system as the common resource ensuring a set of largely private networks cohere into an *Inter*net. Despite working for potentially competing private organizations, this "close-knit yet loosely organized"[1] community *collaborated* to repair "damage" to this common resource. The community of network operators, participants in and stewards of the routing system, applied a number of well-known monitoring mechanisms and operational rules to quickly identify the PTA as the source of the offending route. From the perspective of the network operator community and the common resource institutions (CRIs) that maintain the integrity of the routing system, Pakistan had introduced an illegitimate route, passed on from the PTA to the outside world by Pakistan's upstream provider, PCCW. Once identified by the community, PCCW limited subsequent damage by temporarily blocking route advertisements originating from the PTA. Within *three hours*, network operators' *pri-*

[1]The phrase "close-knit yet loosely organized" is taken from Ellickson (1991).

vate efforts restored the *global* distribution of what stewards of this global common resource considered legitimate routes to YouTube.

Albeit an instance of ad hoc adaptation to resolve a global security externality, the collaboration evident in the Pakistan-YouTube even is not a one-off event. Tacit in this narrative and others[2] is the systematic monitoring and enforcement of well-known rights and obligations. Multiple combinations of formal and informal operational rules, created and maintained by communities of network operators, are at play creating a social order to manage a seemingly uncoordinated Internet infrastructure. The first contribution of this dissertation is to describe these CRIs in terms of the formal and informal collectives that maintain the integrity of the control plane (routing system). What are the critical common resources in play? Who provisions those resources? How are these resources distributed and utilized (appropriated)? How are rights and obligations enforced, and by whom?[3]

To explain the underlying dynamics, the control plane and the attendant institutional complex[4] is framed as commonly managed resource system, derived from Ostrom's work on common pool resource (CPR) systems and later work on common property institutions and knowledge commons.[5] Chapters 2 and 3 operationalize the vocabulary[6] of common resource management, knowledge commons, and resource rights to explain the mechanics of, and externalities endemic in, the jointly provisioned Internet routing system. In this work, the institutional complex managing the this routing system is referred to as the Number Resource System (NRS).

Part II comprises studies describing and explaining the institutional complex that jointly provisions the Internet's routing system. These studies comprise

1. network operator groups (NOGs, Chapter 4), communities developed to share and disseminate operational norms and best practices;
2. the regional Internet registries system (RIRs, Chapter 5), the five regional organizations that delegate the unique identifiers necessary for Internet communication;
3. Internet eXchanges (IXes, Chapter 6) that mutually provision neutral interconnection platforms, facilitating local, non-discriminatory markets for exchang-

[2]Historical instances of operational externalities include the AS7007 leak in 1997 (Bono, 1997), TTNet in Turkey (Underwood, 2005), and the ConEdison leak (Underwood, 2006). More recently, there has been a reported uptick in hijackings (Cowie, 2013), many of which are presumed to have political and economic motivations. In a very recent instance of a politically motivated hijack, Turkish ISPs are hijacking public DNS addresses to limit access to Twitter and other services (Andree Toonk, 2014; Carstensen, 2014). See Section 2.2 for the precise mechanics of a prefix hijacking in terms of routing and security externalities.

[3]Political scientist will recognized these as a variant of the classic question of politics writ large: "Who gets what, and how?"

[4]Institutional complex is a variant of Keohane's regime complex, a loosely coupled set of (function) specific regimes (2010, p. 1).

[5]CPRs refer to the seminal work by E. Ostrom (1990).

[6]As a paradigm experiencing a resurgence since E. Ostrom (1990), the common resource literature has spent some time reflecting on the vocabulary. See Dolšak and Ostrom (2003a, loc. 176–208) and (Hess & Ostrom, 2003) as a general treatment of diverse vocabulary on the commons and the more precise meanings ascribed in the tradition of Ostrom's commons.

ing routes and traffic; and
4. the anti-abuse community (Chapter 7) that develops monitoring and enforcement mechanisms to reduce abuse externalities.

Each CRI comprises function-specific knowledge commons and/or resource management facilities that provision routing system resources, monitor those resources' use, and/or enforce rights and obligations that ensure routing system integrity.

To keep pace with Internet growth and the changes in topology necessary to support services such voice over IP, streaming video, cloud-based services, and the myriad other online services facilitated by Internet communication, the NRS must necessarily adapt CRI rules and capabilities. The source of the NRS's adaptability is operational capacity and credible knowledge assessment. The second contribution of this work explains how a family of consensus processes within this decentralized, sometimes conflicting set of NRS institutions facilitates credible knowledge assessment. Consensus is the means of credibly evaluating and establishing coordination and cooperation mechanisms amongst contentious, nominal competitors, ensuring the NRS remains aligned with participants' *common* interests.[7] Each community has adapted a variant of the consensus model fit to their function-specific resource problems, environment, and institutional arrangements. Each mode of consensus can trace its origin and early influences to the IETF consensus process.[8]

As a mode of problem solving, consensus differs substantively from majoritarian voting processes. Like majoritarian voting, all valid participants can contribute. Unlike voting, each actor's contribution is weighted by how it contributes to solving the problem at hand. If an actor disagrees with a position, they cannot merely say "no," but rather, the actor must defend their dissent on technical, operational, or economic grounds. Contributions in a consensus process are not opaque, equally weighted "voting rights." As such, consensus does not suffer as easily from the fungibility of "votes" and attendant horsetrading tactics that distort voting systems as decision-making systems. Although susceptible to both deadlock and live-lock,[9] consensus processes have proven to be an effective combination of knowledge assessment and collective choice mechanisms.

NRS consensus processes evolved quietly under the hood of Internet infrastructure operations. These institutions are now faced with increased attention from domestic and global regulators. Further, recent issues in the domain of routing security and post-depletion transfers have challenged the capabilities of the consensus process, requiring decision makers supplement deep operational knowledge with evaluations of the political and economic implications of operational rules.

[7]The notion of harmony, coordination, and cooperation are used in the sense of R. O. Keohane (2005).

[8]The conceptual phases of a consensus process, rooted in the IETF process as an instance, are presented in Section 3.2.5. Each study presents the consensus process in the context of collective choice rules used to develop constitutional and operational rules. Chapter 9 revisits and contrasts these to evaluate how varieties of consensus serve as credible knowledge assessments.

[9]A number of interviewees have indicated they prefer a "fail fast, fail often" approach to avoid drawn out discussions of clearly dysfunctional rule proposals or those better suited to evaluation by a dedicated working group.

The simple question for individual CRIs and the NRS is, "Can they adapt beyond technical decision making to continue managing the NRS in a potentially hostile political environment?"

As the NRS comes out from under the hood, engagement between conventional regulators and a common resource management complex with both institutional and knowledge barriers is a major challenge. NRS institutions have demonstrated substantive adaptive capacity rooted in consensus as a credible knowledge assessment process. Further, NRS institutions have limited resource policy and best common practices to their common interests, explicitly avoiding impinging on the domain of public policy, thus avoiding conflict with state authority. While NRS common interests have not run counter to the public interest, this is a weak form of alignment that does not confer the assurances engendered by alignment resulting from explicit coordination and cooperation. Moreover, increased dependence will force them into some form of engagement, regardless of this weak alignment.

The dilemma in Part III explores whether incentives within the NRS are commensurate with the social cost of partial, or worse yet, systemic, failure. Absent assurances that the common interests of the NRS will be more explicitly aligned with the public interest, the answer to this dilemma is no. Would-be state principals also fall short. State-based authorities are severely deficient in basic operational capacity that form the foundation of knowledge assessment capabilities and subsequent adaptive capabilities—states' deficiencies correspond to those capabilities engendered by the NRS.

Although neither the NRS, nor state authorities, have sufficient capabilities and modes of authority to manage an Internet underpinning an ever-increasing array of public, private, and social goods on their own, a mix of their capabilities *is sufficient*. The conclusion frames a discussion of what explicit assurances will look like and the barriers developing them. Chapter 8 ties together the explanatory theoretical constructs developed in Part I with the cases in Part II to explain and evaluate the NRS as a whole. In particular, Chapter 8 highlights that NRS authority is distinctly different from state-based authority. NRS authority inheres in the operational capacity of its participants. Conventional IR that argues any authority that is not rooted in state authority is in competition with the state. This work builds on this distinct character of NRS authority to argue that it is not a threat, but rather, a necessary complement.

Chapter 9 argues that the NRS's operational capability, credible knowledge assessment capabilities, and resulting adaptive capabilities are political capital that can be cultivated to develop durable, explicit assurances with external authorities, in particular state-based authorities. Barriers to explicit assurances draw lessons from the deconstruction and reconstruction of scientific knowledge in political environments, instances of international epistemic consensus, and characteristics of elusive, but effective, adaptation that has survived in conventional regulatory environments. These provide useful analytic constructs and non-Internet cases as comparators. Chapter 9 builds the case for complementarity further by highlighting instances of early cooperation in regional development, IPv6 deployment, and the close cooperation between some anti-abuse organizations and law enforcement.

Drawing on both the analytic arguments from Chapter 8 and early, empirical evidence of cooperation, the dissertation concludes with prescriptions for how NRS institutions can develop explicit assurances with state-authorities. If these two sets of institutions can avoid the pitfalls of previous efforts, in particular short-term usurpation of the others' authority, the global, non-discriminatory character of the Internet may be sustainable.

1.1 Common Resource Management

Resolution of the Pakistan-YouTube incident seems to have, once again, reaffirmed that the network operator community has the ability to sustain the integrity of the control plane. John Gilmore's famous quote, "[t]he Net interprets censorship as damage and routes around it" (Elmer-DeWitt & Jackson, 1993) remains accurate, but wants an explanation of the underlying mechanisms. Embracing the image of an organic, self-healing system is appealing and further romanticizes the cooperative ethos of infrastructure governance, but is ultimately misleading. The Pakistan-YouTube incident is more accurately one of a family of operational, security, and strategic externalities the NRS institutions act to remediate. Framing outcomes as externalities underlines the shared *and* conflicting incentives amongst the distributed stewards of the NRS. Some of the operational rules for resolving these externalities remain informal, or codified in private agreements. Others have been refined into formal strategies that are monitored and enforced by durable, function-specific CRIs.

While function-specific, NRS institutions share three common images of order:

1. resource management decisions are made by consensus;
2. routing system integrity is ensured by the accurate, efficient, and minimally disruptive dissemination of legitimate routing information;
3. end-to-end host communication requires consent, not merely access.

The latter two seem obvious to the engineer, but violations of those common images are the source of a diverse set of externalities as participants succumb to short-term, local benefits over their larger obligations to system integrity.[10] Mechanisms and industry incentives for implementing these image differ across function-specific institutions. Each CRI is presented in terms of how it has refined these images into a set of constitutional, collective choice, and operational rules for managing its component of the routing system. One difference from Ostrom's criteria for CPRs is that the NRS comprises decentralized, yet interdependent institutions.[11]

[10]This is the nice way to frame it. The Internet is also host to a diverse set of malicious, what are referred to by some security communities as "abusive," actors that will exploit externalities for individual gain. Depending on the jurisdiction, the underlying mechanisms may or may not be legal. The discussion of extractive and composite extractive abuse externalities in Chapter 7 illustrates these.

[11]This is in contrast to hierarchical nesting within the jurisdiction of an existing government agency or regulator. While there are layers and modes of delegation of authority, not all delegation

While effective within their respective domains, solutions to these problems *do* yield conflicting outcomes.

In the studies, NRS *mechanics* are described and explained using common resource concepts: *a*) number resources, addresses and routes, are resource units;[12] *b*) number resource *delegation* is a mode of *provisioning*; *c*) number resource *utilization* is a form of *appropriation d*) *types* of *appropriation* can either reinforce or diminish system integrity; *e*) collective-choice rules are used to evaluate operational rules that mediate provisioning and appropriation incentives amongst NRS participants; *f*) communities strive to preserve integrity through well-known operational rules (resource rights and obligations) broadly accepted by credible NRS participants; *g*) combinations of NRS institutions, in increasing cooperation with "external" state authorities, monitor and enforce bundles of rights and obligations; *h*) fragmented bundles create potentially conflicting rights; and *i*) knowledge assessment problems related to system function are endemic. Framing operational rules and externalities in terms of rights and obligations[13] highlights the characteristically essential interdependence amongst constituent network actors.[14]

The multiplicity of (resource) uses may be complementary or conflicting. Empirically observed studies of water resource systems[15] have similar properties and serve as well-understood points of comparison. Common resource systems faced with multiple use problems have developed monitoring and enforcement mechanisms to ensure effective provisioning and to protect particularistic classes of appropriators. This results in fragmented bundles of rights and obligations that align along function-specific provisioning and appropriation boundaries. These boundaries reflect the incentives structuring differentiated, potentially conflicting, uses.

The NRS framing builds on Ostrom's notion of a CPR,[16] but also relaxes key

follows a conventional principal agent model. Rather, Part III argues for a synthesis of a federated model in which authority relations more closely follow the notion of relational authority developed by Lake (2010).

[12]Number resources are information commodities whose value is derived from uses facilitated downstream. The community is wont to say, "Addresses are just numbers," which is one way to highlight they do not have value in and of themselves. That said, the routing information provisioned in the control plane is necessary to orchestrate flows in the data plane. Moreover, stewardship of number resources confers access to the Internet infrastructure, facilitating development of downstream goods and access to the broad set of public, private, and social goods for which the Internet infrastructure is an input.

[13]Throughout, E. Ostrom (1990) implicitly draws on notions of rights an obligations, focusing more on the function and operation of operational, collective choice, and constitutional rules in common pool resource systems. Later work, in particular Cole and Ostrom (2012b) frames rules in terms of rights and obligations to make interdependencies between actors that rely on rights regimes explicit. As per Cole and Ostrom (2012b, loc. 11885), a fundamental notion of rights and obligations is derived from (Hohfeld, 1917) in Cole and Ostrom's elaboration of what constitutes property rights.

[14]Interdependence does not imply harmony. Rather, the combination of interdependence and conflicting interests provide insight into potential foundations and barriers to cooperation.

[15]For instances and discussion see E. Ostrom (1990, pp. 69–87, 103–141), Epstein (2012), and Libecap (2012).

[16]The seminal reference is (E. Ostrom, 1990). This work also builds extensively on studies and essays in (Cole & Ostrom, 2012a) and (Hanna, Folke, & Mäler, 1996).

assumptions: foremost among these is resource unit rivalry.[17] Many CPR framings hinge on the notion of *subtractable* resource units. Number resources—addresses and autonomous system numbers (ASNs) used as unique identifiers—are partially-rival information commodities whose value is derived from downstream demand. Number rights *are* rival in the uses that guarantee the uniqueness necessary for Internet communication. Number resource provisioning and appropriation problems do not focus solely on the *rate* of use (appropriation) relative to some notion of replenishment. Rather, operational rules focus on *kind* of use and externalities' implications for routing system integrity. Within and across the NRS institutions, developing operational rules is an exercise in characterizing, identifying, and limiting negative externalities. This development process, in turn, requires a deep understanding of how the routing system itself functions and adapts—it finds the network operator community, later explained as an operational epistemic community, continuously confronting knowledge problems born of a changing operational, economic, and most recently, political, environment.

1.2 Operational Knowledge Assessment

Historical studies of CPRs—pastureland appropriation, forest management, irrigation, groundwater—are all narratives of CPR participants and governors that shared an intuitive, authoritative image of system integrity on which they based resource management decisions.[18] A shared common image contributes the foundation for more nuanced, context-specific rights regimes[19] that preserve resource system integrity, ultimately benefiting the group as a whole. In small, long-lived CPRs, the common image and nuance property systems developed over time through accumulated knowledge and experience with the system. While the Internet is substantively "bigger," a key commonality is that it grew out of a similarly "close-knit yet loosely organized" community of resource managers that shared operational experiences and developed the tacit knowledge supporting the three common images described earlier. This mode of knowledge sharing is evident in each of the CRI studies in Part II.

To sustain these common images at the operational level, a resource rights regime must also maintain a deep understanding of resource function. Take water resources for instance: what is the capacity of a given aquifer and what is the replenishment rate? What factors in the replenishment rate affect integrity (sus-

[17]Chapters 2 and 3 develop this argument extensively.

[18]Much of the CPR work focuses on notions of sustainability rooted in ensuring withdrawal rates of subtractable, a subtype of rivalrous, resource units do not exceed replenishment rates. As will be developed in Chapter 2, specifically Section 2.1, the NRS discussion focuses on the integrity of the system, in part because NRS resource units are nonsubtractable in many uses. Integrity is arguably a more general characteristic than sustainability tied to harms resulting from rate-based exhaustion. See discussions of rate versus kind in Sections 2.1 and 2.2 for elaboration.

[19]Here nuance and context-specificity should not be interpreted as idiosyncratic fixes. This is an important difference between systematic collective-choice processes and ad hoc solutions that characterize informal solutions.

tainability), and how? The NRS faces similar questions: what are the implications of changes to delegation (provisioning) policies? How will a numbers transfers market affect the integrity of the NRS, both in terms of operational efficacy and institutional obligations? Are there negative implications of making enforcement more durable?[20] In both systems, integrity is based on a common, authoritative image, but operational decisions are confounded by context-specific problems that shape incentive structures.

In the historical settings, government agents, provisioners, and appropriators recognize, and have often directly experienced, the implications of shortages and the efficacy, or lack thereof, of resource management practices. *Existing*[21] resource systems comprise nuanced resource rights configurations and shared understandings (a common image) of the implications of these provisioning and appropriation practices. These are products of participants' experience with the system. Operational rules are, in turn, the product of continuous knowledge assessment processes perceived as credible by both resource system participants *and* conventional (typically government) actors.

The second contribution of this work is a characterization of NRS consensus processes as collective-choice processes used to evaluate operational rules. The variants in consensus processes comprise a family of credible knowledge assessment processes. Studies of existing CPRs speak to "[u]ncertainties stemming from lack of knowledge [that] may be reduced over time as a result of skillful pooling and blending of scientific knowledge and local time-and-place knowledge," (1990, p. 33). Knowledge pooling is evident in the network operator communities, in particular the knowledge commons created by the NOGs (Chapter 4) and the anti-abuse community (Chapter 7). Operational knowledge is developed and cultivated through experience with system operation and the administration of management resources such as the regional number registries, interconnection platforms, and reputation-based blocking lists supporting anti-abuse efforts. Knowledge is disseminated through fora such as network operator groups and other umbrella organizations (IX associations, closed security groups) whose remit is information sharing rather than maintaining a particular management resource.

Cumulative operational knowledge is often leveraged, and more importantly, *updated*, through adversarial, yet constructive dialogue in these consensus processes.[22] Navigating these processes contributes to identifying sources of uncertainty, how these sources of uncertainty contribute to externalities, and the potential operational rules (policy experiments) proposed to remediate those externalities. Like

[20]This seems like a simple question, but it is an ongoing point of contention in the community, especially when considering the potential for authoritarian states to appropriate the more durable mechanisms. See the discussion of RPKI in Section 5.7.4 and later throughout Part III.

[21]Cole and Ostrom (2012b) stress that the lessons for effective resource (property) rights regimes are based on empirical studies rather than abstract game-theoretic constructs. The differences that make these problems challenging are rooted in resource specific incentive structures. The NRS is no exception, and, while benefiting from framing as a CPR, deviates from the mold in interesting and informative ways.

[22]The foundations of consensus processes are described in principle in Section 3.2.5. The chapters on RIRs, IXes, and anti-abuse describe the attendant consensus processes extensively.

other common resource systems, operational failures and deviant behavior by actors in the NRS *is* observable, but not from all vantage points. A more complete picture of the system requires coordination and cooperation that serves the common interest in routing system integrity, but does not fundamentally undermine private interests (i.e., a network's fundamental value proposition).

Consensus differs from voting processes associated with conventional policy development processes. First, consensus requires explicit support of an operational rule (policy) by a large proportion of participants, often greater than 75%. Across NRS studies, participants stress that a 51% to 49% vote *is not* consensus. Rather, simple voting as a rule making process is perceived to limit the full evaluative potential of knowledge exchange. Further, the contingent character of the NRS relies on a large portion of participants complying with operational rules. A "decision" in which 49% dissent, and absent a universal enforcement mechanism will likely deviate, is not sufficient for effective routing system operation. Consensus is not only a decision making mechanism, but also a means to assure credible commitment to operational rules.

The second difference builds on the spirit of operational knowledge exchange engendered by the consensus process and helps explain how credible commitment, even amongst dissenters, is fostered. When evaluating potential rules, the operational epistemic community demands justification for initial proposals *and* justification from those that reject proposals.[23] Consensus is an iterative process consonant with what Ostrom referred to as assessments of "local time-and-space" knowledge. Regardless of whether an actor is a minority constituency, or even a single individual, assuming that actor is offering technically and operationally sound critiques of the (operational) rule at hand, their voice will be heard and taken into consideration. Credible dissent is typically met with a dialogue exploring where a proposal fails and what compromise could be made to satisfy the dissenter's use case. That said, frivolous objections or those based in obviously particularistic or opportunistic motivations rather than interests in the function of the system are discarded. Further, they are often discarded publicly, none too delicately, and with reputational repercussions in the community. The community is also quick to highlight when they perceive the consensus process has been subverted, and with similar alacrity for dealing with frivolous objections.

Given the character of justification above, consensus processes are framed as knowledge assessment processes. The third difference from voting combines the representative character of the operator communities to argue that these are *credible* knowledge assessment processes. NRS institutions create operational rules that either directly or indirectly confer, revoke, sustain, enhance, and/or diminish number resource rights. NRS institutions comprise participants from diverse industry sub-sectors whose value propositions are affected by these changes. Industry subsectors are active in the collective-choice (consensus) processes within each institu-

[23] A long-time member of the IETF highlighted that this is not an explicitly documented characteristic of the IETF consensus process but it is a norm and those that provide justification have a greater influence on the decision making process.

tion, keeping one another in check through processes akin to what Ostrom refers to as "mutual prescription."[24] In the vernacular of NRS institutions, mutual prescription is most akin to neutrality norms: firms administering management resources that affect number rights (and other resource endowments) are expressly forbidden from preferencing particular actors, constituencies, or industry sub-sectors. Credibility is a function of both diverse participation in consensus processes and the capability to affirm rules are being applied fairly via mutual monitoring of both resource use and consensus processes.

Consensus processes also adapt operational rules to current economic and operational realities. This does not imply that uncertainty is completely reduced. Rather, uncertainty "remains even after one acquires considerable knowledge about the resource system itself," (E. Ostrom, 1990, p. 33). Operational epistemic communities are familiar with this type of uncertainty and take it as a given, a recurring, unavoidable element of the operations landscape. Such an approach that embraces continuous change in the operational landscape confounds the conventional image of epistemic knowledge as illuminating the path from one stable (static) regulatory arrangement to the next. Rather, operational epistemic knowledge in the NRS is the basis for adapting a contingent social order apace with Internet growth—continuous, dynamic adaptation to changes in a highly mutable resource system is necessary to sustain routing system integrity. NRS policy entrepreneurs must cope with skepticism born of what appears to the lay observer as severe operational uncertainty as well as conventional interest politics.

Problem solving in the NRS is conceptually and structurally similar to CPRs, but evaluating and conveying the public policy implications of NRS management falls prey to many of the knowledge assessment problems faced by other science and technology issues.[25] Unlike common images developed in historical CPRs, the "integrity of the control plane" is not as intuitive a construct as the implications of subtractive resource unit depletion in fisheries, aquifers, or pastureland management. Communicating NRS dynamics, even in the simplest terms, must translate a foreign mode of consensus building *and* a complex technological construct to external actors that are unfamiliar with, and likely skeptical, of both. Merely transmitting a coarse image of how the NRS works is often considered a victory.

1.3 Making Explicit Assurances

The NRS is facing a number of *credibility dilemmas*: *1*) amongst participants weighing the implications of operational rules *within* the consensus processes of particular institutions; *2*) amongst potentially conflicting operational rules *across* NRS institutions; and *3*) *conveying* the credibility, legitimacy, and durability of solutions to agents tasked with protecting the broader public interest. In contrast to other

[24] See Chapter 4 of E. Ostrom (1990, pp. 103–141) for discussion of CPR participant dynamics, in particular mutual prescription in the process of crafting water rights in California.

[25] See McCray (2003); McCray and Oye (2006); McCray, Oye, and Petersen (2010) for a discussion of risk regulation and adaptation.

global infrastructures, the modern Internet was not born of extensive state development. Nor does the state have substantive regulatory experience on which to base a workable image of NRS function, much less the nuanced function managed by the NRS. The NRS, along with conventional Internet governance models, have historically expected the state to learn a new image of infrastructure mechanics and management. Both the state and the NRS must reconcile their respective roles in minimizing costly failures. The discussion in Part III begins to untangle how these complementary authorities can develop explicit assurances that the NRS's common interest will remain aligned with the public interest.

The NRS has thus far thrived as a "private regime complex"[26] built on a complex web of authority amongst operational epistemic communities animating the CRIs. Historically, private orders are a source of regulatory innovation, but are ultimately absorbed into the state system.[27] Pakistan-YouTube is a compelling narrative, highlighting ad hoc enforcement of well-known rights and obligations. As an assurance of stability, this may be sufficient for those that understand the nuance of informal NRS management, but it is not a credible signal of stability, especially to state actors and those that depend on downstream products and services. Further, NRS participants themselves demand more accountable regulatory institutions.

The CRIs supply management resources for monitoring and enforcing function-specific bundles of rights and obligations that NRS participants, and consumers of downstream products and services, increasingly demand. In terms of the credibility dilemmas above, consensus processes are stable in dilemma *(a)*, exhibit tensions at the operational level in *(b)*, and are working to develop the diplomatic capability to address *(c)*. Individually these institutions have demonstrated credible knowledge assessment capabilities and the ability sustain adaptive collective choice rules informed by partially overlapping operational epistemic communities. NRS institutions now face cooperation dilemmas on the edges, requiring the adaption of constitutional rules to match not only operational demands, but the need to develop the political capital necessary to engage in the global political arena.

Problems facing the NRS are technically challenging, but are not challenges to their operational capability—that capability is not in dispute here. It is unrealistic to believe there is a tractable critical path to a state-based substitute for the operational capability intrinsic in the NRS anywhere on the horizon. Security, engagement with law enforcement, and the stability of critical infrastructure are in the penumbra of technical capability and public policy. The stability and credibility of the resource rights regime that underpins these issues depends on developing *diplomatic* capability in the NRS. Given the dependence of an increasingly broad set of public, private, and social goods for which Internet connectivity constitutes a key input, conventional state actors and the NRS must identify means to cooperate.

As the de facto steward of a critical common resource, the NRS has an obli-

[26]This is a synthesis of a private regime, as per Cutler, Haufler, and Porter (1999); Cutler (2003) and the notion of a regime complex, as per R. Keohane and Victor (2010). In contrast to conventional regime theory, in which the state is the source of authority, private regimes, and by proxy a private regime complex, derives its authority from private actors that have chosen to cooperate.

[27]In general see Cutler et al. (1999) and Cutler (2003). For a specific instance, see Porter (1999).

gation to collaborate with conventional governance actors charged with protecting their constituencies and the public good. The burden of cooperation does not fall exclusively on the NRS. In its charge to protect the public good, conventional governance actors are obligated to protect the integrity of resources valuable to the economy and the society writ large.[28] Simply put, like the preservation of other resource systems, regulators are obligated to not break the institutions that sustain Internet operations. The third contribution of this dissertation frames the strengths and deficiencies of the current arrangements in terms of how the NRS and external authorities can develop cooperative arrangements that align the NRS's common interests with the public interest.

Part III explores these issues and offers prescriptions based on how the substantively different modes of authority engendered by the NRS and the state can be, in contrast to conventional IR notions of authority in the global political arena, quite complementary. As described in the previous sections and elaborated in the studies presented in Part II, NRS institutions have developed substantive, unique capabilities for adapting common resource rules to new technical, operational, and economic realities. Aligning their common interests with the public interest is a different mode of adaptive capability the NRS must develop. Within the NRS, CRI leadership and the resulting resource management organizations have accrued and institutionalized the social capital necessary to sustain contentious knowledge assessment processes necessary to adapt apace with Internet growth while also sustaining routing integrity. Social capital is largely endogenous to the NRS, rooted in the norms of the operational epistemic communities in which it accrues. Aligning the NRS's common interests with the public interest will require the NRS develop strategies for effectively leveraging its unique operational, credible knowledge assessment, and adaptive capabilities as political capital when engaging with external authorities.

A conventional analysis would see this as a simple principal-agent problem. State-based authority is preeminent, be it domestic or international. The NRS has capabilities the state needs, why not simply add the NRS to an existing regulatory portfolio? Recall the basic differences between consensus based decision making and majoritarian voting; as it turns out, the differences in the kind of authority that sustains the NRS as a contingent social order, engendering the capabilities sought by the state, are fundamentally different from what Flathman refers to as the formal-legalistic mode of authority wielded by states. Adding NRS stewardship to a state's portfolio of domestic regulatory interests would expose currently stable

[28]Young (1996, loc. 4603–4608) discusses justifying govern*ment* modes of governance, turning the tables on the common assumption that government is the de facto legitimate option. Of course, this is magnificently confounded by trade-offs between what regulators perceive to be constructive and destructive uses amongst the multiplicity of uses of a given resource. This work assumes the constructive uses outweigh the destructive. Ensuring the integrity of the NRS is one vector to ensure the constructive outweigh the destructive, but ultimately these perceptions are driven by more familiar issue outcomes that play out *on* the Internet rather than the infrastructure issues. Conventional transnational issues that play out on the Internet are much more intuitive images of order than abstractions such as the control plane.

consensus-based resource management processes to powerful short-term interests that would inevitably weaken, if not eliminate, extant credible knowledge assessment and adaptive capabilities. In effect, aggressive predatory rule would likely eliminate precisely the characteristics that make the NRS a valuable steward of a high clockspeed infrastructure.

This is a rather strong assertion; Part III returns to the analytic mode, drawing on evidence from studies in Part II, to back this assertion and offer prescriptions for a way forward. Part I provided theoretically grounded analytic constructs–Ostrom's common resource frameworks, notions of operational epistemic communities, and resource rights—to develop a rich conceptual vernacular for describing, explaining, and perhaps most importantly, comparing and contrasting, the institutions that manage the routing system. The studies in Part II leverages these analytic constructs to explain empirical observations and archival data, explicating common resource function and the role of consensus-based decision making. Chapter 8 returns to the analytic mode to tie these together into an explanation of: *a*) how authority is created in the NRS, *b*) the strengths and weaknesses NRS authority, *c*) how this authority differs from that of the state, and *d*) outlines the characteristics of observed modes of engagement between NRS authorities and external authorities. In particular, differences in authority provide further explanation of the differences between consensus and majoritarian voting. This analysis also further develops the contingent character of the NRS social order as dependent on credibly committed participants. Coercion by an external authority does not engender the same commitment and nor does it have the operational capacity to actually enforce its assertions global. That said, CRIs are organizations in particular jurisdictions that cane be coerced. As stewards of both resources and decision making fora, if the assurances engendered by the NRS social order becomes such that participants no longer garner value sufficient value relative to the cost of participation, [29] that order will dissolve, and with it the operational assurances and adaptive capabilities valued by external authorities.[30]

The story of coercion and the disintegration of NRS social order is not inevitable—this scenario is not a prediction of an ephemeral mode of adaptive management doomed to failure. Chapter 9 builds on the modes of engagement described at the end of Chapter 8 to highlight early instances of successful engagement between NRS and state authorities and the potential barriers to making these ad hoc instances of cooperation more durable. Prescriptions offered in Chapter 9 are rooted in how to leverage NRS capabilities as political capital to more explicitly align NRS common interests and public interests without impinging on state authority. Building on contemporary instances of cooperation and challenges to cooperation in the NRS, prescriptions offer a mapping of NRS capabilities to phases of regulatory development, highlighting how existing ad hoc instances have faired and offering

[29]This does not have to occur due to external coercion, it can happen as a completely endogenous institutional failure

[30]This is elaborated in Section 8.1 in terms of the *contingent* character of the NRS as a social order, rooted in notions of relational authority. For elaborations of relational authority in general, see (Lake, 2009, 2010).

prescriptions to avoid destructive contention between NRS and state authorities.

Part I

Common Foundations

Chapter 2

Routing Mechanics

THE CONTROL PLANE is the commonly managed resource for coordinating routing information in the Internet. Within the control plane, routing information is *jointly* produced and individually consumed by a set of decentralized networks in the process of managing Internet connectivity. Absent centralized regulation, control plane participants have developed operational rules and institutions to *a)* share information about the state of the control plane, *b)* identify harmful (negative) externalities, *c)* monitor for these externalities, and *d)* leverage selective incentives to remediate negative externalities. Explaining these processes as exchanges and externalities is supplemented by framing well-known routing processes in terms of common resource production, provision, and appropriation. This combination of frameworks highlights the fundamental interdependence that motivates cooperative efforts sustaining the integrity of the routing system. Section 2.1 transposes common resource concepts onto the basic elements of number resource utilization, route bridge the gap between purely technical routing mechanics and the NRS's institutional function of coordinating those mechanics in a way that sustains routing system integrity.

Using the NRS vernacular developed in Section 2.1, externalities are first explained purely in terms of routing mechanics (Section 2.2). Contractual relations that bound route exchange are then introduced as economic constraints. Externalities and their remediation are *facilitated* by the flexibility inherent in the mechanics of the BGP routing protocol. Section 2.2 explains *a)* operational rules shaping how routes are produced and consumed; *b)* the operational and security externalities that emerge in practice; *c)* operational rules for remediating these externalities; and *d)* the rights and obligations tacit in these rules. Strategic externalities are *motivated* by interconnection economics, discussed in Section 2.3.

A CPR framing of routing mechanics highlights the essential interdependence that facilitate externalities *and* remediation strategies. A focus on routing mechanics precisely scopes *a)* externalities' effects on control plane integrity and *b)* the actors contributing to those effects. Participants' contributions to externalities and remediation strategies provide the initial frame of reference for *a)* rights and obligations amongst NRS participants and *b)* NRS institutions that make those rights and obligations durable. Section 2.4 concludes with a transition to Chapter 3's ex-

planation of the NRS institutional complex as a resource (property) rights regime created, maintained, and sustained by operational epistemic communities. .

2.1 Common Resource Foundations

To access[31] the control plane, a network actor A requires: *a*) a set of addresses, a prefix, that uniquely identifies hosts A wants to connect to the Internet; *b*) an autonomous system number, an ASN, that uniquely identifies an administrative domain of hosts following a single routing policy; *c*) physical connectivity to at least one other network (autonomous system) connected to the global Internet that is willing to use *its* connections to forward traffic originating in A to other networks and traffic destined for A to A's border. The first two resources—prefixes and ASNs—are referred to as number resources.[32] In this common resource framing of the routing system, number resources are the basic resource units.

Consider Jon Postel's famous quote:

> A name indicates what we seek. An address indicates where it is. A route indicates how we get there. (Postel, 1981, p. 7)

In the NRS, routes are a composite resource unit, comprising a prefix and a sequence of ASNs. Following Postel, that sequence of ASNs (the AS-PATH) is the ordered sequence of networks that "indicate how we get there." Each network in that sequence contributes to forwarding traffic from its origin (a host uniquely identified by an address) to its destination (another host uniquely identified by an address). The first ASN is adjacent to the source network, each subsequent network forwards traffic to the next and, in most cases, the last ASN comprises the destination host. Each adjacency along routes are produced and exchanged in the control plane. Physical connectivity is the means of transporting data, in this case routes, between networks that are "directly connected." For the moment, physical connectivity will be assumed, but revisited in Section 2.3 and is a key element of Chapter 6's discussion of interconnection markets.

Consider Ostrom's general definition of a CPR:

> The term "common-pool resource" refers to a natural or man-made resource system that is sufficiently large as to make it costly (but not impossible) to exclude potential beneficiaries from obtaining benefits from is use. To understand the processes of organizing and governing CPRs, it is essential to distinguish between the resource system and the flow of resource units produced by the system, while still recognizing the dependence of the one on the other. Resource systems are best thought of

[31] Rights of entry, also referred to as access, are a distinct class of property rights mediated by NRS institutions. See Chapter 3 in general and Section 3.4.1 in particular.

[32] This language is community vernacular. For simplicity, the immediate discussion in this chapter will simplify by considering only a single prefix and one corresponding ASN. More commonly, networks are stewards of a non-contiguous set of prefixes and in some cases multiple ASNs to distinguish regional or business differences in routing policy.

> as stock variables that are capable, under favorable conditions, of producing a maximum quantity of a flow variable *without harming the stock or the resource system itself.* (1990, p. 30, emphasis added here)

In much of the CPR literature, the stock of a natural resource comprises *subtractable* resource units some collective would like to consume, or appropriate particular rights of use for. The emphasis above is the conceptual foundation of resource system integrity: resource appropriation "without harming... the resource system itself," and, as will be developed in this chapter, especially as it relates to the NRS, is not simply balancing replenishment and utilization rates of subtractable resource stocks. Resource systems comprise man-made constructs, infrastructure(s), *and* institutions, that mediate consumption of resources to ensure system integrity. NRS resource units—number resources and routes—and the and the routings system in which they are used, are all man-made. NRS institutions comprise knowledge commons and facilities that contribute to participant's efforts at ensuring the integrity of the routing system: traffic is routed to the intended destination, facilities necessary for tracing and remediating routing externalities are available, and these processes are made durable in stable institutions. The next section describes the fundamentals of how these resources are produced and provisioned.

2.1.1 Modes of Resource Provisioning

Number resources and routes are certainly bound together in system function, but their respective provisioning processes differ substantively. Consider Ostrom's distinction between providers and producers:

> The term I use for those who arrange for the provision of a CPR is "providers." I use the term "producer" to refer to anyone who actually constructs, repairs, or takes action that ensure the long-term sustenance of the resource system itself. Frequently providers and producers are the same individuals, but they do not have to be.[33] A national government may provide an irrigation system in the sense of arranging for its financing and design. It may then arrange with local farmers to produce and maintain it. If local farmers are given the authority to arrange for maintenance, then they become both the providers and the producers of maintenance activities related to a CPR. (E. Ostrom, 1990, p. 31)

Number resources were provisioned in the Internet Protocol (IP) specification, RFC 791 (Postel, 1981) discussed shortly. Routes are produced by participants in the course of routing system operation.

Consider IP address, or simply address, provisioning. IP version 4 (IPv4) addresses are defined as having "a fixed length of four octets (32 bits)," in RFC 791 (Postel, 1981, p. 6). This definition of address length provisioned a pool of IPv4 addresses comprising 2^{32} addresses. The IPv4 pool is alternately referred to as the IPv4 address pool or address space.

[33] E. Ostrom (1990) attributes the last sentence to V. Ostrom, Tiebout, and Warren (1961).

In the case of IPv4, RFC 791 was created under the auspices of a US government contract. Some argue that although the IPv4 address space has historically functioned as a common pool supplying demand from global actors, ownership can be traced back to the US government by virtue of the DARPA contract. Similarly, a pool of 2^{16} ASNs was initially provisioned, later extended to 2^{32}. In both cases, the two finite pools of basic resource units were provisioned in a protocol definition process. Using Ostrom's terminology, protocol designers are the providers of number resources.

Providing number resources in the early IETF protocol definition process above and the operational process of number resource delegation and appropriation are processes, animated by separate institutions. *Delegation* from the existing pool is an operational process made durable by the RIR system. Such delegation is framed as a *provider* mediating appropriation in a way that 1. ensures the unique, accurate delegation of number resource rights and 2. has historically served a conservation function.More broadly, NRS institutions such as the number registry (RIRs), IXes, reputation aggregators in the anti-abuse community are *producers* of operational facilities that contribute to "ensur[ing] the long-term sustenance of the resource system itself," (E. Ostrom, 1990, p. 31).[34]

In contrast to the protocol-based provisioning of number resources, consider operations-based provisioning of routes. The specification of the *structure* of a route and *how* it should be provisioned is defined by the BGP protocol.[35] Provisioning the stock from which routes used to direct Internet traffic are selected is a continuous process in the control plane. The complete *stock* of routes sufficient to direct Internet traffic from any source to any destination is dynamic. The stock of routes is produced through mutual provisioning and appropriation practices of control plane participants that: *a*) confirm routes are still valid; *b*) advertise newly provisioned routes; *c*) appropriate new routes as they become available; *d*) remove routes that are no longer available (valid). The subset of this stock used for actively routing traffic is typically referred to as the *global routing table*.[36]

Tacit in mechanisms sustaining the global stock of routes is a set of common, well-known operational rules. Route provisioning is a local process, based on local information. Not all routes in the global stock are available to all networks—and not all the routes available to a network are used to route traffic. A network's *local* stock of routes only comprises those routes shared by immediate neighbors with whom it has physical connectivity. Each network selects amongst available local routes to select those it will use to actively direct traffic based on: *a*) the BGP protocol, *b*) operational characteristics (Section 2.2), *c*) economic factors (costs, Section 2.3), and *d*) more sophisticated interconnection bundling strategies that

[34]Characteristics of NRS institutions as property (number resource) rights managers are described more fully in Chapter 3. Studies of these institutions are presented in Part II.

[35]In general, see RFC 4271 (Rekhter, Li, & Hares, 2006).

[36]Section 2.2 will elaborate this to highlight that because selection of routing announcements is local, it is extremely difficult to know precisely all the routes currently in the global routing table at any given time. This characteristic will be revisited quite frequently in terms of the uncertainty endemic in resource system management.

benefit that network's specific value proposition. In the aggregate, the union of these locally selected routes comprise the global routing table.

Intrinsic in operational rules for provisioning routes are a set of rights and obligations amongst control plane participants. Operational rules strive to ensure global connectivity, Ostrom's "maximum quantity of a flow variable" while limiting externalities that "harm" the control plane as a resource system. In effect, they sustain the integrity of routing information. Corrupt or inconsistent routing information, deviation from accepted operational rules, creates costs and damages, including lowering the efficiency of end-to-end delivery, wholesale redirection of traffic to unintended (and often malicious) destinations, and diminished connectivity to some or all Internet destinations.[37] Some of these operational rules are institutionalized, one of the most significant for this work is the *exclusive* right to originate a prefix, i.e., to claim it is the authoritative steward of those numbers and the hosts they enumerate. Others are informal rules of thumb promulgated in network operator groups, often referred to as Best Common Operation Practices (BCOPs) or Best Common Practices (BCPs).

2.1.2 Joint Use and Externalities

Distinguishing joint use (appropriation) of resource units and joint use of the resource system is key to Ostrom's definition of a CPR. Ostrom focuses on subtractable (rival) resource units in terms of joint use and appropriation:

> A resource system can be jointly provided and/or produced by more than one person or firm. The actual process of appropriating resource units from the CPR can be undertaken by multiple appropriators simultaneously or sequentially. The resource units, however, *are not subject to joint use or appropriation.* The fish harvested by one boat are not there for someone else. The water spread on one farmer's fields cannot be spread onto someone else's fields. Thus, the resource units are not jointly used, but the resource system is subject to joint use. (E. Ostrom, 1990, p. 30, emphasis in the original)

In conventional CPRs, the canonical source of harm is over-exploitation: left unchecked the appropriation rate will exceed the replenishment rate. The resource is eventually depleted and none of the appropriators garner value subsequent value (E. Ostrom, 1990, pp. 30–31). In some cases, falling below a particular threshold not only threatens immediate appropriation, but may threaten the replenishment mechanisms themselves. In other words, over-exploitation my threaten the integrity of the system itself. For instance, over-fishing past the point of population sustainability may cause the complete collapse of a fishery rather than just a temporary shortage. Consider a more nuanced instance. In Ostrom's studies of freshwater basins in California, over-exploitation in certain areas can also lead to damage that

[37] For instance, see the discussions of route-flap, prefix hijacking, and man-in-the-middle attacks in Section 2.2. For the latter, diminished connectivity, see the discussion of reputation in Chapter 7.

results in irreversible saltwater entery, tainting some or all of the broader supply. While both cases are damages to a factor of the replenishment function, the latter is much less reversible.[38]

Number resources and routes are, under ideal conditions, non-rival information commodities. Both are information goods. Nominally, use of a non-rival good by one actor does not diminish the value of the same use for other actors. An IP address, really just an integer, after all, does not have value in and of itself. In contrast, routes do have value relative to one another based on the value proposition of the actor appropriating that route. One route may be shorter, have lower latency, have lower congestion, than another, thus making it more valuable in some applications, like video streaming, but negligible in others, such as casual web-browsing or infrequent file-system backups.[39] That said, this relative value based on derived demand does not alter the non-rival character of routes as information goods.[40] Any actor can use number resources and routes in their routing decisions. Number resources and routes are commodities in the sense that their value is derived from the use to which they are put.[41] Externalities occur when number resource and route uses conflict, where one actor's use diminishes the value of others. In this sense, number resources and routes are partially-rival *depending on* use.[42] Here again, operational rules create rights and obligations to avoid conflicting use.

In the NRS, externalities can arise from *rate* and/or *kind* of appropriation and provisioning. These externalities create costs for actors that have no operational control over the portion of the control plane where that harm originated. Coase defines an externality "as the effect of one person's decision on someone who is not part to that decision," (R. Coase, 1988, p. 23). The next section illustrates the mechanisms by which routes, and externalities, propagate in the control plane. Externalities may be a net positive or net negative—that threshold will be addressed in the following sections. Given a bit more of the operational mechanics of routing in Section 2.2, Section 2.2.2.1 describes the mechanics and remediation of route flap, a well-known externality in the routing system that is a product of rate and kind of appropriation and provisioning in the control plane.

[38]As will be developed in later discussion, the reversibility of an externality is critical to understanding the scope of NRS authority.

[39]These are coarse-grain instances. One can certainly imagine counter-examples, for instance web-applications demand low-latency to ensure responsiveness.

[40]This does not mean the data plane *that carries traffic based on* these routes is non-rival. The data plane is partially rival at the point of congestion. As will be developed later in this work, a diverse set of routes improves actors' ability to select routes best suited for their value proposition, including dynamically routing around congestion.

[41]Number resource and routes can be compared to other commodities with derived demand. Very few individuals buy a chunk of aluminum for non-commercial use. Rather, aluminum is an input into a variety of downstream products, for instance, Apple laptop bodies. The notion of downstream use will be revisited the discussion of the Internet as an infrastructure in Section 3.1.

[42]See Section 3.3 for a discussion of multifunctional resources.

2.1.3 A Common Image of Control Plane Integrity

Negative externalities diminish the integrity of the control plane, and by proxy, create potentially avoidable costs in network operations. The notion of *integrity* used here builds on Ostrom's definition above, referring to provisioning and appropriation practices that do not "harm the stock or the resource system itself," (1990, p. 30). *Control plane* integrity is rooted in *a*) the rights and obligations shaping how routes are provisioned and *b*) the implications of, the externalities engendered in, particular provisioning practices. The first element of control plane integrity is that routes promulgated are legitimate, consensual paths *between ASes*. Paths are legitimate in the sense that adjacencies along the path accurately represent consensual bilateral agreements between those ASes to exchange traffic.[43] Consent to exchange traffic originating from and destined for a particular prefix starts at the origin and is reaffirmed at each subsequent adjacency; this is what the community refers to as the transitive trust model. The second element speaks to negative externalities rooted in the rate and kind of provisioning. Operational rules in the NRS strive to identify, monitor, and mitigate negative externalities.

Recall the distinction between the resource and the resource system—in particular the characterization of the resource system as both infrastructure *and* institution(s). In the early days of the Internet, operational rules were set by a small, "close-knit yet loosely organized"[44] community of academics that often played the roles of network researchers, protocol designers, *and* network operators. In those days, the distinction between providers and producers was much less distinct. Chief amongst these was the late Jon Postel. Postel as the Internet Assigned Numbers Authority (IANA) bridged the role of provider and operational facilities provision by delegating number resources in the small community of trusted academic networks. The modern Internet has grown from this small community of actors to a network of tens of thousands of autonomous systems that, from the perspective of the end user, seamlessly interconnects millions of devices spread across the globe.

As the Internet grew, network operators developed a guild-like community for sharing operational experience. Informal mechanisms coalesced into more formal resource management facilities administered by CRIs. CRI policies are informed by experienced network operators continuously engaging in and negotiating network interconnection in the control plane. Subsets of function-specific operational rules—framed as rights and obligations—have been made durable in the modern IANA and the CRIs discussed in Part II: the regional Internet registry (RIR) system (Chapter 5), Internet eXchanges (IXes) (Chapter 6), and anti-abuse organizations (Chapter 7).

Analyzing provisioning in terms of providers and producers provides clear con-

[43] An important distinction to keep in mind here is that consent is between networks, not the end hosts. The common image of abuse developed in Chapter 7 focuses on consent between end hosts, in particular the users behind those hosts.

[44] This phrase alludes to Ellickson (1991), describing the community of cattle ranchers in Shasta County California. These ranchers preferred to manage common pastureland (a canonical natural common pool resource) through informal processes (informal social institutions) rather than incur that substantively higher transaction costs of engaging existing legal-mechanisms.

ceptual distinctions, but the roles are not always clearly delineated in practice. With respect to maintaining control plane integrity, NRS institutions are producers of *resource management facilities*, such as the number registry, interconnection platforms, and source of number reputation information (reputation aggregators). These facilities serve as infrastructure akin to the irrigation systems presented by Ostrom—they are constructs administered by a distinguished set of system participants to enhance access to resources, and, subsequently, the value garnered by system participation. In the case of irrigation systems, facilities alter the resource topology in a way that supports allocating water resources according to operational rules jointly established by participant appropriators. The irrigation system enhances access and withdrawal rights while modulating these avoid over-exploitation.

CRIs provide function-specific facilities for enhancing access, appropriation, and provisioning of numbers and routes within the NRS. They are the elements of the control plane's "irrigation system," enhancing the flow of legitimate prefix advertisements in the global routing system while filtering out illegitimate advertisements. In the sense of affecting rights and obligations across the control plane, NRS institutions are *producers*. In their role as designers and administrators of particular facilities, individual organizations hewing to the norms of a particular CRI are both *providers and producers*.

A simple interpretation of function-specific facilities is that the close-knit community network operators did what any good engineer is taught to do: divide and conquer, develop modularized solutions with well-developed interfaces. As a results, institutionalized subsets of formerly informal operational rules made more durable; Part II provides the empirical evidence of these processes.

2.2 Control Plane Mechanics

D. D. Clark, Partridge, Ramming, and Wroclawski (2003) succinctly describe the relationship between the data plane and the control plane:

> Most discussions of network architecture recognize two architectural divisions, or planes: a data plane, over which content is forwarded, and a control or management plane, which is used to direct, measure, and repair the data plane. (2003, p. 4)

The Forwarding and Control Element Separation (ForCES) Framework[45] delineates the elements of the control plane and the data plane. The Border Gateway Protocol (BGP) provides the protocol substrate, the NRS comprises the knowledge commons and decision making processes for shaping operational rules. As implied by the epigraph of this chapter, not all actors have access to the same information. An actor's

[45]The relevant RFCs are (Yang, Dantu, Anderson, & Gopal, 2004; Doria et al., 2010; Halpern & Salim, 2010; R. Haas, 2010; Crouch, Khosravi, Doria, Wang, & Ogawa, 2010; Haleplidis, Ogawa, Wang, & Salim, 2010; Haleplidis, Koufopavlou, & Denazis, 2011). This work uses these as reference for precise language on elements of the control plane and it is distinguished from the data, or as referred to in these RFCs, the forwarding, plane.

position within the topology of the control plane, deployment of physical infrastructure supporting access to the data plane, and contracting modes between adjacent actors provisioning routes all affect the diversity of route information available to a particular actor.

This section elaborates the technical elements of routing mechanics in common resource vernacular, namely how routes are appropriated and provisioned within bilateral peering sessions. In terms of the quote above, these are the mechanics of how NRS participants direct and repair[46] the data plane. This section describes and explains the mechanics of *a*) how routes are propagated, how they "flow" through the control plane; *b*) how networks' local stocks of routes are sustained and "replenished;" *c*) how these practices affect the global stock of routes, the global routing table; and *d*) how externalities are remediated. R. Coase (1988) elaborates the notion of an externality:

> Thus, if A buys something from B, A's decision to buy affects B, but this effect is not considered to be an "externality." However, if A's transaction with B affects C, D, and E, who are not parties to the transaction, because, for example, it results in noise or smoke which impinge on C, D, and E, the effects on C, D, and E are termed externalities. (R. Coase, 1988, p. 23)

Routing externalities experienced by a network may by topologically local, but likely network "A" is topologically distant from newtorks "C," "D," and "E".

It is important to stress this section is just about the mechanics as a means to elaborate precisely how both route provisioning and externalities occur without clouding the discussion with economic constraints. The economics of provisioning routes and physical connectivity are discussed in the next section (2.3).

2.2.1 Route Exchange in a Simple Internet

Actors exchange routes, here framed as a form of provisioning and appropriation, in the control plane via the Border Gateway Protocol, BGP.[47] The basic operational rules of route provisioning described here are rooted in BGP protocol dynamics and experience. Figure 2-1 is a "simple multi-provider Internet,"[48] referred to as SimpleNet. Figure 2-1 depicts a set of network actors, identified by their ASNs and delegated prefix, interconnected in a simple "Internet" via direct physical links and exchanging routes via BGP. Here an interconnection relation simply means two autonomous systems have agreed to exchange routes.[49]

[46]Here repair means select alternate routes, not physical, such as repairing an undersea transit cable.

[47]This section provides sufficient detail to illustrate appropriation, provisioning, and externalities. See RFC 4271 (Rekhter et al., 2006) for protocol details.

[48]This phrase is borrowed from Figure 4.4 of (Peterson & Davie, 2011).

[49]The notion of an interconnection relation as a BPG session will be developed shortly; in Section 2.3 it will be further refined to reflect the scope of a contract over multiple BGP sessions between the same pair of ASes.

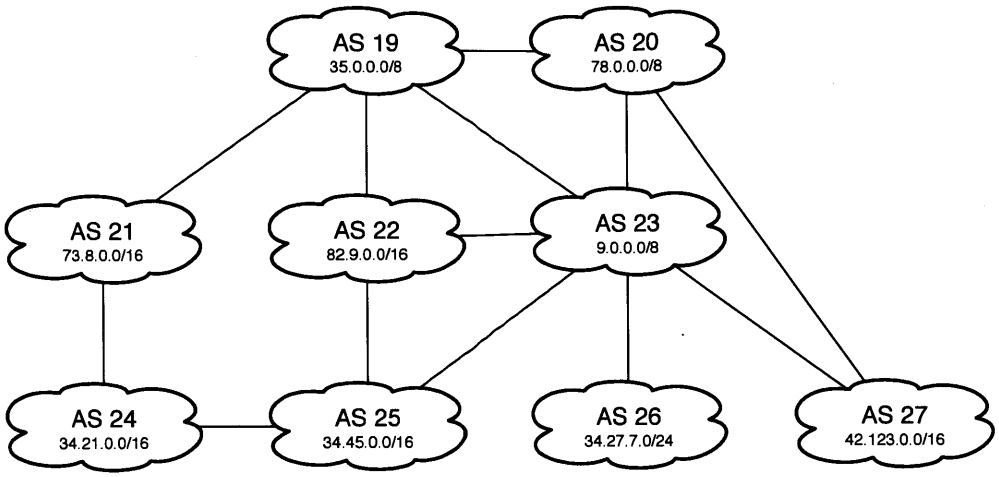

Figure 2-1: Example of a simple network, referred to as SimpleNet. Black lines are physical connections between networks. Arrows and the accompanying annotations represent route advertisements from one AS (network actor) to another. Based on the illustrative scenarios presented here, blue arrows are legitimate advertisements. Red arrows are illegitimate advertisements.

A *route*, or a prefix advertisement,[50] comprises a prefix and an AS-PATH. The AS-PATH is the sequence of ASNs a prefix has traversed from the AS originating that prefix advertisement to the AS currently receiving (appropriating) that advertisement. In common resource terms, the AS-PATH is the sequence of ASes that have contributed to producing the route at hand. Uniqueness is key to Internet addressing and rights to originate prefixes are key to ensuring uniqueness. The origin AS is the only AS that has the legitimate right to terminate traffic addressed to hosts in the prefixes it has stewardship of.[51] ASes are delegated exclusive rights to originate prefixes by RIRs; see further discussion of delegation rights in Chapters 3 and 5. Origination, as well as other route advertisements, may not always be legitimate.

To illustrate, consider AS 26 in Figure 2-1. AS 26 is (only) connected to the

[50] As per RFC 4271:

> For the purpose of this protocol, a route is defined as a unit of information that pairs a set of destinations with the attributes of a path to those destinations. (Rekhter et al., 2006, p. 9)

The set of destinations is a prefix, a set of addresses. The "path to those destinations" is a sequence of ASNs indicating the path that set of destinations traversed. Note that the definition states it is a unit of information; here it is refined into a composite resource unit with distinct characteristics salient to managing the integrity of the control plane.

[51] In reference to delivery on the internal network, routes described here end in ...X I where X is the origin AS and I indicates the remainder of the path to hosts addressed by the prefix is internal to X. More specific still, it indicates the boundary between (e)BGP, and external routing protocol and iBGP, the originating AS's *internal* routing protocol. This is not a simply an idiosyncrasy: the I delineates the point at which private resources are being appropriated, distinguishing between a route providing access versus the sanction of unfettered appropriation. This is especially important in the discussion of anti-abuse strategies remediating abuse externalities.

rest of SimpleNet via AS 23. The blue arrow represents an advertisement of prefix 32.27.7.0/24 from AS 26 to AS 23. AS 26 provisioned the prefix advertisement 32.27.7.0/24: 26 I for appropriation by AS 23. In this example AS 26 originated 32.27.7.0/24. This prefix is a relatively small prefix.[52] A BGP session is the channel[53] for coordinating the exchange of routes between two ASes. A pair of ASes with an established BGP session are referred to as peers.[54] The BGP session is an ongoing TCP connection between AS 26 and AS 23 over which routes are shared, updated, and removed.

In this very simple example, AS 23 does not have any other options to reach AS 26 and thus adds 32.27.7.0/24 26 I to its routing table, the local stock of routes it actively uses to direct traffic. In terms of the BGP protocol, this route has been selected for use to route traffic. Not only is it in AS 26's local routing table, it is now in the global routing table. Any traffic AS 23 receives destined for 32.27.7.0/24 or any traffic created in AS 23 for 32.27.7.0/24 will be routed over this link.

Information exchanged via BGP also serves to reaffirm routes provisioned earlier in the session are still valid. A route provisioned by a peer is considered *valid* until it is either updated, replaced, or withdrawn by the peer that provisioned it. Valid routes are added to the local stock, but not necessarily *selected* by the AS as the active route to the prefix. An *update* changes the AS-PATH or some other attribute of the route. For a given prefix, if the AS-PATH is updated, this effectively rescinds to old AS-PATH and offers a newly provisioned AS-PATH. A *withdraw* indicates the route should no longer be used. In CPR terms, the route is removed from the local routing table.[55] Even though the peer knows the old route existed, and may still exist, it should not use it even if it is still live and delivers traffic.

Keepalive messages are used to ensure the BGP session is still active. Modulo the changes above, routes remain valid as long as the session is confirmed live. In terms of rights and obligations, the live BGP session serves to continuously affirm that peers have the rights to use routes provisioned in that session. If the session goes down, all routes provisioned in that session become invalid. Finally, in the absences of a session, routes shared in previous sessions are invalid with respect to that pair of peers.

---

[52]Generally, a /24, 2^8 addresses, is the smallest prefix most networks on the real Internet are willing to route. This is an effort to avoid routing table bloat. This is an instance of a generally well-known operational rule.

[53]The "channel" is a TCP connection.

[54]The term peer is overloaded in the community vernacular. This definition refers to two ASes participating in a common BGP session as defined by RFC 4271 (Rekhter et al., 2006). RFC 4271 describes the mechanics of the BGP-4 protocol. Colloquially, peering also refers to settlement free peering, a contractual mode in which two ASes interconnect (peer in the sense of BGP) but do not exchange payment for traffic that results from that interconnection relationship. Typically the rationale is that the traffic flow is sufficiently balanced that the transaction cost of determining the difference is than the value of the difference. In this work, the term interconnection relation is used unless a more specific contractual mode is necessary in context.

[55]For example, consider the prefix advertisement 42.123.0.0/16: 27 I in Figure 2-1. Withdrawing this route in the BGP session between AS 23 and AS 27 removes the route from AS 23's local stock. The route 42.123.0.0/16: 27 I remains in the global routing table because it is still being advertised to and used by AS 20.

Returning to AS 26's advertisement, for AS 26's prefix 34.27.7.0 to be reachable from *all* other ASes, routes provisioned from its prefix advertisement need be available to the rest of the actors in SimpleNet. AS 23's immediate adjacencies are AS 27, AS 25, AS 20. To initiate the chain of provisioning, AS 23 provisions 34.27.7.0/24: 23 26 I and offers this advertisement to its adjacencies for appropriation. Each adjacency *can* further provision routes to this prefix based on those routes or other advertisements that AS has appropriated elsewhere.[56]

The appropriation and provisioning processes illustrate operational rules structured by the BGP protocol.[57] Peers' respect for provisioning changes signal rights and obligations tacit within a BGP session. When network n provisions a route to a prefix, h signals the willingness to contributed to delivering traffic to that prefix by forwarding traffic to the network h appropriated that route from. That said, provisioning a route from h only guarantees the hop to h, there are not strong guarantees for the entire path. Advertising a route confers the right to *a)* incorporate a route into local routing decisions and *b)* forward traffic for that prefix to the AS provisioning that route.

Tacit in the obligations attendant to prefix advertisement (route provisioning) is an expectation of due diligence in verifying legitimacy. It is important to note this is an expectation—many of the negative operational and security externalities

[56]Whether routes are further advertised, and to whom, is based on each network's value proposition and contractual obligations. Contractual obligations are described in Section 2.3; a more generalized discussion of rights in the NRS is presented in Section 3.4. Consider AS 20. AS 20 is adjacent to AS 19, AS 3, and AS 27. AS 20 could further propagate the route it received from AS 23 by advertising 34.27.7.0/24: 20 23 26 I. Consider further propagation of the route to 34.27.7.0/24 by AS 22. One result is that AS 19 hass a route 34.27.7.0/24: 22 23 26 I. AS 19 may then propagate the route 34.27.7.0/24: 19 22 23 26 I to its adjacencies, including AS 20.

Given the advertisements discussed thus far, consider AS 27. Recall that AS 23 provisioned a route to 34.27.7.0/24 and offered it to AS 27: 34.27.7.0/24: 23 26 I. AS 27 may also appropriate an advertisement for 34.27.7.0/24 from AS 20. In this case AS 27 has a choice: select the path provisioned (adverised) by AS 23 with the fewest hops (2 hops, 34.27.7.0/24: 23 26 I) or the longer path provisioned by AS 20 with (3 hops, 34.27.7.0/24: 20 23 26 I). Both are valid routes. The BGP algorithm would choose the shorter path, presuming fewer hops means lower latency. Fewer hops correlated with lower latency is commonly, but not always, the case. For instance, three hops in New England will be lower latency than two hops that cross the Atlantic.

Before returning to legitimate advertisements and integrity, consider AS 27's possible views of AS 24. The following routes are *possible*, but as will be discussed in Section 2.3 they are not *likely*:

- 34.21.0.0/16: 20 19 21 24 I
- 34.21.0.0/16: 23 20 19 21 24 I
- 34.21.0.0/16: 23 25 24 I

Assuming length of a path is an indicator of quality, the best path is the last, the three hop path. As will be discussed in Section 2.3, costs and conventional contracting modes mean AS 25 will probably not share its path to AS 24.

[57]These rules and obligations are not always followed. For instance, weaker limitations imposed using the language of SHOULD and SHOULD NOT in the BGP protocol may be violated on a regular basis. For instance, the value of the origin parameter, documented at (Rekhter et al., 2006, p.), is a factor in selecting routes from the local stock for use in the active local routing table. A number of discussions have considered the practice of "origin re-writing" in an effort to influence BGP route selection.

discussed in the following sections occur from lack of this "due diligence." Notions of "transitive trust" or "routing by rumor" is the community vernacular for these expectations. When AS X produces a new route, unless it is the origin, the route is provisioned from some other route[58] in AS X's local routing table. In the weakest (and common) form, AS X simply trusts the routes appropriated from its neighbors are legitimate. In stronger forms, AS X uses various facilities, typically the numbers registries managed by the RIRs, perhaps one of the many more Internet Routing Registries (IRRs), to verify routes provisioned by its neighbors are legitimate.

Verification is a cost that endogenizes potential externalities. Origin and path verification is an operational norm. Another recurring externality, prefix deaggregation, is the case when *too many* routes are provisioned, potentially taxing the CPU and memory limits of routing elements (hardware). Routes may be corrupted as they are propagated. At best corrupted routes create limited local effects, such as AS padding used to reduce the chances a network will select a particular route for use. At worst, corrupted routes undermine end-to-end delivery in terms of assurances that the destination is who they say they are or traffic is purposely re-routed through networks where it can be observed by malicious actors. The next sections move from basic mechanics to a discussion of these problems in terms of positive and negative externalities.

2.2.2 Updates and Convergence

Local and global routing tables are far from static. BGP sessions go up and down, operational issues require updates to advertised routes, and the economic factors described in the next section all result in a dynamic global routing table that is in constant flux and, to the casual observer, seemingly unstable. Although the stock of routes is continuously provisioned and rescinded, *a*) the integrity of the resulting routes should be preserved and *b*) all prefixes should remain reachable via legitimate routes. RFC 4277 indicates that

> the driving force in CPU and bandwidth utilization is the *dynamic* nature of routing in the Internet. As the Internet has grown, the frequency of route changes per second has increased.[59]

That was in 2006. Since then content delivery has given rise to much more sophisticated bundles of interconnection relations.[60]

Temporary local instability that results from necessary updates is an unavoidable cost of control plane management. When a network updates a route, networks formerly using that route must recalculate their local routing tables to accurately

[58] In terms of a production function, among other parameters, non-origin routes are paremterized by some pre-existing route.

[59] Quoted from (McPherson & Patel, 2006, p. 9). Discussion of CPU and bandwidth utilization reference Section 6.1 of RFC 4274 (Meyer & Patel, 2006, pp. 7–9, emphasis added here).

[60] These bundles facilitate lower latency, reliability, and load balancing. These offset costs and increase efficiency. They also create a more complex interconnection ecosystem. These are discussed extensively in Chapter 6.

reflect the change. The recalculation process is called *convergence*—all the routes in a domain, here an AS, are expected to have the same topological information and it should be consonant with that AS's routing policy. An AS with a stable local routing table has, through the application of the BGP protocol and its local routing policy,[61] executed a set of decision rules about which set of routes (of those for which it has valid advertisements) it will use when routing. When this decision process is complete, the network is said to have converged on a stable local routing table.[62] Convergence is a replenishment process. This recalculation process affects routing table stability—when routes are being updated, there is a period of instability.[63] When *any* update is received, part or all of the convergence process must be performed. Convergence creates a cost for the AS and is known to affect end-to-end efficiency and reliability.[64]

The scope of that instability may vary with the size of the network and how far a route propagates. For example, if the change only affects a local interconnection relation, for instance a privately shared route between two small networks is changed, it only affects those two networks. In contrast, if the change propagates to most of the Internet, a wide array of actors may be affected by an externality created by a single actor. Historical instances of operational externalities include the AS7007 leak in 1997 (Bono, 1997), TTNet in Turkey (Underwood, 2005), and the ConEdison leak (Underwood, 2006). All of these had a global impact. Not all operational externalities have global impact. In reference to a 2008 leak in Colombia, Cowie (2008) asserts that "[i]n practice, most leaks tend to fall into two categories,

[61]Local policy is set via a combination of router features, typically expressed in Route Policy Specification Language (RPSL) (Alaettinoglu et al., 1999; Meyer, Schmitz, Orange, Prior, & Alaettinoglu, 1999; Blunk, Damas, Parent, & Robachevsky, 2005), localpref attributes, and in some cases manipulating the attributes of routes used by the BGP protocol as well as the existing rules of BGP set out in (Rekhter et al., 2006).

[62]Labovitz, Ahuja, Bose, and Jahanian (2000, p. 177) defines "a *steady-state network* as one where no BGP monitored peer sends updates for a given prefix for 30 minutes or more." The choice of 30 minutes is based on empirical observations documented in (Labovitz, Ahuja, & Jahanian, 1999).

[63]Consider the discussion of route updates in RFC 4274:

> During periods of network instability, BGP routers within the network are generating routing updates that are exchanged using the BGP UPDATE messages. The greatest overhead per UPDATE message occurs when each UPDATE message contains only a single network. It should be pointed out that, in practice, routing changes exhibit strong locality with respect to the route attributes. That is, routes that change are likely to have common route attributes. In this case, multiple networks can be grouped into a single UPDATE message, thus significantly reducing the amount of bandwidth required[.] (Meyer & Patel, 2006, p. 8)

This defines instability in terms of updates. Here instability does not imply link instability, but rather that the information used to select amongst available links is in flux, and thus for the time unstable. In this sense, following the claim that the routes are in the frequency of seconds, that that routing table is always unstable. That said, this does not imply the entire routing system is unstable. Rather, localized changes may create instability within a particular prefix cone, but not the entire Internet. See Labovitz et al. (2000) for a detailed analysis of the latency of routing convergence.

[64]This is considered well-known; for quantitative studies see (Labovitz et al., 2000) and (Wang, Mao, Wang, Gao, & Bush, 2006).

depending on whether they propagate: well-localized to a single regional customer space, or planetwide." As will be discussed in Section 2.2.3, Pakistan-YouTube is an instance of a security externality that had global reach.

Some sources of change, are unavoidable aspects of network operations. For example, *a*) networks are expected to propagate changes in routes they use; *b*) link failures inevitably happen; *c*) maintenance is necessary and may require temporary routing changes.[65] Ideally, the network will experience the minimal number of updates necessary for accuracy. Minimal translates to, as will be discussed in terms of route flap below, both rate *and* kind. Redundant updates (rate) should be avoided where possible. Updates should be packed (kind) to avoid inefficient sequential convergence. What differentiates these changes is the potential to endogenize excess costs. Behaviors that create costs for external actors are scrutinized by the community in terms of whether those are legitimate operational practices necessary to *maintain* an accurate global routing table or if some or all of the cost (externality) can be reasonably avoided.

In an ideal scenario, there would be a clear threshold. The operational externalities above are uncertainties that can be described and explained analytically, but whose solutions are often empirical, a product of institutionalizing operator experience in operational rules and norms. This type of uncertainty is characteristic of CPRs. For instance, in the case of underground reservoirs, the exact capacity and impacts of certain types of use may not be certain. Historical experience is transposed into analytic heuristics for determining the volume (threshold) that is safe to withdraw on an annual basis. Similarly, riparian rights are affected by uncertainty

[65]To illustrate the implications of updates, consider AS 19's view of AS 23. AS 19 may see at least the two advertisements for 42.123.0.0/16 42.123.0.0/16: 23 27 I and 42.123.0.0/16: 20 27 I. One is provisioned by AS 23, the other provisioned by AS 20. Both have been appropriated into AS 19's local stock. At the moment, both are legitimate route advertisements and both are the same length. Consider the scenario where AS 23 decides to no longer exchange traffic with AS 27. When AS 23 "tears down" the BGP session, all of the routes previously exchanged between AS 23 and AS 27 will be invalid.

In the data plane, traffic to and from AS 27 will no longer legitimately traverse the link between AS 23 and AS 27. AS 27 may still know the "route" exists and that there is working physical connectivity there, but it is now marked as invalid. In the control plane, AS 23 rescinds its obligation to carry AS 27's traffic. The routes offered by AS 23 have been removed from AS 27's local stock and it must converge based on the remaining stock—those exclusively from AS 20. AS 23 must also converge, changing its local routing table to send traffic for AS 27 through AS 20. AS 23 must now also propagate its withdrawal of routes with the 23 27 adjacency to its neighbors. "Repairing" the global routing table (global stock) to reflect this change will be complete when all networks with this adjacency in their routing tables have converged to remove it and select alternative routes where necessary. As implied by the language "repairing," this is a *necessary* update to ensure consistent traffic delivery that reflects the local change.

Another reason for updating may be if the link between AS 23 and AS 27 temporarily fails. AS 23 can no longer fulfill the obligation to route AS 27's prefix. In terms of BGP, a nontrivial link failure will kill the BGP session, causing a withdrawal if AS 23 does not recognize the failure and withdraws the affected routes explicitly. Yet a third reason may be maintenance or a scheduled outage of a portion of AS 27's network. In contrast to terminating a relation, these updates, and the subsequent convergence costs, are potentially recurring.

in upstream volume and the coordination of upstream and downstream uses. In the control plane, operators coordinate to share common knowledge and develop heuristics, operational rules, for navigating these uncertainties. A well-known instance referenced earlier is route flap, discussed in the next section.

2.2.2.1 Route Flap as an Operational Externality

Route flap mechanics illustrate how negative externality manifest and propagate. This externality also helps drive home root causes of externalities as they relate to operational costs, scope of an externality, and various general and specific mitigation strategies. Route flap may be caused by a technical failure and/or operational practices. Canonically, route flap occurs when a route is updated more frequently than necessary to ensure an accurate local and global routing table. As per the previous discussion, this threshold is not clear cut. Route flap is characterized by costly convergence processes considered redundant by those experiencing the externality. Route flap is also known to degrade end-to-end connectivity.

Consider an AS, say `AS 24` in SimpleNet. `AS 24` has a router with a faulty interface that causes the physical link to go up and down. The BGP session goes up and down with the link. The BGP session, and routes provisioned in that session, will be built and torn down repeatedly. Technically, this does repair the failed link each time, but it does not alter the intended topology. Rather, it unnecessarily oscillates between two states, creating convergence costs on each change.

In the extreme, "the load [route flap] processing placed on the control planes of routers caused further instability as the routers were not able to process other BGP updates or they dropped traffic transiting the device. This could produce cyclic crashing behaviour," (Bush et al., 2013). Route flap is attributed to both technical errors (router hardware failures, software failures, configuration errors, communication link errors, etc.) and operational failures[66] such as releasing a sequence of updates rather than a single "packed" update.[67] The former scenario, hardware and software failures, may not be avoidable.[68] The latter, operational failures, including not recognizing and limiting observable cyclic failures, are within the purview of operators' activities.

Operational externalities such as updates can be classified as either positive or negative externalities. When an update that preserves the integrity of the control plane and Internet communication in general is promulgated, that is considered to be, on balance, a positive externality.[69] Convergence costs may be considered the cost of maintaining the control plane as a common resource. In many common

[66]Failures is a polite general term. Failure to implement packing may be due to the cost of coordination, sheer laziness, or ignorance of the implications of sending individual updates. Avižienis, Laprie, Randell, and Landwehr (2004, p. 17) makes a similar assertion about classes of faults, identifying incompetence faults should be highlighted, not glossed over for politeness.

[67]See RFC 2439 (Villamizar, Chandra, & Govindan, 1998) for a discussion of update packing and the Route-flap Damping Protocol (RDP).

[68]This assumes these failure modes were unknown. If these are known but considered "acceptable" this problem falls into the category of operational failures.

[69]It is assumed that the accuracy of the global routing table is pre-eminant.

resources, the community either pays into a common pool to finance common resource maintenance *or* contributes labor and materials directly to maintenance. The control plane is akin to the latter case. For instance, in the Zanjera irrigation communities, local irrigation system appropriators periodically build dams necessary to support CPR objectives.[70] The parallel between convergence and irrigation is not only in the labor contribution structure, but also in resource system topology. Both alter the topology of the resource system to allocate resources to appropriators.

A combination of protocol and operational rules have emerged to regulate route flap. Protocol solutions are documented in RFC 2439 (Villamizar et al., 1998). A technical solution is a route dampening protocol (RDP) which assigns a score to a network based on how frequently it sends updates, i.e. how frequently if flaps. Once that network's score passes a particular threshold, route updates are monitored, but ignored. Over a period during which that network does not flap, the score decays and, modulo additional flap, route updates are accepted when the score falls below the threshold. RDP was not an initial success.

Finding the correct scoring parameters has been an ongoing operational issue.[71] In terms of network operations, update packing discourages the simple approach of sending one update per route change. Rather, it encourages actors to "pack" changes into a single update message. This limits flap and, subsequently, the costs of multiple convergence processes. Packing can be more costly than a series of simple updates depending on the complexity of the network actors' updates, the value proposition, and the technical sophistication of the actor.

Different value propositions give rise to more frequent route provisioning (updates) than others. Some networks, such as geographically concentrated access networks, may not need to frequently provision new routes. Other networks, such as CDNs, change routes frequently for load balancing, modulating the tradeoff between users' quality of experience and transport costs for routes to geographically distributed caches.[72] As such, the frequency of updates that constitute a negative externality for an access network may not constitute a negative externality for a CDN. Middle range flap is the situation where periodic flap creates convergence costs and related instability, but does not crash routers or BGP sessions. Following the threshold of maintaining accuracy and integrity, this also constitutes a negative externality, but may not garner the attention necessary to act.

2.2.2.2 Route Aggregation, or the Lack Thereof

Another operational externality is route de-aggregation. Networks allocated aggregate address blocks may choose to provision a set of smaller blocks rather than aggregate. This is a long standing problem that has periodically garnered attention

[70] See (E. Ostrom, 1990, pp. 82–88) for details.

[71] In the academic literature, see Mao, Govindan, Varghese, and Katz (2002) and later Pelsser, Maennel, Mohapatra, Bush, and Patel (2011). The RIPE community has also addressed the application of RDP, most recently in (Bush et al., 2013).

[72] To be precise, CDNs are in the business of having a deep understanding of that tradeoff space and identifying routes that balance quality and cost.

in the operator community. Provisioning many disaggregated prefixes is a form of over-provisioning advertisements, where routes to multiple small prefixes could be aggregated into a few aggregate prefix advertisements. Even with packing, a common argument is that the growth of disaggregated prefixes will overwhelm limited router memory. In the worst case, this may lead to crashing or operating on an incomplete routing table. In both cases, the local routing table is incomplete and the global routing table may be inaccurate.

The operational rule to solve this problem is route aggregation. Networks are encouraged to aggregate their route advertisements as much as possible. For instance, a common problem is LIRs that have grown over time and their total address allocation is not contiguous. These LIRs may suballocate to their customers, further disaggregating their already fragmented block. These actors could minimize fragmentation by re-aggregating routes advertised my their customers. They do not necessarily do this, simply passing prefix advertisements on from their customers. Again, this is an externality—the LIR could aggregate, but it comes at a cost.

Aggregation is also a potential solution to route flap. The following illustrates both the general benefits of aggregation and how it can be used to mitigate (endogenize) the costs of route flap. Consider the scenario in which AS 23 from Figure 2-1 is an LIR with four customers:

- AS 401 with prefix 9.0.1.0/24
- AS 402 with prefix 9.0.2.0/23
- AS 403 with prefix 9.0.4.0/22
- AS 404 with prefix 9.0.8.0/24

AS 403 is a medium sized network with competent, experienced operators. AS 401 and AS 404 are smaller, less experienced networks. AS 401 and AS 404 are having problems with a faulty interface, causing the link to fail periodically, the BGP session with AS 23 to fail, and, when the link is re-established, the BGP session must be restarted. This is a canonical instance of the route flap externality. Rather than advertising each of these networks to the rest of SimpleNet, exposing all of the network to the route flap caused by AS 401 and AS 404, AS 23 could advertise just `9.0.0.0/8: 23 I`.

Advertising the aggregate route, `9.0.0.0/8: 23 I`, has two immediate benefits for the control plane: *a*) the total entries in the routing table are reduced, saving appropriators' router memory, and *b*) AS 23 shields the rest of SimpleNet from the route flap. The latter means that only AS 23 has to converge on each flap. In effect, AS 23 has endogenized the cost of AS 401 and AS 404's potential operational externality. Rather than all of SimpleNet converging each time AS 401 or AS 404 flaps, only AS 23 must converge.

A more nuanced view distinguishes the BGP session in which a *potential* externality is introduced from *subsequent* route provisioning outside that BGP session. In the example of route flapping, the would-be externality created by AS 401 and AS 404 is endogenized by AS 23 before it promulgates beyond the original announcement (scope). In contrast, it would be an externality to the rest of SimpleNet if AS

23 *had not* endogenized the cost, but rather, had engaged in simple localized optimization (avoided the cost of aggregating) and passed the cost of its transaction on to its adjacencies. In effect, it would have exposed those actors to "noise" that would "impinge" on those actors and the actors they passed those updates on to.

Operational externalities assume the source of the externality is a non-malicious, local optimization of operating costs. Free riding is considered bad, but not necessarily actively malicious. This is a common failure mode in collective action problems. As above, some operational externalities are rooted in unavoidable failures. Others are rooted in the operational choices, often local optimization, that are under the control of the AS. In the case of the latter, RDP scoring mechanisms may be leveraged as a selective incentive. In the next section, intentional route manipulation and the attendant externalities are discussed. In some cases, albeit intentional, route manipulation is a form of operational externality. In others, the intent is to subvert the intended end-to-end integrity of traffic exchange. These externalities are characterized as security externalities.

2.2.3 Route Manipulation and Security Externalities

BGP facilitates dissemination of routes amongst networks, but it does not assure routes are legitimately provisioned and appropriated. Assuming neighbors are simply appropriating routes without verification, any network can, at least mechanically, provision any route it wishes. Any network participating in the provision of a route along what would be the legitimate path may also manipulate that advertisement. Thus far only legitimate advertisements in Figure 2-1, have been discussed. In BGP peering between any network A and B, in the case of a route provisioned by A for appropriation by B, with respect to the AS-PATH, A expects that: *a*) B will not alter the portion of the path advertised by A and *b*) in subsequent provision of routes based on the one just received from A, B only pre-pend its AS number before subsequently advertising the route to its adjacencies.[73] In general, this is part of the *transitive trust* model discussed earlier. This is documented in the BGP protocol, RFC 4271 (Rekhter et al., 2006), but there are not mechanisms in BGP to check legitimacy, what is referred to in the community as origin and path security.[74]

Route manipulation, such as AS padding or path truncation, may be a mechanism for localized optimization. Both of these can create operational externalities in the sense that one actor's independent local optimization decision affects a wide range of others. Localized goals via route manipulation, malicious or otherwise, can subvert the expectations of legitimate advertisements. Routing integrity thus far has assumed that all actors along the path are who they say they are and that the routes they use and promulgate are legitimate. Two violations of these assump-

[73]There are a number of other BGP parameters that should be similarly respected. See RFC 4271 discussion of path attributes in terms of well-known versus optional and mandatory, discretionary, and transitive (Rekhter et al., 2006, p. 23–25).

[74]Origin security will be discussed extensively in Section 5.7.4. Path security will also be discussed, in particular as it relates to security protocols fit with provisioning rights network operators have come to rely on.

tions will be discussed in terms of route manipulation. First, the end points are not who they say they are—typically the destination. Alternately, intermediaries are either subverted or artificially injected to facilitate a malicious actor observing traffic along a route that is illegitimate, ultimately delivering, traffic to the intended destination.

2.2.3.1 Operational Externalities via Manipulation or Error

Route manipulation can introduce changes to routes that do not alter end-to-end integrity. Rather, the utility of the route is altered to suit the needs of the producer. This makes the route less useful to immediate and downstream appropriators. In some cases, it may subvert the intent of exchanging routes. One instance is AS padding. An AS may advertise a route, but when adding its own AS number to that route, it adds multiple copies to make that particular route less likely to be selected. In the scenario where a network originating routes interconnects with two networks but prefers one to act as a primary and another as a backup, it will may pad the routes advertised to the backup to discourage traffic.[75] In other scenarios, padding may be used exclusively to attempt to divert traffic to lower cost connections. The positive or negative effects of these scenarios have been observed in the wild and debated on operator lists.

Butler, Farley, McDaniel, and Rexford (2010, p. 105) discuss various modes of manipulating addresses by truncating, adding additional hops, or introducing particular ASNs. An AS-PATH may be truncated to varying degrees to make it look shorter, eliciting preference by BGP. In a seemingly innocuous situation, an upstream provider of an access network may choose not to pre-pend its AS to shorten the AS-PATH, encouraging selection of that path. More aggressive truncation may result in choosing paths that have much higher latency than expected.[76]

On the other end of route manipulation, additional hops may be added to make a path longer, and less likely to be selected, such as with AS-path prepending. A malicious actor could add an additionalal ASN of another actor, making it look like that actor engaged in prepending, not the actual culprit. Yet another alternative is to remove a subset of the AS-PATH in the middle of a legitimate route. By preserving at least the two legitimate ends of the route and pre-pending the malicious network's ASN, it appears the organization represented by the second hop has invested in infrastructure necessary to interconnect with the destination.

While this does not interfere with the ultimate delivery of packets, any networks that appropriate the illegitimate prefix advertisement over legitimate prefix advertisements may likely unintentionally select a suboptimal route. Unless appropriators perform (costly) experiments to affirm the route is as advertised,[77] they

[75]This also depends on the rights conferred in the contract governing the exchange of routes. As long as the "customer" is paying, the peer may not care. That said, padding is observable and has been discussed on mailing lists as obstructing potentially more efficient routing.

[76]Again, AS-PATH length is correlated with, but does not *guarantee* lower latency than a longer path to the same prefix.

[77]Heuristics for determining the path traffic follows *can* confirm an AS-PATH, but not in all cases.

must trust the route as advertised (provisioned) will perform as expected. Again, this is a manifestation of the transitive trust model.

Butler et al. (2010) also describe the scenario where an attacker injects another network's ASN to cause that ASN to drop the route when it would otherwise accept the route. When that actor receives the route, it sees its own ASN in the AS-PATH. BGP loop detection discards the route, assuming it has already seen this route. This misleads the targeted ASN to disregard what would have been an otherwise legitimate (and potentially less costly) route. In the worst case, the target now has an incomplete routing table.

2.2.3.2 Security Externalities via Route Manipulation

Butler et al. (2010) describe a number of threats to routing integrity. Two classes of threats are discussed: those intrinsic to BGP and those that threaten control plane integrity by proxy, by threatening the integrity of the underlying transport layer. Link layer failures can be a source of failures in the control plane. Failure modes that are exclusively rooted in the transport layer are not considered control plane externalities per se.

To clarify, consider an instance of this rationale. Butler et al. (2010) reference a link cutting attack described by Bellovin and Gansner (2003). A proper subset of the legitimate routes between an origin and a destination is severed at the transport level.[78] Selected link "failures" force traffic over a set of links and/or routers compromised by the attacker. From the perspective of BGP routing presented here, a compromised route is not illegitimate.

A vulnerability is "an internal fault that enables an external fault to harm the system" (Avižienis et al., 2004, p. 18). Based on this definition, link cutting is not a BGP vulnerability per se. Link cutting exploits vulnerabilities in the common infrastructure supporting both "in-band" traffic and "out-of-band" BGP communication. This external fault does not activate an *internal* fault in BGP. Rather, this attack exploits knowledge of an existing topology and how BGP will alter that topology in response to any type of link failure. In effect, the attack leverages, or more accurately, abuses, the correct behavior of one resource to expose traffic to independently compromised resources (routers or links with eavesdroppers, for example).

In contrast, malicious manipulation of BGP routes themselves or altering the value of BGP parameters is considered an operational vulnerability. The Pakistan-YouTube story opening the dissertation is an instance of prefix hijacking. Prefix hijacking occurs when an AS claims it is origin for a prefix allocated to another AS—that AS is violating origination rights. Recall prefix delegation confers the right to originate a prefix to a particular `AS X` and imposes the obligation on all others *not* to originate that prefix.[79] Number rights are discussed at length the next

[78]This may be via DOS or some other means. Regardless of the means, these failures are rooted in transport vulnerabilities.

[79]Or at least not to originate without a temporary transfer of origination rights. See discussion in Section 3.4.5 for a discussion the rights to alienate, or transfer resource rights.

Chapter 3 and 5.[80]

In terms of harm to the control plane, hijacking substantively increases the number of illegitimate routes provisioned in the control plane. One operationalization of control plane integrity is the ratio of illegitimate to legitimate routes that exist in the global routing table. In common resource terms, how tainted is the stock of routes?

Given this background, the Pakistan-YouTube hijacking can be described in more detail and in terms of rights, obligations, and externalities at play.[81] In February of 2008, the prefix `208.65.152.0/22` was used to route traffic for YouTube (`AS 36561`). When provisioned legitimately, users all over the world had easy access to videos of adorable kittens and fantastic exploding jello—YouTube reaped the rewards of the advertising accompanying these videos. A more serious video hosted at the time insulted the Prophet Mohamed. Pakistan requested YouTube remove it, YouTube declined. Pakistan Telecom (`AS 17557`) attempted to block YouTube by rerouting traffic originating in Pakistan and destined for YouTube to a local host serving a message that YouTube traffic had been blocked. `AS 17557` attempted to provision YouTube's prefix, `208.65.152.0/22`, *only* to networks in Pakistan. Unfortunately, the route leaked.[82]

Within Pakistan, this action is legitimately within the power of the Pakistani Telecommunications Authority. From Pakistan's perspective, YouTube was violating Pakistani law. The role of the regulator is to uphold the public's interest, in this case manifest in Pakistani law.[83] In a different context, the route is illegitimate from the perspective of allocating registry: `208.65.152.0/22` was delegated to YouTube, conferring exclusive origination rights to YouTube. In a broader framing, this is challenge to Pakistani sovereignty and may be framed as an externality. A decision outside of Pakistan's jurisdiction imposed additional costs on Pakistan to enforce its laws. A low cost solution seemed to be to simply block YouTube by rerouting traffic within Pakistan.

If Pakistan's perspective is taken, the advertisement was illegitimate *only* when

[80]Consider a malicious manipulation of route advertisements. Assume `AS 24` is a malicious actor. `AS 24` begins advertising `9.0.127.0/9: 24 I`; `AS 24` is now originating `9.0.127.0/9`. `AS 24`'s strategy is to leverage the fact that BGP prefers longer prefixes when selecting routes. Absent other security mechanisms, traffic intended for `AS 23`'s hosts addressed in `9.0.127.0/9` would instead be routed to `AS 24`.

In effect, `AS 24` has hijacked the prefix legitimately allocated to `AS 23`. `AS 24` provisioned an illegitimate (origination) advertisement. Following the BGP protocol, adjacent actors appropriated (used) that advertisement and provisioned subsequent illegitimate routes (by pre-pending their ASN and advertising the route further), perpetuating the security externality.

[81]This instance of prefix hijacking is quite famous and has been documented in a number of sources. This narrative is derived from reports by Renesys, in particular Brown (2008).

[82]A route leak, in contrast to an intentional hijack, is an operational externality rather than a security externality. There is some debate regarding whether Pakistan leaked the route or the "hijack" was intentional.

[83]It is a massive assumption, in *any* state, that the law is a reflection of the public interest. At best it is an approximation, at worst is it imposed by fiat from those with power, but not authority. For a discussion from the perspective of the philosophy of developing legal systems, see Hart (1994).

it propagated farther than intended, outside of Pakistan.[84] The route was advertised further upstream by PCCW (AS 3491), Pakistan Telecom's upstream provider. Brown (2008) reports that the first evidence of the route was observed[85] at 18:47:45. Within a couple of minutes, at 18:49:30 all of the observed ASNs that would carry the route, many of which are large providers, were carrying the route.

This instance of prefix hijacking highlights the difference between valid and legitimate, as well as the missed opportunity to endogenize the externality. The route is valid in the sense that it was seen on an established BGP session amongst other valid route advertisements. As such, it was appropriated into the local stock and used in the route selection process. Given Pakistan's prefix advertisement was shorter than others for 208.65.152.0/22, Pakistan's advertisement was selected by PCCW's neighbors over longer alternatives. That route is illegitimate because AS 17557 does not have the rights to originate 208.65.152.0/22. Had PCCW checked whether AS 17557 was in fact the steward of 208.65.152.0/22, it could have legitimately ignored the advertisement. This would require PCCW expend the operational cost to confirm the legitimacy of prefix advertisements considered for appropriation from its neighbors rather than simply passing them along via subsequent provisioning.

Returning to the narrative, at 20:07:25 YouTube advertised the same prefix that had been hijacked, 208.65.152.0/22, to its neighbors (appropriators in the context of this advertisement) in an effort to stem the impact of Pakistan's advertisement. This advertisement had local effects in the scope for which it was the shorter advertisement. When that route propagated to Asia, over multiple networks, it had a longer AS-PATH than the geographically and topologically more proximate, and thus shorter, illegitimate advertisement from Pakistan. As such, BGP, operating on shorter AS-PATH wins, chose the illegitimate advertisement.[86] At 20:08:30 "40 some-odd providers have stopped using the hijacked route" after YouTube advertised the /24. At 20:18:43 YouTube advertised two /25's knowing a longer (more specific) prefix would be selected by BGP over the /24 advertised by Pakistan Telecom. At 20:59:39 PCCW, Pakistan Telecom's access to the rest of the Internet, withdraws the route.[87]

In this section, operational and security externalities describe how externalities affect routing information and delivery. It was also assumed that actors would promulgate *all* available routing information to *all* neighbors. In SimpleNet, physical connectivity was assumed, but the costs of using that connectivity was not considered. The infrastructure carrying traffic facilitated by interconnection is costly and partially rival—actors engage in a number of contractual relationships to ensure

[84]The route is illegitimate from the perspective of allocation: 208.65.152.0/22 was allocated to YouTube, conferring unique origination to YouTube.

[85]Observed from the vantage point of Renesys's various route collectors around the world.

[86]In a route hijack situation where the legitimate origin responds with a prefix of the same length as the false advertisement, there is a topological distance in hops at which the false advertisement dominates.

[87]There is some discussion of whether PCCW disconnected Pakistan in the sense of its connectivity writ large or simply filtered route announcements. The latter seems more likely.

connectivity to the rest of the network. The next section describes conventional contracting modes that shape and characterize interconnection decisions.

2.3 Interconnection Contracting Modes

Number resources provide the necessary access to the control plane, but access alone is not sufficient to create strategic bundles of contracts that satisfy a network's value proposition. In SimpleNet, some networks have more diverse connectivity than others. This confers advantages when selecting amongst available routes. Simple connectivity may be sufficient for some actors, others require, and develop, more sophisticated interconnection bundles. The contracting modes described in this section are the economic building blocks of interconnection. In the interconnection market, a network's objective is to develop interconnection bundles that provide the routing options necessary to support its value proposition.

In the previous section peering relationships were simplified to a single physical link between two networks. Real networks may have many physical links between them. When two such ASes exchange routes, this is referred to as interconnection. The broad notion of an interconnection relation between two ASes is comprises all of the BGP sessions. It is a comprehensive contractual relationship between two organizations.

Loci of physical connectivity are referred to as an interconnection platform. These loci facilitate markets in which ASes may participate to select actors with whom it would be beneficial to exchange routes with, i.e. with whom to establish, or further develop existing, interconnection relations. Within the interconnection market, route selection is influenced by *a*) the context-specific (derived) value of one or more routes to a potential appropriator (AS) and *b*) the cost of exchanging traffic with the producer (provisioning AS). The value of a route to an appropriator may be based on length, as per BGP in the previous section, the specific prefixes available (for instance a particular end-user market) or both (better quality access to a particular market). Interconnection economics focus on these strategic decision processes. Parameters of this decision process include the relative cost of traffic exchange, access to bundles of prefixes, access to routes, and the strategic externalities that emerge over the course of these continuous decision processes.

A variety of contracting modes between adjacent ASes, that is to say appropriators and producers, have emerged in the interconnection market. Historically contracting modes are variants of bilateral, interconnection or transit. Contracts structure payment for *a*) routes exchanged (provisioned and appropriated) between contract participants, typically a pair of ASes; *b*) traffic exchanges between participants, typically in terms of a traffic ratio between the participants or in terms of cost per unit volume of traffic. Operational updates described in the last section provide the mechanics for maintaining the global routing table. Contracts also impose constraints on the structure (topology), quality, and quantity of the stock of routes available in a particular market. The foundation of this discussion will be provided here. Chapter 6 describes IXes as an interconnection platform that lowers

barriers to interconnection market participation.

Interconnection bundles between ASes range from a single simple contract at a single common point of presence (POP) to a complex bundle of diverse contractual relations across diverse physical POPs. Developing interconnection bundles is both operational and strategic: creating bundles of routes should balance satisfying a network actor's value proposition while also preserving the integrity of the control plane. Part of preserving integrity is endogenizing potential negative externalities. The former is often explicit, the latter tacit in well-known operational rules. The following section provides a review of interconnection economics.[88] In terms of common resource management, static and dynamic efficiency of the decision process are rooted in *a*) the available opportunities exchanging non-rival routing information in the control plane and *b*) the partially-rival economics of traffic forwarding in the data plane.[89]

Faratin et al. (2007) discuss interconnection bargaining games in terms of static interconnection bundles—particular configurations of contracts between actors and the implications of those configurations. These games highlight the shift from simple contracting modes like transit and peering (discussed shortly) to the rise of partial transit and paid peering. Strategic permutations of these leverage knowledge of control plane topology, by proxy interconnection market structure, to create and sustain market power. Faratin et al. (2007) also highlight that the types of actors have changed since canonical notions of transit and peering first evolved—ISPs are but one of many different network actors in the Internet infrastructure ecosystem. Difference amongst network actors' value propositions give rise to differentiated strategic interconnection practices that are not easily captured in simple traffic ratios.[90]

Sophisticated content delivery architectures and large eyeball networks represent a well-known contemporary value network whose business negotiations play out in part through strategic manipulation of bundles of interconnection relations in the larger interconnection market, across a variety of platforms. D. Clark et al. (2011) contemplate the competitiveness of interconnection markets and the role of transparency in the contractual relations that shape how these markets work. Both Faratin et al. (2007) and D. Clark et al. (2011) address contracting strategies of interconnection relations by comparing static bundles and their implications for market power. The following sections provide a primer on transport as a connectivity option, followed by discussions the varieties of transit and peering that are at play in various local, regional, and global interconnection markets.

[88]As per the references, this work draws on existing literature, in particular (Faratin et al., 2007; D. Clark, Lehr, & Bauer, 2011). The summary and discussion draws heavily from Sowell (2013).

[89]This latter will be a focus of Chapter 6, which discusses interconnection platforms. In particular, it frames the value of IXes in terms of general and specific infrastructure asset investments that facilitate better utilization of number resources in service of the actors' value proposition.

[90]Private conversation with a long time control plane participants indicate negotiation was never about traffic, but rather about less tangible business relationships.

2.3.1 Transport

Transport provides a point-to-point connection (or circuit) from the point of presence (POP) of one of an ASes' facilities to the POP of some other facility. A POP may be a facility owned by the AS housing networking equipment or the equipment of an AS housed at a third party facility. Most of this discussion focuses on transport between the POPs of two different ASes, the black links in SimpleNet.[91]

In and of itself, transport does not guarantee an interconnection relation. Transport is one component of gaining physical access to the infrastructure atop which interconnection markets operate. For instance, consider the transport link between the facilities of AS 26 and AS 23 in Figure 2-1. Further consider that the facilities of each only serve the networks of AS 26 and AS 23—each is a dedicated plant. Under transport between these dedicated facilities, the only interconnection "option" available is the mutual provisioning and appropriation of routes between AS 26 and AS 23.

While a dedicated plant is a degenerate form of an interconnection platform, most of the platforms discussed here refer to facilities housing the POPs of many diverse ASes. In the community vernacular, such colocation facilities have been referred to as "carrier hotels." More recently, and developed further in Chapter 6 on IXes, platform participants are no longer exclusively carriers or conventional ISPs.[92] In contrast to transport that connects dedicated facilities, transport to a diverse interconnection platform is a step to gaining *access* to a diverse set of participants, potential route provisioners and appropriators, on that platform. While that diversity is important, one of the factors that makes these platforms attractive is the presence of multiple competing transit providers, discussed next.

2.3.2 Varieties of Transit

Transit means a network actor ensures that customers to whom it provides transit service have connectivity to the rest of the Internet. This guarantee is possible by virtue of transit providers' contractual relations with others. The following describes various forms of transit and how these have been combined.

[91]Transport is also frequently used to interconnect distant parts of a common enterprise networks such as within a large metro region, between metro regions, or even across countries and continents. Network actors have a multiple transport options. One option is to build physical connectivity between POPs de novo. Building infrastructure is expensive and unlikely unless one of the actors is itself a transport or transit provider. Typical options include: *a*) leasing existing dark fiber the AS can light themselves; *b*) purchasing a set of wavelengths in existing fiber; *c*) contracting transport services from a provider that has more extensive infrastructure; *d*) contracting transport platform services, such as IX-Reach or Atrato, that have specialized in facilitating connectivity to interconnection platforms. More exotic transport options include transport over power lines and long distance wireless communication. Both have seen deployment in underdeveloped and/or geographically challenging regions. For instance, wireless has seen use in mountainous regions, hopping from ridge to ridge or in locations with little terrestrial infrastructure.

[92]Faratin et al. (2007) also highlight that network actors have become more diverse than simply conventional ISPs providing network access to end users.

2.3.2.1 Simple Transit

Consider a scenario where AS 23 provides transit services. A *simple transit relation* means a network actor, say AS 26 in Figure 2-1, has: *a*) transport to some facility where AS 23 has a POP,[93] and *b*) the contractual guarantee that the rest of the Internet (here SimpleNet) is reachable by virtue of subsequent connectivity. In effect, AS 23 guarantees AS 26 access to the rest of the world. When connectivity comprises only transit from a single provider, the customer relies *exclusively* on the transit provider for *all* external connectivity. In terms of connectivity assets and decisions, the customer has exported management of off-net connectivity to the transit provider, in particular geographically local routing decisions. This is also referred to as *single homing*.

Consider traffic between AS 26 and AS 25. AS 23 entrusts connectivity to AS 25 (and the rest of the network, to AS 23's connectivity. Following the existing interconnection relations, the only legitimate path is `34.45.0.0/16 23 20 19 22 25 I`. Now consider the scenario where the physical link between AS 22 and AS 23 is active and those networks exchange traffic; the precise contractual mode will be discussed in the next section. There are now two possible legitimate routes between AS 26 and AS 25:

- `34.45.0.0/16 23 20 19 22 25 I`
- `34.45.0.0/16 23 22 25 I`

If AS 23 provisioned these routes for appropriation by AS 26, AS 26 would select the shorter path.

It is unlikely AS 23 will offer both routes for AS 26 to select from. The shorter path is preferred by AS 26 and it is likely correlated with lower latency. That particular route may be more expensive for AS 23, though. AS 23 will offer the lowest-cost route to AS 26. In general, a transit relation provisions the bundle of lowest cost routes. These are lowest cost, with respect to the transit provider, amongst the routes in the transit providers local stock (routing table). As the costs of particular routes available to AS 23 shift, AS 23 will update the routes it provisions for AS 26 in an attempt to ensure lowest-cost route selection. The "sum" of the prefixes result in a full routing table, but the underlying routes may change based on the costs faced by the transit provider.[94]

2.3.2.2 Transit Redundancy

To mitigate the risk of connectivity outages, networks often establish multiple transit relationships. Typically a primary and some number of backup transit relations

[93]This is a simplification. Many networks connect at a number of facilities. This is a simplification for clarity here, the constraint will be relaxed in later discussions.

[94]Having only a single transit provider is considered a risk. Under transit, physical connectivity may be provided by the transit provider as part of the transit contract or a third party. Connectivity via the transit provider compounds customer dependence on the transit provider. A failure in physical connectivity may disconnect AS 26 from AS 23. A failure in AS 23 may result in complete or partial loss of connectivity for AS 26.

are established.[95] An *n-redundant transit bundle* is one in which an actor establishes n transit relations to ensure that if one fails, it has $n-1$ functioning transit relations, each sufficient to reach the rest of the Internet.[96]

Consider AS 25 and the scenario under which AS 22 and AS 23 provide transit services. In turn, AS 19 and AS 20 provide transit services for AS 22 and AS 23, respectively. For AS 25, the *2-redundant transit bundle* provides failure redundancy. It may also provide redundancy in some routes. AS 22 and AS 23 receive their transit routes from AS 19 and AS 20, respectively. Consider AS 25's transit routes to AS 27. AS 27 is a transit customer of both AS 20 and AS 23.

- 34.27.7.0/24 22 19 20 27 I
- 34.27.7.0/24 23 27 I

AS 25 may now *strategically* appropriate *either* route. AS 25 will likely select the second. That said, AS 25 must pay at least the minimal rate for maintaining a transit relation to *both* transit providers. AS 25 is also still beholden to lowest-cost routing decisions made by AS 22 and AS 23, respectively.

Now consider the situation in which AS 27 no longer purchases transit from AS 23, only AS 20. Following the operational rules laid out in Section 2.2, AS 23 must withdraw routes appropriated from AS 27. The result is that AS 25's routes to AS 27 will also change:

- 34.27.7.0/24 22 19 20 27 I
- 34.27.7.0/24 23 20 27 I

AS 22's path has not changed, but AS 23's has. Moreover, from AS 25's perspective, the quality of the "best" route available has been degraded. n-redundant transit is one way to develop a strategic interconnection bundle. While this bundle provides redundancy, AS 25 is still beholden to upstream, lowest-cost route optimization outside its control.

2.3.2.3 Partial Transit

Partial transit relations provide customers routes to a subset of the Internet. A network may engage in multiple partial transit relations to create an interconnection bundle with the same effect as full transit (the entire Internet is reachable). A network may also use partial transit to add redundancy for a distinguished subset of

[95] Having multiple upstreams is known as mutli-homing; one may multi-home for redundancy and load-balancing.

[96] Ideally, each of the n transit relations has independent transport connectivity c_j, where $j = 1 \ldots n$, such that a failure in transport only disconnects one transit relation. Alternately, consider an n-redundant transit relation where a single connectivity provider c is the transport supporting all $t_{1 \ldots n}$ transit relations. Failure of c could result in loss of all connectivity, despite n redundant transit providers. Multiple transport relations $c_{1 \ldots n}$ reduce the risk of failure in overall connectivity; henceforth n-redundant transit relations will be assumed to have this property. This scenario protects against failures in connectivity *to* the transit network, but not failures of connectivity *in* the transit network.

prefixes covered by existing transit. A set of interconnection relations comprising transit and partial transit will be referred to as a *mixed transit bundle.*[97]

2.3.3 Varieties of Peering

Peering is a mode of interconnection in which two networks agree to exchange traffic originating in their networks and their downstreams. BGP peering simply establishes the technical relationship necessary to exchange traffic. The contractual model considers levels of traffic exchanged in terms of on-net traffic and downstream traffic. Networks may peer directly if there is shared transport, or the potential for contracting transport, between dedicated facilities of the two networks. Peering is nominally described in terms of ASes exchanging traffic and the ratios of traffic. The exchange of payment, if any, is, on paper, governed by the ratios of traffic between the two networks. On the ground it is governed by the balance of value garnered by the two actors exchanging traffic, the cost of actually moving the traffic being but one factor.

2.3.3.1 Settlement-free Peering

On paper, settlement-free peering both parties agree that the traffic ratio between the two is sufficiently close to equal that the transaction cost of measuring the difference is higher than the value of the difference to either party.[98] Consider the relationship between AS 19 and AS 20. AS 19 may technically send more than AS 20 (AS 19 $\gtrapprox$ AS 20) but not enough to warrant the cost of monitoring and accounting. Historically settlement-free-peering was also premised on the idea that the two networks are approximately the same in terms of size and per unit of traffic operational costs.

Further consider the potential peering relationship between AS 22 and AS 23. AS per earlier scenarios, AS 22 and AS 23 are transit customers of AS 19 and AS 20 as well as transit providers for their downstream customers. AS 22 must pay transit to AS 19 for its traffic to AS 23's customers, AS 26 and AS 27. Similarly, AS 23 must pay transit to AS 20 for delivery to AS 22's customer, AS 25. Both AS 22 and AS 23 can save on transit by peering at a cost below the cost of transit. If settlement-free, AS 22 and AS 23 may recoup most of the costs of transit between themselves and their downstreams.

In terms of route provisioning, shorter routes between the customers of AS 22 and AS 23 are available. This kind of bilateral peering is associated with the flatten-

[97]For example, consider AS 25. Rather than AS 25 purchasing full transit from both AS 22 and AS 23, it could purchase partial transit from each. For instance AS 25 may purchase access from AS 23 to: *a*) 9.0.0.0/8 *b*) 34.27.7.0/24 *c*) 42.123.0.0/16 AS 25 purchases access to the remainder of SimpleNet from AS 22. Alternately, AS 25 could also purchase full transit from AS 22 and the same partial from AS 23. In the alternate case, AS 25 uses partial transit for redundancy.

[98]There is some ambiguity in the term peering for those new to the network operator community. BGP peering is a technical relationship. Settlement-free is a contractual relationship, but often simply referred to as peering. As may be obvious, this work refers to the technical act of peering as establishing interconnection.

ing of the Internet. Rather than longer, hierarchical routes, shorter "flatter" routes become available. Further, the stock of routes in the global routing table has grown.

2.3.3.2 Settlement-based Peering

In terms of traffic exchange, *settlement-based*, or *paid*, peering has the same technical characteristics as settlement-free. Paid peering differs in that one of the peers pays the other to engage in the peering relation. Consider an oversimplified case of traffic asymmetries. When the ratio is much higher than $1 : 1$ (for instance AS 22 $\gg$ AS 23), AS 22 is sending much more traffic to AS 23 than AS 23 is sending to AS 22. There may be a variety of reasons for this. A simple interpretation is that, all other things being equal, AS 22 and its downstreams are *creating* greater amounts of traffic, and subsequently greater costs, for AS 23 than it incurs for delivering AS 23's traffic. The question then becomes, for whom is the value of traffic, regardless of the volume, greater?

2.3.4 Eyeballs and Content

Interconnection economics is not quite as simple as this model. As per Faratin et al. (2007) and D. Clark et al. (2011), the heterogeneity of "ISPs," here more generally referred to as network actors, means both network infrastructure investment and traffic levels vary substantively across network actors participating in common value networks that require interconnection.

Consider the standard eyeballs and content argument. For instance, end-users in j's prefix cone may place a high value on traffic from i. Under one point of view, j should extract some of that value from those end-users to compensate for the greater costs created by end-user demand. Under a different point of view, assuming i's traffic is valuable and i is deriving value from either delivering to end-users or directly from end-user consumption, i should compensate j for the additional operational costs from that derived value. An interpretation of the former is that j is paying for content. An interpretation of the latter is that i is paying for access to end-users.

Yet a third interpretation is that neither i nor j want to have to risk user attrition by raising prices that cut into consumer surplus. Thus i and j vie for the interpretation (and bargaining position) that places the burden of extracting greater value from end-users on the other. This is a common and well-known problem of determining whether value should flow in the direction of content or counter to the direction of content.

2.4 A Clockwork Internet

The view thus far has explained the technical and economic, asserting a particular set of values, rights, and obligation, but it has not explained the institutions that

ensures the resulting rules and norms are followed. The previous sections have provided an integrated view of the NRS as a resource system that coordinates control plane operations in terms of a common pool resource framing, routing mechanics, and interconnection economics. Within the NRS, management resources have been developed to enhance networks' rights within the control plane while preserving the integrity of the control plane. Operational mechanics and economic constraints described in Sections 2.2 provide the foundation of how the resource functions. 2.3 provides the economic constraints on this operational function.

These foundations provide a conceptual distinction between positive and negative externalities in the control plane. Three classes of externalities—operational, security, and strategic—are identified. The mechanisms animating these externalities highlight the essential interdependence amongst network actors in the control plane. Appropriation and provisioning provide a foundation for reasoning about the institutions that create operational rules.

Conventional externalities are promulgated through function of the resource. In one of the canonical externality parable, that of the polluter and the laundry, black smoke is transmitted into the environment and settles on the laundry. The externality may be worse on some days than others depending on the wind and weather. This characteristic is one source of the externalities in natural resources. The control plane, in contrast, is not only man-made, but its function is exclusively animated by NRS participants. In the NRS, the "physics" of the system experienced by any given participant is a consequence of the joint provisioning practices of the other participants. For all control plane externalities, positive and negative, actors comprising the control plane and experiencing those externalities are to some degree complicit in promulgating the externality. Route flap and whether an actor chooses to aggregate to endogenize the flap or to simply pass it along is a canonical instance. Similarlly, hijacks highlight expectation that those appropriating prefix advertisements verify their legitimacy before provisioning subsequent advertisements, as in the Pakistan-YouTube incident.

Rights and obligations related to positive and negative externalities are tacit, yet well-known by many long-time network operators. The next chapter explicates these more formally, drawing classes of rights and obligations identified in empirical studies of resource systems. Function specific institutions are described in terms of the classes of rights they monitor, enforce, enhance, or diminish. Following the resource rights literature, NRS institutions are described in terms of bundles of rights from well-understood classes of rights, facilitating comparison within the NRS and with other resource rights regimes, in particular water.

Chapter 3

Rights Bundles in the NRS

NUMBER RESOURCE MANAGEMENT relies on rules, rights, and institutions to ensure the integrity of Internet operations. The mechanics and operational rules in the previous chapter provide technical foundations for understanding NRS function. Like in many CPRs, these rules were not obvious at the outset, but are the product of experimentation that lead to effective allocation of common goods supporting diverse uses amongst similarly diverse CPR participants. E. Ostrom and Schlager (1996, loc.2472–2474) indicate that "governance arrangements that have successfully coped with provision, production, appropriation, and use of common-pool resources are frequently complex property-rights systems that do not fit easily into neat and fashionable dichotomies." Moving further from the allure of "fashionable dichotomies," McKean (1996) speaks to the linkages amongst goods, rights, and the actors that provide them:

> part of our problem is semantic: we use the same pair of adjectives, "public" and "private," as labels for three different pairs of things. We use them to distinguish between two different kinds of goods (public goods and private goods), between two different kinds of rights (public rights and private rights), and between two different kinds of bodies that may own things (public entities or governments, and private entities or individuals). (1996, loc. 4206–4209)

This chapter disentangles *a*) the Internet as an infrastructure that serves as a input to public, private, and social goods from the NRS; *b*) the epistemic communities (participants) that both administer NRS resources and develop, update, and sustain rules that shape the NRS; *c*) categories of participants based on the rights and obligations these operational rules engender; and *d*) how NRS institutions make those rules, and by proxy that attendant rights, durable. These theoretical constructs, along with the mechanics in Chapter 2, provide concepts for comparing within and across studies in Part II and provide the foundations for analyses of stability in Part III.

The Internet is an infrastructure that serves as an input into a broad array of public, private, and social goods. Frischmann develops this notion of an infrastructure

in (Frischmann, 2012). The Internet's common, transnational routing infrastructure is a necessary common resource for maintaining global connectivity. The NRS is the institutional complex that manages this common resource. In the words of one CRI manager, Internet governance is not hierarchical, but rather a "complex web of authority." Section 3.1 distinguishes *a*) the Internet as an infrastructure; *b*) the control plane as a common resource managed by the NRS; *c*) direct and indirect participants in the control plane.

Direct participants in the NRS were originally a "close-knit yet loosely organized" collective of protocol designers and network engineers. These actors are framed as an operational epistemic community. Epistemic communities comprise "a network of professionals with recognized expertise and competence in a particular domain and an authoritative claim to policy relevant knowledge within that domain or issue-area," (P. M. Haas, 1992, p. 3). A critical element of CPR management is that participants livelihood is dependent on sustaining the CPR; this is also what distinguishes direct versus indirect users. Given operational epistemic communities' investment in control plane function and integrity, these actors evaluate rules based both equity and expert evaluation of observable and measurable outcomes.

A subtle but key element of the NRS is the artificial character of NRS resources. In contrast to CPRs that mediate the use of natural resources, the NRS comprises man-made number management resources that mediate the allocation and use of artificial[99] resources, numbers and routes. Within such a system, rules not only shape the resource (management) *system*, but they may also alter the structure and topology of the *resource* itself. The section of operational epistemic communities segues from a discussion of the different levels of rules developed (constitutional, collective choice, operational), to categories of rights that help distinguish types of participants.

E. Ostrom and Schlager (1996) offer classes of property rights—access, withdrawal, management, exclusion, and alienability—that shape the dynamics of CPR governance. These categories help disentangle the "fashionable" public-private dichotomy to highlight many systems are, as per McKean (1996), better framed as common resources. Configurations of Ostrom's rights highlight prototypical types of CPR participants—authorized entrants, authorized users, claimants, proprietors, and owners. In Ostrom's work and others[100] these types of participants and categories of rights facilitate reasoning about bundles of property rights in CPRs and comparing these bundles across CPRs. A critical element of these analyses is to recognize that rights configurations across CPRs are far from universal. Rather, they are contingent on, among other characteristics, the nature of the resource. Ostrom's classes of rights are used to *a*) reason about rights and obligations in the NRS as a whole; *b*) understand the rights and obligations *within* CRIs in terms of types of participants; *c*) compare bundles *across* CRIs; and *d*) establish a foundation for evaluating NRS stability in Part III. Section 3.3 describes rights and obligations

[99]Artificial is used here in the sense of Simon (1996).

[100]For interest, see the framing of political analysis by Blomquist (2012).

in the NRS in terms of Ostrom's classes of property rights. It also introduces the high-level rights and obligations engendered in NRS institutions; Part II describes NRS institutions in terms of *a*) constitutional, collective choice, and operational rules and *b*) bundles of rights from Ostrom's classes.

Each of the NRS institutions are guided by a set of constitutional norms rooted in their function specific domain, in particular elements of neutrality and exceptions for security issues. Section 3.1 links constitutional norms of resource provisioning to Frischmann's discussions of infrastructure. Before diving into institutions and bundles of rights, notions of an epistemic community are introduced in Section 3.2 Section 3.4 then introduces notions of resource rights at the NRS level, identifying Ostrom's classes of rights and bundles manifest in the CRIs. The focus of this chapter is to establish analytic frames and describe the NRS as a whole. A brief discussion of high-level CRI functions in terms of how they contribute to NRS function and integrity, the role of epistemic communities in resolving uncertainties in control plane use, and how CRIs make the resulting operational rules, rights, and obligations more durable. The chapter concludes with a transition to CRI studies in Part II.

3.1 Infrastructure

Intrinsic in CPR governance are notions of equitable, or neutral, allocation of resources. In terms of infrastructures, Frischmann discusses this in terms of nondiscriminatory access. In particular, it draws on non-discriminatory resource allocation and provisioning in the RIRs and IXes. Frischmann's notion of infrastructure also makes exceptions for security. A key exception to nondiscriminatory use is manifest in the application of IPBLs.

Global dependency on Internet communications grew rapidly, especially in the late 1990's and early 2000's. In April 1995 the NSF decommissioned the NSF backbone, transitioned the management of the Internet to the private sector. The Internet quickly became a critical input into the global economy. As dependence on the Internet grew, so too did attention to both performance and stability. Within the operations community, operational rules evolved and operators recognized the need for more durable means of monitoring and enforcing the attendant rights and obligations.

NRS institutions fill a regulatory gap by making operational rules governing the routing infrastructure of the Internet more durable. Regional Internet Registries (RIRs) ensure unique allocation and conservation of scarce number resources via the maintenance of common numbers registries. Internet eXchanges (IXes) enhance the utility (value) of number resources and route provisioning by lowering barriers to interconnection at commonly managed interconnection platforms. IPBLs monitor and reduce forms of network abuse by imbuing number resources with reputation that facilitates preserving utility for abuse victims and limiting utility of abusive actors.

Each NRS institution is itself a common resource system that enhances function-

specific rights and obligations with respect to the facilities or arena it administers. These institutions fall into domains within Internet operations that are both historically path-dependent and function-specific. Each CRI impacts broad NRS rights writ large as well as its on set of rights for managing function-specific resources. Each contributes some combination of effects: they may confer, sustain, enhance, and/or diminish resource rights in the NRS. For instance, RIR number allocation facilitate access to the routing system and maintains its own set of rights for sustaining the registry as a resource in and of itself. Although each institution aspires to ensure the integrity of the control plane, among other goals, effective coordination is not guaranteed.

NRS institutions and attendant facilities and arenas contribute to the integrity and stability of the Internet as an infrastructure. An infrastructure is a common good with substantive downstream value that, for distributional and fairness reasons, should be addressed in a nondiscriminatory manner.

> When feasible, society benefits tremendously by leveraging nonrivalry to support nondiscriminatory access to [infrastructure] resources because doing so enables the public to participate productively in a wide range of socially valuable activities. (Frischmann, 2012, p. xiii)

Frischmann goes on to argue infrastructures' spillovers, or externalities, may have positive or negative effects depending on the context and how they affect existing power relations.

Frischmann root the value of infrastructure in how it enables public participation in "socially valuable activities." The NRS supports the Internet as such an infrastructure, but participation in the NRS is limited. Direct users of (participants in) the NRS are those that provision either physical Internet infrastructure or NRS resources. For example, network actors in SimpleNet are direct users. Indirect users are those that consume downstream products such as web applications (Facebook, Google, VoIP). More indirect still are those that leverage lower barriers of communication to enhance the broad array of public, private, and social goods that benefit from Internet communication.

Frischmann (2012, p. xiv) provides the following characteristics of infrastructure resources:

1. The resource may be consumed nonrivalrously for some appreciable range of demand.
2. Social demand for the resource is driven primarily by downstream productive activity that requires the resource as input.
3. The resource may be used as an input into a wide range of goods and services, which may include private goods, public goods, and social goods.

The Internet is certainly an infrastructure resource.[101] The third point above should be immediately qualified:

[101]Frischmann discusses the Internet as an infrastructure in Chapter 13 of Frischmann (2012). Claffy and Clark (2013) discusses various platforms in that infrastructure.

> If the topic of Internet governance were taken as the investigation of the regulation of all these activities when they took place on (or were significantly affected by) the Internet, then "Internet governance" would be more or less equivalent to "law and politics" at least in the "wired" and "wireless" (or more developed) nations. (Solum, 2008, p. 49)

Going yet a step further, the Internet is a *transnationally managed* infrastructure resource. As per the distinction above, the NRS collective is the set of direct users that ensure stability and integrity. Transnationalism in international relations argues that advances in communications and travel catalyze the formation of transnational coalitions of actors that challenge state authority. Such catalysts are the spillover effects of increasingly low barriers to communication. While direct users certainly benefit from these effects, most of the actors leveraging these effects for social and political ends are *indirect* actors with motivations outside the scope of control plane integrity. Many of the societal issues addressed in conventional transnational relations overlap with the societal issues captured in broader definitions of Internet governance. NRS issues pertain to governance *of*, or governance *in*, the Internet infrastructure. Broad notions of Internet governance typically pertain to governance of issues *on*, *facilitated by*, or *catalyzed by*, Internet communication.

This work returns to, and refines, notions of direct and indirect users in terms of *a*) Ostrom's categories of CPR participants, *b*) how these affect credible commitment in CPR management, and *c*) the implications of politics of scope for regime stability. For the time being, the distinction helps distinguish immediate value derived from NRS participation for direct users and the breadth of downstream value for indirect users. The following operationalizes Frischmann's three criteria as they relate to NRS resources that contribute to Internet operation. In particular, these help distinguish the Internet as an infrastructure from the NRS as a common resource system supporting that resource.

3.1.1 Pseudo-Rivalry in Number Resources

Frischmann's first criteria is that "[t]he resource may be consumed nonrivalrously for some appreciable *range* of demand," (2012, p. xiv, emphasis added here). This criteria is applicable to number resources in the control plane, but can be made more precise:

> Number resources are not purely rival. Number resources are non-rival in some uses and may be partially-rival in other uses.

Rather than speaking to a *range* of demand, this articulation highlights *categories* of *use*. Refining to identify categories can be combined with ranges of demand. Ranges within specific categories provide a more systematic explanation of resource dynamics within the control plane.

Number resources are partially rival depending on how different actors leverage different bundles of rights. Data plane resources, whose topology is influenced by the routing decisions in the control plane, *are* rival under congestion. Moreover,

congestion is exacerbated and mitigated by non-rival provision and appropriation of routes in the control plane. Thus, flows in the data plane are better explained by Frischmann's original articulation, recognizing a demand threshold at which valuable characteristics of the resource begin to degrade.

This distinction also highlights the interplay between use of common resources and private. The distinction between number resource rights and the largely private resources that facilitate data flows is a subtle and important distinction. This distinction confounds the conventionally "fashionable" dichotomy of arguing whether the Internet, as a whole, is a public *or* private good. This does not question the utility of ideal types, rather, it challenges the use of ideal types as the source of compelling "intuitive" policy recommendations.

Focusing on infrastructure through the lens of non-rivalry in appropriation and provisioning helps untangle the difference between resources, goods produced, rights, and bodies (actors) that contribute to these processes. In the Internet, constituent networks are largely private resources. In contrast, the abstractions used to coordinate traffic amongst these private networks—the technical substrate described in Chapter 2—are common. Recalling Chapter 2, the global pool of numbers was produced by a collective of protocol designers. Routes are provisioned by private actors, but are only useful if promulgated, to varying scopes, to other actors in the Internet. As a result, the Internet is neither a public nor a private good. Rather, it is an infrastructure that leverages a commonly managed routing system to facilitate coordinated traffic delivery amongst disparate private networks. That said, demand for this infrastructure, and by proxy for control plane resources, is driven by downstream social demand.

3.1.2 Social Demand

The early Internet was a research endeavor—there was little "downstream" social demand for services atop the Internet. In terms of NRS resource production, protocol provision and operational provision was performed by the same actors. The community was tightly integrated. Since then, the two communities have diverged to a certain degree. In effect, the two communities have disintegrated into protocol and operations. The latter can be further categorized into network operations and security.

Modern Internet infrastructure development is also a disintegrated industry story. Deployment of infrastructure is driven by a disintegrated value network ultimately financed by downstream commercial use. Two coarse-grained deployment modes characterize demand for deployment. Low-value density, bandwidth intensive applications such as streaming video drive initial deployment of transport infrastructure.[102] Given stable infrastructure, more diverse sets of high-value density applications—e-commerce and business-to-business (B2B), financial data networks, and others—are fast followers that pay for infrastructure development. Along with the commercially high-value density applications, public goods, such as e-

[102]Here transport infrastructure means structures for point-to-point transport.

government services also follow.

The control plane is critical to the Internet's functionality and not all actors that use the Internet contribute to maintaining and sustaining the control plane. Relative to the end user base, the community of network operators is a very small subset of actors. When any end-to-end connection is made, the routing tables used to direct packets from source to destination and back are a product of ongoing (continuous) activities (exchanges of routing information) in the control plane. In this sense, end users "use" the control plane in the sense that their activities *on* the Internet create demand for the provision, management, and maintenance of a control plane *in* the infrastructure necessary to ensure Internet connectivity. That said, end users only shape the topology *indirectly* by their demand for particularistic services, such as streaming video delivery or low-latency access to web applications.[103] Such services want for value-network specific topologies in the control plane such as relationships that facilitate content caching at IXes or by IX participants, the development of overlay networks, or selection amongst fail-over options for extremely latency sensitive activities. These are instances of resource policy choices facilitated by the control plane.

In general, end users' demand for control plane "services" will be referred to as *indirect demand*. In contrast, the network operators directly shaping control plane topologies exhibit direct demand for control plane "service." For instance, operators make decisions about what routes advertised within the control plane they will use, with whom to establish interconnection relations based in part on topological information from the control plane. Operators may also use services like IPBLs that adorn resource units (such as prefixes and routes) with reputation that, as per the discussion above, allow them to mitigate externalities and maintain quality of experience for their users.. In both cases, operators translate indirect (derived) demand for particular services into mechanisms that enhance the corresponding value that can be elicited from control plane resource utilization.

Distinguishing between direct and indirect use helps understand how social demand is transformed into demand for infrastructure and control plane services. Frischmann offers that the infrastructure is driven by downstream productive activity. Rather than tailor infrastructure to satisfy specific demand, CRIs provide generic services that can be further refined downstream. Utilization patterns in the control plane are derived from demand patterns, but are not directly shaped by them. The value of the infrastructure is in the breadth of goods it serves as an input to, not specific goods themselves.

[103]Users demand quality of experience. Consider the two examples. Users want streaming video that is delivered quickly and plays smoothly. This translates to high bandwidth with low latency. It is arguable that these technical attributes do not mean much to the end user other than they are associated with the quality of experience that is an explicit component of their preference set. Similarly, users want interactive web applications to be responsive—again, this is technically a function of low latency but users perceive responsiveness. The control plane is similar to (and to some extent influences) these characteristics, thus it is indirectly shaped by explicit user demand for quality of experience.

3.1.3 Input Into a Wide Range of Goods

The Internet is clearly "an input into a wide range of goods and services, which may include private goods, public goods, and social goods." This breadth of goods is consumed by indirect users of the control plane, namely end users. In contrast, the NRS supports that infrastructure. A simple analysis partitions responsibility for those goods accordingly. Rather, there is a complex web of hybrid relations that govern the production of this wide range of goods. Like many other value networks, production relies on expert knowledge and potentially costly information.

The scope of goods produced, the actors that can make use of them, and the knowledge necessary to maintain the NRS relies on expert capability of direct users. Expert capability will be described in the next section, along with elements of credible commitment. Although NRS function is not directly consumed by indirect users, as per the discussion above, it is responsive to derived demand. Moreover, as an increasingly critical infrastructure, the range of goods that depend on the Internet, and by proxy the NRS, creates an obligation on direct users, NRS participants, to ensure the integrity of the infrastructure.

Recall the scope and character of externalities. It was generally argued that externalities in the control plane were promulgated by participants. For instance, actors that promulgated a hijack externality were complicit in that they did not check the legitimacy of route origin. Framed as such, this essential interdependence intrinsically creates obligations amongst network actors. Nevertheless, a common argument in a subset of the operations community, especially when faced with security externalities, is that infrastructure providers are *only* the communication channel. It is implausible to expect actors provisioning routes and private transport networks to police the breadth of goods that may be generated from the Internet as a general purpose communication infrastructure. This creates some degree of tension between regulation of Internet infrastructure, including the NRS, and regulation of the goods facilitated by the Internet communication. The breadth of goods produced is another lens through which to consider the difference between creating order in the Internet and, again invoking Solum, claiming "Internet governance" has some causal bearing or authoritative position in creating order in International *society*.

Rather than falling into another "fashionable" false dichotomy, distinguishing between the resources produced to order the NRS and how they affect the breadth of public, private, and social goods helps further refine these two scopes of governance. Moreover, this helps identify where they are complementary rather than competing. NRS participants *do* have public policy obligations. They are not of the same character as conventional public policy makers—NRS participants are not making social policy. Ideally, industry and resource policy should avoid confounding public policy efforts at policing the breadth of goods engendered by Internet communication and, where efforts do not limit the general value of the infrastructure, should contribute to the knowledge and tools necessary for authoritative governance actors to continue legitimate regulatory efforts. This is a subtle balance that has eluded this mode of cooperation in the past.

NRS institutions have an obligation to provide information to those whose public policy obligations are increasingly dependent on Internet communication, and NRS function in particular. The structure of indirect (derived) demand discussed in the first two of Frischmann's criteria provides insight into the set of paths actors in downstream markets have to influence the topology and performance of the control plane. Non-market actors, such as law enforcement, also have a legitimate claim on facilities' services. An accurate registry is not only an infrastructure resource (good), but it is also a public good that facilitates, among other benefits, effective law enforcement investigations.[104] A balance must be struck between the current operational practices that facilitate efficient and efficacious Internet functionality, what will be referred to as the core function of CRI institutions, and obligations to the public policy implications of these infrastructure decisions. To this end, four conceptual classes of policy will be discussed—protocol, resource, industry, and public policy. These helps conceptually identify the sources of authority and obligations while also scoping issues to collectives currently involved.

CPR management often relies on operational and experiential expertise to make effective policy and design decisions. Evaluating the scope and effects of CPR management on adjacent actors also requires a deep understanding of resource function and resource system dynamics. In general, the interface between the broad set of goods produced and general-purpose infrastructure service requires credible assessment from those with obligations to downstream goods and those provisioning infrastructure. The next section describes the notion of an operational epistemic community and their role in internal and external policy processes.

3.2 Operational Epistemic Communities

Knowledge problems are endemic in CPRs. The deep understanding of how the control plane operates in situ that facilitates development of operational rules is not the product of conventional scientific studies. Rather, it is the product of collaboration amongst a community of professionals vested in developing and sustaining a domain of knowledge around routing operations. These types of epistemic communities contribute to continuously solving knowledge problems. They also contribute to the validity of consensus processes. In terms of Ostrom's layers of analysis, the various consensus processes are collective choice mechanisms used to create operational rules.

3.2.1 Operational Knowledge

Epistemic communities in this context are operators whose knowledge-base is derived from operational experience. E. Ostrom (1990) refers to these types of actors in discussing resource systems as those experts necessary to solve knowledge problems related to resource management. Ostrom evokes similar references to the

[104]This is developed further in Chapter 5, in particular in discussion of a recent policy proposal on privatizing part the ARIN registry.

uncertainty in how many resource systems work and the trial-and-error processes that some communities have used to develop rules for governing CPRs:

> Uncertainties stemming from lack of knowledge may be reduced over time as a result of skillful pooling and blending of scientific knowledge and local time-and-place knowledge. (1990, p. 33)

The network operator community engages in this form of pooling: they are experts in network operations but leverage a consensus process to navigate uncertainty and decide which policy experiments to undertake. If these policy experiments fail, they revisit that policy. Operators have distinct incentives to do so as direct users of the resource. Failed policy has implications for their value-propositions. As noted by both Cole and Ostrom, uncertainty is never eliminated and "remains even after one acquires considerable knowledge about the resource system itself,"(1990, p. 34).

Solving knowledge problems is critical to understanding both CPR function and the effects of operational rules that shape rights and obligations. In some systems, knowledge accumulates with use and experience. For instance, Ostrom indicates that:

> Rules devised by resource users are based on years, decades, and sometimes centuries of experience in using a common-pool resource. Such information is gleaned while engaging in everyday harvesting activities. Fishers learn which spots in a fishing ground are most productive and which areas of the grounds are most compatible with various types of gear, by fishing day after day. Consequently, the rules that resource users devise are well matched to the physical environment in which they will be used. (1996, loc. 2758–2761)

Experience is a valuable tool referenced again and again in the CPR literature, but it does not necessarily have to be ad hoc.

Consider a seemingly different knowledge problem, improving the efficiency of hydraulic turbines in the mid-19^{th} century. Layton (1979) tells a story of the development of the hydraulic turbine in late 19^{th} century America. Layton offers three phases in the development of industrial research in the United States. Most salient here is the flawed notion that technology flows "directly" from scientific theory.

> One Swiss engineer argued that the Europeans' devotion to mathematical theory inhibited innovative design on the continent. Designers tended to limit themselves to cases that could be handled by theory. (Layton, 1979, p. 76)

Layton argues hydraulic engineers' "antipathy" for mathematics is not "based on ignorance," (1979, p. 76). Rather, it was a pragmatic response to models that did not account for problems encountered in existing[105] system contexts.

[105]The collected essays in (Cole & Ostrom, 2012a) stress the difference in existing rights regimes and ideal types. Here, this manifests in the difference between systematic inductive prescriptions derived from empirical evidence versus the much-praised eloquent, parsimonious theoretical construct.

Layton offers the hydraulic engineers' critique:

> Perhaps more serious, an idealized mathematical theory encouraged the neglect of factors that were left out of the theory. That is, physical justifications were found for mathematical conveniences. (1979, p. 76)

Layton goes on to quote one of the leaders in the emerging design paradigm:

> [T]he turbine has been an object of deep interest to many learned mathematicians, but up to this time, the results of their investigations... have afforded but little aid to the hydraulic engineer. (Layton, 1979, p. 77)

This is not a denial of the utility of mathematical modeling or reasoning about a system. Layton argues that "[i]dealization is fundamental to understanding the interaction between science and technology," (1979, p. 77). It is important not to mistake parsimony and simplifications necessary to internal consistency with knowledge contributing to pragmatic application. Layton further affirms, indicating "the elements omitted in an idealized theory are often of vital importance to the technologist," (Layton, 1979, p. 77).

A key element of Layton's narrative is a Baconian approach to industrial research.[106] In contrast to Newtonian research, Baconian focuses on applied approaches that contribute to a pragmatic problem facing an industry or society as a whole. In contrast to deductive models of knowledge generation, an inductive approach rooted in systematic experimentation drives efforts to improve particularistic outcomes. Layton's hydraulic engineers are operating in what is claimed to be the early stages of industrial research (Layton, 1979).

---

[106]Holton (1999) distinguishes between Newtonian, Baconian, and Jeffersonian modes of research. Newtonian research is defined as

> The concept of pursuing scientific knowledge "for its own sake," letting oneself be guided chiefly by the sometimes overpowering inner necessity to follow one's curiosity, has been associated with the names of many of the greatest scientists, and most often with that of Isaac Newton.

Holton (1999) argues Baconian research is

> popularly identified with "mission-oriented," "applied," or "problem solving," we find ourselves among those who might be said to follow the call of Francis Bacon, who urged the use of science not only for "knowledge of causes and effects and secret motion of things," but also in the service of omnipotence: "the enlarging of the bounds of the human empire, to the effecting of all things possible."

Jeffersonian is what Holton argues is a conscious mixing of the two:

> The specific research project is motivated by placing it in an area of basic scientific ignorance that seems to lie at the heart of a social problem. The main goal is to remove that basic ignorance in an uncharted area of science and thereby attain knowledge that will have a fair probability—even if it is years distant—of being brought to bear on a persistent, debilitating national (or international) problem.

Both Layton's hydraulic engineers and the network engineers discussed as an epistemic community in Section 3.2.3 are engaged in modes of Baconian knowledge creation and, perhaps more importantly for this work, credible knowledge assessment.

Network operators are arguably in a similar phase of developing industry knowledge of Internet operations as Layton's hydraulic engineers. Network operator groups, the topic of Chapter 4 are fora for sharing current operational experience within the operator community. Simple behavior modes are useful for conveying gross trends, but often fail to inform pragmatic decision processes. Moreover, the community has argued there is a dearth of formal network operations education. Most network research is dedicated to the low-level communication theory, often shoe-horning network operations into simplified economic models.[107] The result is that substantive operational knowledge remains in guild-like structures—accessible to new actors but requiring ongoing engagement, time to learn the community's technical vernacular, experience with the phenomena.

3.2.2 Natural and Artificial Systems

In her introduction, E. Ostrom and Schlager (1996) asks:

> Are locally developed institutions, which rely on knowledge acquired over time, effective when modern science provides better ways of managing local resources than those of individuals who are illiterate and lack knowledge of modern scientific techniques?

E. Ostrom and Schlager (1996) makes a seemingly stark contrast between "modern science" and the rural communities that are often the subject of CPR studies. In this work, the NRS is a commonly managed resource *supporting* a complex technical infrastructure. Moreover, its participants have substantive technical capability. Although some cases of rural development imply early ad hoc knowledge development, knowledge about these systems is not always trial and error. Rather there is a mix of serendipitous discovery over the course of use and systematic operational experimentation that manifests as experience. In mature, professionalized systems, this knowledge is often codified for use by other system participants. In applying the CPR framing to largely man-made resource systems such as the NRS, this work distinguishes amongst three coarsely categorized "epistemic" communities: *a*) those most often considered epistemic communities, academics and scientists; *b*) Ostrom's rural communities;[108] and *c*) operational, or Baconian, communities of technically knowledgeable professionals working to improve system outcomes.

Although there are differences in the cultures of epistemic communities managing CPRs, they share *institutional* problem solving constructs evidenced across the CPR literature. Following the spirit of Baconian research illustrated in Layton's narrative, institutional problem solving arenas started with informal mechanisms.

[107] Over the course of fieldwork multiple actors have described academics as attempting to fit network operations into their favorite model, often simplified economic models. This has engendered skepticism of academic models.

[108] This category plays off the earlier quote, it is not intended to imply that Ostrom's studies focused exclusively on less educated rural communities. Rather, Ostrom's work spanned a variety of communities. That said, the following section will develop the notion of knowledge derived from operational capability and experience.

	Natural	Artificial
Ad hoc	rural pastureland	early volunteer IXes
Systematic	irrigation, basins	Internet

Table 3.1: Modes of communities and resource systems differ in terms of how systematically the system characteristics are evaluated and whether the system is natural, man-made, or some mix.

These became arenas for debating operational rules in terms of their technical implications and the value of outcomes. In Layton, the turbine industry served as the arena and outcomes were easily measured in terms of resulting efficiency. In the NRS, institutional decision fora are the loci of collective-choice processes driven by technical outcomes. These decision processes are each argued to follow the spirit of the IETF consensus process but have different degrees of formality and structure.

Further consider the difference between man-made resources and natural resources. As per earlier discussion, allocation processes are mediated by the resource system—man-made infrastructures and institutions enhance resource rights by making the attendant operational rules more durable. In the NRS, rules not only ensure a common image of equitable allocation of resources, these actors are also reasoning about the nature and function of the resource itself. Conventional CPRs managing natural resources have mutable resource systems that facilitate extraction: various constructs such as irrigation gating systems, netting, etc. CPRs in which the resource itself is man-made have additional degrees of freedom. Operational rules can have much greater influence over resource dynamics, but must still cope with the uncertainty intrinsic in making such changes.

E. Ostrom and Schlager (1996) note that, like other collective action problems, credible commitment to rule making processes is necessary. Like Layton's hydraulic engineers, NRS participants are engineers that have a vested interest in the operational efficiency of the networks they manage as well as the larger control plane. Routing mechanics, scope of externalities, and the essential interdependence of participants by virtue of the control plane demonstrate dependencies that foster a commitment to the integrity of the control plane as a common resource. This is one element of credible commitment to developing effective institutions. Credibility is also rooted in the collective judgment of the community. In particular, collective-choice rules manifest as consensus processes in the NRS. The next section (3.2.3) introduces the notion of an epistemic community as the those that generate domain specific knowledge. Section 3.2.5 develops the baseline notion of consensus and transitions to the studies of existing NRS institutions in Part II.

3.2.3 Epistemic Communities

P. M. Haas (1992) presents the notion of an epistemic community in the context of international issues faced with technical uncertainties. Here, the notion of an epistemic community is developed as one factor in credible commitment to CPR institutions.

According to P. M. Haas (1992):

> An epistemic community is a network of professionals with recognized expertise and competence in a particular domain and an authoritative claim to policy-relevant knowledge within this domain or issue area. Although and epistemic community may consist of professionals from a variety of disciplines and backgrounds, they have (*1*) a shared set of normative and principled beliefs, which provide a value-based rationale for the social action of community members; (*2*) shared causal beliefs, which are derived from their analysis of practices leading or contributing to a central set of problems in their domain and which then serve as the basis for elucidating the multiple linkages between possible policy actions and desired outcomes; (*3*) shared notions of validity—that is, intersubjective, internally defined criteria for weighing and validating knowledge in the domain of their expertise; and (*4*) a common policy enterprise—that is, a set of common practices associated with a set of problems to which their professional competence is directed, presumably out of the conviction that human welfare will be enhanced as a consequence. (1992, p. 3)

P. M. Haas (1992) goes on to stress that the notion of an epistemic community is often presumed to refer to scientific communities.

Haas indicates this need not be the case. Additional elements of epistemic communities are noted:

> [M]embers of an epistemic community share intersubjective understandings; have a shared way of knowing; have shared patterns of reasoning; have a policy project drawing on shared values, shared causal beliefs, and the use of shared discursive practices; and have a shared commitment to the application and production of knowledge. (1992, Footnote 5 on p. 3).

Taken together, these characteristics form what has been referred to as a *common image*. Ostrom speaks to the notion of a common, or authoritative image of the dynamics of problem, derived from either community experience or the analysis of third party experts.[109] Such a common image facilitates negotiation amongst common resource participants, here members of an operational epistemic community, from a common set of operational principles. Haas' elements of epistemic communities above add elements of process, i.e. how the image of *system* characteristics and dynamics is managed. The notion of a common image will be used as defined here throughout the studies in Part II; in Part III the common images of each CRI will be reconciled with both a notions of social order in the NRS writ large and models of authority to evaluate both the stability of the NRS and challenges to engagement in the global political arena.

[109]See E. Ostrom (1990, p. 112) and the discussion in Chapter 8.

Returning to the discussion of epistemic communities, these notions of an epistemic community provide a foundation for evaluating NRS participants' role in creating knowledge about the control plane, the NRS, and CRIs. The four criteria will be elaborated below with respect to the NRS writ large. In many cases, particular instances are better presented in the appropriate study chapter. Part II provides empirical evidence of these based on processes for developing operational, collective-choice, and constitutional rules in NRS institutions.

Following Haas' framework, the network operators participating in the NRS are more akin to a professional group. Participants sustaining CRIs, on the other hand, arguably constitute particular epistemic communities. Each CRI hews to particular normative and principled beliefs, typically linked to their specific function within the CRI. These will be discussed in terms of constitutional rules and norms. Haas' criteria have distinct conceptual overlaps with Ostrom's layers of analysis. These will be drawn out in the CRI analyses in Part II.

That said, the set of principled beliefs *common* to the three NRS institutions evaluated here is the commitment to consensus-based decision making. In general, the nondiscriminatory character of RIR and IX governance result in non-normative evaluation of NRS management. In contrast, anti-abuse rules derived from industry policy have a distinct normative character. While consensus-based decision processes differ across each of the studies, the common image of the meritocratic ideal typically affiliated with the IETF consensus process has been articulated in each community. Shared causal beliefs, especially as they relate to control plane mechanics, dynamics, and economics, is derived from shared experience managing Internet connectivity and information presented in fora such as NOGs and other community fora. As implied by previous discussion and the character of the resource system, validity is rooted in comparative discussion of empirical instances of phenomena in institution fora. In these fora, evidence of phenomena confirmed by other respected members of the community often vetted for legitimacy and may later become authoritative as principles are further affirmed. Combinations of repeated exposure to similar experiences and systematic analysis by either individual actors in the community or CRIs with a non-core remit for research further confirm shared causal beliefs.

Similar to normative and principled believes, the common policy enterprise has both an NRS image and specific CRI images. Across the NRS, the integrity of the control plane is a broad common policy enterprise, but is approached differently by each of the CRIs. This difference is in part a function whether the policy is protocol, operational resource, or industry policy. In the RIRs and IXes, this common enterprise is manifest in ensuring operational rules (resource policy) sustains control plane integrity. In the anti-abuse communities, industry policy is the normative source for rules that shape IPBLs choice to imbue numbers with reputation. Although there are more nuanced differences between RIRs and IXes, the different motivations for the policy enterprise between overlapping RIR and IX communities and the anti-abuse community results in competing normative beliefs across CRIs.

Currently, this divide between the two epistemic communities largely manifests within the NRS. That said, tension within these communities is not being addressed.

In the best case, it will continue to be dealt with on an ad hoc basis by informal social networks. In the worst case, tensions will mount and differences in the normative foundations of these two epistemic communities will begin to become apparent to outside regulators.

3.2.4 Credible Commitment

A superficial view of epistemic communities might imply that these actors operate in unison. In contrast, the operational epistemic community leverages modes of validation to confirm principled and causal beliefs that serve as the source of their authority. Two interdependent modes of credibility contribute to this process. The first is the character of the epistemic community and the focus on outcomes over advocacy (P. M. Haas, 1992, p. 10). Haas provides a succinct characterization of the balance of "scientific knowledge" and authority. Haas invokes Barnes and Edge[110] arguing:

> science is near to being the source of cognitive authority: anyone who would be widely believed and trusted as an interpreter of nature needs a license from the scientific community. (P. M. Haas, 1992, p. 11)

Haas goes on to indicate that

> policymakers and leaders typically expect to remain in control even when delegating authority... even in cases involving what is regarded as a technical issue, policymaking decisions generally involve the weighing of a number of complex and nontechnical issues entering around who is to get what in society and at what cost. Despite the veneer of objectivity and value neutrality in achieved by point to the input of scientists, policy choices remain highly political in their allocative consequences. (P. M. Haas, 1992, p. 11)

Haas also indicates that this is especially the case where scientific evidence is ambiguous.[111]

Participants in the NRS engage in a form of credible knowledge assessment. McCray (2003) provides a compelling discussion of knowledge assessment:

> A full knowledge assessment is taken to have three essential characteristics. First, it undertakes to consider, in part or in full, the available *scientific and other factual* [e.g., economic, enforcement-related] *basis* for policy choice. Second, it reflects the *collective judgment* of a set of experts in the pertinent domains of knowledge, and not just the views of a single individual. Because most policy decisions impinge on several fields of expert knowledge [physics, human biology, economics, etc.] there is a need to synthesize them, and knowledge assessment that is collective

[110] P. M. Haas (1992) attributes this quote to Barnes and Edge (1982, p. 2).

[111] For an elaboration on this situation, see Jasanoff's discussion.

> can do this. Third, the assessment admits both what the experts take to be the objective evidence and their subjective, but informed views—courts have called this the *"reasoned feel of the expert."* Because of the presence of the second and third factors, there would be hazard in typifying knowledge assessment itself as strictly "objective" or "scientific" process, even when it is controlled by scientists of broad renown. (2003, Footnote 1 on p. 1)

These characteristics are present in the evaluation of operational rules.

The first criteria, "scientific and other factual bases" is derived from the activities of the NRS as an epistemic community. Each of the CRIs engages in its own processes of collecting and evaluating CRI data. Following the parallel with Layton, these are not necessarily scientific in the sense of derived from strict deductive mathematical models or theoretical constructs. Processes vary in how systematic data collection and analysis are. Like the value of the water turbine, they driven by the value proposition of the CRI.

The second criteria maps onto the validation characteristic of epistemic communities. As evidenced in SimpleNet, no singular individual or firm has sufficient purview over the system to administer the Internet holistically. The necessity to coordinate and share information has also manifest in evaluating and promulgating best practices. While no singular actor has a complete purview, operational knowledge does facilitate critique and evaluation of claims made by others. Operator lists are rife with discussions of operational behaviors modes and discussions of their implications. In many of these arenas, the collective does have a say, but following the meritocratic spirit inherited from the IETF model, there are distinguished actors considered experts in the community. Some of these are well-known actors in the community, others are employees of CRI firms.

McCray's notion of collective evaluation includes both validation within the community and outside the community. Within CRIs, credible assessment is argued to function effectively; see discussions of consensus processes and credible assessment in Part II. Across the CRIs there are instances of coordination, but also instances of tension. As noted in the discussion of infrastructure, in particular the breadth of public, private and social goods provisioned downstream, combined expertise regarding NRS infrastructure and downstream goods is necessary. This collective expertise is especially salient for reasoning about public and social goods. In Part III discussions of transnational policy development will be evaluated in terms of credible knowledge assessment and credible commitment to ensuring a neutral, nondiscriminatory infrastructure.

In the NRS, a combination of systematic measurement processes[112] and experience contribute to the "reasoned feel of the expert." Layton offers four phases of industrial research development:

[112]For the time being, this work will use the term "scientific" sparingly given it of connotes conventional top-down, deductive modes of inquiry. As may be obvious from the invocation of Layton, this work does not privilege that mode of over other systematic forms of developing epistemic knowledge.

1. Inefficient and cheap turbines; simplified or idealized scientific principles applied
2. Set a scientific style for industrial research that laid the foundation for engineering science
3. Large-scale production and expansion of industrial research and development
4. Academic theorists recruited and development of turbine science

Generalizing these stages, the NRS and Internet infrastructure management writ large seem to be in-between phases 2 and 3. It is argued that the reasoned feel of the expert is part of an informal inductive epistemic process within the community. This process is manifest in the variety of informal and formal information sharing practices.

3.2.5 Consensus and Operational Rules

Consensus is defined as "a position or an opinion reached by a group as a whole," (Pickett, 2002, p. 304). The dictionary goes on further to indicate "*general* agreement," (Pickett, 2002, p. 304, emphasis added). The dictionary definition is presented as a starting point to highlight the breadth of the concept. Breadth of the concepts is checked not by providing strict positive guidelines for what consensus *is*, but rather, what consensus *is not*. Consensus is generally framed as a process. Moreover, it is often presented in pragmatic terms and relates the experience of those attempting to create consensus on an issues. Framing it as what it is not is a way to highlight pitfalls experienced by consensus facilitators. Although the process is the primary topic here, the parallels between articulating consensus in terms of experience with the processes and similar elements of operational epistemic communities should at least be noted.

The first and foremost of assertion is that consensus is not majoritarian voting. Voting is an expression of preferences. Voting as a process does not constitute mechanisms for sharing collective knowledge of an issue, topic, or proposed action. Sharing collective knowledge is ideal, but voting as a mechanism does not ensure that. In contrast, consensus processes are framed as a continuous process from the point at which a proposal is made to the point at which a group agrees that consensus has been reached.

That key phrase, *at which a group agrees*, highlights a key element that differentiates consensus from voting. In voting, minority views need not consent to the outcome. Minority views may continue to be vehemently opposed to the outcome. In consensus processes, building from the definition of consent is again a good starting point: consent is "acceptance or approval of what is planned or done by another," (Pickett, 2002, p. 304, n. 1). In a consensus process, minority dissent is met is encouraged as a means to ensure views that may be over looked by those supporting a position are heard.

A number of interviewees across the NRS have suggested that their consensus process is not the same as, but was derived from experience with and the influence

of the IETF consensus process. Consider the view of the minority position:

> [U]sing rough consensus avoids a major pitfall of a straight vote: If there is a minority of folks who have a valid technical objection, that objection must be dealt with before consensus can be declared. This also reveals one of the great strengths of using consensus over voting: It isnt possible to use "vote stuffing" (simply recruiting a large number of people to support a particular side, even people who have never participated in a working group or the IETF at all) to change the outcome of a consensus call. As long as the chair is looking for outstanding technical objections and not counting heads, vote stuffing shouldnt affect the outcome of the consensus call. (Resnick, 2014, p. 14)

Resnick's discussion highlights a number of a characteristics of the consensus process. A key antidote to strict majoritarian positions is the requirement for a valid technical objection.

The notion of validity will be addressed subsequently, but in the face of such an objection, the group must accommodate the objection. Accommodation may include a number of strategies. The portion of the proposal at the heart of the objection may be reconsidered. This often invites the facilitator to ask for further elaboration for why the objection stands and if the objector has potential solutions to the pitfalls presented. Moreover, it is the duty of those supporting the proposal to consider this objection, present counter arguments and/or compromise strategies that will satisfy the objector's concerns.

Resnick gives two instances to drive home the complementary ideas that *a*) consensus is not voting and *b*) rough consensus can exist when objections have been addressed. The first instance is the case where a facilitator has 95 for a proposal and 5 against. If none of the 95 for can address the legitimate issues of the 5 against, the group cannot come to consensus. If on the other hand, the some subset of the 95 do present an accommodation that either *a* resolves the issue or *b* is a valid explanation for why the objection has been dismissed. The latter, valid explanation, requires that

> the group must have honestly considered the objection and evaluated that other issues weighed sufficiently against it. Failure to do that reasoning and evaluating means that there is no true consensus. (Resnick, 2014, p. 9)

The second instance, 5 people or and 100 against, is what Resnick suggests is the "real mind bender for most people, and certainly the most controversial," (Resnick, 2014, p. 15). The degenerate scenario is 5 active and an active member of the group objecting. Resnick offers the scenario that the objector has an alternative protocol than that with more support, and that alternative works especially well on the hardware produced by the objector's company. The objector's proposal to use the alternate protocol is dismissed on grounds of an existing sufficient solution, that the alternate is not sufficiently better, and that it is not yet as proven on the diversity

of hardware as has been the supporting protocol. The recalcitrant objector does not accept this dismissal and recruits previously inactive observers to join his cause, Resnick indicating that some of these may be from marketing and sales, creating a form of vote stuffing.[113] Even if the mass of objectors materialize, they are all making the same argument. The solution on the path to consensus is to ask if these actors have additional arguments beyond the original argument presented by the objector, those that address the original dismissal. Resnick argues this vote stuffing finds a mass of objectors that is not prepared to make further argument. Consensus here is "rough consensus in the extreme" but still holds to the spirit of finding the best technical solution.

Although controversial, this illustrates consensus's resistance to vote stuffing (and other modes of gaming, such "horse-trading") and the focus on "counting" legitimate arguments for an against rather than counting votes. The horse-trading argument is subtle, but very pertinent here. Under such a system, one's support for or against a proposal is not fungible like votes in a majoritarian system. Returning to the beginning of this discussion, voting is a discrete expression of preferences, it is not a process of evaluating solutions. Other elements of majoritarian voting systems serve the evaluation function. The process of finding consensus is a constrained exploration of the compromise space within a technical issue that weighs arguments for or against but does not privilege the number of actors simply for or against.

Consider the more confounding version of the 5 for and 100 against problem: the 100 against are not sales and marketing, but rather are engineers that can articulate nuance of the objector's argument. The facilitator must now determine if there are new issues raised here or if it is the same argument over again. Resnick indicates that:

> new people to the conversation are developers or others who are directly involved in creating the technology, or even folks who have been participating the entire time whos knowledge of the technology is not in question at all, the principle is still the same: If the objection has been addressed, and the new voices are not giving informed responses to that point, they can still justifiably be called "in the rough". (Resnick, 2014, p. 16)

Again, even if the participants are knowledgeable, the arguments for and against within the compromise space has not changed—no new compelling issues have been presented that the supporters of the existing algorithm have not yet adequately addressed have been presented.

Finally, in both Resnick's narrative and NRS studies, the role of facilitator is key. Returning to McCray's reference to the "intersubjective" , intersubjective reasoning and evaluation are key. Across the board, actors refer to the experience of the facilitator, their understanding of catalyzing a consensus process, guiding it without

[113]As an aside, across the communities in these studies, actors from marketing and sales are not especially welcome in discussions if they cannot discuss engineering details and do not make operational decisions. As such, this form of vote stuffing violates multiple norms, but this discussion will stick with the act of vote stuffing itself.

unintentionally biasing or creating the impression of premature consensus, and finally knowing when to "call" rough consensus. In some communities, this is the role of experienced actors and is an informal evaluation on the part of longstanding actors in the community. In others, facilitation is actively encouraged and in some cases, training in meeting facilitation is provided and encouraged.

Resnick's articulation of consensus processes is presented as both a well-developed discussion consonant with other articulations of consensus. The forms of consensus identified in the NRS studies follow the spirit of IETF consensus. On first pass, there are structural differences. As will be elaborated below, some consensus processes involve both discussion and e-mail list engagement, as well as eliciting participation through various informal public channels and back channels. While the IETF also uses these channels, consensus processes in the following focus on different elements of the consensus process.

3.3 Premises of Property Rights Regimes

Rights and obligations in CPRs are not static sets of "natural" rights. E. Ostrom and Schlager (1996) discusses the distinction between property rights and rules:

> The term "property rights" and "rules" are frequently used interchangeably in referring to uses made of natural resources. The way that we use these terms is to recognize that "rights" are the product of "rules" and thus not equivalent to rules. E. Ostrom and Schlager (1996)

As the Internet grew, epistemic communities of network operators saw the need to continuously update rules and, subsequently configurations of rights and obligations. E. Ostrom (1990) describes CPR decision processes in terms of *a*) constitutional rules (norms), *b*) collective choice rules for deciding constitutional and operational rules, and *c*) operational rules that shape day-to-day operational decision-making. These rules shape the configuration of rights and obligations within the NRS as a whole and individual NMRs.

Cole and Ostrom (2012a) reviews property regimes rooted in empirical studies of *existing* collective resource rights management systems. Cole and Ostrom (2012b) argue that much of the work on property theory relies on "simple models of private panaceas" and that there is a need to "develop a more descriptively accurate and analytically useful theory of property systems and rights in natural resources," (2012b, loc. 1070). Cole and Ostrom, via Eggertsson, argues that Demsetz offers a "'naïve theory of property rights,' according to which the entire history of civilization is an inexorable, unidirectional movement toward private-individual ownership of land and other natural resources," (2012b, loc. 1102). A common theme in Cole and Ostrom's collection is the identification and analysis of existing (empirical) resource systems that exhibit the variety[114] of characteristics that

[114]Hence the title of (Cole & Ostrom, 2012b), *The Variety of Property Systems and Rights in Natural Resources*.

deviate from the panaceas and ideal types offered by Demsetz (1967) and Hardin (1968). Rather, the collection explicitly elicits deviations from theoretical "ideal" types. This work draws on a number of those generalizations, in particular, the deviations observed in water rights in comparison to land rights.

Earlier it is noted that number resources and routes do not have intrinsic value themselves. Rather, they derive value from the wide variety of uses they facilitate for actors in the Internet infrastructure industries. The content and eyeballs problem is as old as the Internet and is rather intuitive. Focusing only on that generalization can obscure the variety of value networks in the infrastructure industry and the diversity of uses *facilitated* by number resources and routes. Some of these value-networks add nuance to the eyeballs and content generalization, others supplement: *a*) content delivery networks, *b*) infrastructure-as-a-service (IaaS), *c*) online retail services, *d*) online payment services, *e*) social networks, *f*) messaging services, *g*) financial exchanges, *h*) specialized data transactions, *i*) gaming networks, *j*) security mitigation networks, *k*) government services.

Among others, Blomquist (2012) argues that the notion of a multifunctional resource

> signifies that a particular resource yields more than one stream of value ... [f]or example, a forest may be simultaneously a source of timber, shade, open space, and habitat, a buffer between communities, a place of recreation, and a place of spiritual significance, as well as providing oxygen and consuming carbon dioxide. (2012, loc. 9482)

Not surprisingly, these uses are not always compatible, and specific appropriators vie for rules that privilege their use over others. Blomquist (2012) goes on to indicate that "[t]he essential point is that property rights in a natural resource can be, and often are, defined or limited *functionally*," (2012, loc. 9482). The result is a common resource system, here the NRS, administered through fragmented bundles of resource rights.

Fragmentation occurs along functional-boundaries. Fragmentation can, but does not necessarily, lead to instability amongst those managing bundles in the larger system. As will be discussed in the conclusion of this chapter and later the analysis of Part III, one solution is a federated model, distinct from conventional hierarchical models produced by the state system. Ostrom's categories of rights are used to describe classes of *property rights holders*: authorized entrant, authorized user, claimant, proprietor, owner.

Comparison with water rights is a common theme throughout this dissertation. The next section roots notions of fragmentation and multi-function[115] resources in outcomes from water resource management. Section (3.4) reviews the framework offered by E. Ostrom and Schlager (1996) to characterize sets of rights. These sets of rights speak to the NRS writ large. In the course of this characterization, the role

[115]In conventional framings, multifunction speaks to the variety of uses of a given resource. These uses often conflict. Frischmann's framing of infrastructure is a variant of multifunction. Rather, the general purpose character is an input rather than having direct uses.

of function-specific institutions in the NRS are identified. Section 3.5 completes the characterization by describing types of rights holders—common sets of rights identified in existing CPRs—and identifying the NMRs that affect these rights.

3.3.1 Fragmentation and Multiple Uses

Cole and Ostrom highlight the narrow perceptions of property rights, challenged by existing, empirical regimes:

> Resource economists were preoccupied with private-individual property rights and often equated ownership with the right to alienate.[116] The focus on just a few specific private property rights was, of course, myopic and limited understanding of the wide variety of existing property systems for a long time. (Cole & Ostrom, 2012b, loc. 1148–1150)

Preoccupations with narrow notions of ownership and particular modes of alienation occur in the NRS. In particular, transfer rights related to number resources in the RIR system are becoming increasingly important with the impending depletion of previously unallocated IPv4 space.[117] Section 3.4.5 will present the conceptual origins of transfer (alienation) rights; the discussion of transfer policies in Section 5.7.3 will highlight evidence of this preoccupation in discussions of ownership and the limitations and prescriptions related to transfer the transfer of number resource rights. Cole and Ostrom stress that simple solutions are not only insufficient, but they obscure the fit of fragmented, albeit more complex, systems of rights and obligations that have evolved in communities managing common resources.

Epstein (2012) provides a compelling contrast between the provenance of early reasoning in land rights, and resource rights in general, with those of water. In discussing notions of possession of land as societies grew, Epstein invokes Blackstone:

> But when mankind increased in number, craft, and ambition, it became necessary to entertain conceptions of more permanent dominion; and to appropriate to individuals not the immediate *use* only, but the very substance of the thing to be used.[118]

The need for more permanent[119] dominion is a manifestation of stable property rights over a relatively static resource. Dominion does not necessarily mean all use

[116]Cole and Ostrom attribute this to Becker (1977). See the section entitled *Well-Defined Property Rights* in E. Ostrom and Schlager (1996) for a discussion.

[117]The last /8 blocks were allocated on 3 February 2011. The RIPE and APNIC regions have already run out. Other regions are expected to run out within a year of this writing. As per Chapter 2, addresses are rival in terms of rights of origination, but remain information goods that are, modulo rights limitations, are easily transferable. Primary concerns relate to the implications of transfers on the integrity of the registry. In particular, see Section 5.7.3.

[118]Epstein (2012, loc. 7999), citing Blackstone (1979 [1766], 2:3-4).

[119]In few cases in the NRS is dominion permanent. Legacy allocations in the RIR system is an instance where informal, underdeveloped rights have led to the perception of permanent allocation of number resources. In other NMRs, rights are often explicitly limited to adhering to operational and constitutional rules of the system.

rights. For instance, here the "substance of the thing" refers to the properties that make a resource valuable. In the case of number resources, that is uniqueness. This does not mean absolute control over all possible uses of number—that would substantively limit use of number resources as well as routes.

In the NRS, numbers are the static (finite) resource. Routes are much more dynamic, being provisioned and appropriated in many different parts of the system. Similarly, water is a much less static, easily bound resource. Epstein goes on to state:

> The difficulties in the context of water are greater than they are with land because of the far greater diversity of circumstances that any comprehensive system of property rights has to address. Some of these problems relate to the highly varied settings in which water rights have to be organized. That physical diversity in turn increases the need to balance competing uses, so that much of the major litigation on the question deals with whether this system governs in the first place. (Epstein, 2012, loc. 7999)

The "far greater diversity of circumstances maps" onto a greater diversity in uses Blomquist (2012) highlights in his notion of multifunction resources; it also speaks to the breadth of downstream uses. Differences in these uses yields unique functional constraints on usage rights conferred on nominal competing users.

As alluded to earlier, there is a tension between a "comprehensive set of rights" and the mutually reinforcing processes of specialization and fragmentation.

The creation of a comprehensive system in the global arena is a challenge. Young (1996) provides a succinct summary of the current approaches:

> Three broad options are available to those concerned with the governance of international commons: enclosure through the extension of national jurisdiction, the creation of a supranational or world government, and the introduction of codes of conduct analogous to common property arrangements in small-scale stateless societies. (1996, loc. 4659–4661)

The former have seen extensive attention in the literature.[120] The former often assume singular bodies, national jurisdiction or a supranational institution, that rationalizes the body of rights amongst potentially conflicting uses. In Hobbes' classical terms, it assumes a Leviathan. Young goes on to argue that while "groups of interdependent actors can and often do succeed in handling the function of governance without resorting to the creation of governments is now well established," (1996, loc. 4612–4613) this does not displace governments as the seat of authority in international society.[121]

[120] Jurisdiction problems were a frequent point of discussion as the breadth of good facilitated by the Internet grew and these goods gained "depth" as they increasingly supplemented and displaced conventional modes of communication and information sharing. For discussion of "Internet jurisdiction" problems, see Zittrain (2005) and Kohl (2007). For a recent collection on delegation and authority amongst international actors in general, see Hawkins, Lake, Nielson, and Tierney (2006).

[121] The implications of this balance is a rejection of conventional notions that alternate modes of authority diminish state sovereignty. Conventional IR argues that any authority not rooted in

The NRS is presented as a comprehensive system that relies on common property arrangements. Historically, coordination amongst NMRs was sufficient to create a comprehensive system. Given the breadth and depth of goods facilitated by Internet communication, a comprehensive system requires developing relationships with organs of conventional authorities. These organs are responsible for maintaining social order but, as per earlier discussion, lack the technical skill necessary to extract information from infrastructure resources necessary to fulfill their commitments. Relative to complex bureaucracies of large states, most existing, empirical common property arrangements are relatively small.[122] That said, these must still balance competing incentives amongst constituencies, often creating what Ostrom has referred to as a polycentric governance structure involving various organs of government and non-government actors.[123]

Returning to Epstein above, he refers to both "highly varied settings" and "the physical diversity that increases the need to balance competing uses," (Epstein, 2012, loc. 7999). In the case of water, physical diversity shapes two modes of access. Access to a small tributary of a water source is substantially different from a large river in the same system. Whether one is upstream or downstream is another factor affecting access, in particular potential volume that may be extracted. In the NRS, physical diversity is affected by physical connectivity to PoPs and the interconnection platform. As noted in Chapter 2, one may have physical access to exchange routes in the control plane, but access may be enhanced by having diverse physical access or access to a POP with diverse interconnection options. More particular still, the notion of neutrality in IXes is a specific instance of ensuring points of access to a resource does not privilege one participant over another.

Key to evaluating rights systems is to avoid what Cole and Ostrom (2012b) refer to as a "myopic and limited understanding" based on focusing on one type of right, such as "treat[ing] the right to exclude as the '*sine qua non* of property'," (2012b, loc. 1144, emphasis in original). Epstein offers two common "doctrinal errors," one with respect to land, the other water. In the case of land:

> [T]he mistake is to fragment the bundle of rights so that strong protection is given to the right to exclude but weak protection is given to every other element within the bundle of rights. (Epstein, 2012, loc. 8029)

---

state authority is competing with state authority. Alternate notions of coordination and cooperation amongst governance bodies are evidenced in the literature: *a*) polycentric governance discussed by E. Ostrom (1990); *b*) the potential for what is referred to here as a federated model but drawing on the conclusions of (E. Ostrom & Schlager, 1996); *c*) framings of governance bodies as a pluralistic set of organizations by Cerny (2010); *d*) historical (Cutler et al., 1999; Cutler, 2003) and more recent (Mattli & Woods, 2009b; Büthe & Mattli, 2011) notions of private authority; *e*) modes of relational authority (Lake, 2006, 2009, 2010).

[122]The small stateless collectives often map onto what has been elsewhere referred to as rural natural resource managers. This work presumes a spectrum that ranges from ad hoc arrangements managing natural resources to formal governance (not govern*ment*) arrangements comprising systematic knowledge assessments of technical (artificial) common resources.

[123]Water basins in California is the prime instance (E. Ostrom, 1990, pp. 133–136), failure modes of fisheries is another (E. Ostrom, 1990, pp. 149–157).

In the NRS, mechanisms to exclude include 1. limiting number allocation, 2. route manipulation, and 3. blocking uses of particular prefixes (IPBLs). Each of these mechanisms take place in different NMRs. In contrast to resource economists' focus on the right to alienate, Epstein argues courts dealing with land rights focus too much on the right to exclude.

Epstein goes on to contrast land with water:

> With water, the courts make the opposite mistake. They treat it as a single unitary thing for public law purposes, even though as a matter of private law, water rights are highly fragmented to reflect the underlying sets of multiple inconsistent uses. (Epstein, 2012, loc. 8029)

Focusing on water as a unitary thing denies the essential interdependence of uses. A variety of functional water rights and relations exist: riparian rights, up-stream and down-stream relationships, in-stream rights, surface rights, and elements of alluvium and avulsion. For instance, not every riparian along a river can establish a mill driven by the flow of the river; this is also a function of physical diversity discussed above. The necessity to dam the river to increase the height of falling water creates constraints on changes upstream and limits the distance to the next dam.

These decision rights are referred to by E. Ostrom and Schlager (1996) as management rights, and are fundamental to CPRs. As implied in earlier discussions, number *management* resources, NMRs, are designed to enhance access and withdrawal rights. E. Ostrom (1990) highlights that within a CPR, the resource *system* may be jointly provisioned and appropriated.[124] In particular, the NMRs are jointly provisioned and have well-developed management rights for coordinating patterns of use. NMRs are infrastructure elements of the resource system and are provisioned by participants. Management rights shape how these are used and their effects on the resource itself.

3.4 Types of Rights in CPRs

The notion of an institutional complex is consonant with Ostrom's notion of polycentric governance. The institutions that manage the NRS are aligned along the boundaries of early problems that emerged in NRS governance and Internet infrastructure management policy. In some cases, these institutions have a history of coordination. For instance, one organization is created from a subgroup of the other.[125] In other cases, institutions have a more inconsistent track record for coordination, most notably between the RIRs and the anti-abuse community. Each

[124] In contrast, individual resource units are rival and cannot be jointly appropriated.

[125] For instance, one narrative of the creation of Euro-IX is that it emerged from the RIPE community IX Working Group. The IX Working Group working group was recently retired as redundant with Euro-IX. Another instance Euro-IX encouraging the creation of IXes (Section 6.3.2.1). In the Anti-abuse world, outreach by M^{3}AAWG to create a M^{3}AAWG chapter in India is an instance of regime proliferation. The creation of LACNIC and AFRINIC is yet another instance.

institution manages a set of primary rights that shape access, withdrawal (appropriation and provisioning).

Hart (1994) describes a legal system as the union of primary and secondary rules. Primary rules are those "thought important because they are believed to be necessary to the maintenance of social life or some highly prized feature of it," (1994, p. 87).

Secondary rules are:

> all about [primary rules]; in the sense that while primary rules are concerned with the actions that individuals must or must not do, these secondary rules are all concerned with the primary rules themselves. They specify the ways in which the primary rules may be conclusively ascertained, introduced, eliminated, varied, and the fact of their violation conclusively determined. (1994, p. 94)

Primary rights and secondary rights are used in this sense. Secondary rights include means of exclusion that limit rights and the ability to transfer rights. E. Ostrom and Schlager (1996) offers a similar distinction. Operational-level property rights comprise access and withdrawal, corresponding to primary rules (rights). Collective-choice property rights comprise rights to management, exclusion, and alienation. Hart also notes that in real (existing) rule systems, primary and secondary rules are conceptual distinctions. Rules are not neatly partitioned into primary and secondary. Rather, they are often tightly interleaved. These classes will be described in the next few sections, disentangling operational-level and collective-choice rules, further disentangling access, withdrawal, management, exclusion, and alienation rights.

Rights related to management of resources draw on both primary and secondary rules. Rights of management refer to "a right to participate in decision making, identify those who can deliberate about and help decide the regulation of use patterns, the necessity and provision of improvements/repairs to facilities, and the like," (Blomquist, 2012, loc. 9428). Management rights are those often conferred to members in NRS institutions—NOGs, RIRs, IXes, anti-abuse organizations.[126] As alluded to in the conclusion of the previous section, the resources individual organizations administer to regulate and/or enhance operational rules (conferred by primary rules), are referred to as number management resources. The classes of rights described here provide a more formal framework for evaluating CRIs. To ensure the durability of these management resources, these institutions have collective-choice rights (conferred by secondary rules) to administer use rights with respect to that management resource.

Ostrom's 8^{th} CPR design rule indicates that "appropriation, provision, monitoring, enforcement, conflict resolution, and governance activities are organized in multiple layers of nested enterprises," (E. Ostrom, 1990, p. 89, 101) Each study in

[126]The distinction between members and customers is a common difference in CRI management. This section elaborates types of rights in CPRs as they apply to the NRS. In Section 3.5 common sets of rights will be used for formally differentiate between members and participants.

Part II corresponds to an institution whose management resources contribute to the institutional complex that regulates the larger NRS. Ostrom's rule above indicates "multiple layers of nested enterprises" but does not necessarily imply hierarchical power structures. Rather, CRIs can certainly be explained in terms of organizational layers, but their lateral relations more resemble a federated collective rather than actors delegated power from a single point of authority.

Rights and obligations in the NRS will be presented at two levels: the NRS as a whole and institution-specific rights and obligations bound to specific management resources. The categories of rights below transpose rights and obligations described thus far (control plane and interconnection economics) into the categories offered by E. Ostrom and Schlager (1996). These categories are not mutually exclusive, nor do any of these categories of rights guarantee any of the other rights. These categories form a typology for effectively delineating *a*) who confers those rights; *b*) who holds those rights and their implications; *c*) intended implications of those rights for NRS integrity. The studies in Part II detail how these categories of rights related play out with respect to particular management resources—a number registry, an IX platform, an IPBL—and how they are modified and sustained by collective choice rules.

3.4.1 Entry

The right to enter, or access a resource, is referred to as entry. E. Ostrom and Schlager (1996, loc. 2484–2485) defines access as "[t]he right to enter a defined physical area and enjoy nonsubtractive benefits (e.g., hike, canoe, sit in the sun)." (Blomquist, 2012, loc. 9428) "cognizable and potentially enforceable claim to be able to be present within that domain." Access rights do not guarantee withdrawal rights. Ostrom's definition draws on an analogy to paying a fee to enter a state park. Access rights let one enjoying walking through the woods, but on may not harvest the trees for exclusive use. Similarly, surface rights related to recreation in water allow one to enjoy a day of boating, but does not account for conflicts with riparian rights of those trying to enjoy a peaceful day along the shore.

Recall from Section 2.1 that, in order to provision and appropriate routes in the NRS, network actors must have *a*) connectivity to an existing control plane participant and *b*) number resource rights. In most cases, access is a vehicle to participation in the control plane, exchange of routes ("withdrawal", here appropriation of routes provisioned by others) in order to send and receive traffic. One case of access referred to in Section 2.2 is monitoring, such as a route server or a firm that offers route analysis services. Such a monitor is more akin to strolling through the park or water surface rights—both are nonsubtractive to a point and the authorized entrant garners value from that use. In the NRS, monitoring routing dynamics from a variety of vantage points is not subtractive but is a valuable input to the observer's value proposition.[127] *Active* participation that shapes the control

[127]A number of organizations in the NRS engage in monitoring and analysis as a security service: Packet Clearing House, Renesys, and BGPMON are instances. Other actors engage in monitoring as

plane requires conferral of rights to bind a prefix to a public ASN in an origination advertisement, implying an allocation (withdrawal) of numbers and *exchange* of routes that potentially alters the global stock of routes; both modes of withdrawal will be discussed in Section 3.4.2.

3.4.1.1 Physical Access

The data plane (and by proxy the control plane) has as many entry points as there PoPs at which existing participants offer interconnection and the option to exchange routes. The point of entry comes with limitations on the types of and range of appropriation and provisioning. Some of these limitations are constraints on degree of access. The point of entry, the facility housing the POP, is controlled by some actor. It may be the dedicated facility of a transit provider, a colocation provider, an IX, or even some combination of these.[128] The potential for active participation may not be the same at different facilities by virtue of a combination of participation at the facility or limitations imposed by the actor managing that facility. Some of these are limitations on exclusion rights, discussed in Section 3.4.4. That said, strictly speaking, physical connectivity provides access to a POP where one or more network actors are available for interconnection. As discussed above, relations in which monitors consume but do not produce routing information, and thereby do not alter the global stock of routes, is a form of access. An interconnection relation that facilitates exchange of routes (provisioning *and* appropriation) constitutes "withdrawal" and does potentially alter the global stock of routes.

Rights to access at the physical layer vary across legal jurisdictions. This is one clear point of intersection between national jurisdiction and the NRS. Typically, rights to access are controlled by the same rules for establishing an Internet communication service. This varies substantively across jurisdictions, some requiring certain types of network actors be licensed. As with any regulatory mechanism, this is susceptible to the classic principles of Stiglerian regulatory capture.[129] As a po-

---

part of their larger portfolio to analyze their own traffic and optimize interconnection contracting. For instance, CDN caches often have logging and diagnostic tools that allow for collecting data. Academic efforts such as RouteViews deploy route collectors to periodically collect route dissemination information from a variety of vantage points, archiving this data for analysis. The RIPE RIS provides a similar service. The community also provides a variety of "looking glasses," typically route servers with open telnet access that allow actors to interactively inspect local routes exchanged with the looking glass and perform diagnostic tests such as ping and traceroute. At a technical level, these require consuming routes *provisioned* by local actors. In this sense, technically this is a form of withdrawal. That said, appropriation by monitors is passive in the sense that a pure monitor, one that is provisioned exclusively to collect data, does not provision routes to be consumed by others. This passive mode of participation intended for data collection but not altering the global stock of routes is a degenerate form of withdrawal from the global stock of routes but is considered more akin to access in terms of system outcomes.

[128] Combinations of these are detailed in the discussion of archetypal IX architectures in Section 6.2. This discussion distinguishes between the options available through various interconnection platforms.

[129] See Stigler (1971). In particular, Stigler describes a trucking licensing regime as a mechanism to limit competition. Telecommunications infrastructure development is licensed in a number of

tential barrier to entry, licensing is a state regulatory mechanism that does impinge on otherwise largely privately managed infrastructure coordination.

With the exception of route monitoring, access to and withdrawal of routes in the NRS are coupled. A tight coupling of access and withdrawal is not always the case, especially in the conventional cases that distinguish access in terms of nonsubtractive versus subtractive uses. Moreover, access differs for routes and numbers. In the case of routes, monitoring is a form of passive participation. In contrast, within the community, access to the use of numbers typically refers to the right to unique origination, which is a form of withdrawal.

3.4.1.2 Number Resources Facilitating Participation

Number resources are the other component of entry. In order to enter—participate in—the control plane, an actor needs at least an ASN. Actors also typically hold stewardship of an address block. Historically allocation of address blocks (and ASNs) have relied on needs-based criteria. Stated simply, needs-based criteria are intended to ensure an actor requesting number resources has *a*) business or infrastructure uses for those number resources and *b*) that scarce[130] number resources are conserved. Needs-based criteria is the historical bar for determining access; withdrawal is a change in numbers registry.[131]

Each subsequent allocation is also subject to needs-based criteria. Subsequent allocations may come from *a*) the common pool or *b*) via transfer. Historically both have been subject to needs-based criteria in order to (re)allocate. In *subsequent* allocations, an additional criteria is typically applied. A measure of whether that actor is utilizing its existing delegation efficiently, HD-ratios,[132] have been used in the evaluation process.[133] Subsequent delegation from the common pool comprises entry via needs-based criteria and number utilization. Entry via transfer requires allocation criteria, but also involves alienation rights discussed in Section 3.4.5. In this sense, entry rules are a factor shaping an organization's engagement in the control plane and decisions regarding deployment of resources that garner value.

3.4.2 Withdrawal

E. Ostrom and Schlager (1996, loc. 2485–2486) defines withdrawal as "[t]he right to obtain the resource units or "products" of a resource (e.g., catch fish, appropriate

states. For instance, licensing regulation was used to temporarily limit access to interconnection access rights in Kenya (Kende & Hurpy, 2012).

[130]As elaborated in the next section, number resources are scarce by virtue of unique origination. Many uses of number resources are non-rival.

[131]Although developed further in Section 3.4.2 and Chapter 5, recall that numbers are information goods. Allocation is not a physical process as with subtractable commodities that need to be transported. It is a very literal conferral of rights.

[132]The HD-ratio provides an idea of what proportion of the existing allocation has been utilized. It is $\frac{\ln u}{\ln d}$, where u is the number of addresses utilized and d is the total addresses in a particular delegation.

[133]See Section 5.7.2.

water, etc.)." Cole and Ostrom (2012b, loc. 1150) define withdrawal as "the right to harvest and take some resource units out of the resource system." The definitions of withdrawal above are rooted in the notion of a rival, subtractable resource unit. Withdrawal in general maps to resource unit appropriation for the purpose of rival consumption.

Resource units in this study (numbers and routes) are information commodities. Nominally, information commodities can be used by anyone. Blomquist (2012) indicates that "withdrawal rights are conditional; that is, they include restrictions on when, where, and how much the rights holder may use the resource," (2012, loc. 9428). As per earlier discussion, the actors shaping the conditions of withdrawal in numbers and routes are quite different. Number resources withdrawal is conferred by RIRs; route exchange (withdrawal from the global stock) is a largely bilateral decision, but as developed below, options for withdrawal are shaped by the point of entry.

Although the definitions above speaks to appropriation, or harvest, it can also be interpreted to speak to types of use. In the context of discussing management rights, E. Ostrom and Schlager (1996, loc. 2507–2508) further refines the scope of withdrawal rights: "[t]he right of management is a collective-choice right authorizing its holders to devise operational-level withdrawal rights governing the use of a resource." Withdrawal rights are assertions of types of use; see Section 3.4.3 for management rights that confer the authority to set (or rescind) withdrawal rights. In this sense, scope, the "when, where, and how" give insight into the CRIs that affect opportunity for route withdrawal and which actors have management rights in which context. In terms of withdrawal, recall from Chapter 2 that number resource and route use is conditioned on kind of use. Further recall the discussion of externalities—those that threaten the integrity of the system arise from *kinds* of use, not *frequency* of use. In effect, negative externalities are a function of violating withdrawal rights.

3.4.2.1 Number Resource Use

RIRs shape rules of number withdrawal. Number use in the construction of routes and routing decisions is a function of bilateral interconnection contracts. Kinds of number use largely focuses on origination (here framed as "when", from the quote above) and use in routing decisions. "Where" refers to the interconnection context: is the number being used in an origination context or a subsequent advertisement? In more nuanced cases of Anycast, it can refer to the variety of physical locations where a particular number is advertised. Origination is governed by both RIR allocation and contractual relations. Subsequent use is governed only by contractual relations.

Number resource coordination leverages two distinct rights granting global exclusivity for particular uses of prefixes and ASNs that ensure uniqueness of addressing. ASNs are allocated by RIRs to identify organizations by common routing policy; an organization may have multiple ASNs but multiple organizations with

diverse routing policies do not share the same ASN(s).[134] RIRs allocate ASNs from a common (global) finite pool; once allocated, they are "withdrawn" from that pool for the term of the allocation. RIRs confer an organization with the right to use an ASN to uniquely identify itself and the right to use that ASN in the provisioning of routes.

RIRs also confer the right to originate prefixes from a specific block of addresses. The right to originate prefixes is bound exclusively to a particular ASN.[135] Address blocks (identified by prefixes) are allocated from a common (global) finite pool. Once allocated, blocks are "withdrawn" from the available pool for the term of that allocation. Recent debates in the RIR system have focused on the operational rules for transferring origination rights and ASNs from one organization to another and the role of the RIR in that process.[136] In terms of the classes of rights discussed here, transfers is a debate over the alienation rights, discussed generally in later in Section 3.4.5.

As it relates to uniqueness of origination, addresses and ASNs are rival in this particular use. Within an organization, addresses may be used to uniquely identify whatever resources in the AS the organization sees fit, but another AS generally should not originate prefixes in a block unless it has been allocated that block. Operational rules are intended to shore up complementary obligations. Under ideal conditions of unique origination, prefix hijacking would not occur. In contrast, numbers are information and thus can nominally be used by anyone. The RIR provides a facility (number registry) for documenting origin rights. Here, origin rights are essentially the right to use (withdraw) a prefix for route origination.

Uses, withdrawal, is limited when negative externalities become sufficiently costly to warrant developing operational rules to mitigate them. Operational rules and mechanisms for exclusive prefix origination are intended to either mitigate or eliminate externalities such as prefix hijacking. Best practices for prefix aggregation are an instance of withdrawal rules that speak to both number use and route provisioning. Given the notion of withdrawal, many of the uses of numbers documented in Chapter 2 are different uses and corresponding externalities.

Consider the means for making origination rights more durable. RPKI[137] uses cryptographic signing of origination advertisements to allow appropriators to confirm the right of origination was respected. RPKI is managed by the RIRs; the RIR allocating the number resource manages the signing process binding the organization, ASN, prefix tuple. RPKI is also an issue of management rights; RPKI has implications for enhancing the enforcement of origination rights by RIRs and other

[134]There are temporary exceptions, especially in the case of mergers. That said, if all actors merging have an ASN, they typically use their respective ASNs until their routing policies are fully integrated.

[135]There are exceptions. For instance, RFC 1918 (Rekhter, Moskowitz, Karrenberg, de Groot, & Lear, 1996) describes the set of private addresses that may be used by any actor in a private network, but they should never be routed on the Internet. Accidentally advertising private prefixes is a known network hygiene failure.

[136]See Section 5.7.3 for discussion.

[137]See Section 5.7.4 for a discussion of RPKI and issues in the RIRs. For details on RPKI *protocol* development, IETF SIDR group.

would-be principals, potentially shifting some management rights. A variety of other origin and path security protocols have been proposed[138]; RPKI has seen the most operational debate in terms of management and exclusion rights. In Chapter 5 RPKI will be evaluated in terms of its implications for number resource rights and the consensus processes that have contributed to RPKI deployment (or lack thereof) in the RIR system.

3.4.2.2 Route Use

The life cycle of a route is shaped by origination rights. Bilateral contracts[139] between network actors determine the scope of conditions of route exchange and whether operational rules bind. In terms of the "when, where, and how" above, the *options* for route exchange are strongly influenced by the type of interconnection platform (where) and whether interconnection is direct or mediated by various types of switching fabric (how). Both scope of exchange between actors and where these exchanges occur both affect route withdrawal (use) rights.

Origination is a distinguished mode of provisioning. Route exchange, the mutual provisioning and appropriation of legitimate routes, is documented in Section 2.2 on technical routing mechanics, is non-rival in the control plane.[140] Any network that appropriates a route uses constituent number resources to make routing decisions. Thus, modulo externalities such as injecting another actor's ASN into a route, number resource use is generally non-rival in route advertisements subsequent to origination.

Recall externalities are generally caused by kind of use rather than frequency. For instance, the root cause of disaggregation is how the block is partitioned for use, not necessarily the frequency of advertisement.[141] Some instances of route flap are periodic, creating a cycle of update and withdraw messages at a rate that causes convergence problems for subsequent appropriators. Other modes of flap are due to efficiency of advertisement, not rate of advertisement. When changes are not "packed" into a single advertisement, each appropriator must converge on each change sequentially. Again, the problem is one of *kind*, not rate.

In conventional CPRs, appropriation fits the subtractable characterization of withdrawal. Appropriation is regulated to avoid over-exploitation, withdrawing resource units faster than they are replenished. In the NRS, route appropriation in and of itself is not subtractable. Rather than some notion of depletion, the externalities created when provisioning rules are violated degrade the integrity of the system. Conditions limiting appropriation as a form of withdrawal are intended to monitor, mitigate, and/or prohibit modes of appropriation that damage control

[138]Butler et al. (2010) provides a recent survey of routing security measures, including RPKI.

[139]Mechanically contracts are bilateral, resources such as route servers facilitate multilateral peering (interconnection). See Section 6.2 for elaboration.

[140]Subsequent traffic may create congestion in the data plane if those using the routes are heavy users relative to available capacity. Congestion (rivalry) in the data plane may affect the route selection process (traffic shaping), but it does not affect the non-rival character of routing *information*.

[141]The volume may increase but frequency of each remains the same.

plane integrity. Integrity can be effected by kind, frequency, and volume of legitimate route provisioning. It can also be affected by the legitimacy of routes. Finally, it can be affected by the reputation of resource units.

The operational scope of route provisioning and route appropriation is the interconnection relation. As per Section 2.2, an AS appropriates the routes its adjacencies choose to provision within a mutual BGP session. These routes may be selected for use in the local routing table of that AS or the AS may simply store them for later decision processes. In both cases, routes are withdrawn and the scope of this process is the bilateral contract, governed largely by informal operational rules.

In terms of classes of rights, Construction of routes to be withdrawn is a combination of route use and number use. ASes provision new routes by pre-pending its AS number onto the AS-Path of a route it has appropriated elsewhere. The provisioning AS then decides whether to advertise that route to a neighbor based on the contract with the neighbor and the routing policy of the provisioning AS. Advertising a route is effectively conferring a set of use rights, constrained by operational rules related to contractual mode, such as settlement-free peering, transit, or other more specific contexts like IX infrastructure routes. An adjacency in a legitimate route may be interpreted as evidence of a formal or informal contract between those two actors.

Some use rights are well-known operational rules, others are explicit in operational rules. For instance, a well-known operational rule is that an AS X typically does not provision routes for a settlement-free peer Y from its transit routes. This would effectively offer routes X is paying for to Y for no cost. The notion of transitive trust in routing applies the general obligation to route traffic to the next hop. Transitive trust does not imply any contractual relation extends beyond a bilateral interconnection relationship. As such, effects of use dynamics are externalities as per Section 2.1.

3.4.2.3 IPBLs and Orchestrating Traffic

IPBLs notion of abuse, developed in the next section, creates withdrawal assertions with regard to data plane use. As operational rules, access networks use IPBLs to shape traffic entering their networks. Historically, this has limited spam traffic, but more recently it has been extended to traffic underlying malware and other more broadly abusive behaviors. In terms of the scope of operational rules, the anti-abuse community has general operational heuristics for what constitutes a good IPBL. That said, as elaborated in Chapter 7, the specific implementations of those are operator specific.

To elaborate, the operational rules for *how to use* IPBL data (resource units produced by the IPBL) to manage traffic entering a privately managed network lies with network operators. As such, orchestration of the data plane based on number resource reputation is performed by network actors. In contrast, the operational rules (heuristics) shaping what constitutes legitimate IPBL reputation data and the methods for collecting that data is more broadly shaped by the anti-abuse community. In terms of an CRI, use of CRI resource units are at the discretion of the

network operator, but production of those resources, and subsequently the quality of those resources, is a joint effort. Here there is a distinction between the provenance of operational rules for using reputation data in the NRS writ large versus CRI-specific rules for how that data is produced (as well as how infrastructure is deployed to disseminate that data).

3.4.3 Management

E. Ostrom and Schlager (1996, loc. 2504) define management rights as "[t]he right to regulate internal use patterns and transform the resource by making improvements." As alluded to in the previous section, management rights carry the authority to set or rescind operational rules, in particular withdrawal rules. Note also that it speaks to *transforming* the resource. (Cole & Ostrom, 2012b, loc. 1174–1176) make transformation explicit with particular examples, indicating management rights constitute the "[r]ight[s] to change the physical structures in a resource system, such as building an irrigation system or a road, changing the shoreline of a fishery, or developing a variety of physical infrastructures for any particular resource." The examples in this definition are changes to the resource system. Each change introduces purpose-specific infrastructure that enhances rights to withdraw resource.[142] Infrastructure changes also interleave mechanisms that ensure monitoring and enforcement of operational rules. These collateral mechanisms make operational rules more durable and the overall system more sustainable in terms of both resource integrity and costly information necessary for sustaining regulatory (management) institutions.

CRIs not only shape primary rights, but mechanisms may enhance or diminish primary rights. Structures introduced to facilitate management of resource system integrity are referred to as *management resources*. A physical manifestation of a management resource in a conventional CPR is an irrigation system. Ideally, the irrigation system as a composite infrastructure and institution enhances usage rights by *a*) lowering the barriers to accessing a share (units of withdrawal) of the water resource and *b*) introducing collateral mechanisms and loci of monitoring and enforcement. A combination of use rights and rights related to the management of the irrigation system shape the rate of appropriation, who can appropriate when, and how appropriation is monitored and enforced.

The resource system in the NRS is shaped by a common image of control plane integrity and implemented through protocol and operational provisioning of resources. The studies in Part II on number registries, IXes, and IPBLs are studies of management resources in the NRS. Each management resource is itself a commonly managed (jointly provisioned) resource. Each has its own access, withdrawal, management, exclusion, and alienation. Each also has its own collective choice

[142]Frischmann offers a broad definition of infrastructure that includes natural resources. Under this definition, all of the changes, even those to the shoreline of a fishery, constitute changes to the resource system. Following the distinction between natural and artificial resources, these are man-made (artificial) changes intended to enhance access and facilitate withdrawal from the resource. As earlier, these artifacts are elements of the resource *system*.

(secondary) rules that create operational and constitutional rules for administering those management resources. Secondary rules within and across institutions vary in scope, kind, and formality. CRIs also participate in umbrella organizations, serving as a federated collective, that facilitate coordination amongst organizations that share common institutional norms.

These management resources enhance or diminish primary rights, such as rights to entry, rights to withdrawal, and rights to exclusion. Moreover, each management resource will be described in terms of its own set of entry, withdrawal, management, exclusion, and alienability rights in its respective chapter in Part II.

3.4.3.1 Number Resource Management

The rules structuring allocation of number resources are set by the RIRs, the organizations that manage the number registries. The number registries are the common record of which networks have been allocated number resources. Registries' contribution to control plane integrity is the centralized documentation of origination rights that ensure uniqueness. The RIRs maintain contact and administrative information for organizations allocated number resources. For given number resources, users of the registry can query to determine which actors hold which resources and acquire contact information for resource holders, among other things. The number registry is considered *authoritative* based on a combination of traceable history of delegation in the RFCs, constituency (epistemic community support), and mutual reinforcement within the Number Resource Organization (NRO). Chapter 5 provides details on the function of RIRs as CRIs.

As it pertains to number resource management rights, the RIR system as an institution is the arena in which withdrawal, exclusion, and alienation rights are formed and conferred. Within the RIR, epistemic communities engage in a formally documented and managed consensus process to establish, evaluate, and rescind operational rules. As documented in detail in Chapter 5, the RIRs' consensus *processes* have a similar structure: active consensus, passive consensus, process evaluation.[143] Within this general structure, there are some nuanced and some substantive differences. RIRs are regionally independent—the NRO serves as a loosely organized federating agent. Management rights and operational rules are largely consonant with one another, in part because of overlapping constituencies and in part because of intra-institutional information sharing.

Historically, enforcement of operational rules related to the registry and allocation processes were backed by repeated engagement. As per the discussion of access, many growing number resource consumers return to the registry for subsequent allocations. The criteria for initial and subsequent allocation described in Section 3.4.1 are established through the exercise of management rights. RIR policy dictates the criteria that decide allocation processes. Each RIR comprises a community that creates policy and a firm that implements the policy and related infrastructure. The RIR system provides the arena in which consensus processes elicit

[143]For detail, see Section 5.6.2.

policy proposals, evaluates those proposals, and adjudicates whether the consensus process has been followed. Chapter 5 elaborates these processes. For the purposes of the discussion here, these processes shape management rights, manifest as secondary rules of change.

Shifting gears from management of allocation and origin rights, IPBLs also apply management rights. Blocking lists (BL) are management resources intended to limit abusive behavior. In terms of management rights, IPBLs provide information[144] necessary for networks to more effectively limit traffic deemed abusive. Individual networks do not necessarily have sufficient information to credibly identify the sources of abusive traffic. An early articulation justifying BLs is that "[a]ll information exchange on the Internet is consensual," that BL maintainers "are exercising [a] right to refuse traffic from anyone [they] choose," and that it is "within anyone's rights to make the same choice (or a different one, so long as only their own resources were affected by their choice)," (MAPS, 2004, p. 2). Under this articulation, abusive traffic is unsolicited traffic.[145] Under the original quote "[they]" was originally "we," comprising BL providers and their users, both of which have common interests. The common good is lack of abusive behavior.

Any network provider may choose to reject traffic from any address or collection of addresses. In terms of management rights, IPBLs imbue network resources with reputation. Acting on that reputation maps to "regulating use patterns" in the data plane. Recall from Section 2.2 that the control plane "is used to direct, measure, and repair the control plane," (D. D. Clark et al., 2003). In the case of IPBLs, number reputation is used to direct, or shape particular types of traffic. Historically, IPBLs have been used to limit spam, unsolicited e-mail. IPBL use has developed since then to address a variety of abusive mechanisms animated by network connectivity—actors can and do use IPBL data drop all traffic from sources with poor reputation.

Nominally, BLs attempt to credibly identify the origins of abusive traffic. BL providers use abuse monitoring tools developed by the anti-abuse community. These tools are distributed at various vantage points to identify the sources of abusive activity. Based on sources identified, IPBLs assemble lists of prefixes that regularly produce abusive traffic. Some IPBLs also provide information regarding why the number resource is listed.

IPBLs effectively imbue sources perceived to produce abusive traffic with negative reputation. IPBLs that are considered credible (low false positives) are used by network actors to limit abusive traffic create network effects. These network effects diminish the value that can be derived from using the prefixes listed. In terms of number resource rights, being listed on credible IPBLs diminishes the *range* of use for the listed prefix. For instance, the historical use of IPBLs is to limit the delivery

[144]As will be developed in Chapter 7, for *individual* actors, reputation information is costly. A subset of the anti-abuse epistemic community has developed mechanisms for identifying abusive actors.

[145]This is a very strong definition of abuse that would, if perfectly enforced, would eliminate quite a lot of traffic that some users do in fact derive benefit from. More nuanced notions of legitimate (non-abusive) traffic can be framed as various modes of opt-in versus opt-out consent to receive classes of information at the discretion of the sender, not the receiver. These will be discussed in more detail in Chapter 7.

of spam.

IP reputation creates a form of rivalry. Abuse monitoring tools are developed in the context of anti-abuse industry norms. In terms of the externalities developed in Section 2.2, the collective action of IPBL users creates an externality that limits the use of the listed prefix. As a collective action problem, IPBLs serve to facilitate and coordinate decentralized selective incentives.

Participants in the RIR community and the anti-abuse community have some overlap, but there is also conflict. Organizations that find themselves on IPBLs are often members of the RIR community. Some of these have representatives that participate in collective choice processes in fora such as M^{3}AAWG. Others are not aware of anti-abuse practices. Others still are regularly listed on IPBLs. A number of RIR community members contest the legitimacy of the IPBLs. The result is that there is tension over the application of IPBL practices. In terms of rights and obligations, there is conflict between the *a*) operational rules developed by the anti-abuse community limiting abusive traffic and *b*) the absence of enforcement mechanisms in the RIR communities. This contention will be discussed from both perspectives, introduced as a mode of rights revocation in Section 5.2 in the Chapter on RIRs (5) and extensively in the Chapter on anti-abuse (7). The fundamental tension is over rights to exclude actors from particular number resource uses.

3.4.3.2 Route Management

E. Ostrom and Schlager (1996, loc. 2507–2508) defines "[t]he right of management [as] a collective-choice right authorizing its holders to devise operational-level withdrawal rights governing the use of a resource." In the RIRs, consensus processes shape number management rights. Management rights for routes are largely operational norms rooted in informal and formal bilateral contracts. Some management rights, referred to by some as network hygiene principles, described operational mechanics for avoiding common technical failure modes and avoiding being "complicit" (as per discussion in Section 2.2) in control plane externalities. Apart from the mechanics of network hygiene, contracts shape withdrawal rights on a per contract basis. As per section Section 2.3, there are well-known contracting modes, but, aside from prefix origination, particular operational rules pertaining to route exchange are not conferred by any particular body.[146]

The operational rules and externalities described in Sections 2.1 and 2.2 are informal operational norms. This does not mean that there are not institutions in which these norms are discussed. Fora such as network operator groups (NOGs), IX membership meetings, RIR meetings, and meetings such as the Global Peering Forum (GPF) and the European Peering Forum (EPF) are all arenas in which actors disseminate, discuss, and evaluate operational norms. These fora, in their capacity

[146]RPKI has the potential to protect origination rights. Path security protocols would attempt to provide guarantees on each bilateral relationship. The question of rights is who has control over the keys that make those keys authoritative. One solution is that the RIRs would act as a certificate authority. Another is to have a market of certificate authorities that the RIRs may or may not participate in.

as arenas for sharing costly information, will be referred to in general as NOGs, specifying function specific elements where necessary. NOGs will be discussed in Chapter 4.

The management resource offered by IXes is a commonly managed interconnection platform. In contrast to RIRs and IPBLs, IXes do not confer or rescind use rights. IXes enhance access rights, effectively increasing options to establish contracts that, in turn, shape withdrawal dynamics.

In this sense, IXes are access and withdrawal hubs. A variety of infrastructure analogies apply. As hubs, these resemble common markets for interconnection options. A more nuanced view is the now well-trod analogy with water resources. Here, IXes shift the loci of route exchange to more neutrally managed facilities that, for successful diverse IXes, offer a variety of sources "farther upstream." Drawing out the irrigation analogy, small to medium networks are not relegated to backwater tributaries controlled by more powerful actors upstream. Rather, these actors are invited to develop bundles of upstream options and cultivate these as necessary to fit their value proposition.

In the broader scope of NRS rights, IXes' role in entry rights was discussed in Section 3.4.1. IXes enhance entry rights by provisioning access to a diverse set of ASes at a single facility, lowering barriers to the market of interconnection options. Early in their history, development of and investment in IX platforms were motivated by potential transit costs savings and the availability of higher quality bilateral interconnection options. Modern IXes are loci for interconnection; a diverse set of interconnection options facilitate potential to appropriate a more diverse set of routes. In the developing world, IXes are seen by some as catalysts for infrastructure development and challengers to integrated incumbents (or multinational would-be incumbents).

It stressed, here and in Chapter 6, that most IXes provide the option, not the obligation, for platform participants to interconnect and exchange (bilaterally provision and appropriate) routes. As per the previous section, route exchange is conventionally a bilateral decision process scoped to an interconnection agreement.[147] Sowell (2013) frames IX-facilitated interconnection as a (real) option. Chapter 6 further develops the option framework and the benefits described in Sowell (2013), supplementing this framework with explanations of how management processes and characteristics of IX platforms serve to sustain and further enhance these benefits.

There are exceptions to IXes' role facilitating withdrawal through enhanced access. Forced multilateral peering is an operational rule in IXes that forces all IX participants to engage in settlement free peering, at least for the prefixes allocated to that network. In this case, IXes do set particular minimum withdrawal rules with respect to routes exchanged over the IX platform. Another form of this is the use of route servers. Some IXes require all interconnection over the exchange occur via a route server. Others provide a route server to facilitate multilateral peering with ac-

[147]The use of route servers within IX infrastructures are an exception. Voluntary route servers facilitate multi-lateral interconnection for willing route server participants. Some IXes, such as CABASE, require all participants interconnect with all other participants through the route server. See discussion in Section 6.2.

tors that elect to interconnect with any other actors that are both on the route server and that elect to interconnect with others willing to engage in settlement-free peering. As will be discussed in Chapter 6, in membership-based IXes, operational rules are decided through collective choice rules amongst members.

In Section 3.4.2.3 the distinction between operational rules regarding blocking based on IPBL reputation was made distinct from the CRI process of determining what constitutes legitimate information. The latter is an exercise of management rights within the anti-abuse community. In contrast to other CRIs, the RIRs and the IXes, IPBLs do not have a membership per se. IPBLs have customers or users. Operational rules shaping legitimate IPBL behavior is shaped by *a*) false positives and false negatives in reputation data; *b*) how reactive IPBLs are to remediation; *c*) market demand for the IPBL as a product; *d*) operational rules of abuse forensics developed by the community. In this sense, management rights are shaped by a combination of market and institutions—in Williamson's terms, they lie between markets and hierarchy.[148]

3.4.4 Exclusion

E. Ostrom and Schlager (1996, loc. 2505) define exclusion as "[t]he right to determine who will have an access right, and how that right may be transferred." (Cole & Ostrom, 2012b, loc. 1177) argue that exclusion confers"the right to determine who else [can] use [a] resource and what their specific rights [will] be." Blomquist (2012, loc. 9456) elaborates this to highlight "[r]ights of exclusion identify those who have authority to determine—on their own or in concert with others—who cannot have access to or make withdrawals from the resource." With the exception of origination framed as the provisioning of a route, rights to exclusion of number resources and rights to exclusion of routes are managed differently. Exclusion rights confer the privilege of dictating particular access rights to particular actors within the CPR.

In the NRS writ large, RIRs control access criteria (needs-based criteria and HD-ratios), the (secondary) operational rules of allocating number resources. In terms of physical access, facilities maintain their own criteria for connectivity between network actors at a given facility (POP). Although on first consideration, it would seem IPBLs also have a role in exclusion. IPBLs do not. IPBLs provide information that affect orchestration. They do not affect criteria for allocating number resources or the exchange of routes. Recall that operational rules dictating how IPBL data is used is in the purview of the network operator.

In contrast, as alluded to earlier, rights to exclusion with respect to routes are held by the actors that provision those routes and are decided in the contracting process. Origination is unique in that the *right to* origination is conferred by the RIR, but the decision *for whom* to provision an origination route is held by the steward of the number resource. Exclusive provisioning of address blocks and ASNs

[148]To be absolutely precise, all of the management rights discussed in these studies lie in the interstices of markets and loosely hierarchical *federations*.

facilitates: *a*) non-rival appropriation of number resources and *b*) provisioning and appropriation of routes. In this sense, number resources are withdrawn from the global pool with respect to origination, but many other uses are non-rival. That said, the kinds of uses documented as creating externalities make number resources partially-rival depending on type, or kind, of use.

The difference between management and exclusion is subtle. Note Blomquist's elaboration above speaks to "who cannot have access to *or* make *withdrawals* from the resource," (Blomquist, 2012, loc. 9456, emphasis added). This indicates whether one can or cannot withdraw. Withdrawal rights are assertions of acceptable rights to patterns of use. As per the definition of management rights from E. Ostrom and Schlager (1996), management rights confer the right to "regulate internal use patterns," in effect articulating precisely which withdrawal behaviors are enjoined by withdrawal rights and to whom those rights are conferred.

Exclusion limitations can be controlled by the actor managing the POP at which an interconnection relationship occurs. Exclusion limitations also manifest in the contracting modes at a particular facility. Consider the dynamics of transit and peering discussed in Section 2.3. Transport provides connectivity to a POP, but does not guarantee any form of interconnection relation. A single transit relationships does provide access, but the routes available for appropriation are limited by a single producer. As per discussion of transit route provision in Section 2.3, transit providers exclude routes from its customers that are not to its benefit. Multiple transit relations may enhance access in the sense that a more diverse set of routes may be available, but exclusion decisions are still ultimately made by transit providers. Adding peering relationships also provides diversity. Arguably, the more diverse the routes available for appropriation, the greater the access to the control plane. Greater access translates to potential value derived from additional possible data plane paths available. Further, these options offer greater opportunity to appropriate bundles of routes that better suit the network actors' particular value-proposition.

Amongst the studies presented here, IXes, commonly also referred to as Internet exchange points, are platforms that enhance entry/access rights. The historical reference to an exchange *point* underlines the role of the IX relative to entry. IXes provide a common interconnection platform that lowers the barriers to interconnection to a diverse set of network actors. Among other benefits, it creates a loci of access to a diverse set of participants and, by proxy, a diverse set of contracting modes. The criteria for access to (participation on) an IX parallels the entry criteria of the control plane above: an ASN and connectivity to the IX.[149]

[149]In general, an ASN and connectivity are the most basic requirements. Many IXes still require participants also have transit that is not provisioned across the IX fabric. Variants in IX peering policies can also be described in terms of access, in particular the guaranteed concentration of access; see Section 6.4.3 for further details.

3.4.5 Alienation

Alienation is defined as "[t]he right to sell or lease either or both of [management or exclusion] collective choice rights," (E. Ostrom & Schlager, 1996, loc. 2506–2507). In contrast to later articulations in (Cole & Ostrom, 2012a), this is a narrower set of rights. Above, alienation is limited to the transfer of secondary rights. Later articulations broaden this to any of the types of rights discussed thus far. Blomquist (2012) indicates rights of alienation,

> also known as rights of transfer, are possessed by those who can legitimately confer their other property rights on someone else. Although rights of alienation may seem inherent in other rights, they are distinct, and the distinction is very important. One may possess rights of access to a resource, for instance, by virtue of group membership, residence, or some other characteristic and thus be unable to transfer that access right to another person. (2012, loc. 9456)

Rights of alienation are most notable in the RIRs. IP address transfer markets have been a contentious topic in the RIRs, especially in the context of IPv4 depletion.

IP address transfer is an instance of alienation rights being distinct from other rights. More precisely, IP address transfer is the transfer of origination rights. Tacit in origination rights is the use of those addresses to uniquely identify hosts managed by a network. Transfers, like initial allocations and subsequent allocations, have historically be subject to the needs-based criteria.[150] A number of discussions of the transfer market have been offered and will be discussed in Section 5.7.3.

The transfer debate is animated by a number of NRS issues. The obvious first is the depletion of IPv4 addresses, more precisely those for which origination rights have not been allocated. This creates scarcity, limiting growth in subsequent allocation and entry under initial allocations. Historically there has apprehension in the community that a "frictionless"[151] transfer market would give ride to hoarding of addresses and a spike in the market. Others argue that a frictionless market will naturally allocate addresses to their most efficient use.

The role of the RIR in transfers has been questioned, in particular needs-based criteria versus the idea that markets will allocate resources to their most efficient use. As information commodities, addresses do not have value in-and-of themselves; their value is derived from uses facilitated. Currently prices seem to be hovering at approximately $10 per address. One concern is the effect of transfers on registry integrity. Transfers that occur outside the registry system may yield conflicts over uniqueness. Absent credible documentation of origination rights, *a*) multiple actors may claim origination rights without means to validate those rights and *b*) it

[150] Experiments with removing needs based requirements have occurred. For instance, Proposition 50 in APNIC briefly removed needs-based allocation for transfer in the APNIC region. Recently, needs-based criteria have been removed from the current RIPE region allocation process (RIPE, 2014).

[151] Frictionless markets are a theoretical fiction. Here frictionless refers to minimizing friction, not an existing frictionless market.

x	Owner	Proprietor	Claimant	User	Entrant
Access	X	X	X	X	X
Withdrawal	X	X	X	X	
Management	X	X	X		
Exclusion	X	X			
Alienation	X				

Table 3.2: Bundles of Rights Associated with Positions, adapted from Table 7.1 in (E. Ostrom & Schlager, 1996)

will become more difficult to attribute abusive behavior to prefixes originating that behavior.

In the former, a series of "frictionless" trades may obfuscate provenance of number rights. In the best case, provenance can be reconstructed, but at higher transaction costs. In the worst case, "origination" by competing stewards of the prefix will be appropriated by ASes in the Internet. Mechanically this would be the same as hijacking but in terms of rights it would be unclear which actor was the rightful steward. Absent authoritative documentation of stewardship, conventional enforcement mechanisms, selecting the authoritative origin, would not be available. Transfer policies currently in place *a*) require transfers be reflected in the registry and *b*) may impose needs-based criteria.[152]

3.5 Categories of Rights Holders

The five property rights in the previous section "are independent of one another, but are frequently held in [a] cumulative manner," (E. Ostrom & Schlager, 1996, loc. 2520–2521). The cumulative character is depicted in Table 3.2. This section will describe NRS participation in terms of these rights bundles. The corresponding roles in each of the CRI studies will be briefly described. Full explanations of CRI dynamics are developed in their respective chapters. For the time, Table 3.3 depicts a coarse-grain view of the types of rights holder in the NRS. These will be described briefly here to establish the concepts and foreshadow how these bundles of rights are are operationalized. In the corresponding CRI study chapters in Part II, each CRI is further unpacked, depicting rights holders for each resource provisioned by the CRI, where variety in rights bundles exists within instances of a particular CRI and within particular CRIs. Analytically, these classes of rights and rights holders as meaningful bundles of rights facilitate evaluating the contribution of each CRI individually. As per Blomquist (2012), referring to this framework,

> For researchers, it provides a means of comparing resource-management regimes, even when the resources themselves (not to mention their geographic, historic, and social settings) are divergent, which is a great

[152]See Section 5.7.3 for a characterization of policies.

	NRS	RIR	IX
Authorized Entrant	X	X	
Authorized User	X	X	X
Claimant			
Proprietor		(<X>)	(<[X]>)
Owner		(X)	

Table 3.3: Common categories of resource rights holders in the NRS and CRIs. An "X" denotes individual participants hold these rights, and "(X)" indicates the participant collective determines rights under collective choice rules established in that institutional complex, an "<X>" indicates the firm has been delegated the authority to implement a set rights devised by the collective, an "[X]" indicates a distinguished subset of individual participants hold these rights.

> advantage in being able to move beyond individual case studies to comparative qualitative and quantitative analyses. (2012, loc. 9472–9474)

This comparison facilitates evaluation across NRS institutions and evaluations of systemic stability.

3.5.1 Authorized Entrants

Authorized entrants are roles that only "hold operational-level rights of access," (E. Ostrom & Schlager, 1996, loc. 2530–2531). Authorized entrants in the NRS in general are those that have been allocated number resources from an RIR or an LIR that has been allocated number resources from an LIR. Recall the discussion of monitors. These actors hold the minimal bundle of NRS rights. In terms of numbers, these actors have satisfied allocation criteria necessary to be allocated some set of numbers, indicating both access and withdrawal. As such, monitors are numbers users as defined in the next section. In terms of routes, monitors are simply entrants—as per earlier discussion they do not exchange routes, thus not affecting the global stock of routes.

The RIR system prides itself in access to anyone that wishes to participate in the RIR as a forum for *discussing* number policy. Participation in the e-mail discussion lists has the lowest barriers to entry. Participation in the RIR meetings requires registration, a fee, and travel to the venue. That said, it is not limited to members. Entry does not guarantee withdrawal. Anyone that uses the Internet has indirect access (as defined in Section 3.1) to the use of number resources.

The primary resource maintained by the RIR is the registry, implemented as a database. A registry entrant may perform limited queries that do not diminish the ability for the system to service other entrants. At a particular load, quality of service suffers and the resource becomes rival in computational load. The RIRs offer a variety of membership options for creating copies of the registry to offset load on public databases. The threshold between nonsubtractable access and subtractable withdrawal is the computational threshold at which service diminishes.

IXes do not have a class of participant that fulfills the role of an authorized entrant. Monitors in the context of the NRS are not entrants in an IX. Rather, monitors at the minimum, are users in commercial IXes and potentially proprietors in membership based IXes. In either situation, both IXes and federating organizations limit access to IX participants with select exceptions.

IPBLs have limited entrants. Many of the IPBLs provide a query interface that allows limited queries to the IPBL database. Here, query load determines the threshold between access and withdrawal. Access is defined earlier as utilization of nonsubtractable resources. Limited queries are those that do not tax the system. When utilization becomes rival a load threshold. At this point, utilization is no longer access, but withdrawal.

3.5.2 Authorized Users

As per Table 3.2, authorized users, or simply users, comprise bundles of access and withdrawal rights. Authorized users of numbers in the NRS are those that have been allocated numbers by either the RIR or suballocated numbers by an LIR to whom numbers had been allocated by a RIR. Authorized users of routes are those that have access to BGP session established between actors and that engage in an exchange in routes. They are authorized in the sense that they follow the attending contract dictating behavior within that BGP session.

In the RIR system, authorized users of the registry follows similar thresholds for nonsubtractive access and subtractive withdrawal from the registry. In the IPBL, authorized use follows a similar pattern. In the case of the RIR, authorized copies of the registry require a contractual obligation in all but the RIPE region, and with explicit limitations on commercial use. In the IPBLs, legitimate real-time copies typically require a contract and payment. In both cases, users are common.

In membership-based IXes, authorized users are relatively new and the specific rights bundle varies from IX to IX. As implied by membership-based, IX participants have historically been members whose bundle comprises primary rights, access and withdrawal, and secondary rights of management and exclusion. More recently, membership-based IXes have created alternate bundles of participant rights that correspond to authorized users. These actors typically have access to the same IX *services* as a member, but do not have what IXes refer to as "voting" rights.[153] The introduction of users has been a contentious issue in some IXes and is also a distinguishing characteristic in IX growth patterns.

3.5.3 Claimant

Claimants are defined as

[153]As detailed in Chapter 6, these actors may also be distinguished by the means by which they physically connect to the interconnection fabric, such as through "remote" interconnection or at remote (outside the original home metro of he IX) nodes.

> individuals who possess the same rights as authorized users plus the collective-choice right of management. With the right of management, claimants have the collective-choice authority to devise operational-level rights of withdrawal. (E. Ostrom & Schlager, 1996, loc. 2546–2548)

No single actor has universal secondary, collective-choice rights for the entirety of the NRS (or Internet for that matter). The distributed character is the hear of this discussion—different institutions manage different bundles of numbers and route rights. In the case of numbers, RIRs control most secondary rights in numbers that are conferred to actors. IPBL reputation has the potential to diminish rights. This action is considered a form of management, as discussed above. As such, the IPBL as a firm is a claimant within the NRS and the data plane.

In terms of CRIs, claimants are included here for completeness. As noted above, in terms of bundles of rights, IPBLs participants are either entrants and users. IPBL as a firm holds management, exclusion, and alienation rights. In RIRs and IXes, members have rights of management *and* exclusion. RIRs differ in that number rights are also alienable.

3.5.4 Proprietor

As per E. Ostrom and Schlager (1996)

> "Proprietors" are defined as individuals who possess collective-choice rights to participate in management and exclusion. Proprietors authorize who may access resources and how resources may be utilized; however, they do not have the right to alienate either of those collective-choice rights. (1996, loc. 2568–2570)

Amongst the CRIs, proprietors exist in the IX and the RIR systems. In both cases, the firm is the proprietor, although members shape management and access rules. Early on some of these were shaped by the type of connectivity outside the IX a particular network maintained. More recently, exclusion rights have distinguished customers from members. These rights have evolved differently in different IXes. More recently developed IXes, such as France-IX and those to be based on Open-IX standards, have been able to streamline rights bundles with the benefit of hindsight rather than incremental changes to more conservative bundles of participant rights.

3.5.5 Owner

Per E. Ostrom and Schlager (1996):

> If in addition to collective-choice rights of management and exclusion, individuals also hold the right of alienation, that is, they can sell or lease their collective-choice rights, then they are defined as "owners." (1996, loc. 2584–2586)

Cole and Ostrom offer a broader (updated) notion of alienation:

> the right to sell one or more of the first four rights permanently or for a given time period. Most attention has been given to the right to transfer full ownership of a segment of a resource that would involve having all four of the other rights. Some forms of alienation are not that general, but still assign the right to sell *some meaningful subset of the rights that are held by a participant*.(Cole & Ostrom, 2012b, loc. 1178–1181, emphasis added)

The emphasis draws attention to the key update from Ostrom's articulation presented in Section 3.4.5. In Ostrom's earlier articulations, alienability of access and/or withdrawal was tacit in the rules constructed under the application of rights of exclusion and management, respectively. Under the updated articulation, alienability is a (secondary) collective-choice right that may be bound to any of the other rights.

Amongst the CRIs, IXes and RIRs have participants that may be qualified as owners, but in very different capacities. A superficial analysis of the membership-based IX model would seem to imply ownership, but that is not the root of ownership as defined here. In the membership-based IX model, participants and firm leadership often explain the model firm as "having a single stakeholder, the collective membership." Colloquially this is often referred to as members having "ownership" of the IX, especially given their membership fees fund purchase of equipment and facilities. Rather, this is more akin to delegation than ownership based in the rights discussed here.

Under the early definition, alienation only pertained to secondary rights. In the RIR system, like other systems, secondary rights are typically tied to resource allocation. In IXes, membership rights are also tied to allocation, but given IX capacity is not the same kind of resource unit, the binding of resource unit allocation and membership rights differ. In the RIR, IPv4 addresses are strictly finite—the pool cannot be expanded. At any given instant, capacity on any IX is finite, but the growth model ensures that capacity is added at discrete points as participation and/or overall utilization grows. In this sense, there is no need for capacity to be alienable between participants on the IX. In contrast, allocation of origination rights from the global stock of IPv4 addresses, especially given depletion is quickly approaching, is a useful right. Moreover, the precise character of alienation of origination rights amongst participants within a given region and across regions has been a longstanding point of contention.

Distinguished IX participants have a narrowly scoped form of ownership. Resellers are IX participants, typically members, that delegate IX access and withdrawal rights to their customers. This requires coordination with the IX firm management, but it is a sanctioned mode of IX participation. Typically, remote participants are customers—only primary rights are delegated. In this sense, only these actors may be considered "owners" in the sense established by this rights framework. That said, any member with sufficient infrastructure could arguably become such an owner.

3.6 Loosely Organized

NRS institutions have evolved from "close-knit yet loosely organized" communities into durable institutions for managing number resources and routing. Operational rules are the product of operational epistemic communities that continuously update management practices and number management resources to reflect the current operational landscape. The NRS has been presented as a resource rights management regime, in particular a man-made CPR in which operational rules shape both allocation resources and fundamental elements of the resources themselves. In such a system, the resource and the resource system, the infrastructure and institutions that mediate resource allocation, are both mutable. This chapter further disentangles stocks of resource units from the management infrastructure and institutions.

In Chapter 2, protocol provisioning of number resources was distinguished from operational provisioning of routes. Ostrom's classes of operational and collective choice property rights further enhances that distinction. Abstractly, the numbers and routes that comprise the control plane are a global resource.[154] NRS resource units are distinguished from local and function specific CRIs that enhance or diminish resource rights (through access to function) and system integrity. CRIs are themselves common resource systems—each provisions a resource used to manage particular rights in the NRS. The studies in Part II describe each CRI and the NRS institutions that maintain those facilities and arenas.

Framing NRS institutions in terms of resource rights sets the stage for evaluating the stability and durability of these CPRs and the NRS as an institutional complex in Part III. As noted in the introductory chapter, operational epistemic communities have in the NRS have historically operated in a relative regulatory vacuum. As the Internet, and in turn the NRS, have become a critical resource in the global economy, states have become increasingly interested in how this resource is managed. P. M. Haas (1992) considers this from the perspective of the state and its own knowledge problems:

> How states identify their interests and recognize the latitude of actions deemed appropriate in specific issue-areas of policymaking are functions of the manner in which the problems are understood by the policymakers or are represented by those to whom they turn for advice under conditions of uncertainty. (1992, p. 2)

NRS institutions are the products of operational epistemic communities and are positioning themselves as "those to whom" conventional policy makers turn to. An essential element of this process is to convey the stability

[154]Do recall that access to that resource is conditioned on local infrastructure and institutions, hence the NRS framed as an institutional complex.

Part II

Management Resource Studies

Chapter 4

Arenas and Knowledge Commons

CLUE IS THE currency of network operator groups (NOGs). At UKNOF 1, NOG activities were described as the "[d]istribution of clue," (Mitchell, 2005). Network operator vernacular comprises a variety of technical terms, but the use of "clue" speaks to the a fundamental ethos of having knowledge about the system. Note this does not denote a deterministic, predictive mode of evaluation. Rather, it implies one has knowledge gained from experience and from learning in the guild-like structure of the network operator communities. When one is said to have useful operational knowledge, they are referred to as being "clueful."

In contrast to the RIRs, IXes, and anti-abuse communities, NOGs and similar knowledge arenas do not confer or enhance rights the same sense that RIRs confer, IXes enhance, or anti-abuse. NOGs provide arenas in which participants, members of operational epistemic communities, can present current issues and learnings in their epistemic domain. In NOG vernacular, it is a market for clue. In certain regions NOGs also provide training services. These arenas comprise face-to-face meetings, conference-style presentations, working groups, Birds of a Feather (BoF) meetings of varying formality, and closed meetings amongst business partners. These are also supplemented by online mechanisms such as instant messaging (IM) chat rooms, Internet relay chat (IRC) back channels to various face-to-face engagements, e-mail lists, and interactive videocasting of events in some fora. Taken together this variety of communication mechanisms are either provisioned by the organization or jointly provisioned by the participants.

The effect is a set of communication mechanisms that facilitate and augment engagement amongst the operational epistemic community, enhancing the exchange of ideas and knowledge in those groups. Some of these groups emerged with the Internet, such as NANOG(NANOG, 2015b) in North America and RIPE(RIPE, 2014b) in Europe. Others have emerged as their region begins to have increasingly locally deployed Internet infrastructure. For instance, some are produced by existing CRIs, such as LACNIC fostering the development of LACNOG in the Latin American and Caribbean region.. Another instance is AfNOG, developed by ISOC, in Africa.

Regular NOG meetings include NANOG (North America), RIPE (Europe), PLNOG (Poland), Apricot (Asia Pacific), PacNOG (Pacific), UKNOF (United Kingdom), among a variety of other regional groups. NOGs serve to promulgate operational epistemic

knowledge within the community and reduce information asymmetries about the state of the system and operational practices. Historically, NOGs in developing regions have served as both fora for experts *and* loci of low-cost training workshops with often week-long education tracks in various core routing, network management, and security domains.[155] More recently NANOG has begun developing an education program and a program to reach out to college students in an effort to develop network operations skills early. Archiving mechanisms, such as e-mail list archives, vidoecasts, and conference transcripts, make the these semi-formal proceedings, and the knowledge generated therein, more durable.

NOGs are not nearly as complicated as the other CRIs, each of which maintains number resource management facilities (RIRs and IXes) or a knowledge commons documenting the navigation of a sophisticated ecosystem of abuse mitigation mechanisms (anti-abuse).[156] Section 4.1 describes NOG structure and function: general objectives, role promulgating the routing norms discussed in Chapter 2, promoting best operational practices, facilitating the presentation and documentation of operational epistemic communities' members' experience, and how they engage with adjacent communities. NOGs are the substrate of the NRS. Section 4.2 describes the how NOG-like arenas that sustain the social information sharing component of the other three CRIs. In each case, these are descriptions of how typically special purpose communities, rooted in the norms of the NOGs, support and facilitate CRI function. Finally, Section 4.3 builds on these instances of the substrate to highlight how NOGs characteristics in terms of the characteristics of operational epistemic communities. This not only highlights the common social and engagement structure, but also illustrates how NRS level norms such as the esteem of consensus decision making. Consensus processes is most accurately portrayed as a diverse family of consensus processes rather than just a single canonical model, the considering the NOG-like arenas a substrate helps understand how such a family developed from overlapping communities.

4.1 Structure and Guidelines

NOGs serve as one convening arena for regional, and in some cases global, operational epistemic communities. The norms of these arenas are quite simple: the objective is to share knowledge about the state of network operations without the noise of commercial sales and marketing activity. In this sense, these groups form what Hess and Ostrom (n.d.) refer to as a knowledge commons. In their work the term knowledge commons is used "to describe the complexity and variability of

[155] For instance, consider the recent PacNOG workshop on routing (Smith & Meynell, 2014) supported by the Network Startup Resource Center.

[156] It should be noted that while the anti-abuse community is rather complex, as an arena, it is much more akin to a NOG and the CRIs managing facilities. This will be discussed in Chapter 8 in terms of relational authority, differentiating between managing equilibrium when mediating access to a valuable resource or, in the case of commons, providing the operational knowledge that facilitates self-help.

knowledge and information as resources." (n.d., loc. 92). In terms of the operational epistemic communities described in Section 3.2, the knowledge commons is where these communities "store" and share domain knowledge.

Network operator groups or network operator forums[157] generally claim three general objectives:[158]

1. sharing technical knowledge and experience,
2. carrying out technical coordination within the Internet,
3. providing education services to the community.

NANOG "is an educational and operational forum for the coordination and dissemination of technical information related to backbone/enterprise networking technologies and operational practices" (NANOG, 2012). RIPE's history states that "Réseaux IP Européens (RIPE) began in 1989 when a group of IP network operators based in Europe began a series of regular meetings to share experiences and carry out technical coordination work," (RIPE, 2014b).[159] PLNOG, the Polish NOG, is one of the largest in terms of participation, and was relatively recently formed with education as its explicit remit:

> The first edition of PLNOG was held in January 2008 at the initiative of the PROIDEA Foundation for Supporting IT Education. The genesis of the creation of the conference was a discussion with representatives of the local ISP market and a willingness to chart new standards of cooperation between companies providing ICT services. ... Our mission is to:
>
> - support cooperation between network operators
> - facilitate an exchange of experiences between representatives of the ICT sector and the telecommunications education network
> - create an opportunity to work on the development of the ICT industry and ISPs in Poland
> - PLNOG achieves these objectives during a conference organized twice a year, as well as workshops and joint mailing lists. Over seven years we have managed to create a community that regularly gathers at our meetings and establishes new initiatives. (PLNOG, 2015)

NANOG started as, and remains a forum for sharing operational knowledge. The RIPE community engages in that role, but forums are more explicitly linked to the

[157]Mostly referred to as NOGs, although a number of NOFs, such as UKNOF, fall into the same category. For simplicity, all will be referred to as NOGs.

[158]This generalization is by the author, not a common generalization espoused explicitly by the NOGs. In other words, to the knowledge of the author there is no document akin to a IETF informational RFC representing any form of consensus on these objectives.

[159]There is a subtle difference in "coordination and dissemination of technical information" and "carry[ing] out technical coordination." One is purely information sharing, the other has an element of ordering.

RIPE NCC (which, among other tasks, fulfills the role of RIR for Europe, the Middle East, and Russia). The differences provide interesting insights into the varying degrees of coupling between the NOGs and the RIRs as two institutions with substantive participant overlap, but distinct remits.

4.1.1 NOG Structure

In terms of structure, NOGs are essentially conferences with well-maintained archives of presentations and community e-mail lists. As noted above, as a CRI, the NOG is a knowledge commons. The NOG facilitates communication amongst actors that hold and/or manage NRS resource rights, but the NOG itself does not confer or enhance rights. Rather, the administration of NOGs are typically volunteers that perform standard administration functions such as vetting presentation, managing the finances of the corresponding organization, activity planning, and collecting and managing membership fees. Most of these administrative roles are fulfilled by volunteers, although some NOGs, such as NANOG, have a paid executive director.

Volunteer administration ranges from a single "coordinating committee" such as with PacNOG (PacNOG, 2015) or a trade association with a board, program committee, development and membership committee, communications, fellowship, and education committees such as the case with NANOG.[160] There are various mixes in-between. For instance, SANOG comprises a core active management committee that is a subset of its advisory committee, a program committee responsible for program development and content, and a fellowship committee (SANOG, 2015b). SANOG is also an interesting instance of a NOG that was founded out of Nepal IX meetings (the founder of Nepal IX and SANOG is Gaurab Raj Upadhaya) and has a mix of internal management and adjacent, external support. For instance, SANOG lists a number of partnerships, (SANOG, 2015a, paraphrased below) many of which provide development partnerships and operational resources:

APIA (Asia and Pacific Internet Association) to coordinate SANOG and APIA Internet development in the region. This is an instance of a sub-regional NOG working with a larger regional group.

APNIC to integrate APNIC discussions and presentations into the SANOG program, this is part of the regional outreach efforts common to many of the RIRs.

NSRC (Network Support Resource Center) works with SANOG on some education efforts, donating books, coordinating fellowship funding, and developing the IP services track.

ISC (Internet System Consortium) provides web and e-mail hosting.

ISOC has provided funding for the fellowship program as well as helping develop workshops and training efforts.

[160] See the discussion of NANOG governance (NANOG, 2015d) for a listing of NANOG committees and links to full descriptions.

Netnod provides supporting secretariat services, in effect, backoffice supporting the core management committee.

SANOG is an instance of a sub-regional NOG whose support services, in terms of infrastructure (web and e-mail), knowledge dissimenation, and outreach is jointly provisioned by parties in the region (APIA, APNIC) and internationally (ISC, ISOC, Netnod).

In many cases, NOGs are stand-alone entities. NOGs are considered stand-alone even when jointly provisioned, such as the instance of SANOG above. The distinction is whether the program continues absent external motivation, if there is sufficient inertia on the existing management committee(s) to marshal resources on its own or is, as above, jointly provisioned. In other cases, they are largely provisioned by a particular existing CRI, typically a RIR. For instance, NANOG pre-dates ARIN by a number of years. While the two have substantive participant overlap[161], they are largely independent.[162] As noted earlier, the broader NRS operational epistemic community maintains a distinction between operations policy and number policy; this does not mean the meetings are distinctly segregated, though. For instance, RIPE as a community and the RIPE NCC as an RIR are distinct institutions, but their membership meetings are colocated, funded by the NCC, and there is little distinction between presentations focusing on community issues versus RIR issues (modulo content explicitly about the RIR, such as delegation statistics updates and reports from RIR firm leadership on the status of the firm). For instance, one *RIPE* Working Group is the Database Working Group, focusing on issues related to the number registry database; in effect, the community has a working group dedicated to the discussion of a facility managed by the firm.

In other cases, typically RIRs, have "spun up" new NOGs where they see a need or perceived demand for dedicated regional or subregional groups. LACNOG is a product of this process in the LAC region. MENOG (MENOG, 2015), the Middle East NOG, and ENOG (ENOG, 2015), the ENOG, the Eurasia NOG, have both been developed by the RIPE to support distinct regions in the rather large and culturally diverse RIPE region. Although there have been discussions of these NOGs becoming independent institutions, they remain largely efforts by the RIPE NCC. Another instance of a relatively new NOG is AfNOG, with partial support from the ISOC. As noted above SANOG, Southeast Asia NOG, developed by leadership of the Nepal IX. PLNOG was developed by an education foundation (PLNOG, 2015). All of these NOGs serve a similar function: provide a durable knowledge commons for network operators.

[161]ARIN meeting *participation* (different from membership) is a subset of the broader NANOG participation, but ARIN and NANOG are colocated once a year. Like outreach efforts between APNIC and SANOG, ARIN has developed "mini" policy sessions that occur during the NANOG program to garner as much contribution to the policy development process as possible. This will be discussed in terms of adjacencies in Section 4.2.

[162]It is important to note that when NANOG became an entity independent of Merit, ARIN provided NANOG a loan to get it started. That said, ARIN does not dictate NANOG activities.

4.1.2 Information Sharing

In terms of characteristics of institutions writ large, namely reducing uncertainty and the challenges of monitoring and enforcement of rules, the NOGs can be considered vehicles for reducing uncertainty. The common three objectives speak to information sharing, namely reducing the information asymmetries that drive uncertainty in transactions, and ultimately, increase overall transaction costs. The types of information sharing to reduce uncertainties in operations include questions and discussions about the performance of particular equipment, sources of upstream operational services, questions regarding the operational parameters of particular services, upgrade requirements for business expansion, interpretation of data from sources such as RouteViews (Advanced Network Technology Center, University of Oregon, 2015), searches for operational data, general routing mechanics puzzles that emerge in the wild, and a variety of other issues relating to operating a network.

A variety of technical topics are covered in informal conversations, working groups, and presentations:

- experience managing Internet connectivity,
- experience deploying new technologies and services,
- reports on resource utilization,
- reports on operations phenomena observed,
- best practices in network operations and problem solving strategies,
- tutorials and reviews of common operations tools, and
- (recently) more in-depth operations training.

Consider one of the many technical (and social) activities coordinated at the NOGs (among other fora), interconnection. As conferences comprised of network operators, many of whom manage interconnection arrangements for their employer, a substantive amount of the conference content relates to peering and routing. Conference content comprises tutorials, vetted conference presentations,[163] and perhaps most importantly, hallway conversations during the coffee breaks. For new network operators, the tutorials are often an introduction to best practices. For new operators and veterans, presentations by researchers (both academic and commercial), vendors, and experienced members of the community often help clarify complex contemporary topics in the field. Tutorials and presentations are certainly formal mechanisms that contribute to knowledge sharing. In developing regions, tutorials typically short, one-session courses.[164] Other regions provide week-long workshops that may be dedicated to one or a variety of technical topics. For instance, Apricot has hosted week long workshops with topic-specific tracks that span

[163]For instance, the NANOG program committee solicits and vets (peer reviews) conference presentations for technical quality, interest to the community, and timeliness.

[164]NANOG provides a dedicated listing of tutorials and presentation materials (NANOG, 2015e). Other NOGs provide archives of their programs, presentations, in some cases recordings of the actual event. Others, such as PacNOG (PacNOG, 2005, See section entitled "Available Items") provide an archive of the web pages for one time download rather than returning to the web site again and again.

the entire week. One the other hand, early PacNOG sessions had a single week-long track that covered a diverse set of topics in the network operations domain.

Informal mechanisms are just as important. For instance, one longtime community participant familiar with this work frequently tapped the author on the shoulder at conferences, circled their finger in the air indicating the activity in the social at hand, and says "This is what makes the Internet work" as they move on to their next conversation. Another longstanding member of the NANOG community provided a narrative whose variants has been repeated by others, indicating this general phenomena has been observed many times over: developing cooperative inter-firm relationships amongst operational actors. The original narrative describes the situation in which two peering managers are sitting against a wall during a rather crowded session. The two have discovered a routing issue between their networks and decided to resolve it then and there, in person. Working side-by-side the two solve the problem, but also gauge one another's technical capability. In effect, they are gauging whether, if this problem or a similar problem emerges, is this a person they believe will solve the problem effectively and can be counted on in an emergency situation. As scale, this is the process of developing a social network for skipping the standard network operations call center for resolving critical problems, but rather having an informal escalation network available.

Turning this to peering arrangements, sharing interconnection data within the social network is a valuable, and often lively topic. Dissemination of on-the-ground experience with network interconnection dynamics, often the topic of coffee-breaks and socials, is where uncertainty, and arguably transaction costs are substantively reduced. Sharing information about interconnection policies, the reputation of different networks and their operations centers, knowledge of the Internet topology, and the value proposition of interconnecting with one network over another is extraordinarily valuable. This information improves the bargaining position of actors in the interconnection market.

The social network also contributes to promulgation of norms and best practices. The social network is manifest in face-to-face meetings as well as NOG and other operator e-mail lists, and private channels that are often vetted based on reputation. The e-mail lists are rife with instances of individuals reporting problems to the community, sharing information collected from topologically different vantage points, and leveraging that information to find the root cause of the problem. Further, these also include reflections and recommendations.

In both these information sharing functions, sharing interconnection information and coordinating the resolution of routing mishaps, it is important to reaffirm the precise role of the NOG as an institution. Both cases rely on information sharing that is arguably already present in the community. As an institution, the NOG *facilitates* that information sharing and the dissemination of norms established in the community (through informal communications and best practices).

The interconnection narrative above implies monitoring and a question of enforcement; the Pakistan-YouTube story is an instance of monitoring and enforcement. A key difference is that responsibility for monitoring and enforcement is not in the remit of the NOG. Interconnection agreements are monitored and en-

forced by the participants [165] and are often informal.[166] In the Pakistan-YouTube case, monitoring and enforcement are clearly at play, but are a manifestation of the norms and informal communication mechanisms developed in the community, and, in part, promulgated in NOG communities. In this case, adapting John Gilmore's famous quote, rather than anthropomorphizing"the Internet" into an organic thing, "[network operators] see censorship as a damage and route[] around it," (Elmer-DeWitt & Jackson, 1993). More precisely, as developed in terms of the mechanics of Chapter 2, network operators see censorship as damage to the integrity of the routing table and act to correct it based on their norms regarding the accurate and legitimate provenance of routing information.[167]

4.1.3 Durable Contributions to the Knowledge Commons

Nominal descriptions of the NOGs' activities, mainly technical presentations, tutorials and social networking, may not distinguish them from the run-of-the-mill trade conferences. That said, the observed role of this information in day-to-day operations, the critical interdependence of the corresponding network-of-networks in the Internet, and the premium placed on quality of information sharing is evidence of significant differences between network operator communities from the conventional trade conference. In particular, the network operator norms certainly do not eschew gains by their individual organizations but multiple actors have referred to the idea of "co-opetition." Presentation guidelines seem like a rather mundane topic, but are quite telling of the distinction between a commercial trade show and NOGs functioning as operational epistemic communities.

Consider NANOG's presentation guidance under the heading "What Kinds of Topics Are Appropriate?"

> Attendees are quite sensitive to keeping NANOG presentations *non-commercial, and product pitches are strongly discouraged.* Repeated audience feedback shows that the most successful talks focus on *operational experience, research results, or case studies.* Presenters who are organizing a panel or BOF are encouraged to *include speakers from several (perhaps even competing)* companies and/or a neutral facilitator. (NANOG,

[165]This is widely the case in the US and EU economies. A counter-example is ARCEP in France, proposing the regular documentation and reporting of interconnection agreements in France, with French companies, or affecting French traffic. The closure captures quite a few more actors than merely those with physical presences in France. A broader, contemporary topic is the renegotiation of the ITRs in the ITU, in particular proposals for regulation of interconnection agreements. While adjacent to this work and important in the discussion of bottom-up organizations as governance arrangements and their engagement with their global "peers," the substance of ongoing ITR proposals will not be addressed in detail.

[166]A recent study by PCH indicates that 99.51% of interconnection arrangements in the sample are informal "handshake"agreements (Woodcock & Adhikari, 2011).

[167]A recently heated and ongoing policy debate in the community focuses on RPKI as a technical solution to the integrity of routing information. It is also a good illustration of the consensus process at play.

2015c, emphasis added here)[168]

The points emphasized are those that were stressed in early conversations and interviews. The first point is the non-commercial character. Program committee members have indicated they scrutinize presentations for this, often having to remind presenters of these guidelines. NANOG's presentation guidelines (NANOG, 2015a) provide explicit instructions for commercial content on slides. For instance, "company logos must only appear on the first and last slides." Again, seemingly innocuous, but this is an important balance in a community of private actors contributing to an operational epistemic community, creating and maintaining knowledge in the(ir) common interest.

Presentation guidelines provide well-documented illustrations of the character and ethos of contributing to problem solving. This ethos extends well beyond presentation guidelines though. A number of instances and observations from fieldwork, reaffirmed in conversations and interviews, illustrate the point. The first two points are in regard to the role of vendors at various NOGs. The last points are direct experiences of the author engaging with participants.

Early in fieldwork a longstanding participant was describing the balance of commercial presentations and the character of participants in private conversation between sessions. This actor indicated that sending marketing actors, rather than those with operational authority within their firm, to NANOG was not completely prohibited, but was strongly discouraged. Participation means one contributes to the operational knowledge domain. The extreme form highlighted, in contrast to more commercial fora, is the absence of "booth babes." Booth babes are attractive, suggestively clad women present at vendor booths to "encourage" engagement. Typically these are not even employees of the vendor, simply "eye candy." Despite the old adage the "sex sells," this is strongly discouraged amongst the NOGs and will likely get the vendor a less than friendly conversation with NOG leadership.[169] That said, many of the NOGs do benefit from vendor sponsorship, but the exposure is limited and controlled. Vendors can sponsor particular meals provided to participants, sponsor coffee breaks, and social activities; their banner will be visible as the sponsor, but the aggressive sales pitch is strongly discouraged. Moreover, that

[168]The RIPE community has a *very* similar statement:

> RIPE Meeting attendees are quite sensitive to keeping presentations non-commercial, and product marketing talks are strongly discouraged. Repeated audience feedback shows that the most successful talks focus on operational experience, research results, or case studies. For example, presenters wishing to describe a commercial solution should focus on the underlying technology and not attempt a product demonstration. (RIPE, 2015)

It is not clear which community lifted the other's guidelines, it is also quite possible the two statements have the same individual author. As a third point of contrast, PacNOG's presentation guidelines are a near duplicate of NANOG's and make that point up front.

[169]To the author's surprise, ICANN is not as strict in this regard. While vendors were partitioned into their own "vendor village," where participants had to go out of their way to engage with largely commercial booths and presenters, booth babes were definitely present and aggressive in their sales pitches.

discouragement is enforced.

A number of NOGs have special areas, "vendor villages" that are not in the main causeway but available to participants. An alternative is to have such a vendor village serve as a social activity. The canonical instance is NANOG's "Beer 'n' Gear" session. This is a regular social activity in a large ballroom where vendor booths line the walls, the next inner ring is food buffets, and the center is typically the bar. To attract participants the food and alcohol is sponsored by the vendors, but the large center area is mainly a social area and vendors stick to the area immediately around their booths. As an interesting balance, a number of participants at this activity feel obligated to at least make a cursory circuit around the perimeter to view the booths. Conversations with vendors are focused on product capabilities; recall these participants are engineers and they expect "non-sales pitch" answers to questions.

The final two instances of an ethos of problem solving from fieldwork are instances of engagement between the author and community members. At one regular informal social event late in the evening the author was in conversation with a longstanding member of a closed community that is well-known for substantive contributions. After a brief introduction, this participant exclaimed, "You are just another goddamn academic voyuer!" Since then, after repeated interactions and discussions, the author has a warm rapport with this individual. In another instance, the author was in transit to a social activity with a founding member of a closed community and a few other members of that community. Upon introduction to the other members, one individual politely yet bluntly asked, "Do you have any operational experience?" In effect, challenging the assertions of a relative outsider. The longstanding member quickly diffused the situation by saying, "Jesse and I have had long conversations. He has clue, he understands what we are doing." Both of these instances illustrate both the problem solving ethos of the NOG-like communities as operational epistemic communities and a degree of skepticism of outside actors.

Returning the character of presentations, recall from the discussion of operational epistemic communities that participation in these communities has valuable benefits in terms of prestige. Including some logo information indicates this actor is in fact a member of the community, has operational experience (which should also be born out in the content), and is willing to represent the knowledge base and capabilities of her firm. The guidelines go on to describe the characteristics of a good talk:

- Highlight operational experience, i.e., present a case study.
- Identify anomalies or counter-intuitive (interesting) aspects of your experience
- Educate in your area of expertise (so the audience can learn something)
- Motivate action (so the audience goes out and does something as a result of the talk)
- Entertain (so the audience stays in the room) (NANOG, 2015a)

These guidelines are a refinement of the second emphasized point above, focusing on operational experience, research, and case studies. In community vernacular, demonstrate you have clue and offer some to others.

The final emphasized point, including speakers from potentially competing firms, is common and quite salient to this discussion. Operational epistemic communities that rely on are not in perpetual harmony. Quite the opposite, they are constantly navigating uncertainties in the routing system management and, given both *a*) different sub-sector incentives and motivations and *b*) different vantage points, data, and experience from which to develop rationales for system behavior the network operator community does, and will continue, to disagree on the salience of a number of topics. Superficially, this would seem to create chaotic presentations. Rather, the contrast of differentiated views from respected experts in the community provides the community with contrasting views. This serves to both educate the community and to give the community the opportunity to judge the arguments and merits of the data presented on their own.[170]

In terms of credible knowledge assessment, this is referred to as the adversarial model. Not surprisingly, this can get heated. A good, vehement, constructive debate can be quite instructive and insightful. There is a fine line between that and outright aggression, though.[171] NOGs and other knowledge arenas make engagement norms clear; consider the norms for UKNOF below:

> By participating in UKNOF activities you are confirming your commitment to follow these principles.
>
> - When you participate in UKNOF remember to show others respect and courtesy regardless of who they are and who you are, whether its online or in person at a UKNOF event.
> - That doesn't mean you cant disagree or have a heated debate, just remember to avoid getting personal.
>   - Nobody should feel hesitant or afraid to participate in discussions.
>   - Focus on the idea, and not the individual on the other side of the debate. (UKNOF, 2014)

Often, these norms are articulated at the beginnings of presentations. Occasionally these need to be reiterated during question and answer sessions, especially when recalcitrant community members make statements considered inappropriate. These norms are also invoked on the mailing lists.

[170]For instance, at NANOG 55 in Vancouver a panel of CDN architects discussed content delivery, framed by a moderator then employed by a large European access network (Bargisen et al., 2012, See CDN on Tuesday agenda). Another instance is the discussion of the effects of exogenous factors on network performance (Sharma et al., 2012) at NANOG 56 and a discussion of traffic accounting (Jasinska, Lucente, van Dussen, Vijn, & Hughes, 2012).

[171]An illustrative instance is the dialogue between Malcolm Hutty and Steve Kent at RIPE 63 in Vienna (Karrenberg et al., 2011). While this discussion did not "come to blows," it was rather tense.

The community dynamics seem to support the argument that these private actors engage in public forums, constructed to facilitate the pursuit of private (at the firm level) interests, yet consistently produce knowledge as a common good. Arguably, a functioning interconnection market, rooted in part in this common good, is a contribution to the public good. This is not purely incidental, many justifications of this cooperation begin and end with "for the good of the Internet."

4.2 Arenas Supporting CRIs

As arenas, NOGs, such as early NANOG and RARE, were the first operational epistemic communities to emerge around sharing information and coordinating network operations. In addition to the "traditional" NOGs discussed in the previous chapter, other CRI fora have "NOG-like" arenas that combine information sharing and decision-making processes. These arenas, such as IX federating collectives (Euro-IX and Open-IX) are largely about exchanging information about IX operations, but have a minimal set of decision-making rules for administrating the finances of the meeting forum and information resources such as web pages and databases. The following describe the arenas supplementing each of the CRIs in Chapters 5-7.

4.2.1 RIR Arenas

RIRs conference meetings have many of the same qualities as NOG arenas. In the RIR, general presentations and information sessions operate very similar NOG presentations. Although developed more extensively in Section 5.5, two distinguished arenas in the RIR community are the Policy Development arena (Section 5.5.1) and Government Arenas (Section 5.5.2). The policy arena part presentation, but largely driven to facilitate engagement and discussion of policy issues. It is in these policy arenas the domain knowledge created in the NOGs and through experience is leveraged to develop, articulate, and evaluate policy proposals.

Government arenas differ, in particular because they are an attempt to facilitate communication between conventional government actors and the operational epistemic communities. Although there are distinct rules governing respectful engagement in the various NOGs, the culture and tact of engineers debating amongst themselves is very foreign to many state regulators or representatives. The mode of constructive, adversarial engagement may not be suitable for sharing information with government actors, in particular those new to RIRs and other fora. Government arenas are established to facilitate that engagement. As per Section 5.5.2, they range from an open working group dedicated to cooperation with governments and external authorities, such as the Cooperation WG in RIPE to closed government working groups such as the ARIN Government Working Group in ARIN or the RIPE NCC Roundtable, which closed meetings between RIR staff, invited members, and government officials.

As earlier, the NOGs attract a large number of operators given this is both beneficial for keeping up with the epistemic domain, but also because many negotiations and customer meetings take place. To keep pace with NOG activity, the RIRs send staff to the NOGs that are not colocated with RIR meetings. In many cases staff make presentations updating NOG participants regarding RIR activities, resource statistics, RIR firm activities and changes, and governance activities. One particularly interesting development is the recent creation of the Public Policy Consultations managed by ARIN, held at NANOG meetings. These are RIR policy arenas transposed into the NOG. The objective is to provide actors that would not normally attend an ARIN meeting the opportunity to participate in current policy debates. As will be developed in Section 5.6.2, contributions to the consensus process can come from e-mail lists or active participation at a consultation. The NANOG PPC meetings collect additional input, but it is not the case that the only consultation with the community on a policy occurs at one of the "remote" consultations.

4.2.2 IX Arenas

Amongst the associational membership IXes, membership meetings also have a distinct NOG-like character. These meetings have a similar knowledge sharing quality, scoped to IX participation and development. IXes invite guest speakers, typically participants, to related their experiences to the rest of the community. They also report on the state of the platform and new activities. They may also invite outsiders to present material salient to the interests of the community.[172]

Within the NOGs themselves, the IXes have historically had a number of venues to present platform updates and their services. For instance, during the Peering BOF at NANOG, new IXes would present themselves and briefly describe the characteristics of their platform: location of PoPs, switching vendor, available capacities, prices, particular characteristics such as route servers, and any other distinguishing characteristics. In observed Peering BOFs, there were run adjacent to Peering Personals, individual networks describing who they were, what types of peers they were looking for, under what contractual mode (typically settlement-free), and what colocation facilities and IXes they were located at. In some cases IXes will submit presentations to the main sessions of NOGs. A particularly famous IX presentation is Andy Davidson's presentation on how to build an IX. Amongst the tutorials and trainings at NOGs, there have also been tutorials on IX engagement and route server use. Amongst the best common operational practices (BCOP), the recent NANOG BCOP series includes "public exchange" configuration best practices—essentially, how to connect to an IX.[173] Finally, in some NOGs there have dedicated IX working groups, such as the former European IX (EIX) WG[174] in the RIPE and the dissolved IX SIG that met in the NOG capacity of the APNIC community.

[172]The author spoke at an AMS-IX meeting on the ITU as a would-be principal for the NRS.

[173]Both individual IXes and Euro-IX also provide this information.

[174]See (RIPE NCC, 2013a) for archives. The EIX was replaced by a more general purpose Connect BOF for discussing interconnection issues.

In addition to colocation with NOGs, a number of NOG-like IX and interconnection focused communities exist and either have colocated events with, or make regular presentations at, NOGs.

Euro-IX is described in Section 6.3.2.1 as an IX association, serving as a "NOG" for IX operators. The informational sessions are very much like NOG sessions. The Euro-IX Secretariat regularly attends various NOGs to update those communities on the status of Euro-IX (number of members, new services, special events) and to elicit additional members, namely new or existing IXes in that NOG's home region.

Open-IX is described in Section 6.3.2.2, it is an IX standards development organization promulgating the associational membership model of IXes in the United States. Open-IX is a recent development, with early framing documents circulating in late 2013. . Meetings to discuss the feasibility of creating an organization such as Open-IX were held at NANOG meetings as a special BOF initially for interested parties. Now that Open-IX has become a full-fledged organization, it has meetings and closed membership socials at NANOG.

GPF is the Global Peering Forum, a yearly semi-closed forum for peering managers, IXes, and colocation facility representatives to present on the state of the interconnection market and trends as well as engage in interconnection negotiations. Although it is populated by many of the actors that participate in NOGs, its primary remit is to facilitate business transactions. Currently, Open-IX is proposing an alternative, the Americas Interconnection Summit in April of 2015 as an open alternative to a semi-closed GPF.

In addition to these, individual IXes also make use of the NOGs as convening fora. For instance, a number of IX leadership have indicated that they attend NOGs to meet with members that can justify a NOG, but cannot justify a dedicated IX meeting. Like the RIRs, the IX meeting participation is typically lower than that of NOG participation. NOG participants typically have limited travel budget and time, thus only attending NOGs as the venues where they can have the most meetings. As such, both the RIRs and the IXes send representatives to engage with their constituents.

4.2.3 Anti-Abuse Arenas

As a CRI, in contrast to the RIRs and the IXes, like the NOGs, the value of fora such as M^{3}AAWG is the arena itself as a knowledge commons. In particular, M^{3}AAWG is akin to an anti-abuse NOG that has a very well-developed best common practices development process, documented in Section 7.4.2 as its collective choice process. Similar to NANOG's best practices, M^{3}AAWG best practices are intended to be living documents. In terms of adjacencies with the NOG community, M^{3}AAWG has presented at NANOG and RIPE's anti-abuse WG recently. While M^{3}AAWG has made these outreach efforts, it is a closed forum.

4.3 Operational Epistemic Substrate of CRIs

The NOGs, and NOG-like arenas in the various CRIs are the normative and epistemic substrate of the NRS. These are the social networks in which many of the norms and values that shape the NRS were developed or that served as the catalysts contributing to the creation of CRIs. NOGs and the artifacts created by NOGs, are operational epistemic communities made durable. Recall Ostrom's quote regarding uncertainty in Section 3.2.1, in particular the "skillful pooling and blending of scientific knowledge and local time-and-place knowledge," (E. Ostrom, 1990, p. 33) as a means to reduce uncertainty. NOG conveners and administrators are the agents of "skillful pooling and blending."

For instance, in Haas's definition of epistemic community, he indicates "an epistemic community may consist of professionals from a variety of disciplines and backgrounds," (P. M. Haas, 1992, p. 3). Encouraging that diversity is another epistemic community characteristic, rooted in the NOGs. It is immediately evidenced in the encouragement of NANOG panels to comprise nominal competitors. It is even more evident in the efforts to elicit constructive criticism in the collective-choice processes of the CRIs. Although in some settings this would potentially create conflict, recall Haas's argument that epistemic communities are engaged in "a common policy enterprise" (1992, p. 3). That enterprise is manifest in the common norms of integrity shared by these communities, ensuring legitimate route provisioning and consent-based messaging. This common enterprise limits the negative effects of "parochial interests" that will be discussed further in Chapter 8.

The NRS is not a static technical or social system, it is in a constant state of change and development. As historically informal knowledge commons, many NOGs do archive presentations, some archiving audio/video and transcripts. That said, these are not the easiest to parse or search. In a recent best common operational practices (BCOP) session at NANOG58 Aaron Hughes, Vice Chair of the NANOG BCOP Committee, indicated that:

> The spirit of the best current operational practices track is to have a living repository of BCOP documents. The simple problem statement here is basically that every one of these operator groups, they have presentations globally, they all get archived on their respective operator forums, usually without text and without audio, and as soon as they are presented they are almost useless because they are stale and it is really hard to determine what is current, living, relevant, and vetted. (Grundemann, 2013, introductory comments to BCOP session by Aaron Hughes)

This work agrees that the notion of a living document is good, but given how frequently discussants and interviews have pointed the author back to "seminal" NOG presentations, the argument that the archives are "useless" is considered a bit strong. Like M^{3}AAWG's best common practices, NANOG's recently formalized BCOP documentation process is an effort to effectively codify and maintain "local time-and-place knowledge," (E. Ostrom, 1990, p. 33) useful for mitigating the uncertainties endemic in common resource systems.

In terms of (operational) epistemic communities, best common practices in many NOGs are a codification of "shared causal beliefs, which are derived from their analysis of practices leading or contributing to a central set of problems," (P. M. Haas, 1992, p. 3). In NANOG, the notion of BCOPs has been discussed in special purpose BOFs, but has only recently (in the last year) been formalized into a committee within the NANOG administrative structure. In effect, in terms of formalization, it appears to have only gained sufficient traction to become institutionally durable recently. In contrast, M^3AAWG has a long standing history of actively developing BCPs, but the process has only recently been codified. Returning to the discussion of Layton's phases of industrial research development, Section 3.2.4 indicates that the NRS is between developing a style of knowledge codification and large-scale production. The state of the BCP development process provides further evidence: BCPs are an effective codification, but one step below the rigor of standards that would facilitate routinized application at a large scale.

The notion of operational epistemic communities provide the analytic frame for understanding the network operator community. As developed here, the NOGs and NOG-like structures produce knowledge commons of varying formality. The ethos of these form the substrate of the modern CRIs. The remaining chapters of Part II build on the notion of a knowledge commons in explaining how each function specific CRI contributes to maintaining the integrity of the NRS.

Chapter 5

Regional Internet Registries

REGIONAL INTERNET REGISTRIES (RIRs) manage the stock of number resources. Unique delegation of number resource rights and the accurate documentation of those delegations is the core function of the RIR. Ensuring uniqueness is but one dimension of the RIRs' historic core norms. Effective resource *management* requires understanding the trade-offs endemic in particular delegation strategies. The number system is the foundation for the routing system. Although there has been a historical separation between number resource policy and operations, delegation strategies do have implications for the provision of routes. For instance, managing the *rate* of delegation and monitoring subsequent utilization, historically in terms of projected utilization criteria,[175] has been both a perennial topic and a recent point of contention addressed in this chapter. Fragmentation and reservation of number resource blocks is another concern. Within the operational epistemic community a common image[176] of IR goals considers resource policy, developed through a rough consensus process amongst an operational epistemic community, an act of balancing goals of uniqueness, routability, registry accuracy, and conservation.[177] In terms of the CPR framing, the RIR regime creates, evaluates, implements, and updates meaningful bundles[178] of number rights through a consensus-based resource rights management regime.

Broadly speaking, the "RIR" is an ensemble of actors and management facilities. Among these, the essential elements are the number registry itself, a collective-

[175]This is needs-based criteria in the community vernacular.

[176]Historically, the common image of this system has historically been RFC 2050 (Hubbard, Kosters, Conrad, Karrenberg, & Postel, 1996), but has been recently updated by RFC 7020 (Housley, Curran, Huston, & Conrad, 2013). See Section 5.7.1 for discussion.

[177]Depending on which constituency within the operational epistemic community one asks, conservation is no long a common goal for IPv4. Rather, it is a barrier to IPv6 transition.

[178]Meaningful bundles are used here in the sense established (E. Ostrom & Schlager, 1996). For instance,

> The exercise of withdrawal rights is not meaningful without the right of access; alienation rights depend upon having rights to be transferred. (E. Ostrom & Schlager, 1996, loc 2525-2526)

Meaningful means that a bundle of rights actually comprises the set of rights necessary to claim the value derived from some use.

choice arena, and a firm delegated the authority to implement and enforce resource policy. Together this ensemble jointly manages the stock of number resources. The "core" resource management facility is a jointly managed *number registry* that documents unique delegations of number rights to entities and contact information for those entities. A combination of the firm and the operational epistemic community provision a *collective-choice arena* in which the operational epistemic community develops constitutional, collective-choice, and operational rules that shape number rights management. Third, this ensemble comprises a *firm*, typically the object of references to the "RIR," to whom the community (continuously) delegates authority to implement and maintain the number registry and provide administrative support to community resource management processes such as policy development.

Tacit in each of these elements is the operational epistemic community, here the RIR community. The RIR community is loosely framed as a principal that animates collective choice arenas and delegates authority to the RIR..[179] Within this community, RIR members have effectively delegated custodial duties to the RIR. As this analysis develops, first in the next section (5.1 and later in Section 5.7, the authority more akin to Lake's notion of relational authority will be used to explain the difference between delegation of resource rights and authority.[180]

The RIRs differ in how these elements are operationalized. RFC 2050 (Hubbard et al., 1996) provided early guidance for constitutional rules and some operational practices, but RFC 2050 also left substantial room for discretionary development of regional differences. For instance, as foreshadowed in Section 3.2.5 on consensus in epistemic communities, Section 5.6.2 highlights regional differences in consensus processes. Contemporary resource policy offers compelling instances of these differences. Trade-offs between needs-based criteria and transfers first in the APNIC region and more recently in the RIPE region comprise one point of contention. The debate over RPKI in the RIPE region is another. These issues are discussed at length in Section 5.7.

The RIR (as a firm) plays an *administrative* role: it supports and helps coordinate community policy development, but it has limited influence over the substance of resource policy.[181] The RIR also provides supporting resources. Operators access the registry to determine who holds rights to which resources. Contact information

[179]The RIR and surrounding community, like other organizational and institutional ensembles in the Internet governance ecosystem, tends to place a strong focus on the community. Absent closer inspection, rhetoric espousing the bottom-up model could paint a picture of an ideal principal and the three elements above as agents that, as so frequently indicated by the RIRs, "do the will of the community." Here the focus is on the institutional structure supporting effective operational rules for managing rights related to the use of numbers and routes.

[180]This chapter provides evidence of the dynamics of delegation and authority in the RIR as an empirical setting. Chapter 8, in particular Section 8.1, develops the balance between delegation and authority further.

[181]There are certainly exceptions. Early in the RIRs history employees of the RIR have proposed policies. These policies were still subject to ratification through the collective choice rules (consensus) discussed in Section 5.6.2. The other primary source of influence is evaluation of the process, performed by the equivalent of the RIR board in all but the RIPE region where evaluation is performed by the collective of working group chairs.

is an essential (albeit not complete) contributor to resolving externalities discussed in Section 2.1.2. Registry access and utilization mechanisms have evolved over time, developing their own technical vernacular and access tools. The RIR provides training to ensure the operational community can effectively make use of and maintain the registry; the scope and extent of training is region-specific. To varying degrees, the RIRs provide dedicated analysis of NRS trends and dynamics such as delegation statistics, routing system trends, number resource security updates. Finally, RIRs also represent their respective region and RIR institutional complex writ large to actors in the conventional global political arena, such as governments and IGOs, as well as to other organizations and institutions within the broader Internet governance ecosystem.

As a resource governance regime, the RIR provides a variety of facilities for enhancing member rights and the arenas in which these rights are developed. Various types of facilities are provided: *a*) rights documentation, such as the registry and certification resources; *b*) knowledge services, such as resource utilization reporting,[182] active measurement efforts;[183] and *c*) arenas for community engagement and discussion such as conferences, fora for consensus development, training, and online fora. In addition to these, the RIRs are also investing in the development of diplomatic capability in the form of external relations activities and resources. This latter is an increasingly important representation function.

The five modern IRs were created between 1992 and 2005. Initially APNIC, the RIPE NCC, and ARIN were created to perform regional number delegations, coordinating with the organization performing the IANA function at the time. Later LACNIC and AFRINIC were created as a coordinated effort between both the IANA and the existing RIRs formerly managing those regions. Framing the process in terms of delegation helps explain a path dependent story of resource management institutions. Since then, five RIRs coordinate through the Number Resource Organization (NRO) to manage the global stock of number resources and attendant supporting facilities.[184] In Part III the delegation story will be revisited and revised in terms of relational authority offered by Lake (2010). As the chapter progresses, divergence from strict principal agents relations will be called out to both highlight those differences but also to set the stage for a more nuanced discussion of relational authority both within the NRS institutions, across these institutions, and in a global political arena of would-be principals.

Recall Section 2.1.1, the number resource stock is the result of protocol provisioning in RFC 791 (Postel, 1981). Section 5.1 provides an overview of delegation patterns and structures in the modern RIR system. Section 5.1, in particular Figure 5-1, summarizes global delegation patterns from this stock. Although rather simple when depicted in aggregate, these trends have been interpreted as inequitable. In particular early "legacy" delegations of large "classful" address blocks has resulted in under-utilization in the US, scarcity in other regions, and a heated

[182]Each of the RIRs provides resource utilization reporting at RIR meetings and on their web sites.
[183]The most notable active measurement effort is the Atlas Probes; see (RIPE NCC, 2014g).
[184]See Section 5.4.5 for a discussion of the NRO.

debate over the dynamics of a transfer market (see Section 5.7.3). The structure of the modern registry system and the attendant operational provisioning (delegation) patterns are presented in Section 5.2. In particular, the simple hierarchy in Figure 5-4 depicts the paths along which meaningful bundles of delegation rights are delegated within the federated collective of IR institutions that delegate meaningful bundles of rights within the number resource stock provisioned in RFC 791 (the IPv4 pool). Section 5.2 helps distinguish hierarchical *delegation* from strictly hierarchical *authority* and *operational capability*, a cornerstone of later arguments for relational authority.

Section 5.3 builds on the delegation paths presented in Section 5.2 to present the registry as a jointly managed facility. The registry and supporting facilities are described in terms of structure, implementation, and use. Resource management rules—operational, collective choice, and constitutional rules of the stock of numbers as a CPR—are presented in Section 5.6. Unique delegation, delegation rate, and the implications of delegation strategies are variants of constitutional norms. Historically, the authoritative source of these norms was RFC 2050 (Hubbard et al., 1996), supplemented by RIRs' own documentation (bylaws), and current policies (operational rules).[185] RIR-specific consensus processes, whose guiding norms were presented in Section 3.2.5, are presented as the collective-choice rules shaping the development of operational rules in Section 5.6.2. Finally, operational rules from resource policy corpora are presented in the section on operational rules (5.6.3). Following the discussion of rights embedded in complex rule systems, Section 3.4, the discussion of operational rules concludes by distilling differentiated bundles of number rights at play in the RIR system.

Rule development in the RIRs is a continual process of eliciting operational experience in order to create policy sustaining the integrity of the number resource stock.[186] To further illustrate resource policy as a mode of solving knowledge problems, select contemporary CPR management issues are presented in Section 5.7. Implied above, constitutional norms have recently been contested, in particular the role of conservation; policies to reaffirm these norms are discussed in Section 5.7.1 along with the implications of trade-offs discussed in the debate over number resource and NRS integrity. Origin security, as a means to ensure origin rights, has been a longstanding issue, particularly in the RIPE region. RPKI and its implications for path security are discussed in Section 5.7.4. Foreshadowed in the discussion of alienation as a generic resource right, transfers, in particular transfers of origin rights, has also been a longstanding issue. Stable transfer rights are increasingly important as the depletion of the unallocated pool looms larger. Section 5.6.3 distills existing transfer rules and proposed transfer rules into bundles of rights

[185]During the course of this study a new RFC on RIR norms has replaced RFC 2050. One of the issues presented as an instance of the application of common rule development in Section 5.7 is the development RIR norms articulations in multiple RIRs after deprecation of RFC 2050. See Section 5.7.1 for details.

[186]Section 3.2.5 provides the conceptual arguments for consensus processes as a means of eliciting tacit knowledge from the operational epistemic community. Chapter 8 revisits this argument, drawing on evidence from the studies in Part II.

comparable to allocation. That discussion sets the stage for a comparative discussion in Section 5.7.3, highlighting alienation of origin rights, community conflicts over notions of number resource ownership, and the role of the registry. Finally, needs-based criteria are discussed in terms of both post-depletion access rights, its relation to transfers, and implications for disintermediating the registry in the exercise of alienation rights. These discussions conclude by stressing the increased need for monitoring *and* enforcement, foreshadowing the confluence of resource reputation and alienability on the overall integrity of the system.

5.1 Number Distribution Trends

As a commonly managed resource, the global delegation of IPv4 addresses by region depicts the global stock and regional appropriation. This section depicts the global distribution of number resources by region, explaining some of the gross trends as background for discussing resource management practices and rights allocations that gave rise to those trends. In term of common resource management, this is a depiction of the rate of appropriation from the common pool of IPv4 resources. It also shows the origins of perceived inequities in distribution that still influence discussions of both IPv4 delegation and IPv6 delegation.

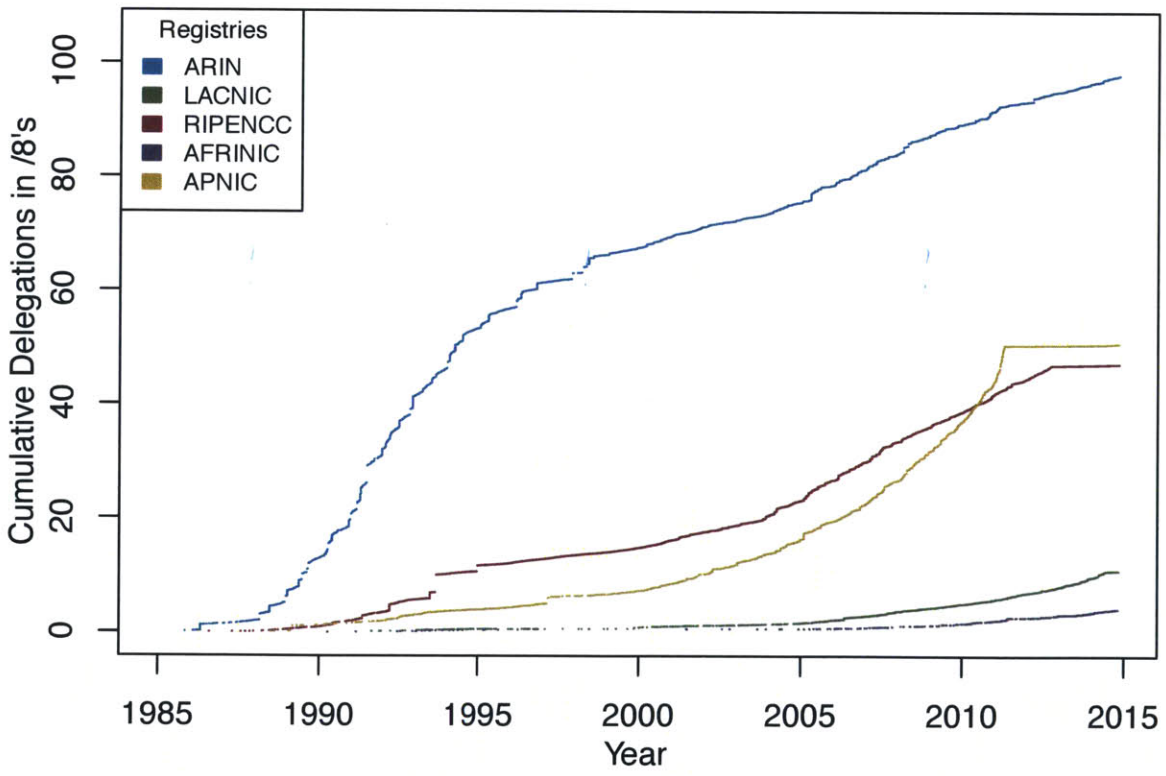

Figure 5-1: Global allocation of IPv4 addresses by the five RIRs. It is important to note that the colors denote the region in which those numbers were delegated. Early delegations were performed by the IANA, so for instance, before 1992, all of the delegations were performed by the IANA.

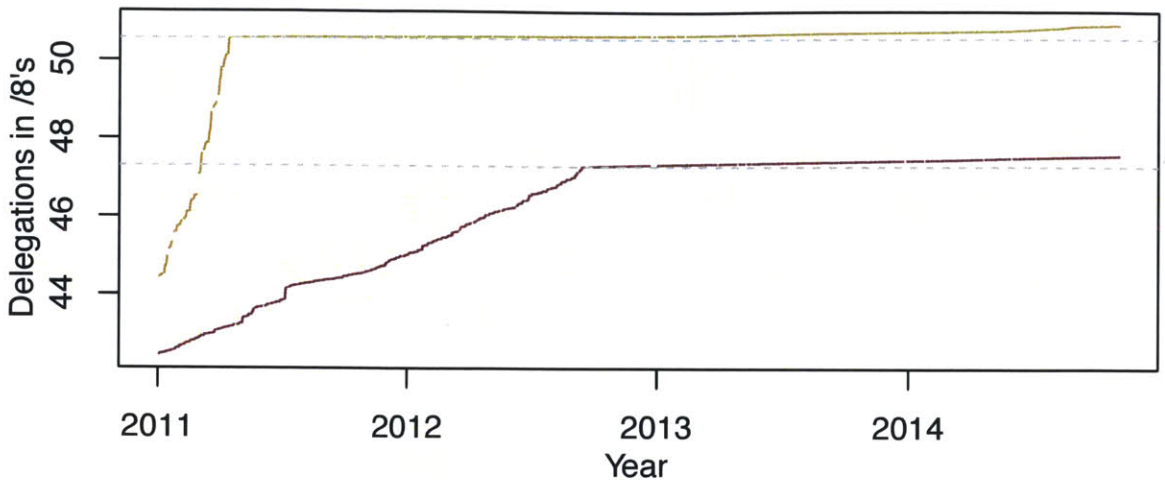

Figure 5-2: Detail view of APNIC and RIPE entering the exhaustion phase of IPv4 allocation. Each of the horizontal grey lines intercept the APNIC and RIPE at the x-coordinate corresponding to the date when that respective RIR's last /8 policy was activated.

As discussed briefly in Section 2.1.3, early number resource allocation was performed by the Jon Postel in his role performing the IANA function. Number rights delegation was later separated from names rights delegation,[187] delegating IR functions to regional IRs (RIRs) such as Network Solutions, APNIC, the RIPE NCC, and later to LACNIC and AFRINIC. Figures 5-1 and 5-2 depict IPv4 rights delegation trends since 1985.[188] In terms of withdrawal (appropriation) from the IPv4 pool, Figure 5-1 depicts the volume of that pool each entity has delegated.

The five modern RIRs are the RIPE NCC, APNIC, ARIN, LACNIC, and AFRINIC. Each is depicted by the color representing it in NRO documentation. In addition to these, Network Solutions, in its role as an IR, and the IANA as implemented by Jon Postel are also represented to depict what are now referred to as legacy delegations.[189] The trends shown in Figure 5-1 provide the backdrop for a coarse history of global numbers allocation.

In the 1980's, allocation proceeded slowly. The Internet was still an academic

[187]See Karrenberg, Ross, Wilson, and Nobile (2001, p. 24) for brief reference to the separation in the context of the Network Solutions contract for IR management and the creation of ARIN. See Mueller (2002) for one of the most frequently cited discussions of the development of the naming system.

[188]These graphs are produced based on data from the extended statistics file (NRO, 2014a) produced by the Number Resource Organization (NRO (NRO, 2014b), discussed in Section 5.4.5). The documentation for this file can be found at (NRO, 2014e). This file is updated daily by consolidating the corresponding files produced by each of the five RIRs.

[189]The delegated-extended file does not delineate legacy delegations form those performed by the modern RIRs. Rather, legacy resources are labeled by the IR in whose region the holder of these resources is based. For instance, MIT holds a legacy /8 and is categorized as in the ARIN region even though the delegation was performed by the IANA under the management of Jon Postel. To distinguish these, the dates when the IANA function was placed under contract with Network Solutions and the inception dates of the RIRs are used delimit rights delegations performed by historical IRs from those of the modern IRs.

experience, nothing like the economic engine it would become in the 1990's and onward. That said, one of the most noted characteristics of this graph, especially considering perceived inequity in number delegation, is the volume of number rights delegated in the ARIN region. In many discussions of IPv4 delegation, this is framed as unfair and in some cases even framed as conspiratorial. The imbalance is largely a function of the path-dependent growth of the Internet. In particular, early growth and allocation practices in the US and classful versus classless inter-domain routing (CIDR) contributes to much of this growth.

The second cause of the substantive allocation in the ARIN region is the transition from classful to CIDR-based subnetting. Under classful subnetting, three classes of resource delegation units were available:

Class A 2^{24}, or $16,777,216$, IPv4 addresses

Class B 2^{16}, or $65,536$, IPv4 addresses

Class C 2^{8}, or 256, IPv4 addresses

As may be obvious, this doesn't scale linearly. There are only 256 Class A blocks. Since the creation of the IPv4 pool,[190] 39 Class A blocks have been allocated. Table 5.1 shows how many Class A delegations have been made.

AFRINIC	APNIC	ARIN	LACNIC	RIPE NCC
0	2	35	0	2

Table 5.1: Number of contiguous /8's allocated in modern RIR regions. It should be noted that one of APNIC's /8 allocations was actually allocated by ARIN (to JPNIC) and later transferred to APNIC under the ERX project. APNIC's other contiguous /8 allocation was in 2005 to Softbank.

CIDR is a mechanism for finer-grained subnetting of IP address blocks. In the language of CPR appropriation, it is a means of more efficient parcelization. CIDR was developed in response to the depletion of IPv4 addresses. CIDR notation has the form *prefix/prefix-length*. The prefix is the dotted notation of the common (network) portion of the address block. Prefix length denotes the length in bits of the prefix. The original classful prefixes A, B, and C correspond to /8's, /16's and /24's, respectfully. With the introduction of CIDR, the size of number rights delegations better match actual expected utilization. In other words, actors that needed more than a class C (256 addresses) but less than a class B ($65,536$) could be delegated the rights to blocks in between those two. For instance, a /20 would yield 2^{12} or 4,096 addresses, which may be just the right size for a medium sized hosting company.

CIDR-based allocation did reduce the average block allocation size. See Figure 5-3. From approximately 1997 onward, address delegation in ARIN was much smoother than the earlier rapid growth periods from approximately 1987 to 1997.

[190]The selection of ARIN inception date as the end of this period is used because it roughly coincides with both the introduction of CIDR subnetting and the end of the steep growth in allocation between approximately 1987 and 1997.

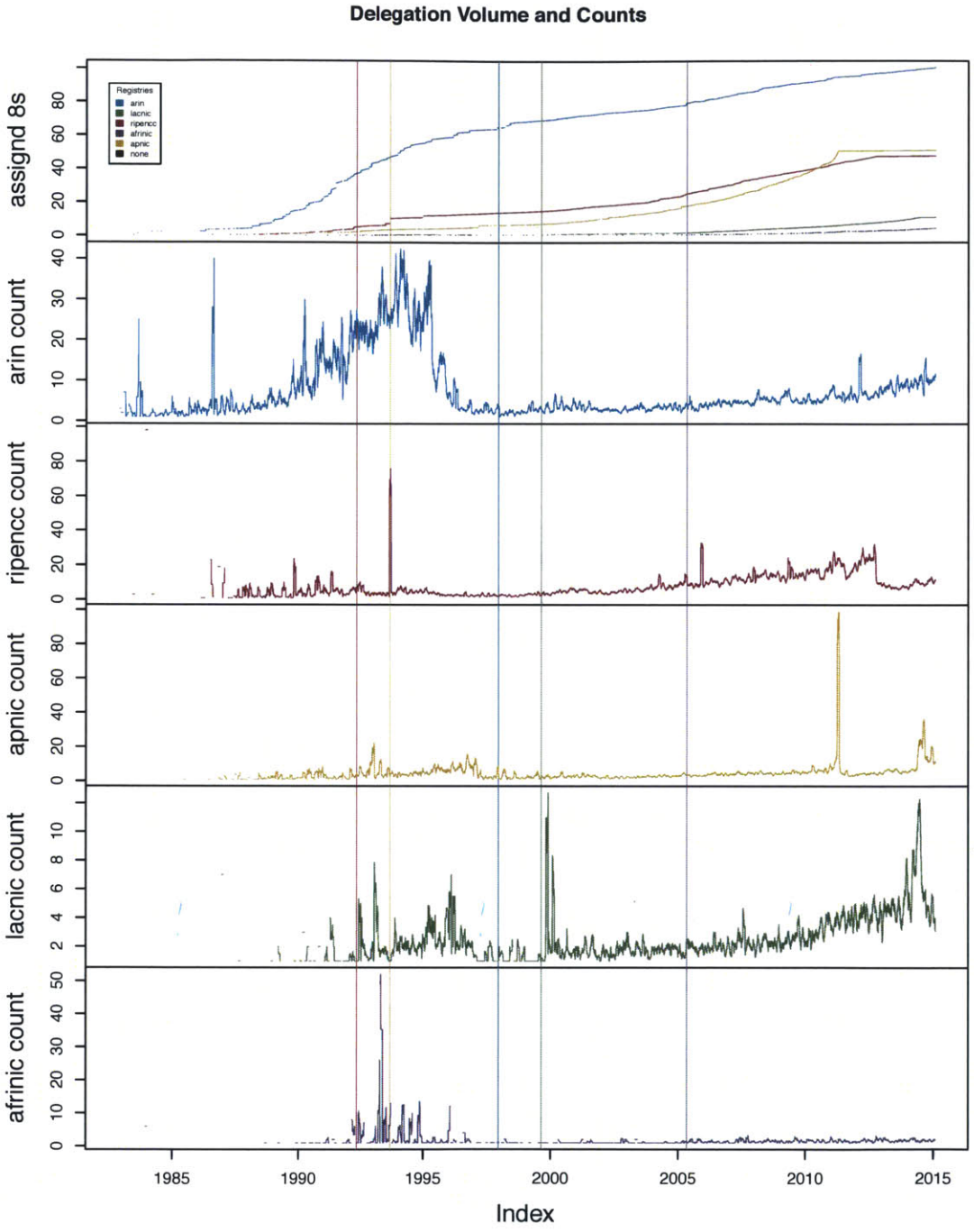

Figure 5-3: The top panel shows total volume if IPv4 delegations for each RIR. Vertical lines represent the dates when each RIR was founded. In the case of Network Solutions (ns) and ARIN, the founding date for ARIN can be considered the end date for NS as an IR. The five bottom panels show the daily average number of delegations. One delegation represents an organization having applied for number resources and having been assigned those resources. Delegation count can be viewed as a proxy for demand for IPv4 addresses in that region.

5.2 Delegation Hierarchy

The RIR system combines hierarchical rights delegation mechanisms and a federated ensemble of authoritative management facilities. This section focuses on the hierarchical delegation of rights bundle to describe the function of the RIR system, tracing the path of rights delegations from the IANA to end users. The delegation hierarchy is depicted Figure 5-4. Each edge along this path represents the delegation of a meaningful bundle of rights between those RIR participants. Addressing structure is hierarchical: blocks of addresses are sequentially fragmented for delegation at the regional (RIR), in some cases national (NIR),[191] local (LIR), and end-user level (ISP). Delegation and exercise of these bundles both facilitate downstream utilization of number rights and reinforce delegation strategies that contribute to sustaining the integrity of the global stock of number resources.

Recall the earlier distinction between the narrow notion of delegation of number utilization rights and the delegation of authority amongst organizations in a federated RIR. The latter is the process by which bundles delegated in the former are distributed. Historically, the IANA was the root of both modes of delegation. Under the modern IR system, until recent exhaustion of the "free pool" the IANA managed alienation (delegation of) rights to the RIRs. Following the recurring theme of "existing" rights in empirically observed CPRs, in the time since the creation of the (R)IR system and its modern form, the seat of operational authority that sustains the integrity of number delegation has shifted from delegation by the IANA to authoritative rights development within a federated ensemble of five RIRs. While this will be alluded to in Section 5.2.2, it will be developed further in Section 5.6.1 on constitutional rules and in Section 5.7.1 on the recent articulation of RIR goals in RFC 7020 (Housley et al., 2013).

Here, the focus is on delegating meaningful bundles of rights. In particular, this section traces what is referred to as the delegation path of these bundles through the RIR system to describe how bundles *a*) accrue additional rights and from whom; *b*) who revokes rights and the general circumstances; and *c*) the inverse relationship between size of block and breadth of rights. Delegations start with the IANA in its role delegating bundles of alienation rights for large IPv4 blocks, subsequent delegation of rights in the IR system, and ultimately the delegation of utilization rights (assignment bundles) to end users.

Figure 5-4 depicts the number rights delegation hierarchy. Blue lines in Figure 5-4 represent the delegation of number rights, starting with allocation bundles conferred from the IANA to the RIRs to assignment bundles conferred from IRs to end users. Red lines represent what the community refers to variously as recovery, returns, or most aptly for this framing, revocations, of number delegations. As information commodities, number utilization and delegation rights cannot be revoked by collecting the actual resource in the same sense a rival good can collected.

[191] The national level only exists in LACNIC and APNIC. In LACNIC, the only two NIRs in operation are Mexico and Brazil, in part because these pre-dated the creation of LACNIC. LACNIC does not have a process to create new NIRs. APNIC has a number of NIRs, the latest being India.

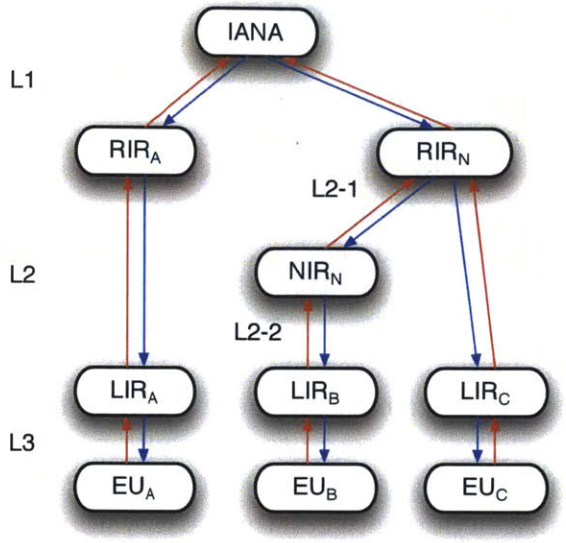

Figure 5-4: Delegation of number resource rights passes through a number of organizations in the RIR system. In this figure, blue lines represent delegation of number rights. Red lines represent what is referred to in the community as "recovery". The IANA and RIRs have been discussed. NIR stands for National Internet Registry. NIRs only exist in LACNIC and APNIC; NIRs are historic in LACNIC and both historic and actively developed in APNIC. LIRs are Local Internet Registries, typically ISPs or NSPs (network services providers). Finally, EUs are end users, actors to which individual addresses are "assigned" such as home users' equipment (such as a home router) connecting directly to ISP infrastructure or servers hosted at a colocation facility or on a business's premises.

Rather, given utilization is ultimately in the provision and appropriation of routes across the Internet, revocation may be asserted by an IR, but to be effective the acquiescence of a critical mass of control plane participants is necessary. This participation in the revocation process is a key element of the RPKI debate elaborated in Section 5.7.4.[192] More precisely for this work, red lines represent a rescinding of number utilization rights.

Edges at each level (designated by Ln) in Figure 5-4 represent a class of delegation or revocation of rights between actors in the RIR system. These bundles are elaborated formally in Section 5.6.3. Here, the focus is on how the number pool is partitioned[193] as rights are delegated along the hierarchy. Delegations at $L1$ transfer rights for a particular block of numbers in the unallocated pool from

[192]In addition to highlighting the general dynamics of the system, this discussion also foreshadows the discussion of RPKI. On the surface, RPKI is a means of security origin rights, potentially reducing hijacking externalities. The resource rights framing alluded to in this Section and developed in the context of the debate in the operational epistemic community (Section 5.7.4) highlights that RPKI does not change the bundles of rights, but rather changes the character of enforcement of existing rights. In contrast to requiring the acquiescence of the regulated when any rights are revoked, under the strong aspirations of RPKI-based automation offered by a subset of RPKI supporters, RPKI substantively enhances immediate enforcement power.

[193]Partitioning can lead to fragmentation, which has the negative connotation that a block has been too finely partitioned, creating deaggregation externalities. Although this is typically associated with the behavior of an LIR, it may also because by resource policy. The instance of fragmentation by KRNIC, an NIR, is discussed in the next section.

the IANA to an RIR. In the RIR vernacular, this is an allocation from the IANA to the RIR. Metaphorical language such as *allocation*, *transfers*, and the common stock of numbers as a finite *pool* of resources all reinforce notions of rival goods. Each transfer in Figure 5-4 shifts bundles of rights from one organization to another.

5.2.1 Delegations Down the Hierarchy

Historically, $L1$ delegations confer allocation bundles comprising alienation and management rights for blocks of numbers in units of /8's.[194] Pre-exhaustion, delegations to RIRs were based on the RIR's delegation rate. As per Section 2.1.2, utilization is a refinement of the general notion of withdrawal. A utilization rate for the RIR was based on the past six months of delegations. Once an RIR's stock fell below the volume necessary to supply the next six months based on the utilization rate, it could request a subsequent delegation of number rights sufficient to satisfy demand for the next year. As such, this criteria is a measure of RIR utilization (withdrawal).

Like the delegation count in the lower panels of Figure 5-3, utilization rate is a proxy (indicator) for demand. Further linking $L1$ delegations to Figure 5-3, delegations along blue $L1$ paths ensure demand indicated by counts can be fulfilled. Moving down the hierarchy, growth in delegation volumes in the first panel of Figure 5-3 are driven by delegations along $L2$ paths. The left path depicts delegation in ARIN, RIPE, and AFRINIC. The right path depicts delegations in APNIC and LACNIC. Structurally, the delegation paths are similar, but the bundle of rights conferred to an RIR differ from those bundles of limited alienation, management, and utilization rights delegated to LIRs in its region.

$L1$ and $L2$ delegations are allocations in the community vernacular. They differ in what rights are conferred in bundles. $L1$ delegations are exclusively delegations of alienation rights. The IANA has the right to delegate rights to the RIR to *further* delegate alienation rights for those blocks to the RIRs. Management rights that ultimately confer utilization rights are created through the consensus process.

In turn, the RIRs themselves do not have utilization rights. RIRs do have the rights to delegate bundles of management rights and/or utilization rights to LIRs. In community vernacular, the difference among *a*) having rights to *just* further delegate, *b*) having rights to delegate or assign (utilize), and *c*) having rights to *just* assign (utilize) distinction between two modes of allocation rights and the latter, assignment rights, potentially conferred onto an LIR. An LIR conferred an allocation bundle at $L2$ may further delegate utilization rights to its customer (assignment rights) that then routes those numbers.[195] An LIR will also typically use some subset

[194]Delegations from the IANA to the RIRs historically governed by the Internet Assigned Numbers Authority (IANA) Policy For Allocation of IPv4 Blocks to Regional Internet Registries (ICANN, 2012b). At the time of writing, the last five /8's had been delegated on 31 January 2011 (NRO, 2011), one to each of the five RIRs, activating their respective exhaustion plans.

[195]If single homed, the LIR is the upstream and more likely originates those numbers for the customer. If the customer is multihomed, those numbers may be originated by the LIR that delegated the bundle and another LIR with upstream connectivity.

of those numbers for its own infrastructure.

$L3$ delegations are exclusively what the community refers to as assignments. $L3$ delegations may not be subsequently delegated. An end user may use a number or set of numbers for its value proposition, or, in terms of Section 3.1, downstream uses. That said, those downstream uses do not include transferring (alienation of) usage rights to another party. In general, the LIR has the right to revoke delegations to end users, "reclaiming" those rights for other potential delegations.

On the east side of Figure 5-4 $L2$ is broken out into $L2.1$ and $L2.2$ to distinguish delegation to an NIR and from an NIR to LIRs in its national jurisdiction. Delegation at $L2.1$ is the case where an RIR delegates the management and alienation rights for a relatively large chunk of IP addresses to an NIR. In what will be referred to as the "unrestricted" form, once delegated, the NIR may further delegate allocation and assignment rights on to its members. In this case, some management authority has been delegated to the NIR. National NIC (Network Information Centers, or national IRs) predate both APNIC and LACNIC in their respective regions. In the LAC region, Brazil and Mexico had number registries before the formation of LACNIC. In the AP region, Japan, Australia, South Korea all had NIC delegating number resources. In the unrestricted case, each had full discretion over how number rights delegated for subsequent national delegation were handled.

To illustrate, consider three consecutive delegation windows to KRNIC. Further consider a rapidly growing access network KPOP.[196] KPOP requests rights for a block to cover growth over the next X months. Similarly, other networks are also growing and deplete KRNIC's current pool before KPOP can request another delegation. KRNIC receives another delegation window and KPOP requests another X months worth of number rights. KPOP now has two delegations contiguous with one another. When this happens for a large number of actors, this creates unavoidable fragmentation.

To resolve the fragmentation problem, APNIC changed the NIR delegation policies such that NIRs do not maintain their own pool of number rights. Under the current model, members of the NIR are vetted by the NIR, pay fees to the NIR, receive registry services from the NIR, and coordinate with the NIR to jointly maintain the content of the registry. Requests for delegations through the NIR to its members (network actors within its jurisdiction) are still processed by the NIR. Actual delegation is from blocks in APNIC's regional stock, not a stock held by the NIR as under the unrestricted model. The actual selection of which numbers will be delegated, in particular from where within the regional pool, is managed by APNIC.

Technically, APNIC has reclaimed some of the duties formerly performed by the NIR and limited its discretion. In particular, it has limited this discretion with regards to the relative positions of blocks within the larger pool. Particular management rights, those related to how the blocks are partitioned in the process of delegation to LIRs, has been reclaimed by APNIC. This is argued by APNIC not as a reflection on the competence of the NIR per se, but rather, as a reflection on opportunities to reserve blocks in a national versus regional stock. APNIC's justification

[196]KPOP stands for Korean Point of Presence, what were you thinking it meant?

is that the particular management rights facilitating effective block reservation are more efficiently exercised in the larger regional pool. In other words, there are more options to exercise block reservation management rights within the regional pool than in the smaller national pools where allocation windows may not be aligned with demand. In contrast, members' rights to delegations have remained the same—the criteria for evaluating delegations remained the same.[197]

This is an interesting confluence of the delegation of number rights and delegation of IR authority. Arguably, NIRs, being local organizations are better suited to evaluate an organization from its home state.[198] In this sense, NIRs reduce the transaction costs of vetting new members as legitimate organizations within a given state. Moreover, the NIR is typically an organ (or function of) are larger organization, such as a NIC, that has existing relations with members, potentially further reducing transaction costs.

That said, fragmentation is better managed in the RIR, here APNIC. In this case, APNIC contracted the role of the NIR to evaluation, revoking certain management rights. Based on interviews with APNIC staff, this decision was based on the effects of fragmentation and efficiency. Revocation of these management rights did result in a consolidation of management rights in the RIR. That said, it was not, as may be assumed under a realist framing, driven by a consolidation of power (authority). The change from what was called the "federated" model to the current model was made through the consensus process.

5.2.2 Pools and Authority

Before starting the climb back up the hierarchy along the red rights revocation links, a discussion of the various number "pools," management of rights bundles, and the broader notion of authority is introduced to begin reconciling the character of authority engendered in this common management regime. As illustrated by the APNIC NIR story, it is possible to unbundle elements of delegation, separating selection from the pool from evaluation of criteria. Structurally, lines of authority seem to flow downward from the IANA. In a strict mercantilistic notion of delegation of authority, such a delegation could be rescinded. This is not the case in the modern IR system. The configuration of rights bundles provide the foundation contravening strict notions of delegated authority. Alienation rights are delegated by the IANA, but the management rights that shape utilization practices are conferred by community consensus.

Consider legacy address space allocated during the period before the IANA function was managed by Network Solutions, i.e. the period before (to the west of) the

[197]The criteria remained the same, but there is not a guarantee that any two evaluations will result in the same outcome. In particular, an NIR may implement those criteria differently, especially where there is subjective reasoning about needs criteria such as network layout and demand. This is less a difference in structure, more a difference in interpretation of rules endemic in any principal-agent relation.

[198]This is an almost textbook instance of operational capability that warrants delegation from a principal to an agent.

black vertical line in Figure 5-3. Much of this space remains what is called legacy space. It is space delegated by SRI under the leadership of Jon Postel. The modern IR system respects legacy space and legacy space holders. Table 5.1 provides the list of all contiguous /8's delegated over the history of the RIR system, with legacy resource delegations highlighted.[199]

In terms of delegation of rights, the old adage that possession, here historic delegation and utilization, is 90% of the law, is appropriate. Legacy resource holders may or may not have a contractual relationship with a registry. Legacy resource holders may not have access to certain services, such as RPKI. Moreover, unlike RIR members, unaffiliated legacy resource holders are not contractually beholden to any actor to demonstrate efficient utilization of number resources they manage.[200] In the context of pools of rights delegated in this section's narrative, one may consider legacy rights as free floating pools of number resources.

Legacy rights holders may claim a form of ownership. Legacy rights holders are not beholden to any contract or entity that limits alienation of number rights currently delegated. Following the community mythos, many of these delegations were supposedly simply delegated by Postel, the registry function implemented in the form of Postel's spiral notebook. Amongst non-RIR members, unaffiliated legacy rights holders and those that do not hold any number rights, legacy rights holders may transfer some or all of their number rights without limitation. The caveat is that these transfers will not be authoritatively recorded. Such a transfer may limit later exercise of rights by the new "holder" of those rights. Thus, even though such alienation is not expressly prohibited, it does not have the guarantees afforded by participation in the larger system.

Consider this in contrast to the discussion of revocation, climbing up the red edges in the delegation tree, discussed at length in the next section. In contrast to the absence of utilization monitoring and enforcement for legacy holders, in all of the RIRs, rights may be revoked if it can be shown that delegations *are not* being utilized according to the criteria under which they were delegated.[201] Thus, even though origin rights have been delegated to a particular organization, those rights are contingent on whether those numbers are put to use in the ways intended in the original allocation. The simplest instance of this is the difference between allocations and assignments—assignments cannot be further delegated. Above, *cannot* is enforced by multiple contracts. Contracts enforcing this bundle exist between *a*) LIR and customer and *b*) RIR and LIR. They are monitored by audit performed

[199]These are the /8 delegations visible from the NRO's statistics file.

[200]Even if affiliated, legacy holders are not necessarily beholden to utilization requirements. For instance, the LRSA (Legacy Registry Services Agreement) guarantees access to registry services but does not impose utilization criteria:

> Note that ARIN will not reclaim unutilized address space from legacy holders who sign this RSA, nor will ARIN attempt to take away legacy resources from organizations who choose not to sign it. However, signing the Legacy RSA contractually locks in a set of rights, and thus reduces the risk of future change to a minimum. (ARIN, 2012)

[201]This is discussed in more detail in Section 5.6.3 in the context of needs-based criteria.

by LIRs and RIRs.[202] Utilization patterns for every block delegated are not continuously monitored by the RIR.[203] Rather, the RIR trusts the LIRs to update their utilization information in an appropriate and timely manner.

Legacy holders are not beholden to any of these utilization criteria. In this sense, authoritative relations in the modern RIR system have been established through number rights delegations. As illustrated by legacy delegations, that authority is not inextricably tied to possession of number rights, though. Taking this analysis a step further, it is technically possible for an alternative registry to emerge, displacing one or more of the RIRs. For instance, it is logically possible that, if it were feasible for some set of existing legacy holders to do this, they could create a registry whose top-level common pool comprises legacy space. The more complicated process would be to usurp the position of an existing registry. Network effects, namely utilization by and reliance on existing registry services by existing members, militate against this kind of attrition in any case other than a critical mass changing allegiance.

As will be discussed further in Section 5.4.5, while the delegation structure is hierarchical, authority is more effectively framed as a federated ensemble. Within the system, dependence on the registry as a common resource binds actors to credible commitment to sustaining it. As will be developed in the next section, enforcement of revocation is currently diffuse, requiring the acquiescence of those routing prefixes. The arguments against RPKI, discussed at length in Section 5.7.4 can be framed in terms of how delegation authority is distributed. In effect, arguments against RPKI frame it as a tool that could be appropriated to impose more concentrated and immediate modes of enforcement that eliminate the discretion currently available in the diffuse model.

Further, again within the system, operational epistemic communities are quite amenable to unbundling authority along functional lines that improve the efficiency or efficacy of resource operations; this was the case with KRNIC. That said, dependence on the registry as a resource also limits attrition that could give rise to an institutional competitor. In other terms, dependence on the registry creates a form of lock-in. Despite this lockin, enforcement activities, such as security mechanisms discussed in Section 5.7.4 and revocation and recovery processes discussed in the next section necessary for the RIR firm to perform its function remain a challenge.

5.2.3 Revocations and Recovery

In contrast to an "ownership" model of number resources, delegations under the modern IR system are not permanent. For instance, consider a network service provider (NSP) that delegates assignment rights to one of it's customers, a metro-

[202]This does not imply that audits are performed on a continuous basis, or even on a regular schedule. Rather, audits are, as elaborated in the discussion of needs-based criteria and utilization, often performed as part of evaluating a request for subsequent (additional) number rights.

[203]To be more specific, monitoring is possible by any number of actors that see a "full" routing table and have bulk access to registry data, either through the registry maintained service or via bulk access. That said, the registry itself does not continuously monitor utilization practices. RIRs do check utilization of previous delegations upon request for subsequent delegations.

scale access provider M. As long as M is a customer of the NSP, it holds those assignment rights, allowing it to provide its customers (home users) with connectivity and "live" IP addresses. Each RIRs' policy corpora reference recovery of delegations in the event that M is no longer a customer of the NSP. For instance, in APNIC, resource policy delegations must be recovered if the "downstream customer no longer receives connectivity from the LIR", (APNIC, 2014d).

As a locally scoped contract, typically bound to a particular national jurisdiction, revocation of number rights is enforceable under state-based contract law.[204] One function of revocation is that it is a means of avoiding losing track of number resources. For instance, if M continues to use resources from an NSP after severing its relationships, the contractual links that enforce updating the appropriate registry are lost. The result is entries no longer accurately reflect resource utilization and/or points of contact for use by the rest of the community.

Revocation at $L2$ becomes more difficult given delegations span multiple jurisdictions. In the modern RIRs, revocation is part of the contract, but may be difficult to enforce. For instance, the LACNIC policy manual provides a clear description of revocation criteria:

- Lack of visibility of the resource on the global routing table.
- Breach of LACNIC policies.
- Breach of the provisions of the registration service agreement or other legal agreements between the organization holding the resource and LACNIC or one of its NIRs.
- Evidence that an organization has disappeared and its blocks have not been transferred.
- Unauthorized transfers under the provisions of the policies.

The first criteria, lack of visibility, is evidence of under utilization. Under utilization may occur if demand expectations justifying one or more allocations were overly optimistic. Such under utilization is one justification for revocation. Resource abandonment by organizations that no longer exist may be another reason for revocation and subsequent recovery. Closing of an organization may be another reason; it is a degenerate form of under utilization.[205] Violation of contract with the RIR is yet another justification for revocation.

In some cases, revocation, here a voluntary return, is part of an exchange of number rights. For instance, when an organization requests its first block of portable addresses, it is often required to "renumber." The typical case is that the requesting organization has a block delegated from one of its upstream providers, like M in the example above. To receive a portable block delegated by the RIR, the requester must "renumber out of" its old block (delegated from an upstream provider) and

[204]There are also alternate bundles of rights, aptly referred to as portable, that solve this problem; see discussion in Section 5.6.3.

[205]This does not account for mergers and transfers, for which each RIR has specific operational rules dictating alienation from the old organizations to the new. Logically this is a form of alienation, but in practice, the actual physical equipment identified by and originating number resources may not change.

into the newer (typically larger) portable block assigned by the RIR. "Renumbering out" is community vernacular for returning an assignment bundle to the IR, typically an LIR, that delegated that bundle, returning it to that LIR's stock. "Numbering into" is the immediate utilization of a subset of the rights in the portable assignment bundle delegated by the RIR. Short-term, this is a transaction costs; long term it contributes to preserving the integrity of the stock of number resources.

Renumbering on initial allocation achieves a number of goals: *a*) the requester is delegated a contiguous block of numbers, *b*) the LIR recovers a block from a customer that can then be delegated to another customer, *c*) fragmentation is avoided, and *d*) the general notion of conservation is preserved. For those with operational experience, the latter two points seem quite intuitive. For instance, allowing the requester to continue using LIR delegated numbers and portable numbers may lower transaction costs in the short term, but could result in allocating smaller portable blocks to bump the existing LIR delegation up to the expected demand. Bookkeeping by the RIR is complicated by having a mix of LIR and direct delegations. Further, if the LIR cannot "recycle" its number delegations, effectively reusing blocks returned by actors that get portable delegations, it will need to request subsequent delegations of its own sooner.

Not returning numbers to the LIR also increases the chances number delegations may continue to be used by the requester after termination of the contract with the upstream. Although not on a national scale such as the Korea narrative above, delegation frequency can have distinct effects on fragmentation. In the discussion of constitutional norms in Section 5.6.1, this will be framed as a trade-off between conservation and routability. Relative to conservation norms, fragmentation can be positive or negative. On the positive, it accommodates delegations that are a better fit given available estimates of demand. On the negative side, fragmentation results in potentially more routing table entries. If the space is not large enough for effective reservation strategies, fragmentation may also yield an unallocated stock of "odd-sized lots" that require delegating multiple disjoint blocks or renumbering to satisfy subsequent demand.[206] Delegating multiple disjoint blocks also increases the size of the routing table and reduces the chances those can be re-aggregated if returned.

In contrast to the revocation conditions above, network abuse *is not* a reason for revocation *within* the RIR system. Rather, revocation of rights because of supposed abusive behavior(s) is a highly contentious issue between LIRS, RIRs, and the anti-abuse community. As will be discussed in Chapter 7, operational externalities created by IPBLs are a means of revoking selective abusive uses or, in strong application, local revocation of all number uses.[207] For the time being, note that

[206]The problem of the "odd-sized lots" is similar to the problem faced by KRNIC. In the face of high demand and relatively small allocation window, reservation policies may not be feasible. In the best case, no reservation policy may be in place and contiguous blocks are allocated until the window is exhausted.

[207]To be more specific, local revocation means denying delivery, typically at the border of the destination network, of traffic that originating in a network known to harbor abusive actors. That said, "local" becomes akin to global when network effects mean a large number of access networks

revocation by the LIR supporting the abusive actor is, as per the discussion of other externalities in Section 2.1.2, an operational externality network actors generally eschew. For those receiving abusive traffic, this is a security externality.

For the LIR, it is an operational externality because such a policing policy would require additional investment in monitoring outgoing traffic, enforcing anti-abuse policies, potentially rejecting customers based on abusive behavior, and dealing with potential discrimination issues. Depending on the type of monitoring, policing may be considered a violation of network neutrality. One mode of monitoring, BCP 38, intends to limit abuse from spoofed addresses. This is a clear instance of abuse—spooking is akin to hijacking in terms of rights violations but claims a single number rather than a block. BCP 38 is embraced and espoused by some actors in the community, but far from all. Enforcing anti-abuse policies would require an established, stable set of anti-abuse policies agreed-upon by a broad portion of the community. Outside of clear protocol violations such as BCP 38, there is little agreement on what constitutes abuse outside select anti-abuse communities.[208]

In terms of rejecting customers, the general idea that connectivity providers are "just dumb pipes" is often used as a justification. Taking the stance that a connectivity provider P should be responsible for their downstreams' behavior means that rejecting actors are considered by some to be abusive places P in a policing role that also denies P customers. If P complies, it is perceived as giving business to LIRs with less stringent standards. Absent a common standard, rejecting customers based on abusive behavior has the potential to create a race to the bottom. Not all actors engage in the race to the bottom, but a sufficient number do. These are the actors whose delegations wind up listed on an IPBL. Revocation by LIRs is thus a mix of revocation in the RIR system and the anti-abuse system.

In terms of authoritative rights delegation, at $L2$ and $L3$, revocation in the RIR system is the prerogative of the delegator. It is unclear if the IANA has any rights to revoke a delegation to an RIR.[209] No documentation of such an action on the part of the IANA itself has been seen. In contrast to revocation by the delegating agent, returning number rights at $L2$ and $L3$ represent the voluntary self-revocation of rights by the recipient.

In the RIR system, the LIR and the RIR do have the rights to rescind delegations, to revoke number rights they have delegated. That said, actual exercise of the revocation requires, as noted before, the compliance of actors routing the corresponding prefixes. This serves as a check on revocation. Each revocation tacitly requires the acquiescence of those actors that formerly provisioned and appropriated routes for those prefixes. Strong forms of RPKI would make revocations from delegating "principals" automatic. In effect, RPKI not only facilitates the enforcement of origination rights, but strong forms make enforcement immediate, vesting authority to revoke

use the same information to limit delivery.

[208]Even within the anti-abuse communities there are constituencies that tussle over what constitutes abusive behavior, what is not, and where the balance lies. These constituencies are described in Section 7.2.

[209]The closest instance of revocation is the controversy over the return of address blocks by ARIN to the IANA's post-exhaustion recovery pool.

in the hierarchy rather than the diffuse set of actors provisioning routes. As discussed in further length in Section 5.7.4, this also shifts decision making authority from a diffuse, peer-based structuring of trust in routing operations to the hierarchical delegation structure of numbers delegation and management. Within this work, this is a shift of authority from a less formal set of institutionalized norms related to routing to a more formalized set of institutions managing numbers.

At $L1$, return and recovery are somewhat confounded. Historically, certain address blocks were returned to an RIR were returned to the IANA for subsequent reallocation from the global pool. For instance, before ARIN Policy 2011-6 (ARIN, 2011), it was standard practice[210] in ARIN to return /8's that were recovered from legacy delegations to the IANA. Under the Global Policy for Post Exhaustion IPv4 Allocation Mechanisms by the IANA (ICANN, 2012a), a recovery pool was established but the language was ultimately relaxed allow voluntary return to the IANA pool at the discretion of the RIR.[211] The result was ultimately a discretionary mechanism for RIRs to return space recovered in their region to a common pool that would be activated after the first RIR fell below a /9 of its last /8.[212]

5.2.4 Modes of Authority

The number allocation hierarchy illustrates two key dynamic in number allocation bundles. First, the delegation hierarchy provides evidence of how rights accumulate

[210]ARIN indicated this was a standard practice in a message discussing implications of then draft proposal 2011-6:

> The wording of the proposal seems to indicate that any legacy space, including a /8, that gets returned to ARIN gets added into ARINs inventory and made available for redistribution within 30 days. In all other instances where a legacy /8 has been returned to ARIN, ARIN has returned that space to IANA. This proposal would change that standard practice. (ARIN Member Services, 2011)

Policy proposal 2011-6 modified this practice to make any space recovered immediately available for delegation in the ARIN region. It should be noted this came in conflict with a Global Policy that, under the interpretation of this work, allowed for the designation, at the discretion of the RIR, to return space recovered in its region to the IANA in contiguous blocks of /24 or greater.

[211]An excellent discussion of these concerns is provided in version 3 of APNIC's prop-097 (Acosta et al., 2011). This highlights ARIN's concern with the requirement that any recovered space be returned to the global pool. As it turns out, the ARIN version (ARIN, 2009c) of the global policy modifies the contested text to introduce language to refer to recovered space that is designated for return to IANA. (ARIN, 2009c) is more specific than the global text at (ICANN, 2009). The explicit text does not imply the stronger form that requires all recovered space be returned. Rather, it is at the RIR's discretion to return space. Similar language is present in RIPE's IPv4 allocation policy documentation (RIPE, 2014), distinguishes between space recovered that is immediately available for subsequent allocation and space that is to be returned to the IANA:

> This section only applies to address space that is returned to the RIPE NCC and that will not be returned to the IANA but re-issued by the RIPE NCC itself. (RIPE, 2014, Section 5.3)

[212]This happened on 20 May 2014 when LACNIC's pool fell below a /9. A /9 is half /8 (ICANN, 2014).

to end users—this accumulation does not follow the typical accumulation depicted in Table 3.2. Rather, the RIR firm acts as an administrative manager. The firm exercises alienation rights to confer alienation and withdrawal rights as defined by the management rights developed by the community. Second, the resource scope, i.e. the volume of the blocks governed by these bundles, decreases as rights accumulate. As an instance of parcelization, this is unsurprising. In relation to notions of conventional ownership, where lesser rights bundles are conferred by an actors holding ownership rights, these bundles provide a stark contrast. Only when the membership is considered a singular collective does their cumulative set of rights resemble ownership.

5.3 Common Resource Structure

Where the hierarchy in Figure 5-4 depicts the delegation process, the registry is the resource management facility maintained to document delegations. Recall from Section 3.4.3 that management (rights) facilities enhance resource rights. In the RIR system, the core resource management facility is the registry itself, implemented as a database. The registry not only documents delegations, but also facilitates enforcement of operational rules. The firm uses registry information in needs justifications. Operators use origin rights documentation to evaluation the legitimacy routes appropriated from others.

The content of the registry is jointly maintained by the firm and its membership. That content comprises number delegations (blocks delegated to entities) and utilization (assignments reported by users). Although the database software and deployment is maintained, and in some cases developed, by the firm, the content is jointly maintained by the firm and the membership. In general, the firm manages documentation of $L2$ delegations and LIRs are responsible for managing the documentation of assignments in $L3$. The following describes the registry structure.

5.3.1 Number Registry

Conceptually, the registry function is quite simple: document the legitimate delegation of number resources. In practice, the registry must facilitate the *joint* access and maintenance (a form of use) of this repository of delegation information. The registry is implemented as distributed database whose core function is to *a*) document the delegation of number rights to organizations by the RIR in a given region, *b*) document delegation of number rights from LIRs to their customers, *c*) document utilization (assignments) of rights allocated, and *d*) facilitate community access to this information. Each function comprises a bundle of use and management rights. The former two correspond to $L2$ and $L3$ in Figure 5-4. These correspond to registry use rights. The latter provides access (but not use) rights to the community writ large. To lay the foundation for a more precise articulation of these rights in Section 5.6.3, this section focuses on the registry function.

In terms of rights holders in Chapter 3, the simplest rights bundle is the registry entrant. Up to the threshold of partial rivalry, the registry is open access. Any actor with an Internet connection may utilize the registry as an entrant. Basic functionality is built on (nonsubtractive) access. Registry data is used by network operators to find contact information to, among other network operations objectives, resolve externalities. Access and distribution is nominally available for any non-commercial use.

Technically, access has been historically based on the WHOIS protocol[213] typically invoked by a command-line WHOIS client.[214] More recently, a variety of access protocols have been provided by the RIRs to facilitate programmatic access to registry information. Both the RIRs and third party websites have developed APIs that abstract away the arguably idiosyncratic WHOIS query protocol to facilitate more user-friendly registry access by less technical users.

Access creates a load on the resource (registry database). RIRs rate-limit entrants to ensure access remains nonsubtractive and, subsequently, equitable. For heavy users that may exceed the rate-limit, "bulk" database download is available from all five RIRs. All but the RIPE NCC require signing terms of agreement. Heavy users that download the database may then use their own resources to access and analyze registry data.

RIR members and the RIR firm comprise registry *users*. Recall background CPR concepts of withdrawal and appropriation correspond to modification of the contents of the registry in the course of joint maintenance of delegation records. Modifying content *is not* a management right. Recall management rights confer the rights to alter the structure of a resource.[215] Registry management rights means modification of the database structure (schema). Within its operational remit, the firm holds some management rights. The community collectively exercises its management rights by creating policy to introduce new elements that will alter number use or adjacent routing operators. For instance, incorporation of abuse contact information into the registry, discussed briefly in this section is an instance of the community exercising management rights through the collective-choice (consensus) process.

The conflict over the abuse contact is over introducing a potential obligation to acknowledge security externalities created by a network's customers. This was considered by some as the first step to policing of more specific resource uses that currently in play. The ultimate consequence could be subsequent incremental changes that would mandate endogenizing security externalities under penalty of resource revocation. As noted in the discussion of the delegation hierarchy, abuse is not a reason for revocation but IPBLs create similar effects. As may be obvious, this is part and parcel of the larger tension between the RIR and anti-abuse communities.

Although both members and the firm are users with access and modification

[213]The WHOIS protocol is described in RFC 3912 (Daigle, 2004).

[214]There are various means to query the WHOIS service: web-based clients, command line clients, and RIR provided APIs. There are also stand-alone applications available on smartphones and tablets. Webclients are provided by third parties as a service and by the RIRs themselves.

[215]Revisit Section 3.4.3.

rights, these rights do not overlap. Firm rights, as a projection of their $L2$ rights, facilitate modifying content that documents $L2$ delegations.[216] Member rights facilitate modifying content that documents $L3$ delegations and may also include rights to alienate those rights, conferring modification rights to customers.[217] Upon delegation from the RIR to an LIR, the firm updates the registry to reflect that delegation. Upon initial delegation, modification means adding new entries for the organization such as address information, points of contact, unique identifiers (referred to as handles) and other descriptive information. Firm modification also links the identifying entries to descriptions of number resources delegated to that organization. Finally, modification also includes providing organizations with credentials that facilitate access, modification, and limited alienation rights.

RIR members are required to update the registry as delegations characterized as allocations are either subsequently allocated to customers or are assigned to equipment. RIR members' rights bundle subsumes the notion of a user, but also includes limited forms of exclusion and alienation rights. RIR firm rights allow modification of any data in the registry, but is scoped to relations in $L2$. In contrast, members' rights are scoped to changes related to their delegations and those subsequent delegations they are responsible for. Violation of that scope have a variety of implications. For instance, compromising a member's account may allow an attacker to change authoritative sub-delegation information to obscure hijacking and other abuse behaviors that require number delegations. RPKI, discussed in Section 5.7.4 introduces additional safeguards, but the trade-offs are still broadly contested.

5.3.2 Supporting Knowledge Resources and Arenas

Supporting facilities comprise those that are not critical to the NMR under its remit, but supplement management of the NMR and enhance participation in the NRS in general. RIRs differ in the supporting services provisioned for access by the community.

5.3.2.1 Training as a Resource

RIRs encourage use of the registry by offering training. RIRs differ in the scope of training. In the ARIN region, training is limited to requesting number resources and training members how to use RIR services to maintain registry content. The RIPE region largely limits itself to registry services training.[218] In contrast, the remaining three regions, LACNIC, AFRINIC, and APNIC, have substantive developing economies. In these regions, training includes not only registry services but also network operations training.

[216] And in the case of NIRs, $L2 - 1$ delegations.

[217] Although registry modification rights may be conferred by the LIR to its customers, it remains the responsibility of the LIR to ensure accurate and timely update of their registry entries.

[218] The RIPE NCC board has discussed the scope of training. It was limited to registry services to avoid competition with network training in the region.

Training serves three general functions relative to the integrity of number resources: quality of the data, RIR recruitment, and constituency engagement. First, training ensures that members know how to make effective use of services provided. RIRs provide training on how to update records and how to use RPKI. The training increases the probability of compliance with registry maintenance required in registry contracts. It also has the potential to reduce help desk tickets. The result is a training regime that encourages effective utilization of the registry.

In addition to training existing members, training serves a recruitment function as well. Non-members may attend to better understand the responsibility and transaction costs involved in holding and exercising number rights. Members may be introduced to new services. In particular, RPKI is another resource management facility in the RIR system that is being advocated by a particularistic constituency. Training decreases the costs of making use of these new services. Part of the controversy over RPKI is the perception that it perpetuates the role and services provisioned by the RIR beyond IPv4 exhaustion. Under this critique, training should be considered in the context of balancing the assurance of effective utilization of a resource paid for by the membership against promulgating a service whose implications are in broad contention and which may perpetuate the role of the RIR.

Training is in and of itself a form of engagement or may be embedded in a larger effort at engagement. For instance, in the RIPE region, both trainers and host masters (staff implementing operational rules on behalf of the community) are often present at training sessions. These training session provide the opportunity for attendees to ask general questions about the RIPE NCC. Further, questionnaires are also used to collect feedback. Together these provide feedback to the RIR that may be used to improve services, training, and rules.

5.3.2.2 Labs as a Resource

A number of the RIRs mission statements include some form of contributing to the good of the Internet.[219] As a supporting knowledge resource, this contributes to the remit of the RIR in supporting the community's understanding number delegation trends, attendant routing implications in the RIR system, and the impact of various operational activities writ large. As an operational epistemic community, knowledge resources are a formalization of collective domain knowledge developed through in engagement with the system.[220] For new operators, this is a valuable resource that represents experience In the context of this work, this non-core resource is an instance of an operational epistemic community delegating authority to collect information about the state of the common resource to a third party. Labs and measurement reports serve as an information resource that is, within the limits of server congestion, non-subtractable and is accessible by anyone.

[219]For instance, RIPE NCC and APNIC in particular, the two largest contributors to the RIR analysis and measurements effort.

[220]The link between operational epistemic community and domain knowledge mitigating uncertainty in common resource management is developed in Section 3.2. The challenges to domain knowledge derived from operational epistemic communities is discussed in Section 9.1.

Compare these services to how knowledge problems are solved in other CPRs. Recall Table 3.1. Ad hoc solutions to knowledge problems evolve in the course of "close-knit" communities developing tacit knowledge[221] of resource dynamics. Following the distinctions in Table 3.1, codification of domain knowledge by these communities is not necessarily offered up as a "rigorous" scientific endeavor. Rather, it is presented as operational knowledge that is pragmatically compelling to the community. For instance, in both the RIPE and APNIC member surveys, labs work is ranked second only to core registry services.

This is the case in many of the historical cases of rural commons management discussed by both E. Ostrom (1990) and in the empirical essays in (Cole & Ostrom, 2012a). For instance, consider the notion of limiting grazing in 10^{th} century Iceland:

> Let them find the number [of grazing sheep] that does not give fatter sheep if reduced but also fills the pasture.[222]

Eggertsson (2012) argues this is an, albeit implicit, application of the price mechanism and marginal analysis. This case falls into northwest corner of Table 3.1. Note the language is not in conventional economic terms, but invokes a common image of the health of the sheep (how fat the sheep are) and resource utilization ("fills the pasture"). Taken together it creates a common image comprehensible by both system users *and* public administrators.[223]

---

[221] Both the institutional economics and organizational behavior literatures draw on the notion of tacit knowledge in organizations. Tacit knowledge is developed in the course of doing an activity and cannot be reproduced by simply reading about the mechanics of how something works. North (1990) likens this to being able to play a musical instrument in contrast to knowing music theory. In this work, operational epistemic communities' domain knowledge may start as tacit knowledge, especially in the face of uncertainty of CPR dynamics. As tacit knowledge accumulates, it is codified into best practices espoused by that community, often becoming intertwined with the norms of that epistemic community. The challenges discussed in Part III of this work is to further elaborate *a*) how which elements of this knowledge creation and assessment process can be demonstrated as credible knowledge assessment; *b*) the anticipatory character (in the sense of McCray et al. (2010)) of these knowledge assessment processes; *c*) how to demonstrate the credibility of these processes to conventional regulatory agents and the global political arena writ large.

[222] From "Grágás, the law book of the Icelandic Commonwealth from 930–1262," cited by Eggertsson (2012, loc. 575–577), attributed to Eggertsson (1992, p. 433).

[223] This common image is rooted in understanding the value of a particular resource *system*. In Coase's early articulation of price theory, a distinction is made between qualitative notions of value and the market assignment of a scalar value. Rather, the common image sets the stage for bounding the market prices based on tacit knowledge of who different pricing levels will affect outcomes. This is evidenced in the quote on grazing sheep above, predicated on a common image of pastureland dynamics and value shared by shepherds and rulers. Although it may appear a top-down edict, if Levi's notions of negotiation between ruler and ruled are considered here, such rules are the product of bargaining that result in both an agreed upon common image and attendant operational rules consonant with that image. To provide a contemporary counter-example, the state of communications regulation in the United States and the substantive disagreement over what constitutes network neutrality is evidence of that lack of a common image of resource value from which bargaining can proceed. As will be discussed in Part III, external policy entrepreneurs efforts are, albeit not expressed in these terms, an effort at developing a common image of the NRS amongst agents of the state in bilateral and multilateral settings.

A more systematic approach can be seen in Ostrom's discussion of Los Angeles area water basins (E. Ostrom, 1990, pp. 111–112). For each of these basins, an association was formed to negotiate rights allocations. Consider the evaluation of the Raymond Basin, the first of three analyzed by E. Ostrom (1990, pp. 111–114). Water basin rights evaluation took place under a suit initiated by the city of Pasadena against the city of Alhambra and 30 other actors (E. Ostrom, 1990, p. 111). To inform proceedings, the Division of Water Resources of the California Department of Public Works was contracted to evaluate resource system dynamics.

Pasadena's suit was initiated in 1937. Six *years* later, in March of 1943 the results were available. Ostrom indicates "[t]he parties then shared a single, authoritative 'image' of the problem they faced," (1990, p. 112). This common image served as the basis of subsequent negotiation and adjudication in the court system. This is a classic case of a CPR developing within a well-defined governance regime, here a state government. Moreover, the Division of Water Resources was both *politically* authoritative (via the judicial system) and was considered technically authoritative.[224]

Pasture rights in 10th century Iceland and 20th century water rights evaluations seem rather distant from RIR labs. Both offer instances of the development of an authoritative common image. Both are rooted in a combination of participants' experience with the system and systematization of that knowledge. RIR labs arguably have a similar character. That said, RIRs' labs contribute to that kind of systematization, but do not pursue the label and legitimacy of normal science.[225] Rather, they are legitimized by continued consumption by, contribution by, and evaluation of the operational epistemic community.[226]

RIRs' labs, internal measurement efforts, and science departments contribute to creating a common image of the routing system. Although methods may have originated in the northeast corner of Table 3.1, modern RIR efforts are structured efforts that fall in the southeast corner. In a reversal of the water basin instance above, state actors and international organizations such as the OECD often invite RIR staff to present on the state of the Internet numbers and routing system. Within the RIR community, labs data is considered authoritative.[227] Knowledge resources created

[224]This is one case. This is does not by any means imply that all state agencies charged with analyzing technical systems are authoritative. This work, especially with respect to modern infrastructures, takes the opposite position. While state agencies remain politically authoritative agents charged with protecting the public good, technical authority and operational capacity increasingly resides in the private sector and the resulting "privately ordered" coordination institutions such as the CRIs.

[225]Normal science is used here in the sense of Kuhn (1993).

[226]This latter, evaluation, is a function of the combination of monitoring capability and domain knowledge. Taken together, these serve as a validation function within the community. These will be revisited in Chapter 8 when discussing the contribution of monitoring to knowledge assessment and adaptation.

[227]Authoritative should not be mistaken with infallible. Rather, it reflects trust in the actors performing the analysis, typically based on reputation in the community. Both the RIPE Science group and APNIC Labs are held in high-regard, both of which are led by longstanding, respected members of the community. Again, this is not to say there are not those that disagree with some of these outcomes and prescriptions from these groups, but rather that they receive substantive support.

by the operator community, here the RIR community, are derived from systematizing learnings from "everyday harvesting activities" referenced in Section 3.2. [228] While RIR labs measures and reports *are* systematized, some are derived from RIR experience and feedback from the operator community. A more accurate depiction of the process is a feedback loop between the northeast and southeast corners of Table 3.1.

5.3.2.3 Atlas Measurement Infrastructure

The RIPE NCC's Atlas probe network[229] is an instance of a jointly managed resource developed to collect Internet infrastructure data. The Atlas probe network comprises small hardware devices that connect to local networks via ethernet. From 5442 vantage points around the world,[230] Atlas probes conduct a number of Internet measurements.[231]

Atlas, as a measurement infrastructure, is, like the other common resources in these studies, jointly provisioned. Atlas probes were designed as a project of the RIPE Science group. Hardware selection, initial measurement specifications and implementations, and testing were performed by the RIPE Science group. Early on, a finite set of measurement types and sampling strategies were available. The project has since evolved into a platform that allows users to appropriate probe network resources for custom measurements on intranets and/or the Internet.

Paralleling the distinction between protocol provisioning and operational provisioning developed in Section 2.1, the Atlas probe network can be described in terms of design, implementation, and deployment provisioning. As above, design and implementation provisioning, referred to simply as design provisioning, was performed by the RIPE Science group. Like protocol provisioning of the IPv4 space, design provisioning established the parameters that bound utilization patterns in the Atlas network as a resource. Deployment of the probes is the product of community interest in the resource, willingness to place probes in private networks, and perceived returns. Probes are deployed in residential connections of RIR community members or within community members' network infrastructure.

As an analogy to Ostrom's irrigation system, Atlas probes can be likened to the distribution of remote measurement instruments to farmers maintaining the irrigation system. In a low-tech setting, this would require trusted reporting by farmers. In a more automated setting, such as the Atlas probes, measurements are automatic so simple deployment and maintenance is sufficient. Depending on the sophistication of the user, probes may or may not provide previously unavailable information.

---

[228] For a general forum, see the Measurement and Traffic Working Group (RIPE NCC, 2015b) in the RIPE region in general.

[229] See (RIPE NCC, 2014g) for an overview.

[230] See (RIPE NCC, 2014e) for various visualizations of Atlas probe distribution.

[231] Atlas probes are capable of performing common measurements such as ping and traceroute. Probes also perform some specific measures such as DNS queries and SSL queries. For more information, see (RIPE NCC, 2014b), in particular the section on Technical Details and subsequent links.

In the case of the sophisticated user, in both the irrigation and Atlas probe case, the user will have their own means of monitoring resource utilization.

In contrast to the irrigation case, access to the probe network in the aggregate gives the user access to data from public probes and the rights to invoke custom measurements. Hosting a probe entitles hosts to credits to perform set and user defined measures using the hosted probe and public probes. The RIPE NCC presents general data produced by the probe network,[232] specialized reports such as the recent view of Turkey's "meddling" with DNS services,[233] and various presentation in the MAT-WG (RIPE NCC, 2015b). The result is a broader view of system indicators such as latency to selected destinations. In terms of domain knowledge, such a measurement infrastructure moves the community from ad hoc tacit knowledge development and sharing of anecdotal best practices to an environment in which systematized experiments may be invoked to test certain hypotheses and confirm community observations. One instance of the latter is the use of Atlas probes to confirm the hijacking of Google's DNS IP addresses in Turkey.

5.4 RIR Constituency

The constituency of the RIR can be broadly described as any actor that accesses resource registration services as an authorized entrant or to whom Internet number resources have been allocated or assigned as a registry user. The rest of this section will be spent unpacking that statement without spiraling into a complete history of number resource management in the Internet.

The following subsections describe the primary roles played by community members in the RIR systems. All of the actors discussed below are community members, but may fall into one or more of the following categories: *a*) registry members (resource holders); *b*) policy development shepherds; *c*) RIR board members; *d*) members of RIR specific groups such as working groups, special interest groups, or task forces; *e*) employees of the RIR firm; *f*) contributors to the Number Resource Organization.With the exception of RIR employees, these classes are discussed below. Particular classes of RIR employees are discussed in line with the RIR function they contribute to.

5.4.1 Community

The RIR community largely comprises, but is not exclusively comprised of, the operational epistemic community. As noted earlier, *RIR* will refer to the firm that carries out the administrative duties of the registry; RIR *community* will refer to the broader community. Across the RIRs, leadership indicates, almost to a fault, that the RIR "does what the community tells it to." In general, this mode of delegation

[232] See the Atlas Results page at (RIPE NCC, 2014a).

[233] See the RIPE Labs article on the view of DNS server hijacking as observed by RIPE probes (Aben, 2014).

from community to RIR most often refers to the policy development process. In the broadest sense, this community comprises registry entrants.

Distinguished non-operator entrants of interest here are law enforcement and IP-address brokers. Law enforcement uses the registry in the course of criminal investigations. Their use is much like that of network operators. Law enforcement finds that it is investigating activity that seems to either originate from or be affiliated with a particular IP address. In order to further investigate, LEAs use the registry to determine who has been delegated use rights for that address. Typically this requires identify the LIR holding the rights for the block that address or group of addresses is in and contacting that LIR for additional information. In terms of the motivation behind registry access, both operators and LEAs are using the registry to contact individuals in order to resolve a conflict that appears to be originating at a given IP address.

IP-address brokers are a constituency that has identified the transaction costs of the emerging IPv4 market as a potential revenue stream. These actors comprise a mix of network operators and finance professionals. Nominally, address brokers attempt to develop specialized expertise in executing IP address transfers within and across RIRs. In terms of operational rules, address brokers have been opponent of the needs-based criteria. Brokers frame needs-based criteria as an impediment to a healthy market, leaving effective delegation of number rights to "pure" or "frictionless" market mechanisms. The operator community has had mixed responses, ranging from: speculation these actors will give rise to resource hoarding, perception that these actors are purely opportunistic and have little regard for the integrity of the system, to support for broker-mediated transfers as a post-exhaustion realty. The role of brokers will be developed further in the discussion of transfers in Section 5.7.3, in particular navigating the transition from needs-based to a transfer-market with institutional protections for the integrity of (in this case the accuracy of) the registry.

There are a variety of communication vectors between actors in their role as community members and leadership. Modes of communication include both public and private communication with leadership. Public modes of communication include public e-mail lists and participation in the RIR meetings. Private modes of communication include *a*) direct e-mails from community members to leadership, *b*) various modes of instant messaging, *c*) private conversations in various NRS management fora, *d*) RIR survey, and *e*) forms developed for collecting community input on various issues and services.

Any individual may become a community member regardless of whether that actor is an RIR member, holding number rights in that region. All community members have access to public modes of engagement. Community membership is a loosely defined in terms of participation. Active participation on e-mail lists is one mode of participation; attending and engaging in dialogue at RIR meetings is another.

Community participation includes rights to engage in policy development. Policy development processes are described in Section 5.6.2. Participation occurs formally through dedicated resource policy development e-mail lists and RIR meetings.

With regards to resource policy development, community members have rights of access that may be used to *collectively* exercise rights of management and exclusion through policy development processes. In general, the community expresses these modes of engagement under the umbrella of open and transparent participation in policy development processes. Refining the statement above, in terms of resource management rights, community members hold a minimum of access rights to the policy development arena through which management and exclusion rights may be exercised.

5.4.2 Number Rights Holders

Two general classes of use rights holders are discussed in the RIR system: RIR members and legacy resource holders. The two classes differ in their obligation to the RIR as the steward of the number stock. Legacy holders were delegated rights before the RIR system was developed. As such some, especially early number rights delegates, have no formal relationship with the modern RIR system. In contrast, RIR members are beholden to contracts that establish their obligations to maintaining the number stock, in particular utilization criteria and registry maintenance criteria.

RIR members comprise organizations that *a*) hold and exercise number resource rights, *b*) are contractually beholden to policies, in particular utilization criteria, and *c*) whose contractual relationship with an RIR defines revocation rights of the RIR with respect to delegated resources. In particular, both the policy manuals and contracts indicate that number rights are tied to the policy under which they were delegated. For instance, needs-based criteria evaluate thresholds based on the criteria at the time those delegations were conferred. The RIRs loosely follow the principle that an actor is bound by the obligations agreed to at the time a delegation was made, but this is not universal.

Legacy resource holders are actors that were delegated number resources by the IANA before the RIR system was established. As such, these actors had no formal relationship with modern RIRs. Under the loose binding principle above, the RIRs have tried to encourage a contractual relationship, but forcing that relationship has been met with resistance. In particular, attempts to limit legacy holders access to services has been met with criticism.

Multiple classes of legacy rights holders exist. Legacy holders that do not have a contractual relationship with a modern RIR are referred to as unaffiliated legacy holders. In terms of rights bundles, unaffiliated legacy holders can exercise origination and numbering rights, but *do not* hold the same set of facilities rights members in the RIR system hold. Unaffiliated legacy holders can participate in the policy development process (any community member can) but they cannot exercise joint registry management rights. In contrast, some RIRs, in particular ARIN, have a contractual mode specifically for legacy holders. These legacy holders have access to ARIN services but are not beholden to utilization criteria. The rationale behind this compromise is that, while legacy space is inefficiently utilized, the RIR does not have the authority to coerce legacy holders and would rather do as much as possible to ensure the integrity of the registry. The resulting legacy contracts are a

compromise.

The principle binding rights to policy governing at the time of delegation does not apply to transfers. As will be developed more precisely in Section 5.6.3, a transfer is a limited form of alienation that is effectively a return of number rights to the RIR and an immediate and subsequent delegation of those number rights to the agreed-upon recipient. If legacy holders engage in transfers, the number rights must pass through an RIR in order to complete the transaction. Currently, this only affects ARIN and APNIC, both of which require recipients be either current members of the RIR or become members. Under this stipulation, transferred numbers are subsequently governed by existing resource policy, including utilization criteria and limitations on subsequent transfers.

In part, RIR policies are enforced by the need of resource holders to return to withdraw additional resources. As a source of authority, this fits a relational authority framework better than a conventional principal-agent framing. Levi (1989) indicates that governors[234] do not necessarily hold absolute power; conferring and rescinding authority is not an instantaneous process. Rather, governors and the governed are engaged in a continual bargaining process—each plays the role of principal in one context and agent in another. Here, the RIR can play the role of a principal that delegates numbers rights as long as it has a stock of numbers rights to allocate. Its management role with respect to those number rights allows it to exert authority over how demand for those rights is satisfied. That said, rights holders play the role of the principal in the collective choice arena. Members *as a collective* hold management rights over the number resource stock and the registry. Members confer specific alienation rights onto the RIR as a firm by exercising those rights in the collective choice arena. There are two outcomes of the exercise of these management rights: *a*) tacit reaffirmation of the role of the RIR to implement those operational rules and *b*) articulation of appropriation rights available to members.[235]

Most legacy rights holders were allocated number resources before CIDR allocations. These allocations were relatively large. In many cases, these rights holders do not need any more resources from the RIR. As such, legacy holders do not necessarily have incentives to either participate in RIR management or follow RIR policy. Given they were allocated before the modern RIR system was created, they are not beholden to utilization monitoring.

In addition to the rights held by community members and prefix origination rights, RIR members also have secondary collective choice rights for managing the RIR as a firm. These management rights are generally manifest in rights to vote for board members and in some regions to vote on the member fee structure. As

[234]Levi (1989) uses the term "rulers," "governors," "peasants," and others in her work to describe those who create and enforce rules, those who are beholden to them, and the bargaining dynamics between them. That work draws on historical narratives such as medieval courts to develop a theory of predatory rule. While the dynamics developed in that work are useful here, using the term ruler would imply a stronger relationship than actually exists.

[235]Recall from Section 3.4.3 that management rights are held by those that determine the parameters of appropriation rights.

such, members have indirect rights to manage the RIR as a firm.

5.4.3 Board Members

Board members are the representatives of the body of members. Members select board members through regular elections. In general, board members' executive authority is over the management of the RIR as a firm. Recall the constituency holds and exercises management rights, conferring the right the implement these rights to the firm. As the elected representatives of the constituency, the board is empowered to monitor firm activities and enforce the bounds of implementation rights conferred. Responsibilities include strategic direction, fiduciary responsibility, oversight of the firm, and travel on behalf of the firm. In all but the RIPE community the board also serves to ensure the policy development process was followed. The board does not evaluate the content of the rules, but rather, ensures the *process* was followed; see Section 5.6.2.

Each of the RIRs describe the general responsibilities of board members. Board members are expected to have both industry experience and senior management experience.[236] These actors are expected to attend meetings; typically someone with experience with the RIR and a track record of participation. Baseline experience both draws on the operational experience of the board member in industry, in the RIR, leverages that in the management of the firm. As part of the epistemic community, board membership is a point of prestige that reinforces board members role as an authority in the community. Board participation has value for the actor and their business.

5.4.4 Policy Shepherds

Each RIR has a group of actors that facilitate policy development by the RIR community. RIR policies are proposed, written, submitted, and evaluated by RIR community participants. Comparisons of these processes are detailed in Section 5.6.2. In terms of constituencies, shepherds are responsible for guiding community members through the policy development process. Shepherds, in conjunction with RIR policy officers,[237] coordinate the policy development process itself. These actors do not hold special number resource rights. Rather, like the RIR firm, the community delegates its aggregate rights to provision and preside over collective rule making arenas to the shepherds collective. Like delegation of rights to implement management and appropriation rules, delegation of rights to manage the coordinative function of the collective action process is not a delegation of substantive rights,

[236]For instances, see (ARIN, 2015) and (RIPE NCC, 2015c).

[237]APNIC, ARIN, and the RIPE NCC each have a staff member dedicated to coordinating policy development between the RIR as the firm and the shepherds collective. These two sets of actors jointly provision the collective action with the administrative and resource support of RIRs' staff contributing to general coordination of dedicated membership meetings and membership meetings held in conjunction with other venues (such as ARIN Consultation Meetings held during NANOG).

but rather rights to coordinate *the development of* substantive rights to be implemented by the RIR firm. In their capacity coordinating policy development, serve to facilitate evaluation of operational rules developed by the operational epistemic community.

Amongst the RIRs, there are deficient configuration of shepherd collectives. In the RIPE community, the shepherds collective comprises of working group chairs from function-specific working groups.[238] In APNIC, shepherds are chairs of special interest groups (SIGs), comparable to RIPE working groups.[239] In both regions, function-specific policies are developed within the corresponding function-specific collectives. The chairs of these collectives shepherd the policies developed in their respective groups. In RIPE, there are eleven active working groups and twenty-five working group chairs.[240] In the APNIC region, there are only two active SIGs, the Policy SIG and the National Internet Registry (NIR) SIG; historically there was a greater diversity but these groups have since dissolved. In LACNIC and AFRINIC, a single working group is dedicated to policy development.

In the ARIN region, the shepherds collective is the Advisory Council (AC), a collective elected by the community. The AC comprises 15 members.[241] Among other differences with SIG and working group based shepherds, a distinguished characteristic of this collective is that the AC assumes ownership of draft policies. Based on interviews, although the AC asserts "ownership" over the policy, the designated shepherds attempt to ensure the original author remains as involved as possible. Assertion of ownership is to avoid delays in policy development when original policy authors decide, for whatever reason, to no longer participate in the policy development process. In Section 5.6.2, this distinction is important to ensuring policy development keeps pace with changes in the operations ecosystem.

5.4.5 Number Resource Organization (NRO)

Authority to operate as an RIR is derived *a*) historically from delegation of number rights from the IANA; *b*) the support of the constituencies described above; and *c*) mutual recognition by other IRs, in particular RIRs within the NRO. Of these three sources of authority, the NRO is the most recent. The RIR system is a federated ensemble that adheres to common constitutional norms, initially codified in RFC 2050 (Hubbard et al., 1996). Policy development in each RIR has adapted operational manifestations to region-specific preferences. This section describes the NRO.

[238] In the RIPE region, working groups are quite active. In the APNIC region, many of the working groups have been retired. It has even been suggested that the APNIC policy development process is no longer needed given IPv4 depletion and the magnitude of IPv6 allocations. This will be discussed in more detail in Section 5.7.

[239] There can be some confusion. In most regions, working groups are not permanent, but are the most long-lived of function-specific collectives. In APNIC, SIGs are the longest-lived, with working groups and birds of a feather meetings (BOFs) addressing shorter term projects. See SIG guidelines discussion in (APNIC, 2014g).

[240] This is as of January 2015.

[241] See (ARIN, 2014) for AC member profiles.

The mission of the NRO is:

> To actively contribute to an open, stable and secure Internet through:
>
> - Providing and promoting a coordinated Internet number registry system
> - Being and authoritative voice on the multi-stakeholder model and bottom-up policy process in Internet governance
> - Coordinating and supporting joint activities of the RIRs (NRO, 2014b)

The NRO was conceived and formed in the early 2000's when the risk of ICANN failing seemed to be an increasing reality.[242] As a risk mitigation strategy, creating the NRO was a way to create a collective[243] that could, among other tasks, perform the numbers role of the IANA. The NRO MOU with ICANN establishes the NRO as the role of the Address Supporting Organization and as a common forum in which the RIRs can engage as a collective and speak with a common voice. The MOU FAQ (NRO, 2003b) makes it clear that the NRO is not intended to compete with ICANN or to create a complete replacement. Rather, the focus is to facilitate coordination amongst the RIRs.

With respect to the NRS, the primary functions of the NRO are *a)* ratifying global RIR policies and *b)* coordinating RIR activities such as RPKI certification. In terms of the form, the global policy ratification process is described in the NRO

[242]The risk of failure is articulated in an open letter from the RIRs to ICANN in 2003:

> the RIRs are aware that ICANN is a private corporate entity, and that its future is one that is not absolutely assured. There is a risk, as with any private corporate entity, that the entity may fail. Failure of ICANN includes the risk of a freezing of the undelegated number pool, which in turn places a significant risk in the continued operation of the registries and the application of their policies. The ultimate risk here is a shift in the number administration from the careful preservation of uniqueness within the assignment of number resources to one of chaotic number movement, with its attendant consequences which appear to inevitably include a breakdown of the coherency of the Internets address realm. Obviously, this is not an acceptable outcome under any circumstances. (NRO, 2003c)

Following the language of the MOU FAQ:

> One of the functions that the NRO is intended to provide is to undertake, if necessary, IANA number resource management. The NRO is not being structured as a potential 'complete' ICANN replacement in the event of the failure of ICANN. The consideration here was that in such an event it was considered probable that the various IANA functions would be undertaken by those with a *direct* interest in each particular functional area.(NRO, 2003b, Response to question "Does the NRO replace ICANN?" Emphasis added)

Note the emphasis on direct interest. Direct interest here maps to direct users (Section 3.1), here described as RIR community members.

[243]In interviews and private conversations it has been stressed that the NRO is a collective, it is not a formal incorporated organization (NRO, 2003a). The NRO remains unincorporated since the 2003 MOU but has since appointed a standing Secretariat (NRO, 2013).

MOU (NRO, 2003a), born of the MOU negotiations between ICANN and the RIRs.[244] Key to the *ratification* process is that the NRO does not make global policy. Global policies may be conceived by community members of one or more RIRs, but to function as a global policy, a variant of the global policy must have been adopted by each of the respective RIRs. The Number Resource Council, comprising a combination of RIR community members from all five regions,[245] *a*) confirms original policies were adopted in accordance with RIR PDPs, *b*) derives a common policy based on the common elements of corresponding policies adopted by the five RIRs, *c*) offers this policy to the RIR community for a final evaluation, and *d*) ratifies that policy if there is no major disagreement. This process is the global process equivalent to the process evaluation phase of the RIR policy development process described in Section 5.6.2. The objective of ratification is not to *make* policy. Rather it is a check on the process by which the policy was made, ensuring (bottom-up) community consensus processes were not subverted.

5.5 Arenas

Number resource policy is the product of both a collective choice process and information sharing within the operational epistemic communities. Operational epistemic communities comprise communication fora that contribute to domain knowledge. That domain knowledge contributes to credible assessment of rules and expected outcomes debated in RIRs' collective-choice processes. As will be discussed in Section 5.6.2, the consensus process has a similar overall structure across the three RIRs. There are differences in modes of engagement, evaluation, and phases of the consensus process with the most salience to collective rule making in a management resource facility. The arena itself is considered a resource. The facilitation and communication mechanisms below are leveraged to enhance access to these arenas and increase participation in the collective-choice process.

RIR meetings are a general forum at which a number of other arenas are convened. Meetings are jointly planned by the RIR staff and a program committee comprised of operational epistemic community leadership. In addition to the face-to-face meetings, e-mail lists are a type of technical forum corresponding working groups, task forces, and other discussion topics. Further supplementing face-to-face arenas, chat forums, both those provisioned by the RIR and privately provisioned forums, are yet another communication vector. Finally, training sessions offered by the RIR are yet another forum.

[244]For instance, see (Wilson, Plzak, & Pawlik, 2002b), (Wilson, Plzak, & Pawlik, 2002a), and (Wilson, Plzak, Echeberria, & Pawlik, 2002) for the dialog that led up to the NRO MOU (NRO, 2003a).

[245]Members of the NRO Council (NRO, 2015) serve as the interface between the NRO and ICANN.

5.5.1 Policy Development Arena

The policy development arena comprises communication fora in which the RIR community introduces, debates, and vets policy proposals. To the conventional policy maker, the most familiar arena will be the RIRs' membership meetings, in particular those that utilize conventional voting mechanisms. The bulk of the policy meetings have similar structure to NOG meetings, but policy sessions comprise either *a*) directed discussions of policy proposals already under discussion or *b*) presentation of a technical issues or evidence pertinent to a particular policy. Proposals are presented either by the author of the proposal or the policy shepherd attending that proposal. This physical space and administrative coordination is provisioned as part of regular RIR membership meetings (by staff, using RIR funds), but its structure and agenda is set and administered by policy shepherds. In each RIR, policy sessions are chaired by one of the policy shepherds. As a final component of policy arena provisioning, ARIN, RIPE, and APNIC archive the transcripts of many meeting sessions, in particular policy meetings.

E-mail lists also serve as a policy development arena. Policy and working group specific e-mail lists are provisioned by the RIR to facilitate discussion of current policy proposals and topics of interest to the community. As will be discussed in Section 5.6.2, the e-mail lists are one vector for coordinating and documenting the process of achieving consensus on a given proposal. In the RIPE region, the e-mail lists are the*primary* arena for consensus.[246] In the course of the consensus process, an actor on the list may engage in discussion of the proposal, arguing both pros and cons of the proposal. To ensure the shepherd is clear where an individual stands, if the actor does not indicate explicitly "I support" or "I oppose"[247] this proposal, the shepherd may send a message to affirmatively confirm support or opposition. Participation via e-mail is low-cost in terms of *a*) maintenance, *b*) archival, and *c*) barriers to entry for those that cannot travel to meetings.

RIRs provision additional channels for participation in the policy development process. RIRs webcast meeting sessions, in particular the policy meetings. Back channels supplement the primary (local) channels. For instance, Jabber (IM) sessions for each meeting session are provisioned to facilitate *a*) asynchronous discussion amongst meeting attendees and remote participants in parallel with the face-to-face session and *b*) questions to meeting chairs and presenters from remote participants. In addition to RIR provisioned back channels, other group communication channels are frequently used amongst social groups within the community. Individuals' Facebook, Google+, and other social media have served as discussion back channels. There are also a number of semi-private IRC channels used by long time members of the community.

[246] In RIPE, policy meetings are fora for presenting the policy and discussion, but determining consensus based on support and dissent occurs exclusively based on utterance of "support" or "do not support" on the mailing lists. For a recent review of the RIPE process presented to the community, see (Steffan, 2012).

[247] The language does not need to be exact, as long as it is a single phrase, word, or known indicator. Some actors simply reply to a thread with the single word "Support" or "Oppose." Others have adopted the Google+ nomenclature, simply replying "+1".

Finally, the informal but oft cited arena for policy discussion is, like many conferences, hallways discussions. Numerous RIR meeting participants have highlighted the role of discussions after policy meetings. Often policy discussions continue during the breaks between sessions. This is expected, but has been explicitly encouraged by the policy shepherd at the end of sessions to encourage further discussion, especially if there will be later policy sessions on that particular proposal. In conversations during fieldwork, participants have indicated that hallway discussions are often the seed of direct participation in primary meeting channels.

5.5.2 Government Arenas

The RIR system has developed a number of arenas dedicated to engaging with government agents and LEAs. Some of these are closed, limited to RIR staff, board members, and government officials: instances of these are *a*) ARIN's Government Working Group (AGWG),[248] *b*) the RIPE RoundTables (RIPE NCC, 2014f), and *c*) AFRINIC's Government Working Group (AFRINIC, 2015). Closed sessions are an exception in the RIR ecosystem. RIRs adamantly support a general ethos of open and transparent proceedings, evidenced by the variety of engagement modes discussed earlier. Government engagement is partially explained by the need to ensure governments have a venue they are comfortable with.

Interviews with actors that participate in these closed fora and bilateral discussions with government officials stress the diplomatic element of government engagement. Recall Section 1.2 concludes indicating that merely transmitting a coarse image[249] of how the NRS works is often considered a victory. The operational epistemic community has a deep understanding of routing resources, protocols, and technology rooted in field experience. In contrast, most government actors do not have this technical background. Further confounding developing a common image is that tacit knowledge is not easily transmitted. Externally facing RIR policy entrepreneurs, typically staff members delegated this responsibility along with CEOs, attempt to convey sufficient dynamics to understand outcomes. Diplomatic engagement within government arenas share the educational components of operator training undertaken by RIRs but also require RIR policy entrepreneurs recognize the cultural and professional differences between the operator community and government actors.

Interviews have indicated that substantive tact is necessary when engaging government actors. Within the operator community, there is an expectation of a lower bound on knowledge necessary for participation. For the uninitiated, in the best

[248] See early discussion by Flaim (2009) at ARIN XXIII.

[249] A task of credible knowledge assessment in this arena is to move from acting as an opaque agent purported to act in the public interest, often manifest as "the good of the Internet," to one with whom actors can reason about the outcomes of a particular outcome based on commonly held axioms of system operation. In terms of political economy, in particular bargaining and measurement costs offered by the transaction costs model, the common image is the kernel of standards. In the global regulation literature rooted in regime development, this development phase can be likened to agenda setting. Here, the question is whether agenda setting is establishing a common image or whether it is building on a common image to develop monitoring and enforcement mechanisms.

case, statements that are technically incorrect or off-topic will be ignored as such or politely disregarded. This is not guaranteed, though. Community guidelines, oft repeated at the beginnings of face-to-face sessions, espouse respectful response to statements one may disagree with. That said, in the broader community the ethos of constructive critique often weighs more heavily than these edicts, leading to less gentle rebukes of those that seem to be offering authority. Rebukes are rooted in the perception that individuals *a*) do not have the expertise to offer the statements they are making, *b*) are offering solutions that "work" but have undesirable side effects, *c*) or in the worst case have clearly not done their homework on a topic or read the appropriate public documents. This is a function of both the ethos of the operational epistemic community furthering its knowledge domain and simple economy of time in face-to-face sessions many actors are attending on company funds.

Government arenas provide a forum to temper the community's potentially caustic knowledge assessment processes with a tactful educational and policy agenda that spans the penumbra between resource and public policy. As a closed forum, engagement is primarily between government agents and policy entrepreneurs from RIR staff. Individuals from the broader RIR community do participate, but usually these actors have developed a reputation as policy entrepreneurs in their own right and have been vetted by the firm for their capability to tactfully engage with external actors. In the words of one external relations officer, it is imperative to avoid making government agents "feel stupid" or look ignorant of issues. This is especially the case in fora comprising representatives of multiple governments. In one-on-one discussions, external relations actors have reported that government agents are much more likely to ask questions about both the technology and the process. Current government fora tend toward educational agendas, providing updates on *a*) clarifying the registry function, its limits, and the sources of those limits; *b*) IPv4 depletion; and *c*) IPv6 deployment and trends. The RIPE RoundTables and the AGWG were both partially catalyzed by early engagement with LEAs, in particular clarifying registry function.

An increasingly important objective across engagement in government fora has been to establish the RIR and the RIR system as the authoritative source of information and tools relating to NRS operations. In early engagement with LEAs, there was a misconception that the registry controlled who could be on the Internet and who could not be. The first round of discussion focused on the simple message that the registry was not a means to revoke Internet connectivity, it was, as has been developed here, simply the documentation of rights conferred. A key element of the number delegation and revocation discussion in Section 5.2.3 is that a resource delegation is necessary and is the legitimate means to NRS participation, but revocation is not centrally controlled. Rather, freezing of registry entries only preserves the information about a given delegation. Moreover, even if revoked, truly malicious actors are not likely to respond to an RIR while "thumbing their nose" as law enforcement. In other words, RIRs currently do not have the enforcement power LEAs initially thought they did.

The misperception by LEAs is in part due to the lack of a common image of

number resource use and enforcement mechanisms. Part of what diffused the situation with LEAs was RIRs' precise explanation of registry function relative to their investigations. In multiple accounts of LEA engagement, the RIR invited LEAs to sit down for what amounted to an RIR tutorial. A theme of this tutorial was that the registry is not a kill-switch. What the registry does provide, relative to investigators incentives and inline with operator use, is documentation of who to contact in the event of costly externalities. LEAs were encouraged to learn how to use the registry in much the same way operators used it. Instead of unraveling the provenance of security externalities, LEAs are collecting information that contributes to the provenance of criminal activity.

RIRs helped LEAs adopt the common image of the registry as documentation and a source of operational information. Reports of initial encounters included a range of threats from LEAs. In some cases, there was a tacit threat of inducement with warrants that would limit or even halt RIR operations. Threats of operational hostage taking were predicated on denial of a regulatory and enforcement tool perceived to be an immediate, low-cost solution. Such "predatory rule" is not uncommon. A key factor diffusing this instance of predatory rule was the common image of the registry as an information resource. Further, cooperation by the RIRs in helping LEAs use the resource in this manner proved beneficial to later investigations.

The common image of the registry as a tool that does not encompass enforcement power is changing, though. RPKI, in its strong form, may offer LEAs and other government agents the enforcement tool initially envisaged in early LEA engagement. As developed in Sections 5.2 and 5.7.4, RPKI does not alter the bundles of rights conferred, but it has the potential to make revocation more immediate. If coupled to routing decisions, as alluded to in both RFCs describing the RPKI system and some supporters' policy discussions, RPKI could facilitate stronger enforcement of not only number rights, but broader abuses perceived by any number of regulatory agents. Currently RPKI has seen some adoption, but it has not yet reached what may be considered critical mass. Section 5.7.4 describes the RPKI issue in detail, highlighting the community's discussion of what is here framed as predatory rule, along with the implications for the common image of two distinct spheres of number resource policy and routing operations.

Without engaging in what could be viewed as unintentional naming and shaming, some policy actors have actively fostered competition amongst state participants regarding engagement with RIRs and development efforts. Often this is an effort to highlight the positive outcomes of engagement and foster operational capacity. One issue area in particular where this has been applied is IPv6 deployment. One external relations actor described a strategy of engendering competition regarding degree of IPv6 deployment in a developing region. Within that region, state actors were already competing in terms of general economic development—actively competing in terms of their IPv6 allocations and deployment is another potential comparator. RIR measurement and analysis efforts are leveraged to provide authoritative information on delegation and deployment that fuels this constructive competition amongst state developers.

The RIR is also careful not to present itself as choosing winners or losers. The RIR frames itself as the authoritative epistemic community reporting on the state of IPv6 deployment and Internet infrastructure trends in general. "Me too" politics are at play and may be leveraged by the RIRs to encourage initial engagement. Given the RIR controls measures and reporting that shows relative "standings in the race" states must follow-through on their commitments to garner the accolades. At the moment, IPv6 deployment is remains an issue of low politics, hidden under gross indicators of more general Internet penetration as a measure of economic growth and viability. That said, if RIRs are successful in promoting IPv6 and themselves as authorities on delegation and deployment, their methods may come under increasing scrutiny.

As a government forum, RIRs seem to have developed a strategy of education and information sharing. This strategy was at first reactive but is becoming increasingly proactive. Initially much of the external relations message was about inviting actors to policy making fora. The message was to come to the RIRs' arena. In addition to developing government-specific arenas, RIRs have also engaged in government organized arenas as observers. Increasingly, RIRs have been invited guests.

The more proactive strategy of will be revisited in Part III. First is the question whether this proactive stance is a prelude to anticipatory policy making contributions. Second, how will this serve as an input to downstream issues *outside* the NRS scope? Again, the RIRs do not intend to be a lobbying group. Eschewing support of particularistic social issues, the RIRs must also behave like a public interest watchdog relative to a non-discriminatory infrastructure while carefully avoiding favoring particularistic outcomes. Recall the general question in Section 1.3: are the incentives of network operators commensurate with social risk of failure? This question runs parallel to the previous two. Providing the domain knowledge necessary for anticipatory policy development is currently sufficient for the epistemic community. An ongoing question is whether these institutions have both the incentives, both within the firm as an agent that wishes to perpetuate itself and a principal constituency willing to pay for these efforts, to pursue this role? Subsequently, will this be sufficient to provide the resources necessary to make credible commitments commensurate with the social risk?

RIR communities also have open collectives dedicated to government collaboration. RIPE's Cooperation Working Group (RIPE, 2014a) and APNIC's recently formed Public Policy Advisory Committee (PPAC) are instances.[250] These groups are intended to perform community outreach to external actors, in particular governments, regulators, and other NGOs. Participation is typically much smaller than other more technically-oriented groups.[251] Participation largely comprises actors

[250]Creation of the PPAC was initiated in August 2010 (APNIC, 2014h). Two working group proposals are available: draft version 1 (Chharia, 2010) and draft version 2 (Tandon, Nair, & Lee, 2010). The transcript of the inaugural meeting highlights existing members scoping of APNIC to registry and routing issues versus the breadth of public policy in the broader Internet governance ecosystem (APNIC, 2013).

[251]This is based on observation at RIPE, APNIC, and LACNIC member meetings.

with existing interest in the intersection of public policy and resource policy. Community policy entrepreneurs comprise external relations staff and active community members. Regulatory policy entrepreneurs are often government officials that have developed relationships with community policy entrepreneurs. One vector for developing the relationship between the Cooperation Working Group and regulators is external relations staff engagement with regulators. Another vector is independent community entrepreneurs. The Cooperation Working Group and PPAC are held at member meetings. For government officials and other "external" actors (not part of the epistemic community) this offers access to the broader epistemic community and visa versa. In contrast, closed meetings, such as the RIPE RoundTables go to Brussels to attract the most regulators, potentially creating new contacts.

Consider the description of the RIPE Cooperation Working Group:

> The working group will primarily discuss outreach from the traditional RIPE community to everyone else, especially governments, regulators and NGOs, all of whom we are trying to bring into our community. Topics are **not** to duplicate issues discussed in other working groups. This working group should complement the other working groups and help participants engage in appropriate work. (RIPE, 2014a, emphasis in original)

In contrast to closed fora, open collectives expose regulators to elements of the RIR processes. One set of regulators participating in these fora are actors with a history in the RIR and that are familiar with RIR processes and consensus processes in general. For RIR community members, external actors provide reports on government, regulator, and NGO activity affecting the community. The Cooperation Working Group is also the forum in which

> [t]he RIPE NCC's current outreach activities will be reported, and the RIPE NCC will seek advice and guidance on future activities. This is to make the discussions more focused—currently the only forum for these discussions is the ripe-list mailing list.[252]

As such, the Cooperation Working Group also serves as a forum for outreach to be discussed amongst external actors and community members.

The notion of complementing other working groups is common across the RIPE region and the recently formed APNIC PPAC. Within RIPE, there are currently eleven active working groups.[253] In APNIC, both the dialogue on the mailing

[252]As per point 2 of (RIPE, 2014a). The statement was taken from the proposal for the working group, hence the reference to the ripe-list as the only other forum for discussion.

[253]The bulk of resource policy is, unsurprisingly, developed in the Address Policy Working Group, but there has been a recent uptick in activity in the Anti-Abuse Working Group. As an interesting aside that will be developed in Section 5.6.3, recent discussions in the Anti-Abuse Working Group have given rise to potential proposals that will likely be topical to anti-abuse but a policy issue for the Database Working Group or the NCC Services Working Group.

list (wg-government) created to discuss the development of the PPAC[254] and the inaugural PPAC meeting dialogue (APNIC, 2013) addressed the role of the PPAC with respect to other fora, in particular the Policy SIG in APNIC. In the wg-government e-mail list a number of active APNIC members questioned the premises espoused in initial proposals,[255] in particular the need for a special forum given resource policy development in the Policy SIG is open to all.

5.6 RIRs' Rules

RIRs' initial constitutional rules are the product of operational experience with RIR development. Generalizations of this experience are documented in IETF maintained RFCs. Until relatively recently, RFC 2050 (Hubbard et al., 1996) was considered one of the authoritative sources of documentation of RIRs' *constitutional rules*, but has been more recently superseded by RFC 7020 (Housley et al., 2013).[256] RFC 2050 set out the norms that formed the foundation of RIRs' operational rules; collective choice rules are derived from IETF consensus processes. As the RIRs developed, they documented their own interpretations and operationalizations of constitutional, collective-choice, and operational rules. The constitutional norms discussed in the next section draw heavily on a history of integrity comprising a balance of conservation, routability, and uniqueness (registration). RFC 7020 is one articulation the change in what factors constitute integrity as the remaining stocks of previously undelegated IPv4 number rights approach absolute depletion.[257] These

[254]At the time of creation, it was a discussion of an APNIC GAC akin to ICANN's GAC. See the list archive (APNIC, 2014c).

[255]The creation of the PPAC is an artifact of the contention over the creation of India's NIR. The process for developing the Indian NIR was fraught with misunderstandings in terms of how to establish the NIR, what the criteria for establishing an NIR were, and later how to change the operational rules of the RIR.

[256]Recall from Section 3.4 Hart's distinction between primary and secondary rules. One class of secondary rules is rules of recognition, namely where one looks for authoritative documentation of rules. The deprecation of RFC 2050 illustrates the recognition, at least within the IETF maintained RFC series, of RFC 7020 as the most recent authoritative statement on RIR operations. The shift in actual constitutional rules engendered by recognizing RFC 7020 is discussed in Section 5.7.1. The development of operational policy that reaffirms the norms in RFC 2050 provides insight into the community's perception of the balance between the RFCs as authoritative and their own policy corpora as authoritative. Both sets of documentation are recognized—RFCs from the body of documentation developed by protocol producers, resource policy corpora the produce of an operational epistemic community. In a discussion of RFC 7020, one of its authors, John Curran, President of ARIN, made it clear that RFC 7020 should not be confused with assertions derived from community consensus. Interestingly enough, RFC 7020 itself serves a similar role relative to the IANA function. In particular RFC 7020 rescinds the role of the IANA in evaluating global policy in lieu of the NRO Council.

[257]The IANA depleted its stock of previously undelegated IPv4 addresses when it delegated the last five /8's on 31 January 2011. Overall, the stock of previously undelegated rights has not yet been absolutely depleted—at present, all the RIRs but AFRINIC have begun allocation from their last /8. Recall the near flat delegation rates for APNIC and the RIPE NCC in Figure 5-2. As will be discussed in Section 5.6.3, a distinguished set of appropriation rules shifted from a needs-based allocation mechanism based on estimates of demand to a static delegation rate rooted in equitable

more recent changes to constitutional norms highlight the role of collective choice rules in reinterpreting norms that, in the aggregate, contribute to what the CPR framing refers to as resource integrity.

Collective-choice rules in Section 5.6.2 are rooted in consensus-based decision making. Across the three NRS institutions evaluated in this dissertation, actors have highlighted consensus-based decision making is derived from IETF consensus processes,[258] but have evolved to fit the environment at hand.[259] Section 5.6.2 discussed the common three phases of RIR policy development processes (active consensus, passive consensus, process validation), and the differences in the processes across the RIRs. While the three phases remain the same, there are differences in terms of scope, role of the RIR board in process validation, policy fora, and shepherds' collective structure and function. The operational rules produced by these processes are discussed in terms of *a*) the registry as the core resource; *b*) non-core resources such as IRRs; and *c*) arenas such as general resource policy arenas, working groups, and closed groups affected.

5.6.1 Constitutional Rules

One of the first articulations of a registry system is RFC 1174 (Cerf, 1990), which suggests IANA remain the centralized authority, asserts IANA have the power to delegate portions of number resource management, and that "candidate delegated registries meet with the IANA and IR to review operational procedures and requirements and to produce documentation to be issued as RFCs describing the details of the proposed distributed mode of operation," (1990, p. 3). RFC 1366 (Gerich, 1992) further refines the criteria for establishing distributed regional registries. A key premise for distribution is that registries "located in distinct geographic areas may be better able to serve the local community in terms of language and local customs," (1992, p. 2). A subsequent assertion is "that there is just a single regional registry per geographical region at this level to provide for efficient and fair sub-allocation of the address space," (1992, p. 2).

RFC 1366 makes the following criteria explicit (1992, pp. 2–3):

a) networking authorities within the geographic area legitimize the organization
b) the organization is well-established and has legitimacy outside of the registry function

access to number rights. The premise was to ensure new network actors could still get an initial IPv4 allocation. Given the rate of delegation, absolute depletion will not occur soon.

[258]For an introduction to IETF processes, see (Hoffman, 2012). Recall Resnick (2014) and Section 3.2.5 on consensus in the IETF.

[259]Recall from Section 3.2.5 that the consensus process is a systematized concept that gives much greater weight to credible minority dissent rooted in domain-specific knowledge of the issue over more conventional majoritarian decision making based-on black-letter voting thresholds. In each study in presented in this dissertation an epistemic community has adapted the consensus process and, in doing so, stressed one phase or mode of consensus over others. While different in structure, process, and formality, this work argues they each share the generalized characteristics of the systematized concept (especially in contrast to majoritarian mechanisms) offered in Section 3.2.5.

c) the organization will commit appropriate resources to provide stable, timely, and reliable service to the geographic region
d) the commitment to allocate IP numbers according to the guidelines established by the IANA and the IR
e) the commitment to coordinate with the IR to establish qualifications and strategies for sub-allocations of the regional allocation.

In particular, criteria a) and b) provide links between community-based consensus and conventional delegation of authority. A simple interpretation would be that RFC's document to the IANA creating a processes by which to delegate IR authority. Point a) and b) above can be interpreted as a requirement that IRs have authority within the community in and of their own reputation "outside of the registry function." In terms of the operational epistemic communities, the potential IR must demonstrate it is at a minimum (necessarily) a member of that community and has reputation as a legitimate and "well-established" participant, an indicator of not merely proficiency, but adept accomplishment in the operations domain. Many of the organizations that played early IR functions were network information centers (NICs) or evolved out of early protocol and operational communities such as RARE serving as the key academic operator community that fostered the development of the RIPE NCC. Organizations satisfying Points a) and b) intrinsically have their own domain of authority and operational capability. Among other ramifications, they are not necessarily dependent on the IANA for either support or legitimization. Rather, they are complementary to the IR function.

Taken as written, the process is not a simple beauty contest adjudicated by the IANA. Rather, would-be constituents of the registry must legitimize the organization. Points c)–e) set out criteria for coordinating resource delegation with the IANA and cooperation with the IANA to create a uniform and consistent application of general norms and principles. Point c) speaks directly to operational endowment. Commitment of appropriate resources by an RIR is an expenditure of membership fees. Service requirements speak to elements of professionalization discussed in a number of interviews across RIRs. Through both collective choice processes and votes on RIR activity plans by members, operational endowment is contingent on service requirements and reaffirms the constituency in its role as a principal.

Points d) and e) are early articulations of coordination between the IANA and IRs and amongst the IRs themselves. Given the role and function of the NRO, the spirit of d) and e) has been fulfilled by that federated collective. Point d) seems to have elements of simple principal-agent in stating "commitment... according to the guidelines established by the IANA." These roles have not remained static. Rather, the roles an organization plays and the actors that play those roles have evolved dynamically sustain a web of relational authority preserving the integrity of the number delegation system. Chapter 8 explains the relationship between the recent IANA function and the RIRs in terms of relational authority and contingent social order.

The principles and norms for RIRs were established in RFC 2050 (Hubbard et al., 1996) as best practices. Although established in the RFC series, RFC 2050 linked

the authority of an established authority in protocol production with operational epistemic authority. The principles described in RFC 2050 were developed through experience in the community[260] and codification of these experiences into a set of principles. The goals (interpreted here as principles) established are:[261]

Conservation: Fair distribution of globally unique Internet address space according to the operational needs of the end-users and Internet Service Providers operating networks using this address space. Prevention of stockpiling in order to maximize the lifetime of the Internet address space.

Routability: Distribution of globally unique Internet addresses in a hierarchical manner, permitting the routing scalability of the addresses. This scalability is necessary to ensure proper operation of Internet routing, although it must be stressed that routability is in no way guaranteed with the allocation or assignment of IPv4 addresses.

Registration: Provision of a public registry documenting address space allocation and assignment. This is necessary to ensure uniqueness and to provide information for Internet trouble shooting at all levels.

As a foundation for understanding constitutional norms, the language in RFC 2050 gives a number of historical indicators.

The notion of needs-based delegation, the topic of many subsequent operational debates, is rooted in the invocation of "operational needs" in the conservation norm. The conservation norm also references the prevention of stockpiling as a means to maximize the lifetime of the address space. Combining the censure of stockpiling[262] with operational needs gives rise to a norm that espouses delegation predicated on demand and historical—expected utilization. This ultimately evolved in to utilization criteria such as the 80% rule and HD-ratios to determine if previous delegations were utilized efficiently. A key element of the transfers debate questions mechanism and parameters that determine "efficient" utilization. Models of utilization based on this notion of conservation privilege demand and fairness. In contrast, recent transfers arguments have argued for a more traditional selection mechanism based exclusively on market price. As will be discussed at multiple points in Section 5.7, this transition disintermediates the RIRs role in utilization-based delegation and

[260]Each of the authors was a CEO or equivalent level of leadership in the existing RIRs. The exception is Geoff Huston, who has held other leadership roles in the broader Internet resource management community, has authored numerous technical RFCs, and whose work in understanding routing dynamics and experience in the RIR system lends credibility to both RFC 2050 and RFC 7020. Taken together, both RFCs are manifestations of operational experience and reputable, credible actors.

[261]The following are directly quoted from RFC 2050 (Hubbard et al., 1996), emphasis added.

[262]The censure of stockpiling later manifested by proxy in reproach of early brokers perceived as agents of organizations that would stockpile addresses as exhaustion became more a more pressing issues. A canonical comparison was with silver stockpiling.

illustrates tension with a background concept in both RFC 2050 and RFC 7020: uniqueness and accuracy.[263]

Routability speaks to both the dependency between delegation strategies and the stock of routes and the separation of number resource policy and routing operations. Hierarchical delegation of number rights can facilitate scalability. It can be incented by prefix delegation strategies, for instance renumbering into a contiguous block and reservation strategies discussed earlier. Aggregation also requires operational diligence, such as endogenizing aggregation costs as discussed in Section 2.2.2.2. CIDR improved conservation by facilitating "right-sized" delegations, but increased fragmentation. As such, it improves the utilization goal tacit in conservation, but creates tension with scalability goals. The routability norm also stresses that a delegation *does not guarantee* routability. In the terms of this work, because route provisioning is a bilateral agreement between adjacent networks, delegation of the basic origination bundles does not, moreover it cannot, guarantee routes for any prefix will be provisioned. The distinction between the two pools and the attendant guarantees, the attendant obligations engendered by their respective institutional rules, is fundamental to understanding precisely where the two are interdependent, and where they diverge.

Registration as a goal offers a number of foundations for registry use rights. The notion of a "public registry" is the source of broad access rights to a partially rival database. The latter of the two functions, "providing information for Internet trouble shooting at all levels" speaks to its role facilitating contact between those creating externalities and those experiencing those externalities. The notion of a public registry indicates anyone may be an entrant—the notion of this registry function as a public service will be revisited in Part III where the registries, along with the IXes and IPBLs, are discussed in terms of private authority.[264] RFC 2050 provides the basic definition of the registry, documentation of delegations. Uniqueness is paramount, although understated in RFC 2050 given most technical readers are expected to understand the value of uniqueness. RFC 7020 refines this assertion to shift the focus to accuracy of the registry function.

RFC 2050 explicitly notes tensions between these norms. It explicitly calls out the tension between conservation and routability discussed above. Discussions of routability have waxed and waned over the history of the RIR system; recently, the milestone of routing tables comprising 500,000 routes has been discussed in terms of certain common router models capability to handle that volume and speculation on further growth. More recently, conservation has become an important issue in discussions over how IPv4 depletion relates to IPv6 take-up and the appropriate

[263]To be precise, uniqueness is perhaps the absolute service of the registry, dependent on accuracy. RFC 7020 refines the registration norm to stress accuracy, especially in the context of IPv4 exhaustion.

[264]Private authority is one explanation of NRS institutions' ordering of the control plane as a common pool resource. That said, as has been alluded to in earlier discussions of authority, Part III will segue from the notion of a strictly independent private authority as a degenerate case to further build the case for relational authority as a better fit for explaining both internal function of the NRS and engagement in the global political arena.

policy response to the dwindling pool of IPv4 addresses in the various regions. A number of additional principles common in RIR practices are evident. One issue is needs-based allocation. The articulation of conservation contains reference to needs of operators and users. The IANA, having "authority over all number spaces used in the Internet," states that "IANA allocates parts of the Internet address space to regional IRs [RIRs] according to its established needs" (Hubbard et al., 1996, p. 3). Further, RFC 2050 states:

> Regional IRs are established *under the authority of the IANA*. This requires consensus *within* the Internet community of the region. A consensus of Internet Service Providers in that region may be necessary to fulfill that role. (Hubbard et al., 1996, p. 4, ephasis added)

Consider the emphasis added here. One reading, through a strict principal agent lens, presents the IANA as the principal and "consensus within" is merely part of the operational capacity supporting IRs' role as agents of the IANA. In the larger context of RIR operation, following the discussion of the delegation hierarchy in Section 5.2, the authority conferred here is rooted in the IANA's role in *evaluating* the process by which new RIRs are created and in delegating management and alienation rights to self-funded, self-directed RIRs. The mechanics of number allocation, both the most recent pre-exhaustion global policy and the post-exhaustion policy were both products of coordinated policy development within the RIRs—these delegation criteria and processes were not handed down from the IANA.[265] Moreover, creation of a new RIR is subject to requirements documented in ICP-2 (ICANN, 2001) and consent to join the MOU amongst the existing RIRs; both LACNIC and AFRINIC were created under this process.

Further note the notion of a needs-based criteria is also referenced. In terms of Figure 5-4, this references $L1$ delegations. This context reaffirms needs-based criteria is not just imposed by the RIRs, but is a system-wide notion of resource distribution that the RIRs adhere to as well as impose on their delegates.[266] Needs-based criteria has received criticism over the years. In the face of exhaustion it has been the subject of renewed discussion on the policy mailing lists, and has been left out of certain some policies altogether. The latter, the initial resource transfer policy in APNIC, was an interesting test of the greater community's support of needs-based criteria as a norm; this is discussed in Section 5.7.2 Other criteria include information such as projected demand and network architecture and topology information used to supplement utilization rates.

As will be developed in Section 5.7.2, although utilization rate is tightly coupled to needs-based criteria, it is used as both *a*) a parameter of needs-based evaluation at the time a delegation is being considered *b*) in audits confirming efficient

[265]These two policies are (ICANN, 2012b) on pre-depletion delegation to the RIRs and (ICANN, 2009) describing distribution of the last five
8's, one to each of .

[266]Global policy (ICANN, 2012b) describes the delegation window criteria for delegations from the now historical IANA pool to RIRs. Recall that, at the time of this writing, the only remaining pool of IPv4 rights available to the RIRs is the recovered pool.

utilization of all delegations to a particular actor. The latter follows the spirit of needs-basis as a notion of efficient resource distribution but can be decoupled from actual delegation practices. The distinction sets the stage for whether some utilization criteria will remain and in what form: *a*) which delegations will be beholden to that utilization criteria? *b*) will utilization criteria will be completely replaced with a market-based price mechanism? *c*) is their the potential for a hybrid that imposes only ex post utilization requirements (audit)?

One of the most important criteria established are the size of the RIRs and the consensus-based process by which they are created. In November of 1996, when RFC 2050 was written, there were three RIRs: the InterNIC in North America, the RIPE NCC in Europe, and APNIC in the Asia-Pacific. RIRs were expected to be of continental dimensions; local Internet registries (LIRs) were considered to be "usually of national dimensions." (Hubbard et al., 1996, p. 4). RIRs are currently of continental dimension. LIRs are typically listed as serving many different states (nations). An interesting structural difference amongst the RIRs are those that have a national Internet registry system in place or not. As discussed earlier, APNIC and LACNIC currently do have NIRs.

The NIR system is an interesting instance of refining RFC 2050. As per above, RFC 2050 assumes LIRs will be nation-state size registries; in practice, LIRs may serve multiple states within and/or across RIR boundaries. NIRs in both APNIC and LACNIC are scoped to national economies, often under the premise of providing more efficient local services. The address allocation process has evolved since the introduction of NIRs. Recall the discussion of KRNIC in the AP region's adaptation of NIR rules, discussed in Section 5.2. APNIC transitioned from a 'confederation' model where NIRs were allocated pools of addresses to be subsequently allocated to NIR members (APNIC, 2008, section 3.5) to one in which allocations are *approved* by the NIR but allocated from the regional pool by the RIR depending on the size of the block (2008, sections 3.2.1, 3.2.2). The change occurred because delegation to the NIRs and rapid consumption of those allocation windows created aggregation problems for rapidly growing organizations.

5.6.2 Collective-Choice Rules and Secondary Rights

Shifting gears from the IANA as a provenance of authority to community-based sources of authority, nearly all leadership articulate the principle that "the RIRs do what the community tells them." Within the framework of CPR rules governing a resource system, the consensus process is a form of collective choice rule making, in particular one that draws on the technical and operational expertise of the operational epistemic community to evaluate potential management practices and changes to management facilities. In terms of resource rights, the consensus process is the collective exercise of management, excludability, and alienability rights. Although stated in different terms, the IXes also express the notion of collective exercise of rights. In the IX vernacular, "the membership is the single stakeholder of the IX firm." In both cases, unlike simple models of ownership that parcelize both primary and secondary rights, secondary rights are governed by a collective

whose decision criteria are rooted in preserving *system* integrity rather than local maximization of parcel resources.

The most frequently referenced arena for engagement is the RIR's policy development process, followed closely by informal engagement. In fieldwork, hallway discussions of actors problems with the current system frequently came back to the asseveration that "if you don't like how the system works, write a policy to change it." The policy asseveration may take the form of encouragement when a good ideas is discussed informally in conversation; this mode will be discussed more in terms of problem identification in the next section. The policy asseveration may also be used to highlight actors that casting unwarranted aspersions on the current set of operational rules. Some are actors that critique the rules exclusively from the perspective of their value proposition, often because rules force that actor to endogenize particular costs.[267] Although much less frequent, other critiques may offer ideal yet unrealistic solutions. Policy asseveration in this case encourages the critic to write a policy in order to see how unrealistic that proposal is—the actor may see the failure modes or may suffer reputation loss for wasting the group's time with such an unrealistic proposal.

Across the five RIRs, policy development arena offers similar tools and processes for engagement, combinations of *a*) mailing lists, *b*) varying degrees of shepherding of proposals through the process by designated community members and RIR staff, *c*) asynchronous back channels, and *d*) face-to-face discussion at policy and WG meetings. Policy development is explicitly framed by the community as a bottom-up process: initial proposed policy drafts are written by one or more members of the community and introduced to the community via an e-mail list dedicated to that resource policy area.[268] Of the three NRS institutions discussed in this dissertation, the RIR PDP is the most explicitly and most similar to the IETF process.

In Section 3.2.5 identified four phases of the consensus process: *a*) problem identification, *b*) active consensus, *c*) passive consensus, and *d*) process review. Recall the notion of active versus passive consensus is present in Resnick's articulation described in Section 3.2.5, but not called out as such. The distinction is most evident in the RIR process. It has been developed from review of the policy development process documentation from each of the RIRs, discussion in RIR presentations, observation of member meetings, and interviews thus far. In contrast to the more open ended IETF processes, RIR PDP's place time limits on these phases, but can return to earlier phases, typically at the discretion of the shepherd(s).

[267]A recent instance of this occurred on RIPE's Anti-Abuse Working Group (AAWG) e-mail list in relation to proxy RIR services currently used by anti-abuse tools. These services were being discontinued because, aside from the single anti-abuse toolkit, there were only three other users of that service. The list moderator encouraged the participant to both write a policy proposal to correct the situation or propose an alternative solution, but none materialized.

[268]Instances are the Public Policy Mailing List (PPML) in ARIN or the various working group lists in the RIPE region.

5.6.2.1 Problem Identification

In Resnick's depiction of the IETF consensus process problem identification is an explicit process. Outside of the creation of specific task forces or a broad, issue-area specific working group such as one dedicated to RIR services, Internet exchanges, the registry database, measurement, or others, specific problem identification has historically been an implicit, informal component of the PDP process. RIR policy officers, RIR staff, and policy shepherds encourage discussion of potential problems on the policy development e-mail lists, but it is not a required point of entry in the PDP. That said, the policy proposal templates in the RIRs do require a problem statement. This has not always been the case. It is typical to presume that problem identification took place during the policy development asseveration. Although developed "outside" the policy development process, shepherds do evaluate the problem to determine if it is legitimately within the scope of RIR resource policy and in the case of those whose process occurs within WGs, within the mandate of the working group to which it has been submitted.[269]

There have been discussions on how to make the problem identification process a more explicit phase of the policy development process. For instance, at the Address Policy SIG meeting at APNIC 34 in Phnom Penh Cambodia, discussion of Randy Bush's "policy to end policy" (Bush, 2012) turned to a proposal by Dean Pemberton to introduce an explicit problem identification phase to the APNIC policy development process. The essence of the idea is to capture the spirit of problem discussions that occur in face-to-face discussions in a phase of the PDP process that occurred on the mailing lists. Very simply, actors could present just the problem statement to determine if their was, in Mattli and Woods terms, demand for a (regulatory) solution.[270] Rather than laying the burden of problem articulation and the drafting of a proposed solution on one actor, problem identification seeks to gain consensus that the problem is in fact a legitimate number management issue and is not merely a particularistic issue. This is in contrast to spending time discussing the solution to a non-problem in the active consensus phase. Further still, a problem identified by a number of RIR leadership in both shepherds collectives and on RIR boards is the need to further accelerate the policy development process—in the words of one actor, the policies need to "fail to reach consensus faster and more often" to allow for more feasible policies to emerge to solve the same problem.

5.6.2.2 Active Consensus

Given the problem and solution have been vetted by the shepherds collective, PDPs enter an active consensus phase. Active consensus means that a sufficient number

[269]In ARIN, the AC evaluates initial policy proposals for scope and technical merit. In the RIPE and APNIC regions the specific WG chairs evaluate policies based on the WG mandate.

[270]Regulatory is included in parens because "regulation" has a pejorative association in the RIR community. In Mattli and Woods work, it is nominally a set of rules that "order social and economic behavior," (Mattli & Woods, 2009b, p. 1). The parens are particularly applicable here given that, as noted earlier, the session in which Dean Pemberton proposed policy identification was the "final IP address policy proposal" (Bush, 2012).

of individuals have shown support for a policy proposal. Support may be through any of the officially documented communication vectors, but most typically via the policy e-mail list or in a public policy meeting, depending on the specific RIR.[271] In simple terms, active consensus implies there is sufficient support in the community for a policy based solution. In Mattli and Woods' terms, this implies demand for "regulation." This is also often a sanction by the community that the problem itself is valid; in Mattli and Woods' framework, there is latent demand for a solution. The ARIN PDP indicates need for policy may be "determined by a change in technology, a change in the operational environment of the Internet, or the *result of experience* of the implementation of the existing policy," (ARIN, 2009b, emphasis added).[272]

The latter, experience with implementation, has been articulated multiple times in other contexts—it provides further evidence of a (tacit) paradigm of policy experiments by an operational epistemic community. Multiple policy proposals have referenced either experience in operational environments or work done by RIR labs, in particular Geoff Huston's presentations on APNIC labs work.

Depending on the RIR, somewhere between active consensus and passive consensus, the RIR firm performs what is referred to in the RIPE community as an impact assessment. Given the RIR must actually implement the proposed policy, it has an incentive to evaluate how difficult the proposal will be to implement, the expected costs, implications for other elements of the policy corpus not considered in the proposal, legal implications, and other factors that may affect implementation. The RIPE NCC is the only RIR that explicitly acknowledges the impact analysis in the PDP. Other RIRs perform the analysis for internal use and share the analysis when appropriate. ARIN will do the analysis when requested on when the proposal seems likely to garner consensus, whichever comes first.

The RIRs have different perspectives on the use and affects of the impact analysis. In the RIPE region, interviews indicate that a cost-benefit analysis is part of the proposal evaluation process. The community may approve of a proposal, it may garner consensus in terms of improving the registry and having no ill effects on particular interests or not pandering to particular interests, but it may be quite expensive to implement. Given membership fees pay for the development, the rationale is that the community should be appraised of how their policy decisions affect current funds and the potential to cause increases in fees in the next fiscal year.

In contrast, other regions prefer not to share cost information. The argument behind this rationale is that a problem and solution should stand on the merits of its effects on the registry system. If the community believes a given service is nec-

[271]For instance, in APNIC and LACNIC, consensus is determined at the end of a session based on e-mail utterances for and against and a tally of support in the face-to-face meeting. In the RIPE region, support for or against only "counts" if it is submitted to the mailing list; it this is later evaluated by the RIPE shepherds collective. In ARIN, a combination of for and against is collected from e-mail list and tally's at policy sessions, but the actual decisions is not made immediately, but in one of the regularly scheduled AC meetings after a policy has been presented at least once.

[272]When the PDP was first reviewed, this version of the PDP was active. It has since been superseded by a new version that does not contain this exact quote.

essary, cost, within reason, should not sway consensus in one direction or another. Harkening back to technical evaluation in the operational epistemic community, obfuscation of costs may see to focus the discussion on technical issues. That said, the *operational* epistemic community is adept at costs-benefit decisions within their own organization and can bring that experience to bear on policy issues. In practice, a substitute for cost in the APNIC region is a time line for implementation. It is a proxy for effort that doesn't necessarily put a dollar sign (an Australian dollar sign in the case of APNIC) on a particular proposal.

In terms of principal agent relations, the impact analysis finds the RIR playing the role of the agent and the collective community that of the principal. In a conventional principal agent relationship, an evaluation such as the impact analysis is a point of substantive discretion. As per (Hawkins et al., 2006), a key reason for delegating is operational capability and domain expertise. Under these conditions, managing discretion is a monitoring costs, placing the principal in the position of acquiring the knowledge necessary to effectively monitor (and subsequently enforce) its goals. This is not only a monitoring problem, but also a credibility problem.

In the case of impact analyses in the RIR, it is more difficult to exercise this discretion. Network operators, especially those that participate in RIR policy development, have a deep understanding of registry management and, having used the registry extensively, have an understanding of its strengths and weaknesses. It is unlikely the RIR will use technical jargon to leverage its discretion. That said, the impact analysis also considers the broader implications of a policy, including those in the global policy arena. Evaluation of transfers and RPKI have both shown to be especially challenging for both RIR and community impact analysis. As will be discussed in Section 5.7.4, both the technical and political impact of RPKI have seen substantive debate and are an instance of a potential weakness in the current collective choice process.[273]

Policies that fail to reach consensus are not necessarily "complete failures." In a number of cases, a failure to reach consensus is an indicator of a gap in the community's understanding of an issue. In some cases, policy failures have led to the development of a task force to further investigate the issue. After multiple presentations to the RIPE community, RPKI policy failed to reach consensus and a task force was formed to further evaluate implementation issues and implications. Similarly, the aftermath of contention over the India NIR gave rise to a policy proposal to create a kind of GAC in the APNIC community. Given contention over this policy a task force was formed to evaluate potential compromises. The result was the PPAC as a SIG, a compromise between creating a "privileged" group especially for

[273]As will be developed in Chapter 8, consensus is currently the primary mechanism by which collective choice decisions are made, but is currently rooted in operational epistemic knowledge, in particular in the domain of number delegation management and routing dynamics. This instance of the consensus process does not yet have explicit mechanisms to evaluate, i.e. perform an impact analysis accounting for, the infrastructure or downstream economic and political implications of resource decisions and foreseeable outcomes. In effect, economic and political factors are a dimension of anticipation that both complements and is in tension with the mode of policy experimentation that characterizes the current consensus process.

governments and providing an open forum dedicated to public policy implications of resource policy and facilitating an introduction to the RIR as a resource policy arena for government actors.

5.6.2.3 Passive Consensus

Over the course of the active consensus process, a policy proposal may have experienced a number of major and/or minor revisions. "Easy" cases such as language cleanup or minor modifications—these rarely see substantive change. In contrast, "hard" proposals are typically instances of larger "hard" policy issues such as routing security, transfers, or anti-abuse. These cases push the boundaries of the operational epistemic community's knowledge base and capability to anticipate implications. These issues and others are discussed at length in Section 5.7.

These hard cases are those that warrant the most scrutiny, both as the proposals evolve, but also as a check that nothing was missed or slipped in during revisions during active consensus. For instance, a canonical instance is RPKI in the RIPE region. RPKI has been a contentious issue for a variety of reasons, one among them is the potential for what is framed in Section 5.2 as more immediate revocation powers in the delegation hierarchy. Malcolm Hutty brought this point up during the passive consensus phase of the RPKI policy proposal,[274] catalyzing a lively debate in the e-mail list. This debate is analyzed, in particular interpreting some of the arguments to highlight the difference between an operational externality and the potential for politically motivated strategic externalities. In the RIPE community, this debate contributed to a plenary between the resident community political analyst Malcolm Hutty and the architect of RPKI, Stephen Kent. The content and implications of this debate are discussed in Section 5.7.4, but the point here is that passive consensus is intended to ensure that uncertainty in hard cases receives due attention, that a policy is not simply rushed through as a result of debate fatigue.

Given passive consensus is an opportunity for community members to react to and challenge the proposal produced by active consensus, the character of consensus differs. Active consensus requires recognition of a problem and "active" support of a solution. In contrast, the absence of objections is sufficient for passive consensus[275] It is important to note that unanimity is sufficient, but not necessary. Mechanically, in multiple PDPs this is described as an opportunity for the community to react to the current proposal as a the product of the active consensus process. In the ARIN region, this occurs after the Advisory Council determines active consensus and moves the draft to Last Call. In the RIPE region, active consensus is at the end of the Review phase (consensus on the documents produced by the proposer and the WG chair and the impact analysis produced by RIPE NCC staff), passive consensus is at the end of the Last Call.

[274] See Section 5.7.4, the last call discussion is at (Hutty, 2011).

[275] An excellent discussion of what is here referred to as active and passive was presented by Steffan (2012). This is also an excellent instance of community leadership addressing questions about the policy development process.

Arguably active consensus is where most of the contention over a policy proposal plays out. Passive consensus is a check on substance of that outcome. It provides the opportunity for those that may not have had the time to follow each incremental change to weigh in on the productive of active consensus. It serves as a consistency check. It also serves to, as noted above, to ensure points of uncertainty are not overlooked. A review of the process itself is the last step of the RIRs' policy development processes.

5.6.2.4 Process Review

Active and passive consensus are called by the shepherds of that particular proposal. In the case of ARIN, two shepherds are typically assigned to a given proposal. In the case of the other RIRs, shepherds are the WG or SIG chairs. Again invoking the language of principle agent, this is a point of discretion. The review phase is a check on that discretion.

In all but the RIPE region, the membership elected board of the RIR reviews policy proposals that have reached consensus. In the RIPE region, although issue-specific WG chairs determine consensus for policies developed in their WG, the collective of WG chairs evaluates whether the PDP was followed. In all cases, the objective of the review *is not* to evaluate the substance of the policy. Rather, the objective is to review *a*) the RIR's PDP was followed and *b*) the policy does not create undue legal risk or obligations for the RIR. For instance, in the ARIN region, the Board of Trustees of ARIN evaluates the draft policy in terms of "fiduciary risk, liability risk, conformity to law, development in accordance with the ARIN PDP, and adherence to the ARIN Articles of Incorporation or Bylaws," (ARIN, 2009b). Review is intended to be final checks on the consensus decision making process and feasibility of the policy before moving to implementation.[276]

5.6.3 Operational Rules

Recall the definition offered by Ostrom:

> *Operational rules* directly affect the day-to-day decisions made by appropriators concerning when, where, and how to withdraw resource units, who should monitor the actions of others and how, what information must be exchanged or withheld, and what rewards or sanctions will be assigned to different combinations of actions and outcomes. (E. Ostrom, 1990, p. 52)

Operational rules are implementations of constitutional rules that balance conflicts amongst those norms. Resource policies are the operational rules developed by the consensus-based collective choice rules described in the previous section (5.6.2).

[276]Recall the discussion of the NRO evaluation of global policy in Section 5.4.5. The evaluation of global policy plays a similar role to review in the individual RIR: the global policy development process was followed and the process for a global policy was followed in each RIR.

Collective choice processes are the means by which the operational epistemic community navigates the compromise space.

As per Section 3.3, rules create rights and obligations amongst actors. Meaningful bundles of rights are embedded in the policy corpus. The policy corpus is structured around types of actors and the corresponding infrastructure uses of number resources.[277] RIRs primary remit is to manage the stock of number resources. That said, delegations strategies' implications on route provisioning and appropriation do have implications on downstream uses.

For the most part, outcomes of the collective choice process are relatively small, function-specific deltas. In the community vernacular, deltas are initially policy proposals, then a draft, and finally are implemented. Implementation means that the policy is incorporated into the policy corpus. Some policies engender changes to resource management facilities, such as the registry or RIR processes, as part of policy implementation. In terms of documentation itself, the policy corpus may be a single policy manual[278] or it may be a collection of "current policies" grouped in coarse grain policy topics such as IPv4 allocation and assignment, IPv6 allocation and assignment, ASN assignment, transfers, abuse, global policies, etc.[279]

Function-specific deltas are the norm for updating the policy corpus. One factor in this instances of incrementalism is the engineering culture and the practice of problem decomposition. As noted in the previous section on collective choice rules, the policy proposal process strongly encourages, if not requires, a clear problem statement to initiate a policy development process. As deltas,[280] consensus process outcomes modify one or more specific elements of the policy corpus. Given the function-specific aspect, these are often narrowly scoped to a particularistic issue or use-case.

Use-cases developed in the following sections map to what Ostrom referred to as meaningful bundles of rights. Policies may address, for instance, types of parameters that are a component of many types of bundles. Specific bundles include utilization criteria, comprised of general parameters. Specific bundles correspond to policies for MDN, infrastructure policy, or various aspects of registry access.

Operational rules describe rights to entry (access, Section 3.4.1) and rights to withdrawal (delegation for subsequent use, Section 3.4.2). Management rules describe who can make what changes to the facilities that enhance the resource sys-

[277]Recall from Section 3.1 the difference between direct uses of number resources to implement infrastructure services rather than indirect, downstream, often public facing services.

[278]Number Resource Policy Manual (NRPM) (ARIN, 2014a) in ARIN or the LACNIC Policy Manual (LACNIC, 2014f). Both of these documents are accompanied by a change log depicting a high-level view of the evolution of the policy corpus.

[279]The RIPE NCC corpus comprises approximately 21 current policy documents (RIPE, 2014b). The NCC is currently engaged in a policy cleanup effort (RIPE NCC, 2014c). The APNIC corpus comprises 9 documents: 6 address number assignment and allocation, 1 addresses historical (legacy) resources in the APNIC WHOIS database, and 2 address National Internet Registry (NIR) policies (APNIC, 2014b). AFRINIC comprises 19 policy documents (AFRINIC, 2014b).

[280]In computer science vernacular, products of the policy development processes maybe seen as diffs, indicating what text of a given policy should be removed, added, and/or replaced. Some proposals are, not surprisingly, structured like a diff.

tem. In the RIR system, management rules shape how the registry and supporting facilities are managed. Policies may shape the structure of the resource. Policies may also set the ranges of parameters that affect integrity of the current structure and the durability of rights to set parameters.

Section 5.3.1 framed the registry as jointly managed. Treating the registry itself as a (management) facility, management rights at the NRS level confer limited management rights to RIR members. Management rights often address *what* types of contact information must be stored in the registry (structural) and *who* can modify that information (parameterization). In contrast, audits shape the integrity of utilization. Alienation allows limited scope of delegation of parameter setting rights, but the obligation binds to the LIR. More nuanced still, bundles for managing registry data often involve delegating (alienability) rights. These bundles facilitate modifying various types of contact information and maintaining distributed information systems, especially as it relates to the hierarchical number resource assignments by IRs on to end user organizations.

The stock of number resources is a common resource jointly owned (in the sense of types of rights in Section 3.4) by the collective membership of the five RIRs. Rules shaping access and withdrawal are developed through collective choice processes at the regional and global level. The facilities for managing these rights bundles—monitoring and enforcing operational rules—are facilities provisioned by (funded by) the RIR membership. Each registry is a resource jointly owned by that particular community and whose structure, as well as utilization rights bundles, are also managed through the consensus process.[281] Bundles of rights to management facilities are different than rights to the stock itself but are managed through the same collective choice rules.

Alienation rights that figure most prominently in the RIR system deal with delegation transfer policy. Given the increasing scarcity of IPv4 addresses, transfers will soon be the *only* mechanism for delegation number rights.[282] In contrast to delegation discussed thus far in the context of Figure 5-4, transfers are delegations of number rights from one EU to another. Transfers may happen within a region or may be inter-region. All regions have an intra-region transfer policy, at the minimum in the form of mergers and acquisitions policy. A rights framing is especially important to understanding transfers. Transfers are framed as one of a family of bundles comprising delegation rights. Such a framing focuses the analysis on the criteria under which delegation rights may be exercised and by which actors. This avoids invoking conventional notions of property and focuses on how those differences affect the integrity of the system. As will be elaborated in Section 5.7.3, there has been substantive debate over the precise property rights of IPv4 addresses,

[281] For instance, ARIN has made it clear that the registry is the property of the RIR firm. In effect, it has established itself as the monitor and enforcer of authoritative secondary rights of management, exclusion, and alienation.

[282] In a very strict sense, at the point of depletion, all previously unallocated prefixes will have been allocated to some actor. Given numbers are, modulo reputation, reusable, actors only means to acquire an initial block of IPv4 addresses or subsequent is to transfer the (meaningful) bundle of rights to use numbers to uniquely identify hosts and origination.

whether they are "owned" or simply held in stewardship, and the implications for varying levels of "friction" in the transfer market.

Policy proposals may have a number of effects on the policy corpus. Each of the RIRs indicates the policy corpus is not static, but is constantly changing in response to industry demands, changes in downstream demand, and changes in technology. A more recent context to consider is the political context, especially as it relates to security and RPKI. The simplest, degenerate policy delta is a clarification that may *a*) corrects language for consistency and clarity, *b*) removes redundancy from incremental policy changes,[283] *c*) re-orders existing obligations to avoid and/or lower transaction costs. The latter is a subtle change. Rights bundles themselves are not changed; rather than the order in which required elements of implementation and conformance are performed is changed to avoid inefficiencies.[284]

Operational rules specify the conditions under which a bundle of rights may be conferred and the conditions under which they may be exercised.[285] The next simplest delta updates parameters of those conditions to reflect the current industry context.[286] Parameters changes alter the range of application of a condition, but does not remove the condition itself. Types of parameter changes include the types of actors affected and/or magnitude of effects. Often these reflect changes in utilization trends, in effect changing the thresholds for evaluating demand and ultimately "need." Particular instances include thresholds for utilization rate, allocation window sizes, or minimum assignment requirements for initial delegations.

In contrast to parameter changes, policy proposals may nominate structural changes. These changes may introduce new criteria, redefine existing criteria, or remove particular criteria. Recall management rights not only determine appropriate rates (thresholds), but also structure. These changes would alter RIR firms' operational practices and subsequently the registry as a resource management facility. Changes in appropriation rates can be likened to changing fishing quotas; irrigation system withdrawal rates or schedules; or timber withdrawal from commonly managed forests. Changes in structure are likened to creating barriers to foster spawning grounds; altering the physical shoreline to shape access; (jointly) building new irrigation channels or dams; or (jointly) building roads that provision access to forest areas. Structural changes enhance or diminish access and appropriation rights. While the structure and parameter distinction is conceptually useful,

[283]Individual policy proposals often attempt to avoid redundancy, but occasionally it does slip through.

[284]For instance, see ARIN Policy 2005-7 (Seastrom, 2006) for a simple instance.

[285]Again note that the scope of these conditions is limited to nominal infrastructure utilization and topology. These are not conditions on downstream use.

[286]For instance, ASN policies in multiple regions recognize the slower than expected uptake of 4-byte ASNs (ARIN, 2009d). Another instance is various changes in IPv6 allocation conditions, such as updating the definition of efficient utilization from an HD-Ratio of 0.80 to 0.94 to prolong the lifetime of IPv6 (Dul, 2006), drawing on research by Geoff Huston at APNIC Labs. Later, HD-Ratio was replaced by the proportion of /48's. This latter does not change the allocation of rights in terms of kinds of actors that have allocation rights nor does it change the existence of a threshold that delineates what it means to have efficiently utilized existing resource allocations. Rather, it does change the definition of that threshold based on community experience.

as noted multiple times thus far, operational rules often interleave these. For instance, one of the most controversial of structural changes has been the debate over needs-based criteria for IPv4 allocation under transfer policies.[287]

Policy proposals may include one of these types of policy changes or may include a number of these to address a coherent resource policy issue. A coherent issue corresponds to a meaningful bundle of rights. As may be obvious, the scope of the issue determines the scope of the rights. For instance, transfers, as an instance of the larger family of delegation rights, has a broad scope, affecting a variety of bundles. In contrast, policies focusing on particular topologies, such as IXes or MDNs (discussed below) are parameterizations of rights that may be conferred and exercised under precise conditions.[288] Each community also has policy related to critical infrastructure, which may include IXes. In some cases, these policies refer to micro-allocations or small allocations necessary for running Internet exchanges, DNS servers, and other networks that provide services that contribute to Internet operations but do not require large swaths of number resources and do not necessarily follow growth trends applicable to those serving end users (indirect users of the NRS). Introducing critical infrastructure policy specifies the use-case, a set of conditions, and how these conditions are to be evaluated by registry staff in evaluating allocation requests.

5.6.3.1 Types of Delegation

Number resource delegation is a parcelization of number rights, in particular origination rights. In the RIR vernacular, there are three general bundles of rights in what is referred to here as the family of delegation bundles: allocation, assignment, and transfers. Allocation "means to distribute address space to IRs for the purpose of subsequent distribution" by those IRs.[289] Assignment means:

> to delegate address space to an ISP or end-user, for specific use within the Internet infrastructure they operate. Assignments must only be made for specific purposes documented by specific organizations and are not to be sub-assigned to other parties.[290]

[287] A recent instance is Mueller's crusade to lower the barriers to transfers in the ARIN region; a recent discussion can be seen at (ARIN-PPML, 2014).

[288] Each RIR has an exception in its policy corpus for delegation of number rights to an IX. Criteria typically include number of actors interconnected (usually 2 or 3, a point of contention in itself) and a public interconnection policy.

[289] This definition is derived from the ARIN NRPM (ARIN, 2014a, Section 2.5) and the LACNIC policy manual (LACNIC, 2014f, Section 1.8). APNIC describes allocation and assignmenyt in terms of delegations and assignments (APNIC, 2011, Section 3).

[290] This definition is also derived from ARIN NRPM (ARIN, 2014a, Section 2.5) and LACNIC (LACNIC, 2014f, Section 1.9) policy statements. Note the use of the term delegation in this definition and that APNIC uses it in a similar manner to allocation. The use of delegation in principal-agent theory confuses this language. Here, delegation is used in the nominal sense, to confer rights. Allocation is used typically used to mean authoritatively delegates (or confers) rights and obligations related to that block. Allocation is used more broadly as well as in RFC 2050 (Hubbard et al., 1996). *Delegation* will be used here to make the reference to the distribution of rights explicit.

Allocation and assignment processes inevitably fragment (parcelize) address space, hierarchically delegating rights to origination from the IANA through the RIRs on to NIRs, LIRs, and ultimately end-users.

Recall from Section 2.1 that the IANA, as the protocol producer, created the IPv4 pool and the original 2-byte ASN pool. Under pre-exhaustion conditions, an RIR requested blocks of number resources from the IANA based on its current stock and delegation (allocation) rate.[291] The global IANA allocation policies designate a /8 as the resource unit; RIRs are allocated IPv4 number rights in /8 units along the blue $L1$ delegation paths in Figure 5-4. $L1$ allocation bundles do not confer appropriation rights to either the RIRs or NIRs. The IANA delegates limited management and alienation rights to the RIRs. This is different than conferring the right to exercise appropriation rights, i.e. for the RIRs or NIRs to use (originate) any numbers in their stock. The RIR collective, through the consensus process, has the right to exercise management rights. The IANA and the RIR membership have conferred the RIR firm with the right to implement management rights defined through the consensus process.

RIRs and NIRs subsequently, in their role implementing management and alienability rights, delegate use and limited alienability rights to LIRs that may then assign those addresses to their own hosts, sub-allocate subsets of the allocated prefix (limited transfer), or delegate those numbers to an actor exclusively for utilization (an assignment). Once a block is assigned to an actor or set of hosts, one may say that origination rights held will be exercised. In terms of SimpleNet in Figure 2-1, most of those network actors would be considered LIRs and end-users. Consider Figure 5-5. Allocation means LIR_A may further delegate allocation bundles for subsets of its number allocation or may delegate an assignment bundle limited to the basic use bundle.

In terms of allocation, the resource policy corpora comprise delegation rules for different infrastructure uses. Uses *do not* mean uses such as whether network resources facilitate communication of social or political issues, whether resources are used for financial transactions, support government activities or private activities, host pornography versus kitten videos, or any other such *downstream* uses. Rather, the uses identified in resource policy corpora are characterize use that supports a non-discriminatory infrastructure and the implications for management of that resource stock. In particular, policy must balance allocating resources necessary to satisfy members' demand for resources, ensuring uniqueness, potential for aggregation in route announcements, and conservation of a scarce resource.

Policy corpora call out a number of different infrastructure uses that have given rise to differentiated delegation criteria. The following lists some of these uses. Note these are not necessarily mutually exclusive; instances will be given where appropriate for clarity.

[291] The delegation rate is a generic term that captures any the rate of number rights transfer from one stock to another. Based on context and the bundles in play, this may be refined to an allocation rate, a utilization or assignment rate, a transfer rate, or a revocation rate. The different types will be disambiguated where the context is not sufficient to distinguish them.

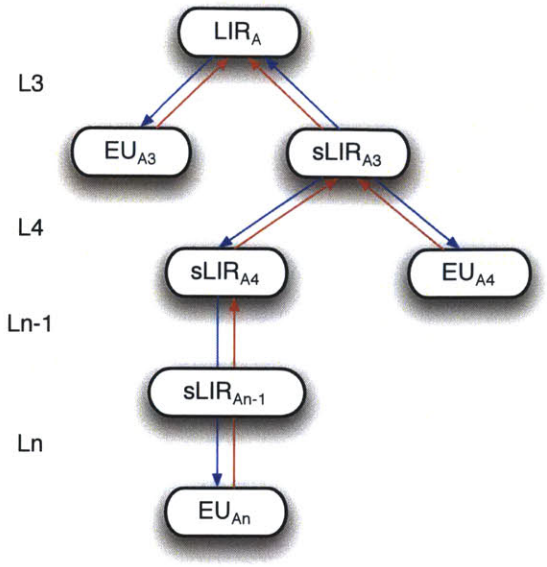

Figure 5-5: The LIR hierarchy elaborates the number delegation hierarchy depicted in Figure 5-4. sLIR are sub-LIRs, networks that are conferred allocation rights for a subset of LIR_A's delegation. It is important to note that that links have the same semantics as those in Figure 5-4 but are enforced by bilateral contracts—there is no direct agreement with an RIR or NIR in this scenario. Rather, it highlights that sub-allocations may occur "on the authority of" an LIR, but within the RIR system, responsibility for allocation and utilization rates and efficiency fall on LIR_A. In this scenario, sLIRs and EU organizations may be firms independent of the LIR or my be subdivisions of that LIR.

Singlehomed versus Multihomed Networks Multihoming is one means to ensure redundancy of connectivity by diversifying interconnection. In some delegation rules, singlehoming has a higher minimum allocation size than multihomed organizations. This is an instance of difference in criteria parameters (not structure). Often, if the singlehomed actor does not meet the minimum demand, especially when requesting an initial assignment, it is required to continue to request resources from its upstream provider. For instance, in the ARIN region, initial and subsequent allocation is predicated on part on whether a network is single-homed or multi-homed, single-homed having a higher minimum before having to request space from their upstream provider. LACNIC differentiates between single- and multi-homed at initial allocation, but not at additional allocation. APNIC does not distinguish between the two. One premise of delegation criteria rooted in demonstrable minimum existing utilization is to ensure scarce IPv4 resources are delegated to networks with sufficient investment in their network.[292] Multihoming has been a historical

[292]The idea of network investment as a criteria is not limited to number delegation. Peering arrangements often have requirements that networks are present at least n colocation facilities over m geographic regions, a proxy indicator of network investment. Early IX membership criteria, in particular at the LINX, limited membership to actors that had capital investment in infrastructure, not merely a bank of modems in someone's garage. As will be developed in Chapter 6, IXes have become a means for small- and medium-sized networks to expand the scope of colocation and geographic interconnection options.

signal of investment, in particular in redundancy. Sufficient investment to reach the higher delegation threshold may be seen as a proxy for investment and a credible signal of expected utilization (demand).

Community Networks Within the ARIN policy corpus community networks are distinguished as serving rural regions and staffed by 100% of volunteers. ARIN currently has an IPv6 direct assignment policy for community networks whose thresholds require 100 existing users with 200 projected in the next year.

Micro-allocations and Critical Infrastructure All RIRs have a policy for allocating number resources specifically for what is labeled critical infrastructure. Critical infrastructure comprises *a*) IXes; *b*) network access points (NAPs);[293] *c*) core DNS service providers such as ICANN-sanctioned root, gTLD, and ccTLD operators; and *d*) IANA, NIRs, and RIRs. A common norm amongst infrastructure management organizations is neutral management of facilities, here referred to as resource management facilities. IX platforms,[294] interconnection platforms at NAPs,[295] core DNS servers, and RIR resources are all elements of non-discriminatory[296] Internet operations.

IXes are often distinguished in policy by either particular conditions within the micro-allocation policy or in its own policy. For instance, under ARIN's micro-allocation policy:

> Exchange point operators must provide justification for the allocation, including: connection policy, location, other participants (minimum of two total), ASN, and contact information.[297]

APNIC's micro-allocation policy includes the qualification that it is at the discretion of the IX whether the block resulting from the micro-allocation is advertised by participants outside the IX fabric.

Constitutional norms of neutrality in the RIR and IX systems apply to notions of administrative neutrality of facilities *and* number resources used to ensure non-discriminatory Internet operations. Delegating number resources directly to the infrastructure facilities manager eliminates potential dependence on, and subsequent influence by, upstream providers with an interest in limiting or altering the character of access and utilization of critical resources.[298]

[293] NAPs can be either colocation facilities or a synonym for IXes. The term is included here for completeness.

[294] Neutrality is a well-developed constitutional norm amongst IXes. Section 6.4.1.2

[295] In many parts of the world, NAP is a synonym for IX. This does not apply to the historical NAPs that were created after NSF transition of Internet backbone to private actors.

[296] Non-discriminatory as per Section 3.1.

[297] Quote taken from (ARIN, 2014a, Section 4.4). Note the ARIN policy indicates an IX comprises a minimum of two participants; other policies indicate an IX comprises three or more. As will be elaborated in Chapter 6, this work considers an IX to require three or more participants.

[298] Note the list above indicates core DNS service providers. Core DNS services are authoritative. DNS servers maintained by particular networks refer to core DNS servers.

One trade-off engendered here is that small allocations that decrease overall routability and neutrality of critical resources. Given the nature of the resources and the value the community places on neutrality, an exception is made. Fragmentation is further limited by allocating micro-allocations and micro-assignments from a common block or set of blocks.

Micro-allocations may also be used for *internal* infrastructure.

Multiple Discrete Networks Private networks are not necessarily concentrated in a single geographic area or may be required by external factors, such as regulatory restrictions on data transmission, from maintaining a single discrete network. These multiple discrete networks (MDNs), maintained by the same organization, may be exposed to different levels of demand. As a policy, MDNs parcelize existing allocations in a way that allows organizations to request an allocation for one of its particular discrete networks. Allocation criteria becomes a balance of evaluating overall utilization *and* per network utilization. ARIN has historically had an MDN policy for IPv4 that required

- utilization of $> 50\%$ of the last block allocated,
- utilization of $> 50\%$ of overall allocated space,
- additional space granted will not be used for a discrete network unless *a*) each block allocated to that network has utilization $> 80\%$ *and b*) $> 80\%$ utilization over all blocks allocated to that discrete network

MDNs are an interesting case of the broad notion of utilization, parameterized to better fit MDN topology.[299] This illustrates that criteria are not bound to particular operationalizations, but rather select amongst, i.e. are parameterized by operationalizations that the community decides best fits the scenario.

Web Hosting Firms may find that the operational costs of in-house server hosting is not cost effective. These firms may outsource server hosting functions to a hosting services company. Hosting services firms provide combinations of space, power, cooling, security, and, in some cases, physical and virtual hosts. Hosting services as a set of functions has grown from early outsourcing of server hosting to ISP hosting websites to data centers specializing in colocation and hosting to cloud architectures offering clusters of virtual machines.

An early concern in the numbers system regarding web hosting was using multiple IPv4 number to reference the same machine. In terms of both efficient and conservative usage, this seemed a potential waste of scarce resources. As such, web hosting as a justification for an IPv4 allocations has been discouraged, but not eliminated. For instance, the LACNIC policy manual indicates that "LACNIC shall consider expectations where applications require the use

[299] In this case the criteria itself is some operationalization of utilization. In the simple form, utilization is the simple percentage of numbers assigned out of total numbers delegated. The HD-ratio is a somewhat more sophisticated form: $\frac{log(a)}{log(d)}$ where a is total number assigned and d is total numbers delegated. MDN utilization builds on simple utilization to

of web hosting base on IPv4 addresses, *which must be duly described and justified*," (LACNIC, 2014f, Section 2.3.2.8). This arguably raises the bar for this particular number use, encouraging the use of alternatives that yield more conservative number resource requirements. Across the RIRs, name-based web hosting (as opposed to number-based) is strongly encouraged (NRO, 2014d, Section 2.5.4).

Facilitating IPv6 Deployment IPv6 deployment has not taken off at nearly the rate desired by the Internet community. Across resource policy corpora, incentives for IPv6 have been introduced. In a number of RIRs, recent IPv4 allocations come with an IPv6 allocation.[300] These policies relax the criteria for initial IPv6 allocation (appropriation) and reduce transaction costs.

IPv6 deployment is not a simple switch, nor is IPv6 simply a bigger address space. Rather, it requires gradual deployment of infrastructure and learning within organizations that may have limited resources to invest deploying a technology that has questionable *immediate* benefits. Delegation incentives above are one mechanism for encouraging deployment and ultimately transitioning away from the nearly depleted IPv4 pool. IPv6 training by the RIRs is another mechanism.

A number of technical mechanisms exist for deploying IPv6 while also facilitating connectivity to IPv4 hosts. The RIPE NCC provides an overview at (RIPE NCC, 2013b). Although there has been substantive work encouraging actors to move to IPv6, a number will not transition until absolutely necessary. Moreover, many transition technologies (as well as infrastructure such as dual stack DNS servers) require IPv4 and IPv6 resources. In addition to encouraging deployment, some RIRs have reserved a block of IPv4 numbers specifically for use in later IPv6 transition efforts.

The conditions for allocation follow common allocation structures but are parameterized for this particular use:

1. applicant has not received resources under this policy in the preceding six months;
2. previous allocations/assignments under this policy must continue to meet the justification requirements of this policy;
3. previous allocations/assignments under this policy must meet the utilization requirements of end user assignments;
4. applicant must demonstrate that no other allocations or assignments will meet this need;
5. on subsequent allocation under this policy, ARIN staff may require applicants to renumber out of previously allocated/assigned space under this policy in order to minimize non-contiguous allocations.[301]

[300] ARIN and LACNIC have implemented this policy. In effect, it relaxes the conditions for an initial allocation (appropriation) of IPv6 resources.

[301] Adapted from the ARIN NRPM (ARIN, 2014a, Section 4.10).

That said, these criteria are instances of common criteria parameterized for specific use.

Consider generic forms of these:

1. Efficient utilization based on a specific operationalization of utilization: the "80% rule," an HD-Ratio, or some topology specific operationalization such as used for MDN.
2. Supply justification period balances conservation with transaction costs imposed on both the RIR and the applicant.
3. Subsequent allocations require earlier allocations continue to meet utilization requirements. Utilization requirements are not satisfied once, but, rather, are continuously evaluated. This is a key distinction that also speaks to the difference between conventional notions of "freehold" ownership and stewardship in the RIR system. Here, the generic operational rule that requires earlier allocations continue to be efficiently utilized highlights that stewardship means, in part, a member is continuing to ensure that resource is being utilized. If it is not utilized, management rights in various corpora confer the right of the RIR to rescind a delegation (reclaim, or revoke an allocation), returning underutilized resources to a common pool.[302]
4. Existing allocations do not or cannot meet the needs expressed under the current allocation request. For instance, specialized MDN criteria are an instance, where individual discrete networks may be susceptible to different levels of demand. In other cases, micro-allocations for infrastructure may be warranted to ensure effective separation of management networks from customers networks. IPv6 transition is yet another.
5. Allocations may require renumbering. In a number of cases the RIR conditions allocation on the willingness of a network to renumber, i.e. to accept contiguous block that is equal to the aggregate of previous allocations and the requested allocation. This reduces potential for fragmentation in the global routing table at the cost of renumbering for the member. In the case of initial allocations, the existing addresses are often

[302]Each RIR has a policy indicating allocations remain valid as long as original criteria remain valid. According to the NRO (2014d, Section 1.3.3), AFRINIC and the RIPE NCC do not actively recover unused space unless the organization closes. In contrast, APNIC and LACNIC have policies for actively auditing members to recover unused allocations. For instance, the LACNIC Policy manual states:

> If there is evidence to suggest that the assigned/allocated resources are possibly not being utilized or are being incorrectly utilized, LACNIC shall verify with the organization that received the assignment/allocation whether the resource is being properly utilized. (LACNIC, 2014f, Section 7, Resource Recovery)

The LACNIC Manual goes on to describe various mechanisms for confirming evidence of underutilization, including "[l]ack of visibility of the resource on the global routing table," and unauthorized transfers. ARIN defines this criteria in terms of compliance, re-evaluating the existing utilization and returning resources "as needed to bring [members] into (or reasonably close to) compliance" with policy (ARIN, 2014a).

returned to the upstream provider and the applicant renumbers from its own allocation.

Temporary and Experimental Allocations Temporary or experimental allocations are special purpose assignments to be used for research or experimental activity. In general these require public documentation of the experiment from a known source. For instance, LACNIC requires experimental allocations have an accompanying IETF RFC designated as experimental (LACNIC, 2014f, Section 10). ARIN stresses that experimental use is temporary and that it is to be renewed on a yearly basis. Experimental space is further limited to non-commercial use and a public report of the results is expected. As a mode of allocation, experimental use represents, in terms of the diversity of uses once allocated, a very narrowly scoped set of downstream uses.

The number rights described thus far are appropriation rights, largely allocations. Direct assignments are a more narrow bundle comprising only appropriation rights—they do not comprise the alienability rights necessary to subsequently confer appropriation rights to others.[303] The definition of an allocation is re-entrant—actors delegated number resources do hold alienation rights but do not necessarily hold the basic appropriation bundle. Allocation is a family of bundles that, at minimum, comprise alienation rights but not always the right to exercise the basic appropriation bundles. The minimum bundle sufficient for allocation is the right to further delegate (alienate) and the right to confer, but not exercise, the basic appropriation bundle.

An $L1$ delegation is a minimum allocation bundle. It confers on the RIR the right to further parcelize (fragment) the /8's for which it has been delegated minimum allocation rights. The RIR can further delegate minimum allocation bundles. It can also confer on those recipients the right to exercise the basic appropriation bundle. An $L2$ delegation to an LIR comprising the minimum allocation bundle and the basic appropriation bundle will be referred to as an appropriable allocation. LIRs can further parcelize appropriable allocation bundles, such as in Figure 5-5. Such customers, sub-LIRs (sLIRs), may then subsequently further sub-allocate or assign those resources. An assignment occurs when the bundle delegated by an actor holding an appropriable allocation bundle delegates *only* the basic appropriation bundle, but *not any* alienation rights.

In terms of utilization and withdrawal, the two types of allocations are more accurately a delegations of rights *recorded* in the registry. The registry confirms these rights have been conferred, but the registry *documents* reservation and utilization, it is not an actual utilization. The resource rights framing developed in Section 3.5 highlights that bundles of (property) rights assure a claim to the value derived from

[303]The earlier definition of an assignment refers to the notion of a sub-assignment. Defining assignment as a bundle of rights that exclusively confers appropriation rights, in particular the basic appropriation bundle and no other rights, the notion of a sub-assignment is a clear violation of the obligation not to further alienate appropriation rights. More specifically, a sub-assignment would be a claim that an actor holds rights it has not been conferred by the upstream IR.

Bundle	Rights	Limitations
Basic Appropriation or *Assignment*	Origination and numbering	
Minimum Allocation	Alienation (subsequent delegation of minimum allocation bundles), rights to fragment (management), and rights to confer and rescind basic appropriation rights (management)	Does not include rights to exercise basic appropriation rights—absence of these rights in this bundle does not preclude those rights a larger bundle
Limited Allocation	Limited alienation, rights to fragment (management), and rights to confer and rescind basic appropriation rights (management)	Alienation and other rights limited to RIR policy in effect at time of allocation
Legacy Allocation	Minimum allocation bundle and basic appropriation bundle	Alienation is unlimited absent a contract with an RIR
Direct Appropriable Allocation	Limited allocation bundle and basic appropriation bundle	LIR has contractual relationship with RIR; LIR is responsible for registry maintenance
Indirect Appropriable Allocation	Limited allocation bundle and basic appropriation bundle	Contractual chain of revocation and utilization reporting back to LIR

Table 5.2: Each bundle is what Ostrom refers to as a meaningful bundle. It is the set of rights that facilitate exercising a claim on some valuable use of a resource unit or set of resource units. In each of the bundles above, the tacit scope is the subset of addresses governed by that bundle.

some use. From the LIR's perspective, the number resource is not utilized until it is assigned to an End User.

5.6.3.2 Registry Management

The RIR manages multiple resources that contribute to NRS integrity. The previous section documented the operational rules that shape number rights delegations, identifying meaningful bundles of rights that help explain precisely who holds which rights. In terms of those bundles, the stock of number resource rights managed by the RIR is the set of resource rights delegated to that RIR in minimum allocation bundles by the IANA ($L1$ delegations in Figure 5-4). Subsequent delegations from the RIR's stock, $L2$–Ln delegations depicted in Figures 5-4 and 5-5 are rooted in the rights discussed in Table 5.2. These subsequent delegations are documented in the registry.

In the RIR system, the registry itself is the resource management facility. As per Section 5.3.1, the registry as a service is implemented and maintained by the RIR firm. The *content* of the registry, documentation of delegations, is jointly produced by the RIR firm, the membership, and customers of the membership. Operational rules discussed in the previous section largely focus on number delegation. A subset of the operational rules created through the consensus process define rules of access and use for the registry—namely which actors may access the registry and which actors may modify which records.

Section 5.6.1 provided the constitutional basis for a registry: uniqueness. Simply recording to whom which prefixes are delegated and whether that delegation is a limited allocation or an assignment is the most basic function of the registry. The basic function documents which organizations are responsible for which number resources. The registry also stores subsequent delegations by LIRs and assignment information. Assignment information is an indicator of utilization[304] used to evaluate number delegation requests and confirm utilization obligations.

Contact information stored in the registry facilitates identifying which organization is responsible for externalities originating from a particular prefix. Various forms of contact information are maintained in the registry. Some contact information is required, other is optional. In particular, contact information for a network's NOC, the registry entry maintainer, and more recently an abuse contact are all used to track down externalities. Rights to modify the registry in order to maintain assignment and contact information generally follow the delegation structures described in Figures 5-4 and 5-5. The registry comprises documentation of $L2$–Ln delegations. $L1$ delegations are maintained in a simple IANA maintained registry that lists /8's delegated to either legacy holders or RIRs

Documentation of $L2$ delegations are maintained by LIR staff. When an $L2$ delegation is conferred and documented, subsequent allocations ($L3 \ldots Ln$) are within the authority of the RIR member.[305] In some RIRs, the assignment window deter-

[304] Routing tables from sources like RouteViews have been used to supplement and validate assignment information.

[305] For instance, APNIC's documentation indicates that

mines the size of allocations or assignments the LIR can make without consulting the RIR.[306] Once an $\geq L3$ delegation is made, the member is still responsible for maintaining assignment information in the registry. In the early days, registry updates occurred via e-mail templates. More recently, APIs have been developed by the RIR to facilitate automated updates.

5.7 RIRs Internal Issues

Within the RIR system, a number of contemporary issues have elicited heated debate within the operational epistemic community. While the following sections draw these out as individual issues, they are certainly not mutually exclusive. The flagship instance is the recent discussion of what is referred to in this work as constitutional norms. Recently, RFC 2050 (Hubbard et al., 1996), discussed in Section 5.6.1 as the longstanding normative foundation of number resource management, has been replaced by RFC 7020 (Housley et al., 2013). Section 5.7.1 summarizes some of those debates, in particular how the notion of conservation, interpreted in this work as community vernacular for one factor contributing to system integrity, has changed in the face of *a*) IPv4 exhaustion, *b*) a relatively massive (yet still under-deployed) pool of IPv6 addresses, *c*) route origin and path security, and *d*) maturing transfers markets.

This latter, transfers, segues the discussion to a longstanding debate recently rekindled by IPv4 exhaustion. Transfers, as elaborated in Section 5.6.3 is one of multiple modes of rights alienation in he RIR system. As a study of resource rights, the debates over ownership versus stewardship highlight differentiated images of resource management within the community. Although some actors consider transfers to be a novel activity, it is, as discussed in Table 5.2, a less restricted form of alienation present in all forms of rights delegation. Transfers policy development is also of instance of a number of issues that approach the penumbra of resource policy and industry market policy.

Origin rights security, offered by RPKI, is another issue approaching, if not squarely in, the penumbra between resource policy and broader public policy concerns over the integrity of the NRS. RPKI has the *potential* to supplement authoritative documentation of route origins in the registry, a means to automatically and immediately validate delegation assertions against registry data, *and* exercise rights revocations.[307] Automatic revocation increases the enforcement power as

> APNIC enters the inetnum objects for portable allocations and assignments in the APNIC WHOIS Database, but Members are responsible for registering the sub-allocations and assignments that are delegated to their customers. (APNIC, 2014f)

[306] APNIC provides a FAQ on allocation windows (APNIC, 2015b). LIRs may delegate rights for blocks less than or equal to the assignment window. Blocks larger than the assignment window require a "second opinion" from APNIC. APNIC indicates that the second opinion will be turned around in one business day. As the LIR gains experience delegating rights, the assignment window will be increased to reflect that experience.

[307] The automatic and immediate application of RPKI is one possible implementation promoted by

one moves up the proposed trust anchor hierarchy. Following the need for solutions to the origin and path security externalities discussed in Section 2.1, such solutions would seem welcome in the operational community. Rather, many challenges to the RPKI system have been made in terms of how it affects routing decisions, potential for abuse of a hierarchical key structure by state actors, and the operational costs of such an infrastructure. Section 5.7.4 discusses these issues in terms of rights delegations. RPKI also serves as a study of the gap between number policy and firm operations in the RIPE community.

5.7.1 Norms

In August 2013, RFC 2050 was superseded by RFC 7020 (Housley et al., 2013). Nearly 17 years later, RFC 7020 is an update of the principles that structure IR operations. In contrast to the prescriptive norms established in RFC 2050,[308] RFC 7020 "describes the present Internet Numbers Registry System," in terms of three "(non-exclusive) goals," (Housley et al., 2013, p. 2–3). These goals do not necessarily conflict with those set out in RFC 2050. Rather, RFC 7020's goals reflect how the goals in RFC 2050 have evolved with operational experience. In the language developed around operational epistemic communities, RFC 7020 is an attempt to articulate tacit knowledge accumulated by RIR managers since RFC 2050 was published.

RFC 7020 frames RIR goals explicitly in terms of number resource distribution. Those three goals are *1*) allocation pool management, *2*) hierarchical allocation, and *3*) registry accuracy. Allocation pool management identifies the IP address and AS numbers as finite pools due to the "fixed lengths" of these integers. This work developed a similar rationale in Section 2.1.3, citing protocol provisioning of IP addresses in RFC 791. The goal of allocation pool management also indicates that

> allocations must be made in accordance with the operational needs of those running the networks that make use of these number resources and by taking into consideration pool limitations at the time of allocation. (Housley et al., 2013, p. 3)

This invokes the spirit of needs-based criteria and conservation from RFC 2050 and existing practices, but is both more precise and addresses the dynamic character of the delegation environment. In terms of the resource rights framing used here, RFC 7020 specifically calls out constraints of the resource system actors whose demand

security hawks. More moderate RPKI supporters highlight that RPKI use is still at the discretion of the RIR member, thus even if RPKI asserts a revocation, the member can ignore this assertion. The problem that has seen the most discussion in the RIRs is not the implementation, but rather the potential for RPKI to be abused by state governments. This is framed as the potential for Levi's predatory rule (Levi, 1989) in both Section 5.7.4.

[308] At the time RFC 2050 was written, two of the five modern RIRs were in operation alongside Network Solutions. RFC 2050 norms are referred to as prescriptive in the sense that they were based on known best practices developed in the early Internet and existing RIRs. Given hindsight, these prescriptions did in fact take hold and, as per the operational policies discussed in this section, are still considered guiding principles today.

should be considered. Moreover, dynamic adaptation is anticipated in closing with "at the time of allocation." Discussion of how RFC 2050 norms have been preserved will return to this dynamic scoping.

Hierarchical allocation argues that "the distribution of IP addresses in a hierarchical manner increases the likelihood of continued scaling of the Internet's routing system," (Housley et al., 2013, p. 3). Hierarchical allocation is argued to "permit[] aggregation ... into a minimum number of routing announcements," (Housley et al., 2013, p. 3). It also goes on to highlight that, like in RFC 2050, these considerations do not guarantee that these addresses will be either announced or routed. The delegation hierarchy in Section 5.2 illustrates how "hierarchical allocation" works in the present system. Of the three goals, hierarchical allocation seems to hew most closely to RFC 2050's goals of balancing the implication of delegation strategies on "routability" with operational factors in the actual dissemination of routing information in the control plane.

The necessary prescription of uniqueness has been refined to modern operational needs in the registration accuracy goal. In particular:

> provid[ing] accurate registration information of ... allocations in order to meet a variety of operational requirements. (Housley et al., 2013, p. 3)

Here, the variety of operational requirements include the coordination to resolve externalities. In addition to coordination, accuracy is a function in the RIRs' role in documenting transfers. Bundles of rights discussed in Section 5.6.3 distinguish transfers from other modes of alienating number rights. A fundamental requirement of an accurate registry is careful management of alienation of number rights, especially in the case of information commodities. In terms of NRS integrity, accuracy is a refinement of the notion of integrity offered in Part I. The challenge of registry accuracy is to ensure delegations are accurately documented while facilitating the maturing transfers market between now and when (if) IPv6 reaches critical mass. The discussion of transfers in Section 5.7.3 builds on the differences in alienation rights and the role of limitations as one means to ensure registry accuracy.

The term protocol provisioning is not explicitly used in RFC 7020, but is further developed in discussing the Internet Numbers Registry System Structure. RFC 7020 highlights the role of the IETF as what this work refers to as the protocol producer:

> The IETF specifies the underlying technical facilities and constraints of Internet numbers administered by the Internet Numbers Registry System. These specifications are aimed at enabling and protecting the long-term viability of the Internet, and adjustments to those specifications are made over time as circumstances warrant. (2013, p. 4)

Explicitly calling out "long-term viability" speaks to sustaining the resource system. In contrast to the early prescriptive character of RFC 2050, framing of the IETF's role in developing specifications speaks specifically to efforts at maintenance, sustainability, and further development of the Internet number system.

Finally, RFC 7020 makes a distinction between the non-policy elements of the system managed by the IETF and operational resource policy developed in the RIR system. RFC 7020 further develops the distinction by indicating "discussions regarding the evolution of the Internet Numbers Registry System structure, policy, and processes are to take place within the ICANN framework and will respect ICANN's core values," (2013, p. 5). While ICANN is established as the forum, RFC 7020 abandons the IANA appeal process. Instead, it establishes the independence of the RIRs' respective rule making institutions, in particular the appeals processes as authoritative. As a step towards relational authority, selecting local appeals processes further preferences existing operational structures rather than the black letter of a simple principal agent hierarchy rooted in the IANA. In effect, it simultaneously designates ICANN as the forum in which the RIR community will develop system-wide rules while also asserting the independence of RIR rule making and adjudication processes from the IANA.

Earlier discussion in Section 5.2.2 alluded to a distinction between hierarchical delegation of authority and delegation of number rights bundles. RFC 7020's abandonment of the appeal process, along with existing MOU's discussed in the context of the NRO in Section 5.4.5, are evidence of the RIR collective reinforcing its rule making authority relative to number resource policy. This should not be interpreted as distancing itself from either conventional governance organizations or other organizations within the Internet governance ecosystem. Although not highlighted, in discussing the ongoing evolution of the Internet Numbers Registry, RFC 7020 speaks to respect for and "recognition of the policy roles of other responsible entities that reflect the interests of affected parties," (2013, p. 6).

5.7.2 Needs-Based Criteria

Both RFC 2050 and RFC 7020 recognize needs-based criteria for delegating number resources.[309] Needs-based criteria are applied as a general condition of delegation. Information contributing to evaluating needs-based criteria are proxies for demand. This information is supplied in part by the organization requesting resources. Demand is interpreted by the RIR firm to determine whether a particular delegation is credibly expected to be utilized within the justification period.[310] Efficacious utilization here means that a significant proportion[311] of that delegation will be utilized—those numbers will be assigned to hosts and routed on the public Internet within the justification period.

The spirit of needs-based conditions across policy corpora is to ensure scarce number resources do not lie fallow (are not wasted) once delegated from the stock

[309]Needs-basis in RFC 2050 was discussed in Section 5.6.1.

[310]The justification period is the amount of time a particular delegation request is expected to fulfill an organization's demand. The justification period is a parameter set in RIR policy. Early on, justification periods differed substantively across the RIRs. More recently, justification periods have been generally aligned across the RIRs.

[311]The significant proportion is the utilization rate discussed in Section 5.6.3. In general this rate is 80%. The utilization rate is a policy parameter.

managed by an RIR (i.e., once delegated by an RIR or an NIR) to an LIR stock. Needs-based criteria condition *a*) evaluation of a request for a rights delegation and *b*) subsequent auditing of existing rights delegations. As noted in the discussion of operational rules, the latter, an audit, occurs upon the request of any subsequent rights delegation. Each RIR has audit criteria in its policy corpus.

Although needs-based criteria appear in both RFC 2050, RFC 7020, and have been enshrined in most, but not all, policies that have reached consensus in the RIR system, it is a contentious policy. Part of that contention has to do with perceptions of how the transfer market will operate. In the context of transfers, discussed more extensively in the next section (5.7.3), needs-based criteria has been argued by some in academia[312] and in the community as creating unnecessary friction in the market. Another argument is that the needs-basis places substantive discretion in the hands of the RIR firm. Consider a recent post to the ARIN PPML:

> ARIN should not be in the business of turning down resource requests if they have the resources to allocate - EVER. Doing so is arbitrary and discriminatory. ARIN should only be in the business of right sizing allocations to match the size of the organization (including their existing network size) making the request - AND - keeping the registry database as accurate as possible. (Ryerse, 2014)[313]

The assertion of "arbitrary and discriminatory" is not an isolated complaint. That said, there is an appeals process[314] for resolving these issues kinds of issues. Stated another way, needs-based criteria ensure resources are delegated to actors based on both current demand asserted and the accuracy of previous needs assertions (a proxy for credible utilization estimates). Also note that many of the RIRs indicate that, while needs-based evidence can be validated, there is also an element of trusting the assertions of network requesting delegations.

One option is whether to provide exact thresholds for acquiring resources. On the face of it, this makes the criteria more transparent and fair—discretion is eliminated and everyone has equal knowledge of what is required to acquire what volume of number rights. The trade-off is that, absent effective validation of informa-

[312]See Mueller, Kuerbis, and Asghari (2013) and the follow-up (Kuerbis, Asghari, & Mueller, 2013) for discussion.

[313]This assertion was made in the context of discussing the role of needs-based criteria in the context of number (rights) transfers. While the focus is on the articulation of "arbitrary and discriminatory," it should be noted that this particular actor was recently denied an allocation.

> When ARIN denies allocation requests like they did to us, they force organizations like ours to make a decision on possibly purchasing resources outside of ARIN. When that happens, the result of the needs based polices may be an off the ARIN books transaction and the registry database becomes less accurate. Thus existing policies cause ARIN to NOT be able to fulfill its Mission of keeping an accurate database.

This is included here to complete the context of the assertion, but also to foreshadow discussion of black and grey markets in Section 5.7.3.

[314]For instance, in the ARIN region, see (ARIN, 2009a). In the RIPE region an arbitration process (RIPE, 2014a) has been developed, with arbiters (RIPE NCC, 2014d) elected from the community.

tion submitted, any actor may claim "demand" corresponding to the volume they want. The countermeasure to complete transparency is strict validation of *a*) demand asserted by a requesting actor, *b*) regular and strict audits of utilization, and *c*) reporting of both allocation and audit data. This increases the transaction costs for both the requesting actor, the RIR, and would require sharing information many firms would prefer remain private for, among other reasons, competitive advantage.

The opposite end of the spectrum is a completely opaque decision process that would easily mask arbitrary decision making. Complete opacity in the decision making process is possible, but would not limit visibility of the outcomes. A feature of the NRS, like many other commonly managed systems, is that elements of the system can be monitored by any number of participants. For instance, analysis of registry data can be performed by anyone that applies for bulk access. In the NRS, resource management facilities are intended as neutral actors whose number rights delegation authority and certain evaluative criteria are subject to changes through collective action rules. Again, like other common resource management systems, the potential for monitoring is a backstop, a mode of accountability, that substantively increases the probability that both arbitrary delegation will be identified and a dishonest or opportunistic participant will be identified.

RFC 2050 (Hubbard et al., 1996), RIR policy corpora, and the content and structure of resource applications themselves, all contribute to documenting the *types* of information required to justify need, the *general* criteria on which they will be evaluated, and sources that may be used to *validate* assertions where necessary. Although not all strict criteria for delegation are provided, the utilization rate is the primary threshold documented in RIR policy. Utilization is presented as the primary factor in determining eligibility for a delegations. For instance, LACNIC indicates that it is the first criteria considered and a request will not be considered further if that criteria is not satisfied.[315] There has been substantive discussion on the list regarding the efficacy of the utilization function. Other factors, such as network topology, subnetting, and growth plans, are arguably private. Moreover, evaluation of these factors is also subjective.

Needs-based evaluation builds on these monitoring characteristics, occupying a space between strict transparency and opaque command-and-control management. In contrast to complete transparency and a strict auditing regime, the potential for monitoring by participants and/or the RIR is a form of deterrence. Needs-based evaluations operate on a similar principle that many other such regulatory mechanisms in common resources operate: the potential for having dishonest practices called out by either members of the community or in an audit by the RIR is the

[315] See LACNIC Policy Manual, which states:

> This utilization percentage shall be based solely on announced networks with IPv4 addresses connected to the Internet. For IRs that have assigned IPv4 addresses to their clients, the method available to prove this utilization is through the records kept in LACNIC's WHOIS database. Consideration of the application shall not continue until utilization of at least 80% of the previously allocated block is verified. (LACNIC, 2014f, p. 19)

deterrent mechanism. For instance, route announcements in the routing system are one way to validate utilization assertions. This is not limited to the RIR. Actors with bulk registry access and access to route announcements can perform the same utilization analyses as RIRs, potentially recognizing and calling out actors that are not effectively utilizing public addresses. Accountability is not a one way street for evaluating the veracity of claims by requesting parties. As noted in earlier discussion, the RIRs have appeals processes that facilitate contesting decisions made by the RIRs, in particular delegations based on needs-based criteria.

The countervailing forces above arguably both strike a balance between transparency, organizations' privacy, and reducing costs. Some degree of privacy is necessary to both protect the interests of applicants and to avoid increasing the costs of evaluating requests attempting to game the system.

Another factor in evaluating the role of needs-based criteria is to understand how it serves as a trigger for auditing processes, and by proxy, the accuracy goals in RFC 7020. Under current registry operations, delegation requests are the primary trigger for audits; independent audits occur, but do not cover all rights holders.[316] Delegation requests typically invoke audits. Audits can and should be conceptually decoupled from the needs-criteria process. Given a goal of RFC 7020 is registry accuracy, a possible alternative to triggering via needs-based criteria is a broader sweeping, more systematic audit process.

This analysis intends to be non-normative, it neither proposes to privilege or deprecate needs-based criteria. Rather, it highlights that needs-based criteria have played a critical role in preserving the integrity of the IPv4 address pool. In a number of policy discussions reference to managing the "allocation pool." As discussed earlier, the definition of this pool invokes a similar notion of "fixed lengths of IP addresses" as the constraint. In context, a number of the policy discussions implicitly or explicitly assume what is referred to as the allocation pool refers only to the pool of resources from which allocations are made. For each IR in the NRS, this scopes their immediate responsibility to managing number rights for those number in their respective allocation pools.

In contrast, the notion of integrity, presented in Chapters 2 and 3, argues the resource being managed is the entire pool. Focusing on "allocation" pools potentially obscures transfers as the soon-to-be only mode of IPv4 number rights delegation. More precisely, the IRs and end users depicted in Figure 5-4 comprise a federated collective of partially hierarchically ordered organizations that manage the entire number resource pool based on localized implementations of the constitutional norms described in Section 5.6.1. Based historically in RFC 2050, the evolution of policy thus far, and recently RFC 7020, needs-based criteria are the primary means of ensuring the integrity of the pool as a whole. The question for an evaluation of needs-based criteria shifts from whether needs-based criteria is appropriate in the current operational environment to what mechanism can effectively fill

[316]This is based on a combination of policy analysis and interviews. Policy indicates that historically requests require a needs analysis, including evaluation of utilization of existing delegations. Interviews of operations officers indicate audits are performed, but their are not complete, i.e., audits follow a sampling strategy rather than auditing every rights holder on a periodic basis.

the role of ensuring the integrity of the pool in the absence of a needs-based triggered system? In other words, what mechanism can facilitate registry accuracy given registry content is jointly provisioned?

Returning to the premise of needs-based criteria offered earlier in this section, the objective is to ensure numbers do not lie fallow—they are not allocated to actors that do not have a credible commitment to using them. A number of actors argue that the price mechanism is a potential alternative. In the face of exhaustion of the unallocated pool, conservation, and thus needs assessment for the purpose of conservation of a scarce resource, is no longer needed. A strong interpretation would be that addresses should be delegated regardless of *whether or not* they will actually be used. It is also unclear what incentive actors receiving such delegations would have to serve as an IR.

In Ostrom's studies of resource systems, a key factor in credible commitment is that participants' "livelihood" is dependent on the integrity of the resource. Dependence on not just the existence of the resource, but on the integrity as a condition of ongoing function, serves as a compelling selective incentive. Moreover, the dependence of participants' livelihood means that, as a selective incentive, it has a clear, direct, and meaningful impact on the participants' decision processes. In terms of the infrastructure framings from Section 3.1, NRS participants, in particular those that hold number resource rights delegation, are *direct* users. They have a credible commitment to maintaining the integrity of the system. Needs-based criteria can also be framed as one means to distinguish direct from indirect users.

All this said, this is not an apologia for needs-based delegation criteria. Rather, it is intended to highlight the function of needs-based evaluation as a means to indicate the precise function of a potential replacement criterion. Following the distinction between credible commitment of direct users, a key function that would need to be introduced to continue joint management of the registry is a means to evaluate and enforce credible updates of the status of delegations.y

5.7.3 Transfers

Within the RIR system, transfers have been a long-standing point of contention. The operator community has argued for and against transfers on grounds of registry integrity (accuracy), potential hoarding, price escalation, and detrimental effects on IPv6 deployment. The simplest argument is whether transfers should be allowed. First, how should alienability rights be scoped? In other words, what are the limitations? Second, how does the firm, acting as the agent of the community, incent documentation of transfers? Building on the differences between delegation bundles, this section illustrates how transfers can confound registration obligations in existing delegation bundles comprising alienation rights. The rights framing helps reason more precisely about the various types of transfers that have been considered in the community. The first rough cut is the difference between *intra*-RIR transfers and *inter*-RIR transfers. As may be obvious, intra-RIR transfers occur between members of the same RIR, transferring number rights delegated by that common RIR. Inter-RIR transfers would occur between members of different RIRs, but require

reconciling potential differences between the operational rules of the RIR, amongst those, differences in needs-based criteria. Within the constraints of the allocation window, allocations can delegate rights to any network but require documentation in the registry. As per Section 5.3.1, the obligation to maintain documentation of $> L3$ delegations lies with the LIR. LIR allocations require the LIR create a contractual obligation with customers to return addresses on termination of the contract. Unrestricted transfers could potentially confound these limitations and obligations.

Recall the basic use rights bundle comprises origin rights and numbering rights. In Section 5.6.3, it was argued that unrestricted transfer rights are broader than the scope of alienation rights conferred in delegation bundles currently available.[317] As such, some obligations are lost and some management authority is reclaimed by members. The result is a fragmentation of management authority. While this does not assure instability, as the transfer market grows, absent a coordinating agent, in particular an agent that assures the basic appropriation bundle is respected, decentralizing this element of the system can reduce the integrity of the control plane.

Consider a brief review of the differences between assignment bundles and allocation bundles. Assignment rights bundles do not include alienation rights: the spirit of an assignment is to confer the rights necessary for effective use of a block of addresses for Internet communication. Assignment thus confers numbering and origin rights. Assignment does not confer the right to subsequently delegate number rights, and the attendant obligation to document (re)assignments in the registry, to a third party.

Allocation bundles do include limited alienation rights. Actors whose rights bundle is an allocation bundle have the rights to *further* allocate or assign use rights to their customers. That said, allocation holders are ultimately responsible for (obligated to) maintain accurate (re)assignment information in the registry. Rights to update reassignment information may be conferred to those to whom an LIR assigns addresses, but, with respect to the RIR, responsibility for accuracy and timeliness lies with the LIR. The result is that this bundle of rights preserves the joint obligation to maintain the registry.

The ongoing debate over transfers tracks the problems of accuracy and integrity in the previous section. In contrast to allocation rights, the question for transfers is, how can the RIR, in the face of free pool exhaustion, facilitate the exchange of number rights that *a*) allows actors to transfer basic use rights bundles, *b*) allows actors to compensate the original rights holder willing to give up number rights, and *c*) ensure the integrity of the pool as a whole, in particular ensuring the ongoing accuracy of the registry. To satisfy the latter criterion, needs-based criteria have been applied to most transfer policies as a trigger for audits discussed earlier.[318] Like the use of needs-based criteria in the previous section, for transfers, the objec-

[317]Legacy bundles have unrestricted transfer rights. That said, legacy bundles are just that, legacy, and are not available in the modern RIR system.

[318]Exceptions include Prop-50 in the APNIC region, which has been revoked to re-introduce needs-based criteria and a modification of allocation rights to facilitate changing the status of provider aggregatable assignments to provider independent assignments without a needs assessment in the RIPE region.

tive is to introduce conditions that ensure *a*) the registry has some record of the transfer, *b*) subsequently it has record of the identity of the recipient, and *c*) the recipient has a contractual relationship with the registry, ensuring registry update and audit criteria hold. The contractual relationship typically obligates the recipient to participate in registry updates, thereby ensuring the accuracy of the registry.

Returning to the types of transfers, intra-RIR transfers are available in all of the RIRs, at a minimum a merger and acquisition policy that has the same delegation effects as a transfer. In most cases, intra-RIR transfers are conditioned on evaluation of need. In the case a rights holder has addresses that are no longer in use (assigned) or, assuming the RIR will recognize underutilization and invoke returns policy, knows it will soon no longer need some or all of its delegation, it may put that delegation up for transfer in all or parts. In the most general case, the basic appropriation bundle is an asset. The basic appropriation bundle and bundles that comprise the basic appropriation bundle are assets in the same sense that property rights are a claim on value—claim to an asset is a claim to the value that can be derived. This does not imply simple ownership often affiliated with ownership or unrestricted alienation. Assuming a source and recipient have agreed to a transfer, under existing policy, the recipient must satisfy the needs-justification. The recipient must also typically sign a membership agreement with the RIR if it is not already a member.

A number of factors affect whether a transfer may proceed. The legacy status of a particular delegation may require the legacy provider establish a relationship with the RIR in which it is based. Recall the legacy rights bundle is the basic use bundle with unrestricted alienation rights: there are no contractual obligations with the RIR to maintain assignment information. Legacy rights holders are also not beholden to utilization requirements. In the situation where legacy rights holders have a larger delegation than necessary, it may be very lucrative for them to transfer, for a fee, some of that delegation. This transfer can occur on the black market—if the legacy rights holder chooses not to route a block of addresses and the recipient begins routing those prefixes, there is little the RIRs can do. Moreover, the recipient may simply claim it is leasing the rights. In terms of claims to rights, this is a bilateral contract with no guarantees other than the court system in that jurisdiction, even if they are in the same jurisdiction. In terms of diffuse enforcement by actors provisioning routes, there is no authoritative source to confirm stewardship. It is possible that once "sold," absent any claim to registration, such addresses would be targets for hijacking and subsequently a variety of abusive purposes.[319] The result is that certain number assignments are not accurately reflected in the registry, confounding a number of downstream uses, including law enforcement.

Returning to arguments of endogenizing costs in Chapter 2, absent needing subsequent delegations, legacy providers can effectively free-ride on registry maintenance by non-legacy members that are either contractually bound to comply with RIR policies and/or historically depend on subsequent allocations. For instance, to be in compliance with RIR policies, a legacy holder would have to invest in mech-

[319]The mechanics of a growing stock of tainted addresses is discussed in Chapter 7.

anisms for updating assignments. This would include retroactively documenting existing assignments and introducing mechanisms to update the registry with ongoing assignments. Depending on the size and infrastructure of the organization, this may be rather expensive.

Inter-RIR transfers further complicate the issue. Registry accuracy issues in intra-RIR transfers remain. The key issues in inter-RIR transfers is the perception that differences in allocation policies will create substantive shifts in delegations from one region to another. Further, there were fears that substantive differences in needs-based criteria (or the lack thereof) could facilitate number resource stockpiling, driving up prices. At the same time, a class of indirect users, namely IPv4 brokers, became active in the policy development process. In general, these actors efforts were aimed at lowering the barriers to IPv4 transfers. Following the language set out by Mueller et al. (2013) one could argue they were trying to reduce the friction in the IPv4 market.

Attempts to move towards a more "frictionless" market[320] have been met with some degree of skepticism. From the perspective of the RIR system, a legitimate transfer market respects the policies of RIR system, including needs-based conditions on transfers. Needs-based conditions place a number of constraints on the legitimate transfer market; all but legacy holder to legacy holder transfers are required to follow policy by virtue of contractual obligation of at least one transfer participant. As noted in the discussion of intra-RIR transfers, typically the recipient is an existing member and thus contractually obligated to follow RIR policy. The second constraint immediately follows: needs-based transfers are limited to a volume justified by the requester, as adjudicated by the RIR's review of the transfer request.

The third constraint is structural. It immediately affects member-member transfers and has implications for legacy that become members. Consider the case where the originator of a transfer is an existing member and is no longer utilizing resources at the level expected, specified in the original resource request for those resources. If this under-utilization is demonstrable, recovery procedures may be invoked by the RIR. In effect, the RIR may recover resources and make them available for subsequent delegation. For instance, ARIN's transfer policy states:

> Number resources are issued, based on justified need, to organizations, not to individuals representing those organizations. Thus, if a company goes out of business, regardless of the reason, the point of contact (POC) listed for the number resource does not have the authority to sell, *transfer*, assign, or give the number resource to any other person or organization. The POC must notify ARIN if a business fails so the assigned number resources can be *returned* to the available pool of number resources *if a transfer is not requested and justified.* (ARIN, 2014a, Section 8, p. 16, emphasis added)

[320] Outside of the fictions used to teach undergraduate micro-economics courses, frictionless markets do not exist in the wild. Rather, every market is accompanied by an institution that orders transactions and moderates opportunism.

Within ARIN's policy corpus, any transfer requires and existing or new RSA agreement, including compliance with RIR policy. Similarly, APNIC's transfer policy indicates any transfers to APNIC account holders will be subject to APNIC policy at the time of transfer.

At the time of writing, the only two RIRs with functioning inter-RIR transfer policies are ARIN and APNIC. It should be noted that, as can be seen in Figure 5-1, the bulk of legacy space is in the ARIN region. Moreover, much of this is not routed. Also, note APNIC was the first RIR to run out of IPv4 addresses, invoking delegation of the last five contiguous /8's from the IANA. Thus it seems rational that the community would develop a mechanism linking an obvious supply (ARIN's stock, in particular legacy space, as a source) with demand (APNIC community as potential recipients).

Here, transfer policy can be interpreted as an institutional response to supply in one region (ARIN) and demand in another (APNIC). APNIC provides an interesting case in transfer policy development, constitutional norms across the RIRs, and the role of needs-based justifications. APNIC's Prop-50 (Huston & Smith, 2010) is well known for creating policy that would remove the needs-based condition from transfers once the last /8's had been delegated and were being actively distributed. As a policy, it addressed many of the issues discussed in Section 5.7.2, in particular insuring the accuracy of the registry and the integrity of the routing system.[321]

> The underlying proposition behind this policy proposal is that the registry of IPv4 addresses operated by APNIC is of general utility and value only while it accurately describes the current state of address distribution. *If a class of address movement transactions are excluded from being entered in the registry, then the registry will have decreasing value to the broader community, and the integrity of the network itself is thereby compromised.* This proposal's central aim is to ensure the continuing utility and value of the APNIC address registry by allowing the registry to record transactions where IPv4 addresses are transferred between APNIC account holders. (Huston & Smith, 2010, pp. 1–2, emphasis added)

A common outcome considered in the IPv4 run-out is the emergence of secondary markets for IPv4 addresses. If such a market's activities were not documented in the registry, the registry would, as described above, have decreasing value.

Although the needs-based mechanism was removed from the transfer transaction, Prop-050 was clear that

> The recipient entity of the transferred resources *will be subject to current APNIC policies*. In particular, in any subsequent APNIC IPv4 address allocation request, *the recipient will be required to account for the efficient utilization of all IPv4 address space held, including all transferred resources*. (Huston & Smith, 2010, p. 3, emphasis added here)

[321] In general, the use of integrity in this work is rooted in the definition developed in Chapter 2. In this case, the term was originally used in the policy, in contrast to other policies that tend to focus on conservation as a norm established in RFC 2050.

As noted in earlier discussions, common resource management often has the goal of delegating resources to those actors that are expected to utilize those resources. In the number resource system, that objective has been tightly coupled with, and arguably obscured by, the debates focusing on needs-based evaluation of delegations. Again, this is a not an apologia for needs-based criteria or a condemnation of those criteria. Rather, the point is to highlight that, as per the discussion of goals being interleaved in rules,[322] powerfully embedded norms of conservation may have obscured more fundamental goals related to assuring integrity, in particular the mechanisms that ensure accuracy.

While Prop-50 is an interesting case in terms of highlighting integrity, it is also interesting in its message regarding experience and anticipation[323] related to resource policy development.

> APNIC has no experience in determining what actions by potential parties to a transfer may need to be constrained in some fashion. Attempting to create policy in anticipation of the need for such constraints is going to be a guessing game that has accompanying flaws, irrespective of what constraints are initially specified in policy, it will be the case that as the levels of experience in this form of activity increases some iterations over the policy of constraints will be necessary in any case. This approach argues to start from a position that is relatively open and unrestricted, and recommends that APNIC impose additional constraints only when all other forms of constraint are inapplicable and there is a clear need and common desire for such constraints to be enforced by APNIC as distinct from using another party for such a role.(Huston & Smith, 2010, p. 7)

Invocation of experience and anticipation here seems at odds with the analysis thus far, in particular the effective decoupling of needs-basis from registry accuracy. Arguing that "APNIC has no experience in determining what actions by potential parties to a transfer may need to be constrained" is incongruous with the fact that APNIC has thus far done just that—constrained the scope of transfers as an alienation of particular rights as a means to achieve conservation and registry accuracy. As such, when considered in the context of fundamental elements of a rights transfer, this is a different type of rights transfer, but one that has historically been limited. The argument regarding anticipation, namely that "iterations over the policy of constraints will be necessary in any case" coupled with starting from "a position that is relatively open and unrestricted" is an instance a economically liberal policy that rejects the precautionary principle.

[322]See the discussion of Hart in Section 3.4. In particular, Hart argues types of rules interleaved in actual articulations, see (Hart, 1994, Chapter 5) for a general discussion. Following that argument, the intersection of transfers and need is an instance of fundamental principles of a system being embedded in and at times obscured by context-specific manifestations. Here, a powerfully embedded norm of conservation seems to obscure the more fundamental norm of assuring integrity.

[323]Here anticipation is used in the sense of (McCray et al., 2010).

5.7.4 RPKI

As a number resource issue, RPKI is a tool for better protecting origin rights. In all but the RIPE region, RPKI was introduced as a service offered by the RIR—RPKI was not a resource policy issue. In the RIPE region, RPKI was discussed as a policy issue, in particular whether the NCC should implement certification services. RPKI in the RIPE region illustrates a number of issues in RIR decision making: *a*) it is a study in balancing consensus processes with other decision making processes in the RIR, *b*) it illustrates the intersection between number rights delegation authority and institutional authority discussed initially in Section 5.2,[324] and *c*) it illustrates a policy that, by straddling the line between number rights management and routing operations, highlights the interdependence between delegation rights enforcement and routing operations discussed earlier. Section 5.6.2 discussed decision making components. This section will further develop the latter two points.

RIRs' number policy has a history of limiting its scope to number, or "addressing" policy, avoiding interference with routing operations. In the broader NRS ecosystem, this can be likened to the IXes norms of neutrality with respect to operational practices that do not affect the integrity of the common resource. That said, as noted in Section 5.2 in particular, delegation strategies do have an effect on routing operation practices and subsequently integrity. Here, security externalities are the threat to integrity. Security externalities such as prefix hijacking and route manipulation have been considered long-standing weaknesses of the routing system. As discussed in Section 2.2, these security externalities are violations of rights conferred by the RIRs (origin rights, but also numbering rights) and tacit rights (expectations and contractual obligations) in bilateral interconnection arrangements. RPKI has been developed jointly by the IETF, RIRs, and ISPs and documented in both IETF RFCs (referenced in discussion below) and by the RIRs providing RPKI services[325] to protect origin rights and provide a framework for later developing path security infrastructure.

5.7.4.1 RPKI Concepts and Structure

RPKI is a PKI infrastructure that "enables an entity to verifiably assert that it is the legitimate holder of a set of IP addresses or a set of Autonomous System (AS) numbers," (Lepinski & Kent, 2012, p. 3). RPKI provides automatic verifiable attestations of number resource delegations. The registry is the authoritative source of delegation information. RPKI provides facilities for credibly attesting to the delegation of these rights, in particular, delegating the authority to originate a set of prefixes. Delegating origination rights here is not alienation in the sense of a transfer that permanently delegates all rights and obligations to the delegate. Rather, delegating origination rights is in the sense of LIR sub-allocations and assignments to customers.

Historically, BGP has been referred to as routing by rumor. Recall Section 2.2

[324] This is later developed in Section 8.1.

[325] See the NRO description of RPKI as resource certification (NRO, 2014c).

described route promulgation as a form of provisioning and transit trust. Given the variety of sources of information, the adage is certainly true, but some actors make better use of information to verify those rumors. IRRs are one mechanism, but are considered unreliable and often out of date.[326] The RPKI infrastructure attempts to offer authoritative attestations that can be used to validate "origin provisioning" rumors that also assures reliable, regular updates. As will be developed later in this section, based on both the externalities framework developed here and private conversations, RPKI represents both costs of endogenizing security externalities and coordination costs amongst actors within the delegation hierarchy.

The most commonly referenced product of the RPKI is route origination authorizations (ROAs, typically pronounced row-ahhs in conversation).

> At a high level, the ROA's content contains (*1*) an AS number; (*2*) a list of IP address prefixes; and, optionally (*3*) for each prefix, the maximum length of more specific (longer) prefixes that the AS is also authorized to advertise. (Lepinski & Kent, 2012, p. 10)

ROAs correspond to EU delegations (blue links at $L3$) in Figure 5-4. ROAs are stored in a repository[327] where any actor in the routing system can access them to verify an origin announcement is in fact legitimate; the signing hierarchy that assures legitimacy will be discussed shortly.

It is important to note that RPKI does not alter rights bundles recorded in the registry. As alluded to in Section 5.2, RPKI does enhance enforcement power. This enforcement power has implications for members discretion in route provisioning.

RFC 6480 argues that:

> [M]anagement of this PKI is a natural extension of the resource-management functions of the organizations that are already responsible for IP address and AS number resource allocation. Likewise, existing resource allocation and revocation practices have well-defined correspondents in this architecture. (Lepinski & Kent, 2012, p. 3).

As implied above, this assertion is certainly the case for number rights delegation discussed in Section 5.2 and depicted in Figure 5-4.[328] In Section 5.2 the conceptual distinction between delegations from the IANA to RIRs ($L1$), RIRs to LIRs ($L2$), and LIRs on to EUs ($L3$) was established the foundation for differentiating the precise rights bundles described in Section 5.6.3. The RPKI offers a certificate hierarchy

[326]IRRs typically rely on network actors to maintain route information. Thus, while it is authoritative in the sense that it is provided by the provisioning actor, it has week auditing and updating mechanisms. As a point of interest, one RIPE policy proposed using RPKI data to create a more authoritative IRR, but was withdrawn due to lack of support.

[327]The details of the repository system are out of the scope of this document. It is documented in RFC 6480 in terms of the structure of the repository and local caching strategies (Lepinski & Kent, 2012, Section 6).

[328]Moreover, the analog in RFC 6480 is Figure 1 (Lepinski & Kent, 2012, p. 11) which depicts the authorization hierarchy structuring certificates attesting number rights delegations. This structure is, by simple inspection, a projection of rights delegation paths depicted in Figure 5-4.

that corresponds to $L2$ and $L3$ delegations. Like most community documentation of numbers delegations, the RPKI RFCs also refer to allocations and sub-allocations.

"Certificates in this PKI are called resource certificates... [r]esource certificates attest to the allocation by the (certificate) issuer of IP addresses or AS numbers to the subject." Two types of resource certificates in this PKI are significant for this discussion: the certification authority (CA) certificates that facilitate issuing resource certificates and end-entity (EE) certificates used to sign ROAs. RFC 6483 indicates that the "[t]here is an extant IP address space and AS number allocation hierarchy, and thus IANA and/or the five RIRs are obvious candidates to be the default [trust anchors] here," (Lepinski & Kent, 2012, p. 8). Each RIR plays the role of trust anchor in its respective region.

CAs are both resource certificates and are used to issue resource certificates. Certificate issuers are actors that have delegation rights. The IANA, RIRs, NIRs, and LIRs will all have CA certificate. As a resource certificate, the CA certificate describes the prefixes or ASNs for which it can authorize subsequent certificates. As may be obvious, a CA holder A can only subsequently delegate CAs for resources within the set of resources A's CA attests it is the steward of. As such, the CA is a mechanism to assert the limited alienation rights corresponding to the actor's rights bundle.

The EE is the mechanism for creating and revoking ROAs. Unlike the CA, which can be used to sign subsequent CA's, the EE is not used to sign additional certificates. As per RFC 6480, there is a one-to-one relationship between an EE and a ROA. The private key of an EE is used only once, to sign the corresponding ROA as an attestation of the rights delegated therein. In general, the CA hierarchy makes the delegation hierarchy durable and the creation of the EE corresponds to assignment given actors can only use it in an assignment, they cannot use it to issue sub-assignments.

A "standard procedure for issuing a ROA" is described in RFC 6480:

1. Create an end-entity certificate containing the prefix(es) to be authorized in the ROA.
2. Construct the payload of the ROA, including the prefixes in the end-entity certificate and the AS number to be authorized.
3. Sign the ROA using the private key corresponding to the end-entity certificate (the ROA is comprised of the payload encapsulated in a CMS signed message [RFC5652]).
4. Upload the end-entity certificate and the ROA to the repository system.

Adapting Latour's notion that technology is society made durable, the process above illustrates how the RPKI infrastructure makes an assignment (contracts) durable. More precisely, it makes that assignment durable conditioned on the will of the resource holder. Recall that earlier it was noted that RIR policy corpora indicated that LIRs should create contracts with customers that clearly state that assignments should be returned in the event that actor is no longer a customer of the LIR. Here, CAs facilitate revoking an assignment, making such contracts directly enforceable

by the right holder. It is the aspiration of strong RPKI advocates to automate revocation. The automatic, immediate revocation is the heart of many of the concerns about RPKI.

This strong, absolute[329] form of enforcement may certainly be appealing at the LIR level. Simple transaction costs will be addressed first, then the more compelling issues of enforcement, in particular what Levi refers to as predatory rule. That said, origination by rumor does not expire—resource certificates do. Moreover, changing an assignment requires an e-mail to the customer and a registry update. RPKI participation would introduce the requirement to replace the ROA attestation as well. In effect, it increases the operational costs. Whether this pays off in the individual benefits of reduced number of security externalities is a function of what constitutes critical mass of RPKI and whether actors are actively acting on attestations. One argument is that critical mass is all or nothing, but this is a strong ideological belief in strong enforcement. That said, it is unclear whether or what level of critical mass would yield benefits. Moreover, there is substantive uncertainty around the potential for abuse by government actors. Durable revocation has been less appealing as one climbs higher in the revocation hierarchy. This discussion is taken up in the next section.

5.7.4.2 Perceptions of Authority

A common mantra of RPKI supporters is that using RPKI does not force an actor to follow its directives. Relaxing the strong aspirations of automation, RPKI attestations are just another piece of information used when establishing local route selection preferences. Registry information, local preferences, and IRR data are other factors. IRR is considered incomplete, but is used in limited cases. Registry data can be incorporated but does not affect the immediacy of revocation. Local preferences reflect knowledge of local interconnection relations including reputation, "rumor" that informs which routes should be trusted, and which can be considered reputable.

RPKI uptake has been slow.[330] A number of simple rationales explain slow uptake. The first is that PKIs are complex and not necessarily intuitive to the average network operator. A more compelling reason is that incorporating RPKI is not mandated and may not be considered necessary. RPKI is a risk mitigation tactic against prefix hijacking externalities for which many actors have never directly experienced or recognized the costs. Absence of the selective incentive can be seen from various vantage points in a hijacking event, discussed earlier in Section 2.2. The first externality is the cost of having one's prefix hijacked; this may be the most immediately measurable cost if experienced. The second externality is the cost to the source of traffic being misdirected because of a hijack. For complacent actors along the AS-PATH, not checking the legitimacy of routes appropriated abdicates tacit obliga-

[329] Absolute assuming all actors use RPKI. This is certainly not the case, but on the margin, assuming some number $n > 0$, ideally $n \gg 0$, network actors are using RPKI, then assignment revocation makes the contract marginally more enforceable.

[330] See the certification statistics page maintained by the RIPE NCC (RIPE NCC, 2015a).

tions. Complacent actors, as discussed in Section 2.2, externalize the costs of these obligations.

Like other risk mitigation factors, RPKI may be seen as a cost center. Caches of ROA repositories need be maintained. New processes need to be introduced to incorporate RPKI into route filtering mechanisms. Exceptions need to be considered and implemented if RPKI is to be used in some cases but not in others. Decisions regarding whether all routes appropriated will be verified or just immediate adjacencies. If all routes, how should the benevolent RPKI user engage complacent upstream providers provisioning illegitimate routes? How will this engagement affect those relationships? The discussion of RPKI is another instance of technical issues for which the community has substantive expertise analyzing running up against political implications of a decision. Technical issues of deployment are well within the domain expertise of the operational epistemic community. As alluded to above, the challenge is the implications for liability and predatory regulation.

5.7.4.3 Security and Routing Operations

Reconsider the implications of Section 5.2's discussion of the RPKI as a projection of the number rights delegation hierarchy. Under ideal use, the RPKI is number resource rights made both more durable and more immediately enforceable.

The RPKI discussion is an instance of a longer standing separation between number policy and routing operations. In the NRS, this corresponds to the scope of number policy and routing operations. Recall the mechanics in Section 2.1 and the distinctions between number resource policy and routing operations throughout this chapter. Number policy shapes the appropriation of numbers and the corresponding policy managing that stock. Number policy influences operations, but does not directly dictate routing decisions. RIR staff interviews and policy discussions both stress that number policy and routing operations are two distinct spheres of authority.

The summary of the debate above highlights concern over concentration of authority engendered by RPKI automation. Also interleaved in these discussions is the expropriation of routing decisions through the CA hierarchy. Above, the routing by rumor metaphor was used to highlight the veracity of information in the RPKI. The worst mode of expropriation is the fear that validation would become blind automation. In the extreme form, expropriation means RPKI information is the exclusive factor in route selection. Under this framing, expropriation is what one actor in the discussion of the RIPE 2008-08 referred to as an abuse-vector.

The typology of NRS externalities helps disambiguate operational externalities from strategic externalities. In Section 2.2, one distinction amongst of operational failures is amongst purely unavoidable failures rooted in faults such as hardware failures or unknown software bugs from negligent failures such as lack of patching from incompetence (lack of operational experience). One point of discussion compared abuse of the RPKI with operational externalities. The distinction is subtle, but salient. Unavoidable faults are just that—unavoidable hazards that threaten the integrity of the network.

Abuse of the RPKI lives at the intersection of an operational externality and a strategic externality driven by political motivations. Under a negligent operational externality, a hazard is intentionally tolerated to avoid costs. Negligence here is predicated on a hazard that may have been at one point an unavoidable operational hazard but, proliferation of information about the hazard, an extant mitigation strategy, and the costs of that mitigation strategy transition the externality to negligence when costs trump mitigation. For the purpose of this specific analysis, removing cost and the initial unavoidable character allows the notion of negligence to be generalized to the *tolerance* of a hazard. In the case of perceived RPKI abuse, the question hinges on the debate over the strategic value of RPKI to LEAS and governments. If RPKI is sufficiently strategically important, a distinct set of actors in the RIPE community consider LEA and government intentions a hazard. Under the interpretation of this constituency, adoption of RPKI identifies the combination of *a*) concentration of authority up the CA hierarchy and *b*) the immediate, low-transaction cost enforcement of rights revocation as tolerance of a hazard. Arguments against RPKI also include the assumption that lower transaction costs for immediate enforcement are perceived to increase the appeal of the instrument to LEAs and governments. In terms of Figure 5-4, the red revocation vectors in $L3$ are also potential abuse vectors. The key distinction is that RPKI hazards are the consequence of political agency, not a stochastic failure. In this sense, RPKI externalities include strategic externalities motivated by downstream political issues.

Currently RPKI only verifies prefix origination rights. As per RFC 6480, the RPKI framers have aspirations for RPKI verification of both origin provisioning and broader path security. A focus on automation and expropriation of not only diffuse authority within the scope of number policy, but also aspirations for expropriation of routing decisions, militate against broad acceptance. Throughout, discussion of the implications focus on both immediate outcomes and potential expansions of power.

Chapter 6

Internet Exchanges

PACKETS DO NOT always follow the path expected from the geographic origin and destination. In the late 1990's, packets originating in London, destined for a network *also* located in London, crossed the Atlantic and back based on routes provisioned by international transit providers. This phenomena is referred to as tromboning,[331] when the path of geographically proximate traffic travels to a distant point and back again, and the shape of that path resembles a trombone slide. At the time, transit was the primary means to gain access to any portion of the routing table *and* the primary operational cost for ISPs. At approximately the same time, a similar tromboning story was playing out in Amsterdam (AMS-IX), Frankfurt (DE-CIX), Stockholm (Netnod), Seattle (SIX),[332] Buenos Aires (CABASE), Sao Paulo (PTT Metro), and, more recently in 2009, Paris (France-IX).[333]

Internet eXchanges (IXes) were initially developed to eliminate tromboning: the effect has been reduced transit costs and "keeping local traffic local." More precisely, IXes improve performance for participants by facilitating the exchange of geographically local traffic in a common, *local* interconnection platform. Increased local capacity and lower latency amongst local participants are two of the most commonly cited performance benefits. Commonly managed, jointly provisioned IX

[331]Early on tromboning often found the "bend" to be an interconnection point in the US. For Europeans, this often found traffic making at trip to MAE-East. For those on the Asian and Australian Pacific Rim, this meant a trip across the Atlantic to interconnection points on the West Coast of the US. More contemporary instances include traffic from Africa tromboning up to major interconnection points in Europe, and traffic from South America tromboning up to NAP of the Americas in Miami. Smaller scale tromboning occurs between Turkey and the AMS-IX in Amsterdam and on a national level, within the UK to the LINX in London. The latter was part of the motivation for developing LINX nodes around the UK, first in Manchester then later in Edinburgh. In the smaller cases, the latency argument is less critical than the cost of transport to relatively remote interconnection platforms and regional development efforts.

[332]Tromboning in Seattle is perhaps one of the most poignant instances: traffic originating in the same building in Seattle followed transit routes to Texas and back.

[333]France-IX is a story of rationalizing a fragmented interconnection market. In terms of route provisioning, France-IX consolidated loci of route provisioning to reduce the transaction costs. Historical associational membership IXes value proposition were rooted in reducing transaction costs, with route diversity a collateral benefit. As will be developed in Section 6.1 and based on interviews, France-IX's value proposition was rooted in rationalizing the Paris interconnection market.

platforms are the norm in Europe. Two of the three largest IXes in the world, the Amsterdam Internet Exchange (AMS-IX) and the London Internet Exchange (LINX) are European, associational membership based IXes.[334] These and similar European associational membership IXes are CRIs that provision capacity in the data plane (local capacity) and options to provision diverse routes that may not be available over transit. In effect, IX platforms facilitate interconnection markets that offer low-cost access to diverse, often low latency routes.

IXes, like RIRs, are an ensemble of network actors and management facilities. This ensemble comprises

- a commonly managed interconnection platform, or in the general terms of this work, interconnection facilities;
- the membership collective and collective-choice fora;
- the IX firm, delegated authority to exercise management rights necessary to develop, adapt, and enforce operational rules.

Note this ensemble parallels the RIR ensemble in terms of a facility administered by a firm, directed by and accountable to a common collective of users. Like many commonly managed resources, organizational structures are similar, but the nuance of management strategies differ—this is especially interesting given the substantive overlap amongst participants in the RIR and IX regimes.

In contrast to the RIR ecosystem, the IX ecosystem as a regime complex is decentralized. Unlike number resources, there is no singular, hierarchically structured source of resource units.Rather, each IX provider produces both interconnection options and L2 (layer 2) capacity through capital investments in a physical switching fabric. Archetypal topologies of IX platforms are described in Section 6.2. Whereas the NRO serves as a coordination mechanism for the RIRs, umbrella associations such as Euro-IX in Europe, Open-IX in the US, LAC-IX in Latin America, and the emerging IX-Federation serve as arenas for sharing domain knowledge within the IX provider community. As the IX regime proliferates, these arenas are taking more formal roles as knowledge commons in which actors can learn IX management and deployment best practices.

Returning to the structure and function of individual IXes, these organizations provision resources facilitating access to both the control plane and the data plane. In the data plane, the common resource consumed by participants is platform capacity. Exchange capacity is a function of the infrastructure (interconnected switches and routers) that support the common logical switching fabric. In the community vernacular, capacity, especially capacity contracted for appropriation by IX participants, is described in terms of *a*) the number of physical ports available, *b*) the various capacities of those ports, and *c*) the total volume of traffic an IX can carry.

As a commonly managed resource, IX platform capacity is the NRS facility that harkens most closely to a canonical CPR with partially-rival resource units. Both the exchange capacity and the number of ports is finite and traffic is partially-rival

[334]The other of the three largest, DE-CIX, started as a member-based IX but transitioned to a commercial model shortly after formation.

under congestion. A management objective of the IX is to monitor utilization in order to provision additional capacity as membership utilization (an indicator of demand) grows. A challenge in the membership-based IX community, in particular volunteer IXes, is operating as a non-profit organization that maintains sufficient capital to make step-wise upgrades necessary to match demand. Tacit within this challenge is whether to shift from ad hoc, volunteer management of the IX to a professional firm.

Interconnection options are the resource unit provisioned by the IX, and when exercised, manifest as a bilateral interconnection relation in the control plane. Interconnection platforms such as transport and colocation require participants commit capital resources to provision capacity to be used between each pair of would-be partners before they can interconnect over that platform. For large actors interconnecting with established interconnection partners, sufficient information is more likely available to evaluate the interconnection proposition.[335] For new relationships, though, such information, in particular traffic growth when connectivity moves from transit to a bilateral connection, may be expensive, increasing measurement costs and potentially increasing bargaining costs amongst these potential interconnection partners.[336] Interconnection options provisioned by the IX help reduce some of these costs.

Exchange capacity and physical paths between an two IX participants are *available* for appropriation from the moment a new participant joins the IX and contracts some level of port capacity. Developed further in Section 6.2, an *interconnection option* is the option, but not the obligation, for an IX participant to develop an interconnection relationship with another participant on the IX platform. The available exchange capacity and physical paths jointly provisioned as part of the IX platform serve as a general asset, facilitating the exercise of interconnection options with as many other participants as an actor may like, limited only by that actor's port capacity.[337] Recall the discussion of route provisioning in Section 2.3.2—transit outsources control over route provisioning decisions to an organization that optimizes on minimum cost routes. Interconnection options provide small to medium sized actors the opportunity to exercise greater decision-making power over the local set of routes available for appropriation, from whom they will appropriate those routes, and under what conditions. Rather than simply taking the lowest cost routes

[335]Sufficient information at the time a relationship is established is a static analysis. It definitely does not mean those conditions will hold. When the value proposition of one or the other changes, the relationship will likely be modified at the next opportunity. In the case of an informal relation, this may be immediate. In relationships structured by a contract, this change may be accounted for or may warrant renegotiation once the contract expires.

[336]The phases of interconnection provisioning are discussed in Section 6.1.1 and depicted in Figure 6-1.

[337]Credible interconnection partners work to ensure sufficient capacity for traffic exchange. This is not always the case though, most poignantly evidenced in the recent conflict between Netflix and Comcast. On a well-managed IX, exchange capacity is sufficient to absorb large traffic spikes. What remains for participants is to ensure port capacity sufficient for the aggregate volume of traffic it exchanges on the IX. See Section 6.4.3 for a discussion of operational rules regarding port capacity and congestion.

provisioned by transit providers IX participants have a market in which they may select routes that improve their value proposition.

Section 6.1.2 formalizes the set of options into a family of diversity indicators. In fieldwork, private conversations, and interviews, network operators consistently returned to two criteria for evaluating IX participation: redundancy and uniqueness. Diversity is a model for evaluate the character of an interconnection market x in terms of how much participants in m participate in *other* markets. Diversity helps evaluate the gains from a portfolio of existing IXes and potential deployment to new IXes. Diversity was initially introduced by Sowell (2013). Parts of that work are reproduced, and refined, in this chapter to explain IX participation dynamics, how IXes contribute to participants' value proposition, and hypotheses of learning effects in IX.

Sustaining the value-proposition is not without contention. Associational membership IXes are governed by three common norms: mutuality, non-compete, and neutrality. Mutuality is IX vernacular for collective management—as described in greater depth in the section dedicated to mutuality (6.4.1.1), an oft quoted statement amongst associational membership IX leadership is that "the single stakeholder in the IX [firm] is the collective membership."[338] Non-compete ensures that IX services do not compete with the services of its participants. Discussion of non-compete in Section 6.4.1.3 begins with the degenerate form: participation on the IX will not adversely affect any participant's revenue stream. Practical application of non-compete has been an exercise in exploring the compromise space,[339] identifying limited exceptions that nominally compete with participants' interests but are selected for their contribution to the IX as a common good benefiting the local interconnection market.

Neutrality is the third norm. Neutrality originated as carrier neutrality, but has since developed into a family of neutrality norms such as data center neutrality, connectivity neutrality, operator neutrality, and administrative neutrality. Carrier neutrality asserted IX facilities should not be housed in carrier facilities, but rather at a third party facility that provided non-discriminatory access to both participants' equipment and transport options to the facility. Neutrality in general aspires to ensure no single actor or constituency can undermine mutual management of IX facilities through control of either platform infrastructure or organizational constructs. Stated as such, neutrality limits parochial interests that may undermine mutuality. Neutrality is developed further in Section 6.4.1.2.

A brief comparison of the RIR and IX institutional complexes helps set the stage for discussing how IXes contribute to NRS integrity. IP address structure is hierarchical. In the simplest form, IPv4 addresses comprise a network component (the prefix) that can be uniquely mapped to an organization that holds basic appropri-

[338] In the course of fieldwork, especially early in on, multiple actors from multiple different IXes repeated this edict in various forms. In the terms of this work, the essential message of collective management and development of secondary rights was the common theme.

[339] Sections 6.4.1 highlights that relaxing non-compete has been an exercise in mutual (collective) decision making. Here mutuality is a form of collective decision making that has been explicitly codified in constitutional rules.

ation rights for the numbers in the host component. The delegation structure in Figures 5-4 and 5-5 is the administrative projection of this hierarchical structure of number resources and the stocks of number resources. That administrative structure and the meaningful rights bundles in Table 5.2 provide the foundation of the institutional mechanisms that order resource delegation.

In contrast, the IX institutional complex is best described as a decentralized set of interconnection markets. The IX firm manages the platform, as a non-discriminatory interconnection fabric. For associational membership based IXes, variants of the constitutional norms of mutuality, non-compete, and neutrality are established and maintained by the constituency of each IX. Taken individually, each IX provider is independently managed. Taken as an institutional complex, IX constituencies have substantive overlap, providing coarse evidence of a broader IX regime constituency.[340] In contrast to the hierarchical coordinating mechanisms in the RIR system, IX coordination occurs through indirect transmission of norms amongst common constituents and engagement in a variety of fora, in particular IX associations framed here as loosely federating agents. This decentralized structure is a projection of, and enhances access to, the global stock of routes.

As a comparison of CRIs, RIRs and IXes are instances of diverse governance modes. The IX and the RIR communities overlap substantively. These operational communities share common culture, a belief in consensus-based decision making as an operational epistemic community, and value-propositions dependent on the integrity of the NRS. That said, despite the overlap and commonalities, these actors do not impose either model as a "golden institutional hammer." This is in contrast to externally imposed rules rooted in authority structured around government or IGO models contending for power in downstream issue areas such as cybercrime and censorship. Part III will evaluate the interaction of internal enforcement mechanisms, how those tools support downstream regulation and enforcement, and the tensions between these.

Like the RIR chapter (5), the structure follows the conceptual framework for describing, explaining, and evaluating resource management facilities in the introduction to these studies (Part II). Section 6.1 summarizes IXes' contribution to the NRS, elaborating the notions of control and diversity introduced in Section 2.3. All IXes provision, at an abstract level, a common switching fabric, but have varying levels of complexity and sophistication. A typology of topology archetypes is presented in Section 6.2. Section 6.2.1 opens with a formal definition of IXes. Following the establishment of a common base terminology in Section 6.2.2, Section 6.2.3 tracks the evolution of increasingly sophisticated IX topologies that explain differences in deployment and foreshadow management rights decisions made by associational membership IXes.

Section 6.3 builds on these archetypes to describe types of participants, how they map on to topological archetypes in different IXes, and which have rights to

[340]Further evidence is presented in this chapter. Regional and global clustering can be seen in Figure 6-3 , where a hierarchical clustering algorithm was applied to patterns of common membership as the feature vector. The diversity metric, developed in Section 6.1.2 and depicted in Figure 6-2, shows the proportion of IX participants (constituents) that are members of other IXes.

exercise management rights (members) and which do not (customers). IX rules summarized in the previous paragraph are elaborated in Section 6.4. In additional to the constitutional norms of mutuality, non-compete, and neutrality, Section 6.4.2 describes the differences in the extent to which consensus is used to make strategic decisions and the potential trend to what is referred to as the "forgiveness versus permission" model, offering France-IX as a contemporary instance discussed in Section 6.5.2. To illustrate the decision processes that animated the topology evolution presented in Section 6.2, Section 6.5 describes debates over multinode deployment, intra-platform transport, diameter of the IX, membership types, and remote peering. These instances illustrate that application of collective choice (mutuality) rules to balance trade-offs amongst constitutional norms are not always in harmony.

6.1 IX Overview

IXes do not confer rights to number resources or routes. Recall from Section 3.3 the notion of a right as the claim to garner value from typically specific uses of a resource.[341] In a commonly managed resource system, the institutions and facilities enhance participants' ability to garner value from available resource units within community standards for maintaining system integrity. Stated as an optimization problem, the goal of the IX as a resource *system* is to maximize the *total potential* value that can be derived by *participants* at a common loci of interconnection. IXes are platforms that enhance the NRS by facilitating more efficient use of numbers and appropriation of higher value routes in a platform that has provisioned exchange capacity sufficient to handle resulting traffic amongst participants.

One class of resource units provisioned by the IX is the set of interconnection options available to participants. In the sense of a real option,[342] an interconnection option is the option, but not the obligation, to interconnect with any other participant on a common platform. An option is exercised when two network actors on a common platform choose to establish a BGP session in order to exchange routes and traffic. An interconnection option is uniquely identified by a pair of participants and the platform on which those participants are interconnecting.[343] Establishing an interconnection relation, i.e. exercising the IX-mediate interconnection option,

[341]Further recall that the claim is not always the simple, ideal form of "ownership" in which an actor has exclusive rights to garner *all possible value and the right to alienate subsets of those use rights*. Rather, it is a claim to a specific set of users.

[342]Real options applied to engineering systems and design are a way of modeling how to invest in flexibility in engineering system design and implementation. For instance, see (De Neufville, 1990). Here, the notion of an option is the *option, but not the obligation*, to make a particular asset investment, namely the provision of an interconnection relationship.

[343]The same pair of participants may exercise interconnection options over multiple platforms, for instance establishing a private cross-connect at a common facility (for instance a colocation facility) where they exchange substantive volumes of traffic and over IXes in other facilities where they want the benefits of a direct bilateral connection but their traffic exchange in those regions do not warrant a dedicated cross-connect. While these a unique instances of exercising the option, the aggregate traffic flows (across all options, at all facilities) are typically managed under a single contract.

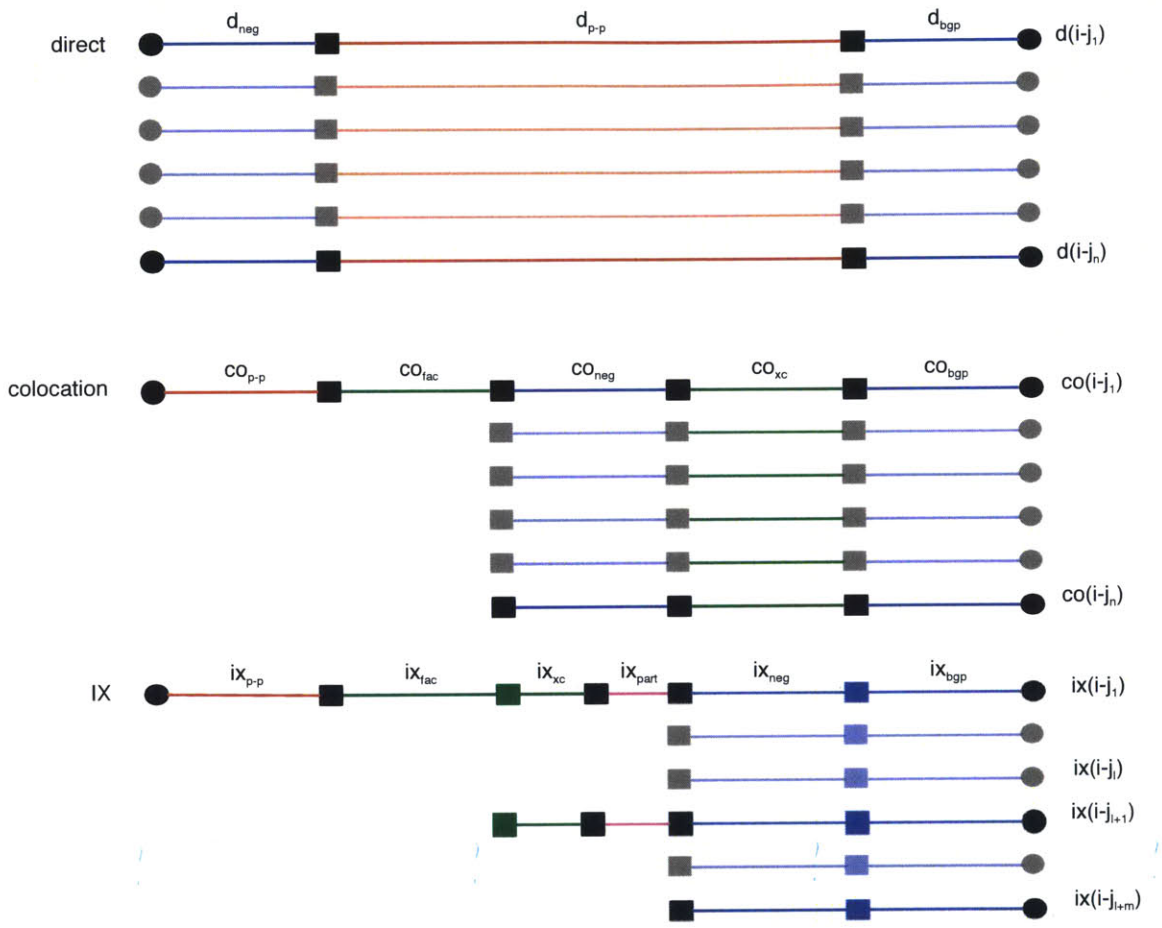

Figure 6-1: Three types of interconnection provisioning are presented: direct connections are denoted d, colocation co, and IX ix. Each horizontal line depicts the sequence of actions necessary to establish interconnection via that particular platform. Each step is denoted by the platform, subscripted by the type of action: for instance, establishing point-to-point connectivity to a colocation facility is denoted co_{p-p}. The actions are: neg, negotiation of interconnection contract terms; $p-p$, point-to-point transport; bgp, establishment of the BGP session between i and j; fac, establishing residence at a hosting or colocation facility; xc, establishing a cross connect at a facility; and $part$, establishing participation on an IX. Colors denote who interacts in a given decision: red indicates interaction with a transport provider; green is interaction with a colocation facility; purple is an interaction with an IX; blue signifies interactions solely between i and j. Circles represent start and end of interconnection relation development; square delimit the different actions.

transitions participants from nonsubtractive access to the market to withdrawing resources from that common platform. In the case of the IX, an interconnection relation consumes (withdraws) data plane capacity on the common interconnection fabric.

The IX platform is, like other congestable common resources, partially-rival: traffic will congest if the traffic participants attempt to exchange exceeds exchange capacity. A key function of the IX as a firm is to serve as a professional, neutral third party that monitors the "health" of the platform. In terms of congestion, health means ensuring there is sufficient capacity to handle average load, peak load, and the occasional traffic burst.[344] In common resource management terms, the IX monitors individual and aggregate utilization levels to ensure the interconnection platform is not exploited. Participants may exercise interconnection options with any actor with any willing partner, but each is limited to the total capacity contracted with the IX. The combinations of these limits and close monitoring of capacity growth allow the IX to make informed stepwise infrastructure upgrades necessary to avoid congestion. As a third party manager of a diverse platform, the IX firm is in a position to have much better information about regional congestion than all but the largest network actors in the region.

IXes contribute to both the NRS as a resource system and to the capacity of the data plane. Framing the health of the IX more broadly, as a CRI the IX is contributing to a functioning local interconnection market by providing loci of interconnection and by lowering barriers to participating in this market. Such markets foster increased production through concentrated access to diverse participants. Similarly, the IX fosters concentrated production of routes through provision of low-cost, general purpose interconnection options. The result is a class of interconnection platforms that foster generating unique interconnection relations and routes.

Ager et al. (2012, p. 2) report that more unique routes were observed in a single large European IX than previously estimated to exist in the entire Internet.[345] Building on this volume, given the option, it seems network actors will invest in appropriating higher quality routes. As alluded to in Section 2.3, the logic behind these markets is that IXes facilitate greater control routing by increased access to this diverse market of interconnection options. For instance, CDN operators are a class of network actor that has invested in diverse, widespread portfolios of IX

[344]For instance, the LINX performed substantive capacity planning and testing in preparation for the 2012 Olympics in London.

[345]The full quote reads:

> in terms of AS links of the peer-peer type that are typically established among member AS pairs we show that this IXP has close to 400 members which have established some 67% (or more than 50,000) of all possible such peerings and use them for exchanging some 10 PB of IP traffic daily. *To put this number in perspective, note that as of 2010, the number of inferred AS links of the peer-peer type in the Internet was reported to be around 40,000 less than what we observe at this particular IXP alone!* (Ager et al., 2012, p. 2, emphasis in original)

It is very likely that many more than the routes greater than those reported in the Internet are unique to that IX.

participation. As depicted in the inset of Figure 6-2, three of the six network actors participating in more than 40 unique IXes are three of the worlds largest CDNs.[346] While this logic seems to hold for established IXes, IX deployment in new regions has historically needed substantive community building.

In terms of local bandwdith, IXes generate capacity at the interstices of network connectivity, where congestion often occurs.[347] Part of IXes' growth is driven by the aggregate demand for exchange capacity from its participants.[348] A combination of the membership model, cost-neutral infrastructure management, and participant capacity management contribute to ensuring IX capacity effectively tracks gross market demand. The introduction of an IX shifts traffic from transit and long-haul transit to remote exchange facilities to local IX-mediated interconnection. On balance, shifting traffic from transit to uncongested IX interconnection increases overall and local capacity.

6.1.1 Interconnection Options in a Commonly Managed Switching Fabric

As will be developed in more depth in Section 6.2, an IX comprises one or more commonly managed interconnection platforms. The community vernacular refers to these platforms using a variety of terms: "the switching fabric," "public exchange," "network access points (NAPs)," "exchanges," and/or "Internet exchange *points*." To illustrate the resource units provisioned by these platforms, the nuanced differences between these labels will be abstracted away. This section builds on the simple definition in Section 2.3. An interconnection platform is a loci of physical connectivity available to or more networks that have built into that facility. Further recall "building in" typically means a network i has provisioned transport to an interconnection platform p in order to establish interconnection relations with other networks also present at that facility. Three different types of interconnection platform are considered: transport, colocation, and IXes. Interconnection on these platforms is referred to as transport-mediated, colocation-mediate, and IX-mediated interconnection, respectively. The differences in these interconnection mediation structures are depicted in Figure 6-1. These are described in terms of private and common[349] platform provisioning in Sections 6.1.1.1 and 6.1.1.2.

[346]Netflix is also gaining ground, deploying in IXes in Europe as it expands its market into the EU.

[347]A recent report by Weller and Woodcock (2013) also argues IXes are source of increased bandwidth in a given market.

[348]Other sources of growth are latent demand in less densely developed regions outside the metro-region a platform has historically serviced and demand for IX platforms in other "remote" metro-regions. These growth patterns are discussed throughout the chapter. Topological archetypes of these growth patterns are discussed in Section 6.2.3. Common management issues, related to the management of these growth trends, framed in terms of IXes constitutional norms, are discussed in Sections 6.4.1.

[349]In community vernacular, IXes are a "public" peering fabric. It is public in that is collectively provisioned. A number of arenas in the NRS use the term public where common is the more accurate term. This mislabeling is not exclusive to the network operator community. Rather, it is frequently observed in descriptions of jointly managed resources. See (McKean, 1996, loc. 4229–4240) for

6.1.1.1 Private Interconnection

Transport is the simplest loci of physical connectivity, comprising a link with dedicated capacity between a network i and some other network j. Consider the interconnection relations depicted in Figure 6-1. Establishing an interconnection relation, here a transport relation, comprises *a*) upfront negotiation of a contracting mode between i and j (d_{neg}), *b*) provisioning the point-to-point link (d_{p-p}), and *c*) establishing the BGP session (d_{bgp})) over which routes are exchanged by i and j. The link d_{p-p} is privately provisioned—once provisioned the capacity is available exclusively to the pair i and j. for traffic they choose to exchange. The link is finite and congestable at either end. In terms of capital investment, d_{p-p} may be provisioned by i, j, or jointly by i and j. In the simple interconnection narratives from Section 2.2, the links between the ASes in SimpleNet (Figure 2-1) are basic transport links.

Transport is a bilateral relationship built atop dedicated, private capacity shared only between i and j.. As a platform, transport directly between the facilities of two actors is the degenerate case. The only "option" available is interconnection between those actors i and j.. For i to establish an interconnection relation with j_2, j_3, on through j_n, it must establish additional transport capacity from its facilities to the facilities of j.. As such, transport links are *specific* investments whose capacity cannot be used to support other contractual relations.[350] Not only are these specific investment, but the typical minimum capacity available is more expensive than that capacity via other modes of interconnection.

The degenerate transport relation requires a dedicated bilateral relationship between every pair of interconnection partners. In contrast, interconnection platforms are *loci* of interconnection amongst a diverse set of network actors. Colocation facilities are commercial firms that furnish housing for network elements (such as routers and switches) and server equipment. Economies of scale in colocation facilities' services contribute to reducing the costs of interconnection by creating a privately managed, typically carrier-neutral, facilities participants build into in order to establish *dedicated* bilateral interconnection relation with others. Professional colocation services include *a*) ample power, including backup; *b*) air conditioning, often including backup; *c*) various types of cross-connect cabling; *d*) space for routers and servers, typically in a hierarchy of units starting at the single rack unit, an entire rack, a cage comprising multiple racks, or suites for especially large tenants; *e*) "smart hands" labor for various simple maintenance tasks; *f*) cages and various other access control to ensure security between tenants (participants); and

---

discussion.

[350] Strictly speaking, specific *assets* have a very limited range of uses. Transport for the purpose of a single bilateral interconnection relations is exclusive to i and j.. For instance, if the link between i and j_1 is congested (capacity is completely consumed), i cannot reallocate capacity from its link with j_2 to reduce the congestion with j_1. Although transport itself is not intrinsically a specific asset, but this provisioning of transport is a specific *investment*. Transport providers offer contracts to provision point to point transport. Under the contract, these point-to-point links remain specific assets, may be upgraded without additional physical provisioning, but short of the situation in which j_1 and j_k reside at the same facility, capacity cannot be transferred.

g) external security for the facility itself. Establishing residence at a colocation facility requires the purchase of some amount of space (at minimum however many units of rack space necessary for i's routing equipment and/or servers). Depending on the price structure and product bundles, this may include costs for power, AC, the type of cage, etc.

Consider the series or interconnection relations labeled colocation ($co_.$). Transport is still required to build into a colocation facility (co_{p-p} in Figure 6-1). The cost of establishing a presence at the facility (rack, cooling, power, etc.) is represented by co_{fac}. Negotiation is denoted co_{fac}, cross-connects are denoted co_{xc}, and the BGP session is denoted co_{bgp}. Under colocation-mediated interconnection, once i builds into a colocation facility, it may negotiate with other participants to establish interconnection through the provisioning of one or more cross-connects.

In the context of colocation-mediated interconnection, transport becomes a more general asset, supporting potential interconnection relations between i and $j_1 \ldots j_n$. In colocation-mediated interconnection, the cross-connect (or a set of cross-connects depending on capacity needed) is the specific asset provisioning capacity dedicated to route exchange and traffic between i and j's equipment at the colocation facility. Transport is now a more general investment. In this context, cross-connects are the specific investments in the same sense that transport is potentially general, but specific in context (when provisioned for dedicated, bilateral transport-mediated interconnection). It is important to highlight that asset specificity here is a consequence of how links are provisioned, in what context, and the length and conditions of the contract.

Contrast contextual specificity with intrinsic specificity. A canonical instance of intrinsic specificity is a capital intensive fabrication facility that, while efficient in the production of a particular product, is costly to modify to produce a broader variety of goods, even within the same family of goods, such as automobiles or computer processors. Like d_{p-p}, cross-connects' capacity is dedicated to the bilateral interconnection between each pair i and one of $j.$; it is not fungible across interconnection relations. A cross-connect may be as extensive as cabling between two buildings at a large colocation facility or as simple as a six foot piece of fiber between adjacent racks.[351] Colocation facilities pre-provision racks, power, etc. that are included in the facilities contract and costs (co_{fac}).

Once beyond the same room, colocation facilities typically design meet-me rooms where actors can interconnect with a variety of other actors hosted across the facility. That said, cross-connects are not necessarily provisioned ahead of time, especially in the case of links between relatively distant rooms or floors in a given facility. Once i and j_k complete interconnection negotiations (co_{neg}) a request for a cross-connect is placed with the colocation facility. The colocation performs the

[351]These may be any number of media such as fiber optic cable or copper ethernet cabling. Cross-connects provisioning is implemented over a range over of cabling and infrastructure arrangements: *a*) a piece of fiber less than six feet long, between different participants network elements in adjacent racks; *b*) between racks across the room from another; *c*) between rooms on the same floor of a building; *d*) between rooms on different floors of the same building; *e*) between nearby buildings of the same colocation facility.

physical provisioning, running the cable from the location of i's equipment in the facility to the location of j_k's equipment.[352] Depending on the contract, the cross-connect may be financed by i, j_k, or jointly by the two.

Both transport-mediated and colocation-mediated interconnection are forms of "private" peering. In both cases, all assets involved are physically provisioned by, owned by, and contracted out for use by private actors.[353] Used in individual interconnection relations, as per the description of transport above, their use is exclusive to the provisioning parties and is partially-rival to the point of congestion. Private provisioning requires actors bear the costs of provisioning (or contracting) as well as the potential uncertainty of whether the portion of capacity provisioned warrants the specific investment.[354] The next section describes the last component of Figure 6-1, IX-mediated interconnection and the IX as a jointly provisioned loci of physical connectivity.

6.1.1.2 Jointly Provisioned Interconnection

IXes comprise nominally private assets described in transport- and colocation- mediated interconnection, but differ in terms of how the direct link by which participants exchange routes and traffic is provisioned and by whom. IX-mediated interconnection, $ix.$ in Figure 6-1 illustrates the mechanics of an interconnection option in terms of asset utilization. The simplest case of an IX, a single switch in a single colocation facility, will be presented to focus on disintermediation rather than IX topology. Section 6.2 unpacks the variety of archetypal topologies that have been created based on these building blocks. All but one of the $ix.$ components in IX-mediated interconnection, ix_{part}, have counterparts in the other interconnection modes.

Consider asset provisioning and utilization in IX-mediated interconnection in Figure 6-1. Like colocation-intermediated, i must first build into a colocation facility (ix_{p-p}). Rather than contracting the provision of a cross-connect to j_k's equipment (such as co_{neg} followed by co_{xc}), i instead provisions a cross-connect to the IX's switching equipment (ix_{xc} to ix_{part}). Like the disintermediation that made transport a more general asset in colocation-mediated interconnection, in IX-intermediated interconnection, the cross-connect is disintermediated to a more general asset class. In contrast to individually provisioned cross-connects, the switching fabric provisions a fabric such that there exists a direct layer 2 link between every pair of ports on that fabric.

[352]The cross-connect cabling may or may not be connected to residents' equipment by colocation facilities staff. Equipment security has various levels of physical security. For instance, some equipment is kept in "cages" that are literally what they sound like, dedicated areas that only resident staff have rights to access. In other settings, equipment may be in simple racks, accessible to any resident in the facility. Some colocation facilities provide what is called "smart hands," staff that can perform tasks such as connecting additional cross-connects directly to residents equipment or making other relatively minor adjustments that might otherwise require costly visits by residents' staff.

[353]This is generally the case, there are certainly possible exceptions, such as government managed colocation facilities and/or transport infrastructure.

[354]Different actors have levels of uncertainty regarding traffic with their peers; this will be discussed in the next section (6.1.1.2).

Under transport- and colocation-mediated interconnection, either d_{p-p} or co_{xc} was a dedicated, specific asset. Under IX-mediated interconnection, all assets utilized for establishing interconnection are general assets. The link established via ix_{p-p} and ix_{xc} may carry traffic from i destined for *any* IX participant j.. In terms of cost, ix_{part} comprises IX membership fees and recurring fees for provisioned port *capacity*. In contrast to transport and cross-connects, ix_{part} is not a point-to-point connection. Rather, it is what some have referred to as a "multipoint" connection. In the terminology developed here, ix_{part} represents the appropriation of access to the set of interconnection options provisioned by the IX.

In most cases, the presence of the path and the definition of an interconnection option are consonant: the path neither guarantees, nor obligates, any participant to interconnect with any other participant.[355] The switching fabric itself is commonly provisioned by the IX firm, under the aegis of the IX constituency. The switching fabric is financed by participant fees, which pay for both capital investments (switching and routing equipment) and operational costs, such as maintenance by IX staff (or volunteers). As a part of capital investments, the IX is upgraded as necessary, based on trends in participant utilization, by the management of the IX firm.[356] Even though the IX switching equipment is legally owned by the firm, recall the single stakeholder in an associational membership IX is the collective membership.

A subtlety here is the relationship between exercising an interconnection option and capacity. The IX provisions interconnection options and capacity as two distinct, but interdependent resource units. Interconnection options are, in and of themselves, non-rival. Exercising an interconnection option does consume router capacity and a small amount of bandwidth.[357] Establishing a BGP session consumes very little capacity. Operator labor is consumed and, depending on the complexity of the relationship, can vary substantively. Short of an environment that is already congested or that is very near capacity, establishing interconnection options is non-rival.

Consider $ix(i - j_l)$ and $ix(j - j_{l+1})$ in Figure 6-1. Network i has provisioned a cross-connect (via step ix_{xc}) as a general asset to exchange traffic with participants on the exchange using appropriated port capacity c. Network i has interconnection relations with j_1 through j_l. The traffic exchanged with $j_i \ldots j_l$ under these relations has average peak consumption at approximately 90% of c. Actor i *could* technically establish an interconnection relation with j_{l+1}—the capacity for the BGP session

[355] Some IXes impose rules such as forced multilateral peering. Instances of this include CABASE and PTT.br. These instances are discussed in Section 6.4.3.

[356] In general, this financing structure is the case, but associational membership IXes do differ in the details. Some, such as the SIX in Seattle, require only a one-time sign-up fee or one-time port fee. Others require recurring fees membership fees. Many of those in this study require recurring port fees. Actors in the one-time fee category tend to be ad hoc volunteer organizations. A recurring theme has been the professionalization of services amongst IXes, one element of which is recurring fees as a revenue stream supporting the income of IX staff. In effect, a revenue stream ensures there are staff participants can hold accountable for IX function.

[357] Note that exercising the interconnection option comprises the completion of the series of activities in each of phase of IX-mediated interconnection in Figure 6-1. In most cases, the critical step is negotiation over the contracting mode between the interconnecting actors.

is not significant. That said, if *a*) peak traffic coincides with the peak traffic of i's other peers, *b*) the traffic between i and j_{l+1} is greater than 10% of c, then *c*) exchanging traffic with j_{l+1} would exceed c, causing congestion for i. In order to exchange traffic with j_{l+1}, i must upgrade its general investment at ix_{xc} from a single cross-connect to two cross-connects, increasing capacity to $2c$.[358]

Note that in this scenario, i would likely see congestion during peak traffic periods if it established an interconnection relation with j_{l+1}. Nothing is said about whether j_{l+1} sees congestion. To play out the scenario a bit more, assume i and j_{l+1} did interconnect, but i did not upgrade its port capacity. Further assume that actor j_{l+1} is a large CDN or content provider that knows it is seeing growth over the next year and has optimized by "over-provisioning" its capacity at all IXes over some time horizon, say two years, rather than face the transaction costs of incremental upgrades. In this case j_{l+1} does not see congestion on *its* port, but *will* see dropped packets at i's hop in a traceroute or by monitoring segment retransmits.[359] In this case, it is likely clear to both j_{l+1} and other actors exchanging traffic with i that it now has a congested port.[360]

Nominally, an actor can exercise the interconnection option with any set of actors willing, but is limited by capacity available to exchange corresponding traffic. Route collectors are an instance of an exceptional, albeit degenerate case in which routes are exchanged, but not traffic. Research firms such as PCH, interconnectivity intelligence firms such as Renesys, and NRS firms such as RIRs all deploy route collectors at IXes. Route collectors exercise interconnection options with any actors willing, on the basis that these collectors will not exchange traffic over the link, but rather are only interested in routes provisioned over those options.

As resource units in the NRS framework, in terms of Ostrom's common rights holders, route collectors are authorized entrants. Appropriation of interconnection options without traffic exchange is effectively non-subtractive access.[361] Appropriation of interconnection options and attendant traffic exchange is effectively non-subtractive in the control plane but partially rival in the data plane. In the case of the latter, the commonly provisioned exchange has a finite capacity at any given point in time. Moreover, it is possible that port capacity provisioned is greater than the exchange's switching capacity. The IX firm is responsible for monitoring capac-

[358]This assumes the capacity c was a single port capacity, say a 10G port. It may well be the case that i had a bundle of 4 10G ports, giving $c = 4 \times 10G$ and the new capacity, $c' = 5 \times 10G$, increasing capacity by 25%. Another option may be that j_{l+1} and subsequent relations will need to satisfy substantive growth, perhaps it is i that is a CDN that has just landed a major content originator. Actor i may upgrade its equipment and replace its $4 \times 10G$ with a single (or even multiple!) $100G$ ports, an increase of 250% in the case of a single $100G$ port.

[359]Depending on the flavor of TCP and options, retransmits can be monitored continuously by keeping track of various types of ACKS coming across an interface or set of interfaces.

[360]Note that it is likely, but not absolutely certain. For instance, if i's aggregate peak is, say, 105% of c and i prioritizes particular sources, some actors may see dropped packets, an indicator of congestion.

[361]Recall the discussion of route servers in Section 3.4.1's discussion of rights of entry; only the small amount of traffic necessary for exchange is consumed. Under modern capacity levels, the volume for route exchange is considered an epsilon—it exists but by itself is not significant.

ity. The firm measures capacity allocated, capacity utilized, and patterns in traffic volume. This information is used to adequately configure and upgrade platform infrastructure as necessary to meet projected demand.

6.1.2 IX Diversity

IXes are interconnection markets in which networks develop diverse strategic interconnection bundles. An interconnection bundle comprises a set of interconnection relations. In terms of a network actor i's stock of routes, the interconnection bundle is the set actors from whom i *may* appropriates routes. Diversity is a family of indicators that describe existing and potential interconnection relations available to i at a given set of interconnection platforms. Conceptually, diversity is what it sounds like: network i benefits from a general asset providing access to a variety of actors. Access facilitates strategic selection amongst these in an effort for i to maximize its value proposition.

Greater diversity does not *guarantee* either greater potential or realized value. The premise of strategic selection is that one has the option to select from routes of various qualities (lower latency, stability, particularistic contractual relations with peer). This is in contrast to simple transit, which provisions the route that is optimal for the transit provider, not the downstream appropriator. Moreover, as long as the transit provider offers routes to the entire Internet, it may rescind one route and provision another from its stock that may reduce its costs but is not necessarily a good fit for its downstream. Here, the strategic value of developing interconnection bundles in a more diverse market is that routes can be selected contingent on i's value proposition. For example, an access network will develop a different interconnection bundle than a CDN. These will be different still from a network specializing in the hosting of financial transactions and different still from a network hosting live action gaming servers.[362]

To illustrate interconnection bundles and develop the notion of diversity, consider simple transit, described in Section 2.3.2.1. Simple transit is a bundle comprising one interconnection relation, between i and a transit provider. As per the discussion of simple transit in Section 2.3.2.1, i has outsourced routing decisions to the transit provider. Network i is beholden to that transit provider's lowest cost routing decisions as to which routes i traffic will take to various destinations.

Upgrading to multihoming, i.e. establishing connectivity with n transit providers, creates a bundle of n interconnection relations, $t_1 \ldots t_n$. Network i has nominally diversified its interconnection bundle. Network i now has a few options: if transit provider t_1 provisions a poor quality route[363] to a desired destination, one of $t_{k\neq 1}$ may provision (offer in community vernacular) a better quality route. Network i's

[362]These comparisons do not imply that each of the different classes of network actor listed will have the same interconnection bundle. Rather, within each class of network actor, instances have similar value propositions and similar *objectives* when developing interconnection relations. These will be developed more thoroughly in the discussion of uniqueness and redundancy being some of the simplest, perhaps irreducible, objective functions when developing interconnection bundles.

[363]Poor quality may be high latency, congestion, frequent intermittent failures, etc.

interconnection bundle gives it a set of options from which to choose. All of n of i's interconnection relations are transit relations $t.$, all of whom make route provisioning decisions based on lowest cost routing decisions that optimize $t.$'s costs (respectively) across all of it's transit customers. In effect, while i does have more options, these options may change as each of $t.$ selects amongst the routes it wishes to provision for i. Network i must actively monitor these routes if the value of the gain in quality is greater than the measurement and transaction costs.

Diversity indicators for IXes are a general indicator of the potential supply of routes available from a network actor, a platform, or a bundle of platforms. Like markets in general, different markets have different quantities and quality of goods. Various operationalizations of diversity will be discussed here, in particular simple platform diversity, network diversity, and compound platform diversity. These operationalizations are depicted in Figure 6-2, Table 6.1, and Table 6.2.

Simple platform diversity is the number of interconnection options available to a participant on a particular platform. Recall an interconnection option is characterized by the identifiers of the two participants (typically ASNs) and the interconnection platform. Simple platform diversity is effectively the number of participants on a particular platform. As may be obvious, transport, as a degenerate interconnection platform, always has a diversity score of 1.[364]

Simple colocation diversity is the number of customers at a colocation facility. *Simple IX diversity* is the number of IX participants in a particular IX platform. Simple IX diversity (number of participants), along with total volume of traffic exchanged on the platform, are the two most frequently cited (and easiest to comprehend) comparative indicators of IXes' size. As will be developed further in Section 6.2, IX platform infrastructure is often housed in several colocation facilities spread across a single metropolitan area. It is possible that simple colocation diversity and IX diversity are the same. In many cases, an IXes common fabric comprises equipment distributed over more than one colocation facility.

In community vernacular, the more participants at an interconnection facility, the more "sticky" that facility is. When exercising colocation or IX options as depicted in Figure 6-1, each step incurs operational and/or capital costs. From the perspective of i, the greater the number of options exercised, the greater the capital and operational investment at an interconnection platform. The prospect of leaving that platform means potential loss of operational, if not also capital, investments. Leaving may also require identifying and developing suitable substitutes for the rescinded interconnection options.

Stickiness is usually applied as a characteristic of a platform p. A nominal indicator of stickiness is simple platform diversity. Qualitatively, a platform with sought after participants, such as large access networks, CDNs, and/or transit providers is considered more sticky than a platform comprising a larger number of generally

[364]Transport may provide connectivity to a colocation facility, or an IX hosted at a colocation facility. In those cases, diversity binds to the last platform over which the BGP protocol is transmitted in Figure 6-1 before the session is established. For instance, for colocation, this is co_{xc}. In the case of an IX, this is ix_{part}, the switching fabric of the IX.

less sought after smaller networks.[365] A network's diversity score is one, albeit incomplete, measure of the extent of a network's investment in developing a diverse interconnection bundle. The *network diversity score*, denoted $d_X(i)$, is the total number of interconnection platforms i participates in out of the set of X platforms of a given type (transport, colocation, or IX). In this work, network diversity will typically be limited to IX participation.[366] A network is considered to be diverse if $d_X(i) > 1$. Given the overlap between IX participation and colocation participation, a network's diversity score is often conditioned on a network's participation in IXes or colocation facilities.

A distribution of simple IX diversity and network diversity is shown in Figure 6-2. The three largest IXes in terms of simple diversity are the AMS-IX (Amsterdam), the LINX (London), and DE-CIX (Frankfurt). Note these each have a large proportion of participants that participate in at least 10 other IXes (signified by the range of yellow to red). Note the next two IXes in terms of simple diversity, PTT.br (Brazil) and MSK-IX (Moscow) both have significantly smaller sets of highly diverse ($d(i) \gg 10$) networks. Both PTT.br and MSK-IX are regional IXes; as such, the majority of their participants are regional actors participating in 1–3 IXes.

To further illustrate regional interconnection patterns, consider the hierarchical clustering of IXes in Figure 6-3. This clustering is based on the number of participants common to any pair of IXes documented in Euro-IX's Peering Matrix (Euro-IX, 2013a). The three largest IXes cluster together—these actors have a large number of diverse participants from all over Europe and globe. A number of the medium-sized IXes cluster by geography. Geographic clusters correspond to *a*) Italy, *b*) Great Britain, *c*) the Caribbean, *d*) Finland, *e*) Russia, *f*) Japan, *g*) France, *h*) Germany, *i*) Scandinavia, and *j*) Switzerland.

Consider the range of IXes with approximately 50–150 participants in Figure 6-2. For IXes in this range, a greater proportion of their participants have a diversity score of greater than 8 in comparison to the other groups of IXes. Consider this range in conjunction with the geographic clustering in Figure 6-3. A working hypothesis is that the higher density of diverse network actors in the 50–150 participant range in Figure 6-2 are regional value networks, such as local CDNs and local access networks. This working hypothesis has been confirmed in private conversations during fieldwork and in selected follow-up interviews.[367]

[365] It should be stressed that this is generally the case. For actors whose value proposition is based on local traffic, platforms comprising small, regional networks may be more valuable than larger networks that may not service those regions. For instance, this is the case for regional CDNs and actors hosting local e-government services.

[366] Given the overlap of transport, colocation, and IX participation, a network diversity score in which X comprises all would over-count participation. In general, X will be the universe of known IXes or, in the case of Figure 6-2, the IXes documented in Euro-IX's database. In more sophisticated analyses, X may be limited to the IXes already in a particular network actor's interconnection platform bundle and those it is considering deploying to over some time horizon or in a particular region.

[367] As a means to test the hypothesis, exploring this hypothesis with the operational community is a an effort to leverage tacit knowledge of the market to interpret patterns in the data. Ongoing work is attempting to further quantify this hypothesis with an eye for identifying regions where IX

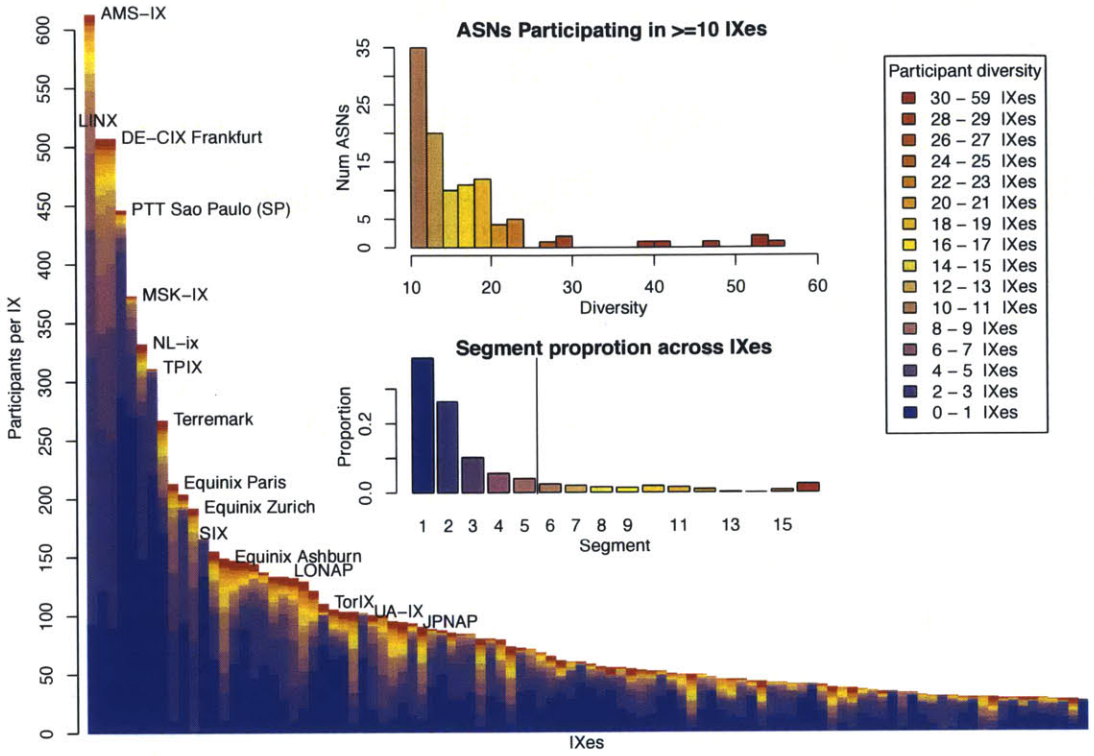

Figure 6-2: Each bar in the distribution is the number of participants at that IX, the simple IX diversity. Each bar is stratified into the number of participants that are participants at another IX, the diversity of the network actor. For instance, the dark blue strata represents networks that participate at only one IX ($d(i) = 1$). Moving along the spectrum, from blue to red $d(i)$ increases. The data in this graph is from October 2013.

Composite platform diversity, builds on network diversity to offer an indicator of platform, and by proxy market, quality. *Composite platform diversity* is the sum of the network diversity scores of participants that are considered diverse ($d_X(i) > 1$). Table 6.1 lists the top 40 IXes, ranked by their diversity score. Note that the first three, AMS-IX, LINX, and DE-CIX, correlate with their simple diversity metric, the number of participants. Moving beyond the first three, the diversity ranking and the participant ranking diverge. For instance, PTT.br and MSK-IX are ranked third and fourth in participation, respectively, but are ranked 35^{th} and 39^{th} by diversity, respectively. Again, a number of private conversations and interviews confirm this ranking matches those operators' intuitive ranking of IXes' "importance."

The notion of diversity helps illustrate differences in the types of interconnection bundles network actors develop. Thus far the unit of analysis has been the interconnection platform. Platform diversity, both simple and composite, are indicators of the stickiness of a *market*. Unpacking this aggregate indicator, network actors look at more than just the popularity of the market. Network actors also look at what

deployment may be beneficial.

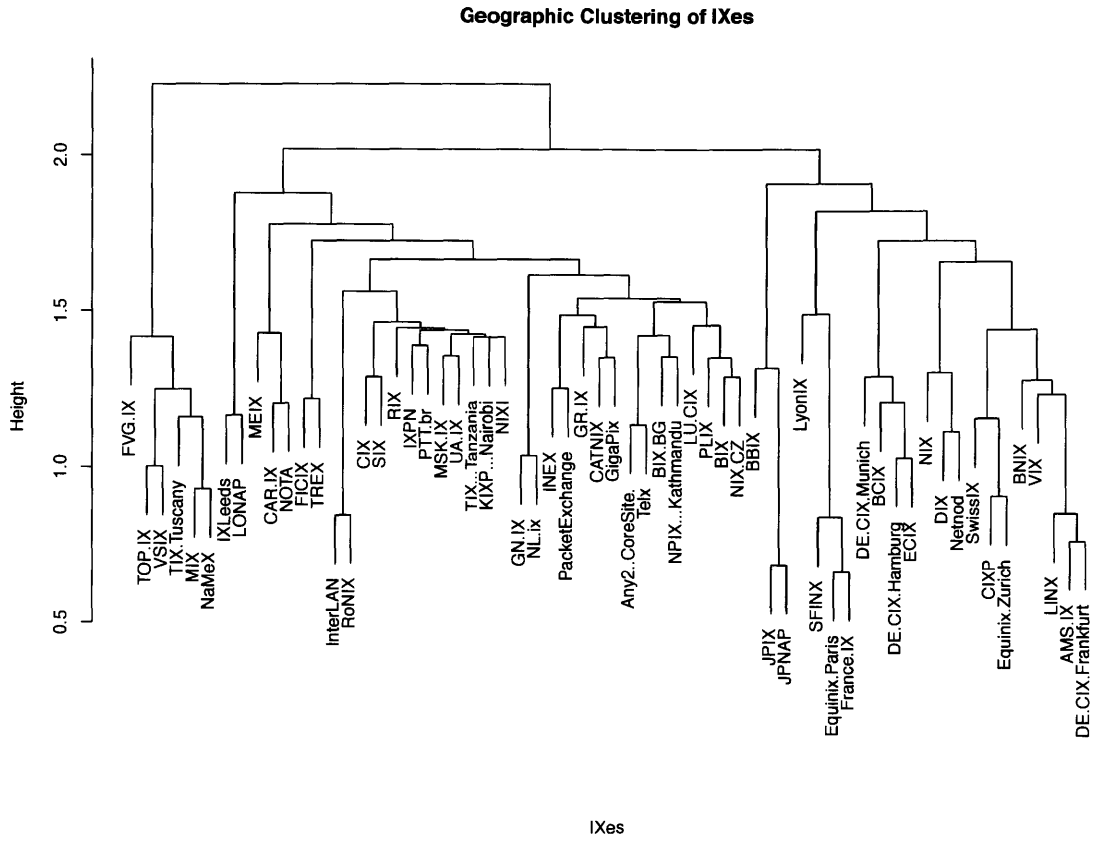

Figure 6-3: Euro-IX provides what is referred to as the "Peering Matrix" (Euro-IX, 2013a). Each side of the matrix is a vector of IXe s whose participation is documented in the Euro-IX database. Each cell in that matrix is the number of members common to that pair of IXes. Using the rows of this matrix as feature vectors, this dendogram shows how IXes clusters based solely on the number of common members. What is intuitive, but verified by this analysis, is that IXes cluster by region and scope. The data shown here is from May 2012.

goods are on offer in the market. As such, the general interpretation of an IX with a high composite diversity score is that its participants offer a diverse set of routes that are valuable to a diverse set of actors.[368]

Network actors have general criteria for selecting amongst interconnection platforms based on the characteristics of platform participants. Interviews indicate two general objective functions: uniqueness and redundancy. These objective functions are conceptually distinct but not necessarily mutually exclusive in application. Redundancy is intrinsically relative to the routes already in a network actor's local stock. Prefix *redundancy* means a subset of the prefixes available on a particular

[368]This is a general interpretation. Sowell (2013) provides a more in-depth discussion of the family of diversity indicators, in particular, that these may be conditioned on a set of IXes, for instance those in one of the regions depicted in Figure 6-3, to determine which actors have greater local diversity relative to actors that may have greater global diversity, but little local presence. For instance, Comcast has substantive global diversity, but it is concentrated in a densely connected US.

Rank	IX Name	Simple Diversity	IX Diversity
1	AMS-IX	613	3317.00
2	LINX	507	2884.00
3	DE-CIX Frankfurt	507	2869.00
4	Equinix Ashburn	149	1555.00
5	Equinix Paris	213	1404.00
6	NYIIX	129	1343.00
7	Terremark	267	1320.00
8	Equinix Zurich	192	1315.00
9	NL-ix	332	1274.00
10	Equinix Palo Alto	103	1142.00
11	France-IX	132	1068.00
12	Equinix Chicago	90	1041.00
13	CoreSite - Any2 Los Angeles	147	1034.00
14	Equinix San Jose	95	998.00
15	SIX	155	996.00
16	PTT Sao Paulo (SP)	446	993.00
17	Telx	94	925.00
18	NetNod Stockholm	80	900.00
19	Equinix Los Angeles	73	890.00
20	VIX	121	874.00
21	MIX-IT	133	859.00
22	LONAP	133	826.00
23	MSK-IX	373	794.00
24	SwissIX	146	793.00
25	SFINX	100	782.00
26	JPIX	144	740.00
27	PLIX	204	679.00
28	Equinix Dallas	49	667.00
29	Equinix Singapore	79	648.00
30	HKIX	103	633.00
31	ECIX AMS/BER/DUS/FRA/HAM	100	612.00
32	NIX.CZ	105	610.00
33	LAIIX	43	581.00
34	KleyReX	137	550.00
35	Equinix New York	38	532.00
36	BCIX	54	503.00
37	Telx - New York	24	501.00
38	Equinix Sydney	85	491.00
39	TPIX	311	489.00
40	Equinix Tokyo	37	470.00

Table 6.1: Ranking of IXes based on their diversity score $d_{\mathcal{X}}(x)$ in contrast with the number of members. Note that all four of the membership size based categories are represented in this set.

platform are in i's local stock, having been appropriated on other platforms.[369] Re-

[369]Prefix redundancy can be further refined to indicate whether these are prefixes appropriated

Rank	Colocation Name	Simple Diversity	Colo Diversity
1	Telehouse London (Docklands North)	289	3185.00
2	Equinix Ashburn (DC1-DC11)	215	2990.00
3	Equinix Frankfurt KleyerStrasse (FR5)	197	2720.00
4	Telehouse Paris 2 (Voltaire)	184	2582.00
5	Equinix Chicago (CH1/CH2)	147	2480.00
6	CoreSite - LA1 - One Wilshire	201	2437.00
7	TelecityGroup Amsterdam 2 (South East)	174	2418.00
8	Equinix San Jose (SV1/5)	137	2290.00
9	Telehouse London (Docklands East)	183	2218.00
10	Terremark Miami	129	2100.00
11	Telx New York (111 8th)	99	1971.00
12	Equinix Palo Alto (SV8)	110	1902.00
13	Equinix Singapore	140	1852.00
14	Equinix Los Angeles (LA1)	98	1821.00
15	MEGA iAdvantage Hong Kong	123	1791.00
16	Equinix Dallas (DA1)	91	1781.00
17	NIKHEF Amsterdam	127	1681.00
18	TelecityGroup London (HEX67)	116	1658.00
19	TelecityGroup Stockholm 1	83	1535.00
20	SARA Amsterdam	98	1530.00
21	TelecityGroup London (Sovereign House)	100	1517.00
22	Westin Building Seattle	123	1497.00
23	InterXion Vienna (Wien)	77	1452.00
24	TelecityGroup London 2 (HEX89)	81	1450.00
25	Telx New York (60 Hudson)	93	1433.00
26	Sitel Prague / CE Colo Prague	76	1394.00
27	Equinix New York (111 8th)	60	1244.00
28	InterXion Frankfurt 1	71	1231.00
29	Telx Atlanta	87	1214.00
30	InterXion Madrid	48	1174.00
31	Global Switch (Amsterdam)	74	1082.00
32	Equinix Paris Saint-Denis (PA2)	72	1076.00
33	TelecityGroup Amsterdam 1 (Science Park)	64	1073.00
34	Interxion MRS1 (previously SFR Netcentre)	33	1028.00
35	Equinix Tokyo (TY2)	38	1004.00
36	Equinix Sydney	93	998.00
37	Netscalibur Milan	41	992.00
38	InterXion Duesseldorf	64	991.00
39	Equinix Hong Kong (HK1)	40	969.00
40	InterXion Stockholm (Kista)	37	957.00

Table 6.2: Ranking of colocation facilities based on their platform diversity score $d_{\mathcal{X}}(x)$ in contrast with the number of participants. The data presented here is based on PeeringDB.

under a transit relation or some other interconnection relation. Given transit routes may change as

dundancy can be refined to include the degree of overlap in the AS path of the corresponding routes.[370]

Uniqueness is a characteristic of the prefix *and* the route—*uniqueness* means either the route itself or sub-path of the route is unique relative to routes to that prefix in the network actor's stock or observed in other interconnection platforms or relations. Uniqueness can be universal or relative. In both cases, the fundamental concept is whether a route is provisioned elsewhere, i.e. how unique is this route? *Universal uniqueness* can be framed as the proportion of platforms in which a particular route may be appropriated. For example, a desired route may only be provisioned at one or two regional IXes, otherwise one must acquire supersets of that route from transit providers. More common is *relative uniqueness*, whether a desired route is available in i's local stock. Given most networks have at least one transit relation in its stock, uniqueness is often conditioned to mean a route appropriated from a non-transit source.

Uniqueness can also apply to IXes. Operationalized to an IX, uniqueness is a measure of how many participants at an IX interconnect only at that IX. Reconsider the top five IXes in Figure 6-2. In Figure 6-2 IX uniqueness roughly corresponds to "how blue" a particular IX's bar is; narrowing this, singleton uniqueness (participant interconnects on only one IX) the relative height of the blue bar signifying participant diversity of $d(i) = 1$. The blue region of PTT.br and MSK-IX's bars are both much higher than those of AMS-IX, LINX, and DE-CIX and represent approximately half in the case of PTT.br and approximately three quarters of MSK-IX's participants. PTT.br and MSK-IX have a higher proportion of uniques ($d(i) = 1$) than the other three large, more diverse IX. Now consider a global CDN such as Akamai, Limelight, or a firm whose value-proposition benefits from proximate, stable interconnection with unique access networks such as Facebook, Netflix, or Google. These actors are participants at the top three IXes, but also seek out more direct access to unique markets than transit provides. If these markets prove valuable, these networks invest in identifying redundancy.

6.2 Common Resource Structure

IXes across the board have a common objective: facilitate interconnection and traffic exchange between participant ASes. At a purely technical level, this is achieved via configurations of switches and transport technologies into what is commonly referred to as a *switching fabric*. These are later refined into a precise, formal

the provider optimizes for lowest-cost route selection, differentiating between transit appropriated routes and those appropriated from more specific interconnection relations is a potential indicator of the stability of the availability of the route (but not the actual path).

[370] One operationalization may be the proportion of common contiguous hops along a particular AS path. A strict subset offers another operationalization: the maximum length common path terminating at the destination. What remains of the different routes is the actual diversity in the paths, thus giving insight into implications of failures along one of those paths and the quality of the remaining redundant paths.

definition of an IX platform. As a resource shared amongst IX participants, platforms are accompanied by distinct management and governance arrangements. The archetypes and terminology presented here abstract away technical nuance and provide a semantics that facilitate explaining the relationship between topology and governance arrangements.

Thus far the discussion has focused on how routes are provisioned, not the architecture of the platform provisioning interconnection options. The following sections refine the notion of an IX platform to disambiguate it from other types of networks. The terminology and archetypes depict a variety of ways interconnection options are provisioned and will provide the foundation for mapping different elements of these archetypes to governance norms.

The test of any nomenclature is whether it depicts both expected (in this case observed) instances of the phenomena and *can depict* future, unexpected, and/or degenerate cases. This topology of IX archetypes' immediate function is to depict empirically observed topologies. The typology arguably admits possible topologies that do not (yet) exist in the wild.[371] IXes emerged in the mid-1990's and have been continuously evolving since. The interplay amongst technical, organizational, and governance components has shaped the evolution of observed topologies. Archetypes abstract away unnecessary details of equipment configurations, replacing this detail with a terminology sufficient to express the administrative and organizational semantics that help explain the relationships between the technical, organizational, and governance dimensions of IX operations.

6.2.1 Defining an IX

This work is far from the first to define what it is to be an IX.[372] An essential characteristic of an IX is that it provides a platform on which participants can *seamlessly* exercise the option to interconnect with any other participant on that platform. Although a seemingly simple concept, the notion of a participant needs some elaboration. As implied earlier, network actors participate on an IX platform to *gain access to* interconnection options. It is a *subsequent* option to exchange traffic. Most

[371] More accurately, they are not in the current sample and have not been highlight in interviews of IX experts that have solicited information about unique, or less common topologies.

[372] Euro-IX provides a longstanding definition developed by the community:

> [An IX is a] physical network infrastructure operated by a single entity with the purpose to facilitate the exchange of Internet traffic between Autonomous Systems. (Euro-IX, 2012)

The networking literature discusses IXes in terms of their effects on traffic and topologies. Ager et al. (2012) is a study of a large IXP in Europe in SIGCOMM'12, building on ideas related to traffic flows around the tier 1's (Labovitz, Iekel-Johnson, McPherson, Oberheide, & Jahanian, 2010), the flattening of the Internet topology (Gill, Arlitt, Li, & Mahanti, 2008) and earlier work on IXes (Augustin, Krishnamurthy, & Willinger, 2009; Stanojevic, Castro, & Gorinsky, 2011). More recent work by Castro, Cardona, Gorinsky, and Francois (2014) discuss remote peering, but with less flattening than building directly into the IX. Remote interconnection and remote nodes are discussed in Section 6.2.3.3.

participants do join an IX to exchange traffic, but there are a few exceptions. Some incumbents have been known to join, but not exchange traffic. Another class of participants that do not exchange traffic are network measurement and analysis firms such as Renesys and PCH described earlier. As alluded to earlier, in terms of rights bundles, there are two classes of IX participant: members and customers. These are management constructs. The full implications of differentiated authorized entrants are not explored until the discussion of mutuality in Section 6.4.1.1, neutrality in Section 6.4.1.2..

A more complete label for the IX platform is a *common logical interconnection platform*. While the term "switching fabric" is community vernacular that rolls right off the tongue, each modifier emphasized above refines the notion of a platform that provisions interconnection options. Both the notion of a common logical interconnection platform and the notion of an interconnection option need further unpacking. Given the platform is built to provision interconnection options, the platform will be presented first.

The term *logical* abstracts away the nuance and detail of how switches and transport equipment that implement the platform are physically connected. Underlying complexity varies significantly. A *simple* platform may comprise a single switch in one site. At the other end of the spectrum, a platform may comprise 10's of switches connected by a combination of cross-connects and transport equipment that span multiple metro-regions, nation-states, or even continents. Constituent switches are connected in such a way to create a single logical switching *fabric*—it is this configuration that maintains the logical L2 path to any other participant. The existence of a path does not imply it is in use, hence an interconnection *option*.[373]

The platform is *common* in the sense that it is a shared resource available to all participants. In the case of IXes based on the associational membership governance model, common also speaks to role of the membership plays in shaping the strategic direction of the IX as a whole. By virtue of establishing a path between any two participants, the switching fabric implementing the platform provides the *technical option* to interconnect with any other participant on that platform. The creation and maintenance of that *path* is what is meant by *provisioning* interconnection options. Reiterating the character of the option, choice to use that path, to exercise the option, is the joint decision of the participants at either end.

Part of the administrative objective is to abstract away as much of the underlying topology as possible so participants can focus on the benefits conferred by IX participation, the development of strategic interconnection bundles. Exploring strategic bundles is low-cost in part because, as illustrated in the discussion of Figure 6-1, all but negotiation (ix_{neg}) and setting up the BGP session (ix_{bgp}) has been disintermediated.. A result, interconnection, especially across fabrics hosted in multiple facilities, is intentionally made as *seamless* as possible, reducing the exercise of an option to setting up a one-hop BGP session.[374] Participants can build and tear down

[373]A switching fabric does not necessarily imply a common L2 domain. VLANS, MPLS, GMPLS, and other virtual circuit technologies are used to isolate L2 domains, often to contain misconfigurations that result in events such as broadcast storms.

[374]The exception is the use of a route server, an L3 service that is often used to facilitate estab-

a variety of interconnection relations, using a common resource, at will.[375]

Given these characteristics, the following definitions are offered:

1. An IX *provider* is an organizational entity[376] that comprises the administrative and technical arrangements necessary to build, maintain, and manage common logical interconnection platforms.
2. IX *platforms* provide Autonomous Systems non-discriminatory access to interconnection options that may be exercised by two or more participants to exchange routes and Internet traffic.
3. IX *operators* are organizational arrangements that may comprise a collective of individuals, a single firm, or a cooperative of firms that collaborate to create and maintain one or more IX platforms across potentially distributed network infrastructure components.

The IX platform is the technical substrate, managed by IX operators, cohering around norms and decision processes that are common to and made durable by the organizational constructs of the IX provider. The remainder of this section focuses on archetypal typologies of IX platforms, mapping administrative constructs to variously scoped IX providers.

6.2.2 Terminology

The terminology introduced here decorates graph representations of archetypal IX topologies with IX provisioning and management semantics. In turn, these decorations facilitate mapping elements of the archetypes to empirically observed interpretations of constitutional norms. Most of these archetypes occur "in the wild" and reference particularistic instances in the notes. The terminology also admits configurations that do not occur in the wild, but are included for contrast and completeness.[377] The terminology used is a basic descriptive notation to facilitate as precise and rigorous a discussion as possible.

As noted earlier, the IX platform is comprised of one or more switches that create a common switching fabric. Subsets of those switches are housed in one or more physical facilities, typically a colocation center, but also in network operators'

lishing many interconnection agreements via one BGP peering session with the route server. While the resulting connections are still technically bilateral, it has the *effect* of facilitating multilateral interconnection policies.

[375]Reiterating, building up requires mutual exercise of the option. Tearing down may be done by either participant.

[376]In the larger work this organizational entity has been referred to as the FMF, the facilities management *firm*. Not all associational membership IXes have delegated the authority to exercise management rights to a firm, retaining a less formal structure for exercising management rights. As such, the term organizational entity is used here. Although this work is only peripherally interested in commercial IXes, organizational entity also captures the resource management structure of that class of IX, as well.

[377]Some of those that *do not* occur in the wild illustrate inconsistent or intractable governance semantics. Others can be shown to have adverse effects on the interconnection market and objectives of the interconnection community.

data centers or other sites that provide sufficient power, cooling, and security.[378] A *node* is a switching fabric (comprising one or more connected switches) located at a particular facility. An IX node, henceforth simply referred to as a node, is uniquely identified by *a*) the IX administering the node as part of an IX platform[379] and *b*) facility in which the corresponding network elements are hosted.[380] A node, in and of itself, can, and often does, comprise a common logical interconnection platform.[381] For the participants connected to that node, it often functions as such.[382] In this typology, the node is the basic building block of a particular IX topology; the details of precisely how switches are interconnected is well-known and outside of the scope of this work.

As may be obvious, multiple nodes can and do exist within the same physical facilities, especially colocation facilities. Typically when multiple nodes cohabitate the same facility they are affiliated with different IXes. Recall the notion of stickiness developed earlier in Section 6.1.2. An IX node with substantive traffic (relative to the region) and whose interconnection options represent access to many of the region's networks and international connectivity, is considered successful.[383] A successful node makes the hosting colocation facility *sticky*. Stickiness is proportional to the value of interconnection relations at a node (or set of nodes) and the transaction costs of altering the capital and operational expenditures related to losing those relations or re-establishing those relations elsewhere.Colocation facilities benefit from stickiness because it increases the cost of, or barriers to, exit. As a result, participants are less likely to leave the colocation facility. This does not imply the actor does not have incentives to take up residence in other colocation facilities.[384] Rather, a subset of participants, such as transport networks, will seek out

[378]Variance in power, cooling, and security is quite high. AMS-IX requires certain colocation standards. In other regions where colocation facilities are not as prevalent, the value of the IX itself is a catalyst for improvements. CABASE in Argentina and PTT Metro in Brazil have a heterogeneous mix of colocation standards. In both cases, the IXes prefer deployment of nodes to meeting potentially overly stringent standards, the expectation being that as those nodes become critical to local economies, standards will improve. Open-IX has elicited community knowledge on colocation standards, codifying these in OIX-2. Open-IX is discussed in Section 6.3.2.2.

[379]This will be later generalized to the administrative domain, which may comprise more than one organizational entity.

[380]Different IXes use different terms for the notion of a node. Some call them IXPs, some NAPs (network access points), some are simply sites, and some also call them nodes. The term node, as per this definition, is used throughout for accuracy and consistency.

[381]For instance, in CABASE, effectively the national IX for Argentina, each "remote NAP" outside of the central node in Buenos Aires serves as common logical interconnection platform for the city, region, or cluster of towns and villages it serves. CABASE requires forced multilateral exchange of routes to prefixes directly managed by participants through the route server.

[382]This becomes and interesting distinction between remote interconnection models and remote node models.

[383]Unsurprisingly, these factors create network effects. In contrast to other markets that exhibit network effects, network effects in one IX do not necessarily compete with the network effects of another IX, even in the same metro-region. Rather, IX are often complementary and have common participants.

[384]In fact this is a strategy of some types of network actors such as transport network and DNS infrastructure providers that benefit from participating in a wide range of IX platforms.

additional colocation facilities, especially if they house nodes of IXes the transport network does not yet participate in.

A collection of nodes will be referred to as a *set*.[385] A *connected set* indicates there exists a path from any node in that set to any other node.[386] A connected set managed by a particular firm is an *interconnection platform*, or here simply *platform*. Consider a connected set comprising nodes i and j, where it may be the case that $i = j$. In terms of interconnection, any participant at node i has the option, but not the obligation, to interconnect with any participant at any node j.[387] Typically such a connected set is part of a common platform, but there are exceptions.

Scoping helps group nodes into meaningful sets. Geographic, administrative, and connection are three scoping domains used in this typology. A set is considered *covering* if it comprises all the nodes within a given scope. A geographically scoped set means the set of nodes within a particular geographic region. A *metro-set* is the covering set of all the nodes in a metropolitan statistical area (MSA, similar to the metro-region in the community's vernacular); instances are the metro-set of Frankfurt or New York.

The administrative domain is another way to scope a set of nodes in terms of what organizational entity manages those nodes. The *administrative domain* of a node is the *set* of IX operators that manages that node. For instance, the *administrative covering set* of Equinix comprises all the nodes it manages, which spans multiple continents. A *partial set* is a proper subset of one or more *covering sets*: a partial set of Equinix's administrative domain comprises the intersection of Equinix's administrative covering set and the Paris covering set. Most nodes only belong to one administrative domain. The France-IX ecosystem comprises well-known exceptions; for instance, nodes of the Lyon-IX and nodes of France-IX are under the administrative domain of both France-IX and Lyon-IX, but they form a connected set. The *node operator domain* delineates the set of nodes technically managed by a particular operator. Typically the node operator domain and the administrative domain are the same, but there are cases, especially remote node scenarios (See Section 6.2.3.4), where the two are different. This creates interesting management implications.

A common archetype is a single connected set that is both administratively covering *and* geographically scoped to a metro-region. More formally, the administrative covering set is a subset of a metro-set. If the administrative covering set is a *proper* subset of the metro-set, multiple IXes are deployed in that metro-region. In

[385] Formally a set is an unordered collection of objects with no duplicates.

[386] This is essentially a connected graph: for any nodes i and j in graph G, there is an $i - j$ path in G. See Chartrand and Oellermann (1993) for further elaboration of graph terminology.

[387] This is typically via a one-hop BGP session between i and j, although many IXes provide route servers that technically introduce an intermediary. There are a number of technical and administrative exceptions. The most common is the use of VLANs to create communities such as those participating in a service level agreement, those using a shared cache, and for facilitating private sessions with one or more other network actors. The objective of defining a connected set is to concisely convey options of a common interconnection platform but abstracting away the underlying switch and interconnect topology. This eases the discussion of management issues, including those above, that are a function of both topology and governance decisions.

either case, that metro-region constitutes that IX's primary market. This is referred to as a *single-metro administrative set.*

The *connection scope*, or in the community vernacular, the *reach*, of an IX platform reflects the area spanned by it constituent nodes. Connection scope, or reach, thus conveys the geographic scope of the platform infrastructure.[388] Above, the reach of a single-metro administrative set *is* the metro-region. It should be stressed that reach refers to a particular platform, not all the platforms in an IX's administrative set. For instance, the LINX has platforms in multiple metro-areas, but the reach of each is that metro-area.[389] The context and connotation of reach is intended to convey the geographic diversity of nodes, and by proxy, how difficult it is for a given participant to gain access to the geographically nearest node and, subsequently, the corresponding set of interconnection options. In other words, are these nodes sufficiently distributed within a region to realize latent demand or unrealized potential interconnection markets?[390] For IX providers with platforms in multiple cities, each platform geographically is scoped to that metro-region (implying the platforms are not themselves connected), and thus the reach of each platform is still the metro-region.

6.2.3 IX Archetypes

Archetypes have evolved from single switch (node) topologies to distributed (virtual) IX platform topologies that comprise nodes distributed over varying distances. Some archetypes, such as single-node IXes, are the historical first step for a number of the successful IXes considered here. Other archetypes, such as remote participation models, are relatively new and have been controversial within the IX community—they are also great exemplars of the interplay between topology and management norms. The following archetypes lay the foundation for discussion of the IX life-cycle, the role of common management decisions in platform design, and how the characteristics of these archetypes reflect various interpretations of, and conflicts amongst, management norms.

6.2.3.1 Single Switch/Node

Many IXes started as a single switch. The LINX history (LINX, 2013) notes the original platform comprised a Cisco Catalyst 1200 with eight 10MB ports. In the LINX of November 1995, Telehouse Docklands was the site of the first LINX node.

[388]Connection scope is different from geographic scope of a set of nodes. Geographic scope captures all nodes in a particular region regardless of the administrative domain. The partitioning of geographically scoped nodes along administrative lines is a mechanism for partially explaining how fragmented the attendant interconnection market may be.

[389]While the reach of the platform is the metro-area, a number of new members of platforms outside the home London metro, in particular Manchester, have, recognizing the benefits of IX participation, also invested in connectivity to the London metro LINX platform.

[390]The expansion of AMS-IX into Haarlem and the expansion of LINX to Slough are instances of expanding the connection set of their respective metro-sets.

In addition to the LINX, Netnod and CABASE, among many other IXes, also started as single-node topologies.

In the simplest form of a single-node IX, as well as mutli-site IXes below, each participant *builds into* an IX node by (*1*) establishing transport from the participant's network to the node facility[391] and (*2*) hosting a router at that facility. That router is connected to a switch in the IX node to access the interconnection options with other participants reachable from that node.[392]

A single node platform has limitations. Physical space at the colocation facility was an immediate limitation.[393] Following the LINX history, soon after its creation the LINX had to navigate the colocation market of the dot com boom to find another colocation facility suitable for the second LINX node.[394] More recently, IX development practices encourage multi-site platforms for a variety of reasons: *a*) redundancy, *b*) proximity to participants in large metro-regions, *c*) physical space limitations, *d*) to encourage competition in the colocation market, and *e*) to ensure diversity in colocation facilities hosting nodes, yielding colocation (or datacenter) neutrality. The first three are technical and physical aspects of node deployment and will be discussed in the next section. The latter two are market and platform management issues to be addressed in Section 6.5. The notion of a multi-site platform is quite general and covers the remaining topologies that will be discussed. In the next three sections, variants of mutli-site platforms that have seen significant deployment over the history of IX development are formally described.

6.2.3.2 Metro Multi-Site IX

A metro multi-site platform is an IX platform within a particular administrative domain and geographically scoped to an MSA. Consider the partial set defined by the intersection of an administrative covering set and a particular metro-set; this is referred to as the *IX metro-set*.[395] For instance, the AMS-IX nodes in the Amsterdam metro-region (metro-set) would be referred to as the AMS metro set. A *metro multi-site platform* is an IX platform whose nodes are a subset of a particular IX metro-set. The IX metro-set may include multiple platforms that, by definition, partition[396]

[391] Facilities typically provide the equipment necessary to connect equipment to transport provisioned to that facility. Transport providers are often already present, so in many cases "provisioning" is really just establishing a contract for transport service.

[392] In some cases the cross connect between the participant router and the node is covered by the IX fees. Stickiness is sufficiently valuable to colocation centers that in some cases the colocation facility waives cross-connect fees. In others, the cross-connect is managed by the IX itself by its own staff.

[393] Colocation availability is still a factor. For instance, space in the primary colocation facilities in New York City are at a premium.

[394] More precisely, this was rather early, in 1996 when both the Internet and colocation facilities markets were relatively young. At that time the LINX offered a tender to build a colocation facility.

[395] Recall a metro-set comprises all the nodes in a metro region. Modifying with an IX narrows this scope further to just the nodes in that administrative domain. Similarly, modifying the term set with a geographic scoping operator implies the set is the covering set of that region.

[396] A platform is a connected set of nodes. For a set to contain more than one platform, it must contain $n > 1$ components that are themselves connected and for which there are no edges (connections) between the components.

the IX metro-set.[397] Continuing with the AMS-IX instance, the AMS-IX metro-set comprises only one metro multi-site platform.[398]

A subtype of the metro multi-site archetype is the single-metro administrative set discussed earlier. As per the definition, the administrative set is a subset of a metro-set. This implies that it has nodes in only one metro-region. Single-metro administrative sets have also been referred to as a virtual IX in the community. "Virtual" is used in a similar manner that logical is used to characterize an IX platform.[399] During the same time period, IXes were moving from single to multi-site, giving rise to non-compete discussions related to transport services; see Section 6.4.1.3.

A number of established IXes have transitioned through the single-metro administrative set topology. LINX, Netnod, CABASE and others were all single-metro administrative sets over the course of their history. LONAP, IX Leeds, and Nepal IX are instances of existing single-metro administrative sets. Until the recent deployment of IX Manchester, IX Scotland, LINX NoVA, and IX Cardiff, the LINX was also a single-metro administrative set.

To further drive home how the geographic scoping of node sets depict the IX landscape, consider London.[400] The LINX set[401] is not the London covering set. The LINX is not the only IX in London. The LONAP set *is* a single-metro administrative set. Moreover, some of LONAP's nodes are sited at the same colocation facilities as the LINX. The single-metro administrative set topology is still the norm for many regional IXes.[402]

Two types of geographic scope are relevant to the discussion of IXes: the geographic distribution of *nodes* discussed thus far and the geographic distribution of *participation*. The two are not necessarily the same. Geographic distribution of nodes refers to the distribution of nodes in a given platform. The geographic distribution of participants refers to the distribution of participants' network elements and/or customers. Reseller topologies and remote node topologies, discussed in the next two sections, have expanded the geographic scope of participation and

[397]This is not common. IXes strive to reach critical mass necessary to yield the network effects that, in turn, attract more participants. Network effects are a function of having a common platform. As such, multiple platforms in the same metro-region is atypical. That said, two IX providers, each of which deploys a metro multi-site platform provide redundancy that is embraced by a distinct subset of participants. More than two platforms in a metro-region, or market, is often considered a recipe for fragmentation.

[398]In other words, the AMS-IX metro-set is equivalent to the AMS-IX metro multi-site platform *in* Amsterdam. This is a proper subset of the AMS-IX administrative set, which includes nodes in Hong Kong and Curaçao.

[399]Virtual is not used here because it has been used to refer to an IX whose administrative scope is limited to a metro-region and an IX that connects its metro multi-site platforms in nearby metro-regions. Further, the term virtual is overloaded in the literature to the point of losing any significantly precise meaning.

[400]Geographic scope is also a proxy for an IX's market. This becomes much more interesting when we talk about how overlaps in participants limit invasion of one region by another.

[401]When a set is referred to as the "IX set", this is shorthand for the administrative covering set of that IX provider.

[402]In this case, regional refers the geographic scope of IX participants, not the geographic scope of the IX nodes.

platform nodes, respectively. The next section describes an archetype in which geographic scope of participation is expanded *without* deploying remote nodes.

6.2.3.3 Remote Participation Models

IXes deploy nodes in multiple sites, typically dispersed in a metro-region, to *a*) reduce access costs for would-be participants in that region, *b*) make it more convenient for participants to connect, and *c*) to increase redundancy. The foregoing discussion has focused on the metro-region. A proportion of IX participants have a network presence in the IX's metro-region. Conventionally, participants outside the metro-region build in by provisioning or contracting their own L2 transport to a node in the desired IX metro-platform. There are also potential participants that do not have network presence in the metro-region. These potential participants are in neighboring geographic (as opposed to topological) regions and would like to participate in the interconnection options market provisioned by the IX platform, but for whom it may be prohibitively costly to build in. This may not be easily done for a variety of reasons, including the cost of transport, the cost of the router, and/or operational experience.

Remote participation models[403] are a class of solution to this problem. One solution in this class is the creation of remote ports. From the perspective of the remote participant, a *remote port* enables that participant to seamlessly exercise interconnection options as if they had built into a node of the platform directly. To realize this goal, IX providers partner with a participant transport provider whose infrastructure extends outside the metro-region. Such participants will be referred to as *transport partners*.

To be a successful transport partners, two general criteria need be met: *a*) infrastructure extends beyond the metro-region and *b*) transport PoPs are close to potential remote IX participants. To provide capacity, the transport provider aggregates ports contracted at a node of the platform. For this discussion the aggregate capacity contracted will be denoted $C = c \cdot n$ where c is the port capacity and n is the number of ports contracted. The transport partner sells remote ports by establishing virtual wires[404] of varying capacities $rp_i \leq C$ to remote participants.[405] The poten-

[403]These are also called reseller programs or partner programs. Each provider has its own partner/reseller branding.

[404]This is typically implemented using VLANs, (G)MPLS, or some other mechanism for managing virtual circuits. The details of this implementation are beyond the scope of this work, see Peterson and Davie (2011) and Davie and Farrel (2008) for general descriptions of these technologies.

[405]Let us say a transport partner contracts $n > 0$ ports at capacity c at a node of an IX platform for a total capacity $C = n \cdot c$. The provider may provision R remote ports rp_i (where $0 \leq i \leq R$) such that

$$\sum_{i=1}^{R} rp_i \leq C \tag{6.1}$$

The IX provider monitors C as it would for any participant at a given platform node. The transport provider monitors each of rp_i. There are at least three variants that make this simple calculation more complicated. First, the partner could contract ports of different sizes $c_1, c_2, \ldots, c_n$ although there is rarely more than four port sizes available. The second is that the partner may contract

tial remote participant connects a router to the reseller's POP to contract a remote port. The transport provider typically receives an IP address on the interconnection (peering) local area network (LAN) for the remote participant and assigns it to the remote participant's remote port (and thus the remote participant's router). At this point, the remote port effectively enables the access to the interconnection option market provisioned by that platform.

Transport partners solve a variant of the remote participation problem. In some scenarios, a number of remote participants may be clustered relatively close to one another, but they all "come in" to exercise interconnection options via a transport partner connected to a node in a relatively distant metro-region. In effect, the tromboning problem has been recreated, albeit on a smaller scale.[406] Moreover, it is not clear that it is in the transport partner's interest to correct this problem—the transport partner is being paid largely for transport. Remote node implementations help solve the tromboning problem but has the potential to introduce subtle rights problems related to neutrality (discussed in Section 6.4.1.2).

6.2.3.4 Distributed Node Models and Mixed Administrative Domains

Topologically, and by historical convention,[407] the geographic reach of a single IX platform has been a metro-region. An IX provider's topology is considered to be distributed if its covering administrative set comprises nodes in *multiple* metro-regions. Here the notion of distributed is relative to the historic geographic scoping of platform nodes to a metro region. Two subtypes are distinguished: disconnected and connected.[408] The disconnected model is more common: an IX's administrative covering set is partitioned into components that correspond to IX metro-sets. Typically, each IX metro-set comprises a single platform. Instances of this model are the LINX,[409] Netnod, DE-CIX, and Equinix. In the case of DE-CIX, platforms are sited in Frankfurt, Hamburg, Munich, and Düsseldorf. Netnod has platforms in five Swedish cities: Stockholm, Gothenburg, Malmö, Sundsvall and Luleå. In terms of the geographic distribution of platforms, *remote platforms* refer to platforms in an IX's administrative domain that are outside that IX's home metro-region, typically

---

ports at different nodes of the same provider. Third, there is a subsector of the transport market that specializes in providing remote participation at IXes, thus the partner may provision capacity at multiple nodes from multiple providers.

[406] At this scale, latencies have shifted from approximately 80ms to cross the Atlantic (depending on where you are, this is based on a recent measure from Boston to London), to ranges of $< 1ms$ to $30ms$.

[407] Convention refers to the non-compete norm, in particular as applied to competing with transport carriers. See Section 6.4.1.3.

[408] In other words, are the platforms components of the graph comprising the administrative covering set of the IX platform. A component is defined as "a subgraph H of a graph G ... if H is a maximally connected subgraph of G," (Chartrand & Oellermann, 1993).

[409] In the case of the associational membership models that have chosen to follow the disconnected distributed node model, in particular the LINX, the mandate of the IX provider had to be modified to sanction the expansion to Manchester.

the first metro-region that IX provider deployed in.[410]

A *connected* distributed platform comprises a single platform that spans multiple metro-regions. Geographically, in both the connected and disconnected case, nodes managed directly by the provider are typically clustered in the MSA. Instances of connected distributed platforms are the AMS-IX, ECIX, and France-IX.[411] In the case of the ECIX, the administrative domain is a singleton: the nodes are exclusively administered by the ECIX. The AMS-IX and France-IX have mixed administrative domains: the administration of some remote nodes are influenced by two organizational entities.

In the discussion of connected versus disconnected, it is important to distinguish the administrative domain of the connectivity infrastructure. In the case of disconnected, the provider neither provisions nor contractually facilitates connectivity between the nodes. In the connected case, the provider provisions or contractually facilitates connectivity between the platforms. In the case of the AMS-IX and AMS-IX Hong Kong, connectivity is provided by Hutchison, a transport partner. This contract is available to any other transport provider that can meet AMS-IX's requirements. In terms of governance norms, this maps to connection neutrality.

Some connected multi-site IXes build on the partner relations in the remote participation model described above to deploy nodes. Under this model the topology is the same but the "remote" node (typically outside established metro-regions) itself is managed by a partner.[412] The partner manages connectivity between the remote node and provider managed nodes; the partner also typically owns the site and attendant facilities. The partner may also be another IX under a different administrative domain. In either case, the remote node is now within the administrative domain of two organizations: the IX for which it is the remote node and the manager of the remote node. In the case the remote node is actually a remote IX, this is typically a smaller IX that benefits from being connected to the larger. AMS-IX has such a relation with NaMeX (NaMeX, 2015) and NL-ix [413] through its Partner program (AMS-IX, 2015).

There are additional exceptions to single administration: the Paris covering set comprises platforms managed by France-IX and other French IXes in that region such as SFINX and Lyon-IX. When a metro-region has more than one IX, it typically

[410]The home metro-region is historically where the founding participants of the IX are based. This does not mean the current participant set is centered there. For instance, in the AMS-IX, participants in the Amsterdam metro-region are a minority in the constituency, but many remain active in membership meetings.

[411]AMS-IX and France-IX are two associational membership IXes that share some of the growth patterns of commercial organizations. ECIX is a purely commercial entity. The LINX has also developed remote platforms, which could be considered a growth model, but it is different in terms of the membership model. This will be discussed in Section 6.4.1.1.

[412]There are variants of this model that have similar management structures but a uniform governance structure. For instance, CABASE's remote node model more resembles a federated governance model with a hierarchical administrative domain than a remote "reseller" or partner model.

[413]NL-ix (NL-ix, 2015) is owned by KPN Wholesale. It is the smaller IX in the Netherlands, akin to the relationship between LONAP and the LINX in the UK. Recall LONAP is an associational membership IX.

has two multi-node IXes, each providing one common interconnection platform for that metro-region. In some interpretations these two IXes are competing. Other participants see this as diversity that facilitates redundancy. Nodes may be housed at the same site (colocation facility) but the switches themselves are commonly administered by their respective IXes, modulo exceptions noted above.

6.3 IX Constituencies

Participating on the IX confers at minimum *access* to the platform market, the set of interconnection options provisioned and maintained by the IX firm and the participation of other network actors.[414]

There are two general classes of participants: members and customers. In most cases, members and customers are both entitled to the same sets of services and resources, sometimes with the option for customers to become members. In terms of resource rights, members and customers are both authorized entrants whose appropriation rights are managed by the firm. Port capacity directly limits appropriation of exchange capacity and indirectly limits the appropriation of interconnection options. Members have additional rights and obligations that includes *a*) their right to participate in consensus processes that shape the strategic direction of membership based IXes, *b*) votes for Board members, and *c*) affirmation of IX activity plans. As noted earlier, management rights by IX members largely affect constitutional norms. As will be developed here and in Section 6.4, the IX operator (the firm) has been delegated discretion to both create and exercise management rights within the scope of operational, but not constitutional, rules.

Associational membership IX participants may be exclusively members, or a mix of members and customers. For instance, *all* participants in the LINX are members—that is a significant distinction relative to other global IXes. In the case of commercial IXes such as Netnod, DE-CIX, and data centers with an IX product such as Equinix, all participants are customers; there is no membership class. That said, commercial IXes employ varying elements of the arenas created by associational membership IXes. Commercial IXes also collaborate to establish and participate in arenas in which both mutual and commercial IXes can solicit input from participants and exchange domain knowledge; such venues include Euro-IX, GFP, EPF, and Open-IX. AMS-IX falls between these two classes, having both members and customers.

6.3.1 Participants and Rights

IX participants are also participants in the network operator and RIR communities. Of the three studies in this work, the scope of the IX regime is considered the most

[414]Recall an interconnection option is characterized by the peers and the platform on which they interconnect. A firm may provision a platform with as much exchange capacity and as many ports as they like, but if there are no participants on that platform, there is no one to exercise an interconnection option with, and, ultimately, no market.

narrow. Non-compete norms and historical geographic scope have worked to limit the scope of the IX—IXes provides a neutral interconnection platform for exchanging routes and attendant traffic, full stop. The previous section offered diversity and interconnection bundle development as an explanation of the value-proposition of IXes (beyond simple savings on transit costs). As a platform, the associational membership IX *facilitates* the development of these bundles precisely because it is a neutral platform that uniformly lowers the cost of interconnection options. IX provisioning is technically and operationally sophisticated, but the resource rights bundles that specify the types of participants are quit simple. The complexity for participants arises in the application of objective functions such as redundancy and uniqueness to develop sophisticated interconnection bundles. By lowering the barriers to participation, IXes have, as argued in Sowell (2013), created a platform on which small to medium sized network actors can develop sophisticated interconnection bundles using strategies previously only available to large networks. In some developing regional markets, IXes have been framed as a potential challenger set to would-be incumbents.

Of the studies, IX participant rights bundles, participant roles, hew most closely to Ostrom's cumulative rights bundles described in Table 3.2. In the RIRs discussion in Chapter 5, Table 5.2 was developed based on operational rules. In contrast, changes to rights bundles in the IXes are a combination of demand from participants, market scope (platform or provider geographic scope), and attendant platform topology. The following sections describe IX participants in terms of resource rights held. Like the RIRs, there is a basic appropriation bundle that defines resource appropriation. Differences in bundles delineate the scope of use, alienation, and management rights by those holding that bundle.

6.3.1.1 Basic Appropriation Bundle

The basic appropriation bundle available at an IX platform comprises access to interconnection options and rights to utilize some proportion of exchange capacity (provisioned as port capacity). In its simplest form, the availability of interconnection options is a form of access. As discussed earlier, exercising an interconnection option is, relatively speaking, non-subtractive.

In terms of the obligation to exercise an interconnection option, a few IXes require multilateral peering. Under "forced" multilateral, every participant must engage in settlement-free peering with every other participant.This case illustrates the non-subtractive quality of exercising an option and different types of bundles. In most settings, forced mutlilateral requires settlement free exchange of routes to prefixes delegated to the ASN of a participant j and the subsequent exchange of traffic. This does not require the exchange of a participants entire stock of routes. This does include hosts assigned rights by j and ASN's delegated or assigned rights by j. Thus, any two participants have settlement-free paths to the hosts numbered by the prefixes delegated to that ASN. Note forced multilateral does not include upstreams of j or other actors j may have an interconnection relationship with via transport- or colocation-mediated interconnection. Forced mutlilateral does not force an actor

to provide others with free full or partial transit.

Forced multilateral does compete with transit routes that participants would have had to appropriate to exchange traffic with other IX participants. That said, forced multilateral does notecompete with transit routes to networks not on the IX. Consider i, a small access network, and j, a large transit provider at IX F that requires forced multilateral. Network i is peered with j, but under this particular relation, j only provisions routes to prefixes it was delegated. If i wishes to purchase transit from j, it must negotiate with j to establish transit. In terms of Figure 6-1, this is ix_{neg}. In effect, i must negotiate with j to provision transit routes. Depending on the traffic between i and j, i may need to upgrade its capacity to the IX (ix_{xc} and ix_{part}).

Depending on the IX, forced multilateral is implemented and enforced by requiring all participants peer to a route server.[415] When any new participant joins the IX, it peers with the route server and peering with every other participant occurs automatically. This enforces multilateral and reduces transaction costs for participants.

Route server policies differ across IXes. In some IXes, the route server is a convenience. Actors that have an open peering policy peer with the route server. As above, these actors are then automatically peered with others on the route server. Some IXes use the route server to enforce forced multilateral. Other IXes require *all* interconnection relations to occur over the route server.

Rights to appropriate exchange capacity is the second set of rights delegated by the IX. Interconnection options are non-subtractive resources that are bound to participation. Exchange capacity is a subtractive resource. Upon joining an IX, participants typically must purchase the minimal port capacity.[416] Contracting port capacity c entitles the participant i to exchange c volume of traffic (in aggregate) over the sum of its interconnection relations.

Depending on the IX, operational rules may dictate upgrades when i approaches its threshold c. In some cases IXes require automatic upgrades when total port utilization reaches some threshold t, typically $t = 0.8 \times c$. Others have stated that if a network wants to run their ports at 99%, they can. Some IXes send a note indicating the actor should upgrade, others assess a fee if the participant congests for some proportion of the time they are exchanging traffic, others still simply allow the connection to congest. Differences and instances of these will be discussed further in Section 6.4.3.

[415] Route server policy is more accurately an operational rule and will be discussed as such in Section 6.4.3. That said, it does bear on the exercise of the basic appropriation bundle and, along with forced multilateral peering, helps illustrate the non-subtractive character of *exercising* interconnection options for the IX.

[416] In many exchanges the minimum is a single port of the lowest capacity. Some exchanges, such as Netnod and the LINX, maintain redundant peering LANs; if one fails, traffic is still exchanged over the other. In the LINX, purchasing ports on both is optional. In Netnod, the customer is required to purchase ports on both LANS.

6.3.1.2 Members

Members are the most common type of participant at associational *membership* IXes. As a participant, members hold the basic appropriation bundle; they are authorized entrants. Members also have rights to *a*) vote in board elections, *b*) vote for changes to the remit of the IX, and *c*) vote on activity plans and ongoing strategy. As will be developed in Section 6.4, constitutional norms are explicit in the IX bylaws.

Bylaws may be changed by a variant of the consensus process or a vote, depending on the IX.[417] Under these collective choice processes, members exercise management rights. Like RIRs, the membership *collective* has the authority to exercise management rights. Collective choice rules comprise *a*) informal consensus processes, *b*) formal consensus processes, and *c*) membership votes. Changing the remit is an act of navigating the trade-offs amongst constitutional norms.[418]

RIRs have a nuanced hierarchy of meaningful rights bundles, complemented by a variety of function-specific criteria that are applied to justify need. Changes to resource policy regularly change the particularistic operational rules of number delegation, and subsequently, the bundles engendered therein. In the IX, the basic appropriation bundle and its derivatives are fairly stable. Rather, for IXes, most changes to rules affect the interpretation of constitutional norms and how they frame the authority delegated to the firm to exercise management rights. Section 6.4 builds on the topology of archetypes developed Section 6.2 tracking how constitutional rules have evolved.

Growth drove professionalization of the firm. Concurrently, the physical footprint of the platform, the reach or the "diameter" of the IX in community vernacular, grew from a small number of "points" to multinode platforms that often cover metro-area equivalents.[419] In a resource *pool* such as the IPv4 space, rival bundles such as the basic appropriation bundle in the RIRs, were scarce precisely because the pool was finite and immutable. Modifying the appropriation rate in operational rules was a strategy for temporary stability. In contrast, IX capacity, in both options and exchange capacity, are *both* mutable; moreover, demand has not yet out-paced technology or the attendant growth strategies. Rather than place conditions on rate and volume of appropriation, growth of the platform and providers gave rise to new types of participants differentiated by whether they held management rights (emergence of customers, see Section 6.3.1.3, next), members with limited alienation rights (resellers, Section 6.3.1.4), and remote nodes (Section 6.3.1.5).

The emergence of these appropriation modes co-evolved with the topological archetypes described in Section 6.2. Customers emerged out of a combination of growth, drawing participants from increasingly far away and demand for remote interconnection. Remote interconnection has been developed in a number of IXes

[417]For instance, in Article 18 of the LINX Articles of Association (LINX, 2015c) a poll is discussed, but it is expressly noted this is not necessarily a counted vote.

[418]Trade-offs will be discussed in Section 6.4.1. The elements of the collective choice process will be discussed in Section 6.4.2.

[419]These correspond roughly to what the US census Bureau designates as a metropolitan statistical area (MSA).

through what is referred to here as the reseller model.[420] Resellers are network actors that have extensive transport networks and offer transparent L2 connectivity to a "remote" IX. In some models, these remote participants are members, such as in the case of the LINX. In other cases, these remote participants may only be customers, such as in the case of the AMS-IX. Remote nodes are a natural extension of the reseller model, and also see instances where nodes comprise customers of the parent IX (AMS-IX), are full members of the parent (LINX), and hybrids in which members have differentiated management rights for their local node and for the parent organization (CABASE).

6.3.1.3 Customers

Customers only hold basic appropriation rights. Customers do not hold management rights. In terms of Ostrom's classes of rights holders, these are authorized entrants. Just as with members, these actors may appropriate interconnection options and exchange capacity within the constraints of contracted port capacity. That said, customers cannot participate the application of management rights. Moreover, AMS-IX leadership indicated this is by choice of the customers; this seemingly diminished set of rights is not pressed upon any particular class of actors. In other words, most customers could become members if they built into one of AMS-IX's platform nodes.[421]

Multiple interviews have cited the mutuality model as confounding to outside actors, especially those from outside the European Union. For instance, the LINX has noted that US members scrutinized the notion of being a stakeholder in a foreign organization. One driver was perceived liability issues. In the AMS-IX, would-be participants from Russia had legal limitations on becoming members, technically stakeholders in, foreign companies. In other cases, AMS-IX members simply did not wish to engage in, or be expected to engage in, collective choice discussions. A number of the AMS-IX staff mentioned that actors that had opted for the customer contract simply wanted the service, with no other obligations or implications of obligations.

As will be discussed in terms of constitutional norms, AMS-IX, LINX, and France-IX are interesting comparators for understanding the balance between an exclusively membership driven IX, an a membership IX that has explicitly integrated commercial elements such as customer participants, and a model derived from the lessens of growth from these early membership models. France-IX has implemented what is referred to here as a "forgiveness versus permission" membership model. In each case, the presence of customer participants (or the absence thereof) is an indicator of the degree of mutual-commercial hybrid. The range of membership-customer hybrids is certainly not a perfectly ordered set, but rather, a mix of consti-

[420] In the AMS-IX, this is the Partner Program (AMS-IX, 2015). In the LINX, this is the LINX from Anywhere (LINX, 2015b) program. In DE-CIX this is the GlobePARTNER (DE-CIX, 2015) program. In Netnod this is the Netnod Reach (Netnod, 2015) program.

[421] Building in does not mean they have to become members. At the time of writing, building into an AMS-IX node was necessary, but not sufficient to become a member.

tutional norms that fundamentally define IXes but accommodate variants of structural and management configurations, along with attendant delegations of authority.

6.3.1.4 Resellers

Resellers contribute to the remote participation models discussed in Section 6.2.3.3. Often transport providers' or transport resellers' infrastructure extend outside the metro-region serviced by the IX. The attraction of remote participation via the reseller is that *a*) one of the reseller's PoPs may be closer to the remote participant, *b*) similar volumes are cheaper than the same volume under dedicated transport, and *c*) cheaper than placing, and dedicating, equipment at facilities in the metro region. Resellers "re-intermediate" transport dedicated to IX connectivity to tap latent demand in regions outside the metro-area, historically considered the diameter of an IX platform. The development of resellers was a point of contention, in particular a relaxation of the non-compete norm. Such an arrangement is akin to non-compete contention over inter-node connectivity within a multinode platform. Both of these issues are discussed in Section 6.4.1.

As an IX member, the reseller holds basic appropriation rights. Participants do not necessarily have the rights to alienate participation rights. For instance, in the General Terms and Conditions established for AMS-IX participants:

> Customer[422] is not entitled to assign and/or sublicense any of its rights under the AMS-IX Connection Agreement, or make the Connection available, to any third party unless it has signed a reseller or reseller plus agreement or other type of contract with AMS-IX B.V. which allows it to do so. (AMS-IX, 2013a, Footnote on "Customer" above added here for clarification.)

IXes hold the authority to exercise management rights necessary to create resellers, conferring the rights to alienate participation rights.

Resellers are a class of IX participant that holds basic participation rights and limited alienation rights. This bundle is referred to as the participant reseller bundle. Resellers offer remote participants a portion of the resellers' capacity and the opportunity to exercise interconnection options over the reseller's L2 connectivity to an IX node. In general, additional rights follow the customer and membership roles discussed in the previous sections. At minimum the reseller holds the participant reseller bundle. In the case of exclusively commercial IXes, resellers are by definition participant resellers. Member resellers hold the joint membership and participant reseller bundle—they hold participation rights and alienation rights.

A number of reseller models exist, and, like participant reseller status, are differentiated by whether remote participant can be or are members. In the LINX model, all participants are members of the LINX association. In the AMS-IX model, one

[422]Amongst associational membership IXes, AMS-IX is the most liberal in the sense that it has integrated elements of the commercial model, such as customers. In interviews with AMS-IX staff, AMS-IX tends to refer to all participants as customers.

can only be a member if connected directly to an AMS-IX platform node. All other AMS-IX participants, via both resellers and Reseller+ remote nodes, are customers.

6.3.1.5 Remote Platform Participants

Recall from Section 6.2.3.4 that a platform comprises one or more nodes. By historical conventional, a platform was limited to the metro-region. In the past few years a number of the large European IXes have begun exploring remote platform development. Regional (EU) instances are *a*) Netnod's remote platforms in Stockholm, Copenhagen-Malmö (COMIX), Gothenburg, Sundsvall, and Luleå; *b*) DE-CIX deployment in Munich and Düsseldorf; *c*) the LINX's deployment in Manchester and Scotland, and *d*) to some extent France-IX's "federated" model with other French IXes such as Lyon-IX and Marseilles. AMS-IX skipped regional (EU) development, deploying in Hong Kong, the Caribbean, East Africa, and most recently, as part of the first round of Open-IX certified IXes, to New York. DE-CIX has also deployed to New York; LINX has deployed to Northern Virginia (NoVA).

Each of these IX providers has a different customer and membership model. DE-CIX started as a mutual IX, but shifted to a commercial model early in its development.[423] Remote nodes under the AMS-IX Reseller+ model are not connected to AMS-IX, but participants may only become members if they "build in" to a platform node. As stated above, LINX remote platforms are not connected, but participants are members. Early IXes, historically influential and now quite diverse, developed in relatively green fields—European IXes did not exist, much less compete with one another.[424] Remote platforms are typically deployed in metros or regions that do not have an IX presence.

The disconnected platform model is characteristic in Europe.[425] South American IXes offers two large instances of a connected platform with national distribution. CABASE in Argentina was developed by CABASE the ISP Association in response to incumbent's limitations on transit and transport options, effectively reinforcing

[423] As per interview with Arnold Nipper.

[424] The notion of competition amongst mutual IXes is interesting, especially given the overlap in membership. IXes do not compete directly in the same sense that goods manufacturers compete. Rather, they provide generally substitutable services that provide redundancy. That said, interviews have indicated that mutual IXes are unlikely to "invade" metro-region that has an existing, stable, well-run IX. Rather than collusion between the IX *firms*, interviewees argue that members of both do not appreciate their fees being spent on unnecessary development or worse, counterproductive competition. As a counterpoint, France-IX was formed precisely because there were X poorly performing IXes in Paris, creating what a number of interviewees referred to as a dysfunctional interconnection market. The community collaborated to determine whether a critical mass of participants would be willing to participate in a mutual, professionally run IX in Paris. In 2008, a group of operators with experience developing IXes created France-IX, originally dubbed the Phoen-IX, to replace the fragmented French interconnection market. As will be developed in the discussion of mutuality, this is an instance of collective provisioning of a more rationalized common resource, in particular a more efficient market. Open-IX is a similar movement to develop mutually managed IXes in the US—early on Open-IX advocated for the development of "European-style" mutual IXes in the US.

[425] The disconnected platform model is not universal. Packet Exchange was an early attempt to develop a regionally connected IX platform. On the subregional scale, ECIX (ECIX, 2015) is a commercial platform present in Germany and the Netherlands.

the incumbents' oligopoly. PTT.br (PTT Metro) in Brazil is a project of the CGI, a joint government and industry Internet governance body in Brazil. Both function as connected platforms distributed on a national scale.

In terms of participant rights, CABASE provides an interesting contrast to the LINX and the AMS-IX remote platform model. All CABASE participants are members of the IX. CABASE comprises 10 nodes distributed across Argentina.[426] Each node has a local membership. The local membership, comprising staff from multiple firms, manages local decisions at the node, such as the purchase of equipment, management of facilities, and selection of members that provide connectivity back to the central node in Buenos Aires.[427] Billing is managed by the central IX firm and platform wide rules are also decided by a collective choice process. This will be referred to as the federated node model. As will be discussed in Section 6.5, differences in remote node management and participant rights bundles are an emerging issue for IXes as resource management firms.

6.3.2 IX Associations

IX associations provide a forum in which IX firm staff from *different* IXes gather to exchange information related to IX operations. As discussed in general terms in Chapter 4, IX associations serve a similar function as a knowledge commons that NOGs serve for the broader network operator community. Euro-IX, one of the oldest IX association, offers an effective definition:

> Euro-IX is an association of Internet Exchange Points (IXPs), promoting an open interchange of ideas and experiences, gained to mutual advantage of the membership, by offering fora, meetings, mailing lists and on-line resources. (Euro-IX, 2014)

IX associations are very much like NOGs for IX operators. In this capacity, IX associations are the forum for perpetuating and developing the IX operational epistemic community. Euro-IX maintains and supports *a*) best practices materials, *b*) databases of information related to equipment performance, *c*) databases documenting IX participation, *d*) processes for pairing experienced IXes with developing IXes,[428] and *e*) development of other IX associations, in particular the IX Federation (a federation of IX associations, IX-F). IX associations facilitate promulgating the membership-based IX constitutional norms described in Section 6.4.1.

6.3.2.1 Euro-IX

Currently, Euro-IX is the oldest and largest of the IXAs. It is complemented by LAC-IX in Latin America, APIX in the Asia Pacific, and AF-IX in Africa. Open-IX,

[426] As per (CABASE, 2015), note this page is in Spanish.

[427] The carrier member is a distinguished member with a transport network capable of providing connectivity from the local node to the central node. The carrier member and the arrangement for billing through the central node are instances of both a unique non-compete arrangement and use of the federated model to ensure billing neutrality.

[428] This is referred to as the Twinning Program (Sanghani, 2013).

discussed in Section 6.3.2.2, has some of the qualities of an IXA, but has focused more on IX standards development, IX deployment under those standards, and the implications for that deployment on interconnection and cross-connect markets in the North America region.[429] Each IXA has a similar NOG-like role for IX operator community development. The baseline service provided is an arena for sharing operational information. In Euro-IX, presentation categories include:

- introductions from new IXes, patrons, and guests;
- technical sessions;
- tools updates such as operations automation and route server implementation and performance;
- updates on public affairs and regulatory issues affecting IXes;
- specific market dynamics, such as video peering;
- development of reseller programs (such as those discussed in Section 6.3.1.4)
- challenges to IXes such as transit price erosion and de-peering by large actors;
- IX value from customer perspectives such as video delivery, analyses to determine new interconnection markets (potential node locations), value to gaming networks;[430]

Newer IXAs address similar topics, but are often meetings embedded in other fora. As instances, LAC-IX typically has its general meetings at LACNIC meetings, APIX at APNIC meetings, and AF-IX was recently formed at ISOC's AFPIF meeting.

In addition to member meetings, Euro-IX provides a number of resources for IXes, existing IX participants, and potential IX participants. One of the most widely used resources is the database containing IX characteristics and participant lists.[431] Euro-IX provides web interface to this database, providing access to tools for the types of analyses discussed in terms of diversity in Section 6.1.2. One view of this data set is the Peering Matrix (Euro-IX, 2013b). The peering matrix rows and columns comprise a list of IXes in the Euro-IX database for which Euro-IX has a participant list. Each cell of the peering matrix indicates the number of members that are at both IXes. For instance, the Peering Matrix indicates that the AMS-IX and the NL-ix (another IX in the Netherlands) share 154 members; AMS-IX share 250 members with the LINX and 296 with the DE-CIX. Figure 6-3 was created based on the Peering Matrix by treating each row as a feature vector and applying a hierarchical clustering algorithm to cluster IXes with similar patterns of common membership.[432]

[429]Current discussion in the Open-IX community is considering expansion of the geographic scope beyond North America.

[430]Gaming networks have been cited by a number of IXes as a sub-industry that is garnering value from low-latency, local interconnection fabrics. For instance, AMS-IX provides a case study (AMS-IX, 2013c) on a gaming network that has benefited from AMS-IX connectivity.

[431]Figure 6-2 is based on a snapshot of that database.

[432]The peering matrix is available for download as a CSV file. Different IXes have different numbers of participants. Each feature vector was normalized by dividing by the total number of participants at that IX. The result is a feature vector whose cells indicate proportion of total participants in common with the corresponding IX. The diagonal of the normalized peering matrix is thus 1.

Another view of the database is the ASN Filter (Euro-IX, 2015). For instance, using the ASN Filter a user can select a set of networks and determine which IXes they participate in. In effect, it is a tool for determining which IX or set of IXes will offer the most value for an interconnection investment. In terms of developing an interconnection bundle, the tool allows participants and potential participants to calculate potential value for exercising a platform option[433] In a number of the early conversations that contributed to developing the notion of diversity, interviews or personal conversations often suggested exploring the peering matrix and ASN Filter as a means to develop an intuition of who interconnects where. Notions of redundancy and uniqueness were stressed; the notion of diversity as a common metric is a product of the analysis in this research, discussed in (Sowell, 2013) and here, in particular in Section 6.1.2.

Euro-IX has been what is referred to here as a federating agent. Federating agents provide fora in which decentralized actors within an institutional complex convene to share information and promulgate institutional norms. Amongst the institutions discussed in this work, the NRO is perhaps the strongest form of a federating agent, acting as the collective voice of the RIRs, fulfilling the role of the ASO in the ICANN framework, and serving as a coordinating agent amongst organizationally independent RIRs. NOGs such as NANOG have some of the qualities of a federating agent, they contribute to disseminating information about routing norms, but are not explicitly promulgating a normative regime. That said, the NOGs have the *effect* of promulgating informal routing norms. Euro-IX lives between the NRO and NOGs, moving closer to the NRO. Development of the IX-F moves both Euro-IX and the other IXA's closer to formal federating agents that both formally promulgate a normative institution complex and the beginnings of engagement with external actors.

6.3.2.2 Open-IX

North America has historically been dominated by commercial IXes. The SIX (Seattle Internet eXchange) and the TorIX (Toronto Internet eXchange) are exceptions. Commercial IXes are typically affiliated with, if not owned by, a colocation facility, the most well-known being Equinix. A number of prominent, large US operators believe the US interconnection market was dysfunction. In particular, early Open-IX documentation and exploratory meetings highlighted the differences in costs between the US and Europe, in particular *a*) differences between fixed and recurring costs for cross connects ($\cdot_{xc}$ in Figure 6-1), *b*) transparency in cross-connect pricing, *c*) standardization of operational processes and quality across colocation facilities.

To solve these problems, Open-IX was formed in XX to promulgate membership-based IXes in the US. An IXA such as Euro-IX currently serves as a knowledge commons and an arena for discussion alongside a functioning interconnection market

[433]The idea of a platform option is developed in (Sowell, 2013). Exercising a platform option involves investing in transport (ix_{p-p}), colocation residence and cross-connect fees (ix_{colo}, ix_{xc}), and IX participation fees (ix_{part}) in order to gain access to interconnection options provisioned by that platform.

facilitated by a diverse colocation and IX market. In contrast, Open-IX was created in a manifestly hostile market, framed by Open-IX founders as rent-seeking monopolies. In effect, Open-IX was created to challenge the perceived rent-seeking character of incumbent colocation facilities. The IX and colocation market in the US was translucent at best. Pricing information was not available publicly from large colocation facilities such as Equinix or Coresite. A frequent question on operator lists and many private conversations was what the cost of interconnection *should* be. No one wants to be the sucker. Nevertheless, many contracts forbid sharing pricing information.

Open-IX is expressly interested in standards development for and certification of IXes. Open-IX was initially focused exclusively on the US and Canadian markets, but has been recently considering expanding its scope to any IX that wishes to apply for certification. The membership of Open-IX comprises experienced members of the network operator and IX operator community to develop IX and colocation facility standards. Open-IX has developed two standards: IX certification (OIX-1) (Open-IX, 2014b) and data center certification (OIX-2) (Open-IX, 2014a).

OIX-1 documents "minimum functionality" of an IX, but also room to accommodate well-developed, multi-node platforms distributed over a metropolitan area. OIX-1 indicates:

> The purpose of the requirements is to provide publicly available information on what participants of the IXP can expect, and not to describe in detail how the IXP is designed, built or operated. (Open-IX, 2014b)

Note the focus on what a participant may expect, not a technical detail. The intention is to establish well-defined standards and compliance mechanisms in service of a more transparent interconnection market. The ultimate intended consequence is a more competitive interconnection platform market, in particular more competitive amongst colocation facility providers. Note standards and compliance does not lower technical barriers to entry such as operational experience.

Specific prescriptions for IX service provision in OIX-1 include public and private VLAN requirements, definitions of physical interface and traffic forwarding requirements, IX platform components requirements,[434] and operational requirements such as maintaining a 24x7 NOC and monitoring the platform. In interviews with IX staff, in particular longstanding IX operators, many of the criteria in OIX-1 are what these actors discussed as outcomes of IX professionalization and maturation processes. Standards development is a codification of knowledge in the epistemic domain even though Open-IX is not famed as a knowledge commons. In this case, standardization does not convey the operational experience necessary, but the outcomes provide advice on where to focus one's efforts. A number of the European associational membership IX histories have a distinct inflection point where the IX

[434]IX platform components include switching platform, IX number resources, and route serves. This *does not* include infrastructure provided by the data center, documented in OIX-2. IX platform components are limited to information assets (such as number delegations, databases) and physical assets (switches, servers) exclusively under the control of the IX. In the case of information assets, this means exclusively delegated. In the case of physical assets, this means owned by the IX.

transitioned from an ad hoc, volunteer organization to a firm structure that could hold some degree of liability and thereby make service guarantees. Management advice facilitates leap frogging the ad hoc volunteer model to identify a stable, professionalized IX platform. The organizational mode of the professional mutual IX emerged out of demand for greater reliability.

OIX-2 specifies requirements for data centers that would like to host IX nodes. OIX-2 data center requirements impose market, operational, and technical standards. These will be briefly presented here and revisited when discussion IX norms in Section 6.4.1, in particular neutrality in Section 6.4.1.2. Consider the discussion of open access:

> The party applying [data center] must have the right to bring any new network provider into the facility and must be willing to do so on a non-discriminatory basis. An Open-IX approved Data Center must also provide non-discriminatory access to any Open-IX approved IXP for a minimum of 12 months after approval subject to available space, power, and cooling. (Open-IX, 2014a)

This requirement seems rather simple, but the notion of neutrality here separates capital investment in infrastructure from rent-seeking in downstream operations. The requirement does not preclude the provisioning of colocation-mediated interconnection. It does establish greater access to L2 infrastructure providers. In effect, the objective is to establish non-discriminatory opportunities to appropriate capacity in privately held L2 facilities.

Here open access harkens to the origins of neutrality norms in carrier-neutral requirements. Non-discriminatory criteria ensure that the parent facility does not limit participation in facilities certified by OIX-2.[435] Non-discriminatory *access* to OIX-1 certified IXes ensures these IXes can host nodes at OIX-2 facilities even if other IXes (OIX-1 or not) are present. For instance, some colocation facilities host their own (commercial) IX but do not allow other IXes to colocate at their facilities. Non-discriminatory access eliminates this limitation.

A founding motivation for Open-IX was to correct opacity in the interconnection and cross-connect markets. Consider the OIX-2 pricing requirements:

> Pricing must be offered to the OIX exchange and the participants of the exchange for all interconnection types on a fair, reasonable, open, and non-discriminatory basis. Open-IX requires transparency of pricing and MMR providers must publicly post pricing of standard elements on their

[435] For instance, this allows carrier meet me rooms to be OIX-2 certified. A key historical premise of carrier neutrality was the problems faced when network A, typically a small ISP, hosted their equipment at the facilities of B, typically a large carrier. Consider if A and B are competitors and it happens that B is the host of the IX both interconnect at. B may limit access to competitors' equipment and, in the event of failures or the need for maintenance, degrade A's service. Anecdotal evidence indicates this was the case for SprintNAP. This was also a known problem at poorly maintained IXes in Paris before France-Ix was created.

> web site, or provide an equivalent open method to all OIX exchange operators and participants. (Open-IX, 2014a, Discussed in Meet Me Room (MMR) requirements under Physical Requirements)

A foundation of a functioning market is the ability to compare the price of two good in terms of the value of those goods to the consumer. With neither price information, nor a standard of comparing services, nor a pricing history, efficient market function is severely impaired. Here, Open-IX is a commonly provisioned institutional mechanism for providing additional information, and subsequently, greater, and more precise, guarantees binding to the services offered. This is an effort to shift engagement in the market from an experience good to a search good.

Amongst membership based IXes, a common element of neutrality is pricing neutrality. Fees and port capacity costs are the same for all actors. A 10G port costs the same for a small actor provisioning one port as it does for a large actor bundling 20 10G ports. Imposing non-discriminatory pricing is perceived as a mechanism for leveling the playing field. Given the strong investment in neutrality, Open-IX's efforts at developing more neutral colocation facilities promotes the promulgation of the mutual IX regime.

6.4 IXes Rules

IXes' constitutional rules have remained stable, but the sophistication of the trade-off space amongst the three core norms has co-evolved with archetypal topologies to account for changes in demand for interconnection markets. Among other factors, IXes are contributing to what has been referred to as a flattening of the Internet. In terms of stability, the constitutional norms of mutuality, neutrality, and non-compete have persevered. That said, to highlight the nuance, each will be presented in an ideal, albeit degenerate, form. These norms are not applied as such in practice. Ideal forms facilitate conveying the spirit of the norm and precise articulations of the trade-offs amongst these norms. Starting with the ideal form, the evolution of these norms is one lens through which to explain and evaluate changes in the IX regime. Such changes will cumulate in the articulation of a trend toward more commercially-oriented, yet still mutually managed IXes.

Collective choice rules in the IX differ from the RIRs. As noted in Chapter 5, RIRs harken most closely to the IETF consensus process.[436] In contrast, IX collective choice rules do leverage consensus-based decision making, but vary from IX to IX. For instance, the LINX explicitly calls out a polling process that is at the discretion of the council, with the *only* criteria being that it is not necessarily a majoritarian

[436]They harken most closely in terms of the mechanics of how one decides if a group has reached consensus or not. They differ in the structuring of when the phases described in Section 3.2.5 occur and how interleaved those phases, in particular, active and passive consensus, occurs. Amongst the CRIs discussed here, the RIRs are the only actors that reify the active and passive phase as distinct, time delimited phases although the process and outcomes of these conceptual phases are evident and observable in the other studies.

vote.[437] AMS-IX's consensus process is largely informal, finding much of the discussion occurring initially in the hallways at meetings, later on the e-mail lists amongst active members, and, if inertia holds, synthesized into a proposal by the Board as representatives of the membership body.. France-IX is different yet still, having been designed with the experience of the LINX, AMS-IX, and other IX's experience as hindsight, operating on a "forgiveness versus permission" mode of resolution approval. While not strictly following the processes laid out by Resnick (2014)'s discussion of IETF consensus processes or those of the RIRs from Chapter 5, IXes, even the largest, are arguably sufficiently small that informal mechanisms are sufficient. The largest IX by participants, AMS-IX, has a total of 697 participants as of late January 2015.[438] Many IXes having between 50-200, and a large number have less than 50.

Operational rules are perhaps the simplest and most stable of the CRIs considered here.[439] Modern operational rules boil down to requiring participants have a public ASN and, in some cases, transit appropriated outside of the IX platform (i.e., transit via a cross connect at the same facility hosting the platform node through which the actor gains access to the broader IX platform). In addition to the requirements, most of the day-to-day rules are network hygiene. In general, the only concern of the IX is the integrity, or in the community vernacular, the health, of the switching fabric. With the exception of the LINX, if a network congests its port, it is up to the network to upgrade or remain congested.[440] The remaining operational rules are largely technical, discussed in Section 6.4.3.

In the RIRs, operational rules establish the structure and parameters of the criteria that scope number rights. IXes do not confer number rights or rights in the routing system. IXes *do* provision a common resource, the IX platform, that enhances the value of number resource as a means to provision more valuable interconnection bundles. Aside from early membership criteria in early associational membership IXes, constitutional rules and operational rules are largely address design and maintenance issues attendant with a given IX topology. In the RIRs, operational rules specify a variety of appropriation bundles, with a variety of criteria stipulating limitations based on type of use. Appropriation rights are heterogeneous, while management rights are largely homogeneous.[441]

In contrast, in the IXes, appropriation rights, access and appropriation bundles, are largely homogeneous under neutrality and non-compete. IX participants contract various port capacities, but prices and appropriation options are non-

[437]This is documented in Article 18 of the LINX Articles of Association and discussed in Section 6.4.2.

[438]Based on data from the Euro-IX Peering Matrix (Euro-IX, 2013b), retrieved late January 2015.

[439]Operational rules are the simplest in action. The development of these rules required delegation of authority to the IX firm and some degree of experimentation over the course of IXes' management and topology cycles of co-evolution.

[440]In the LINX, a fine is levied, often on the order of magnitude of the necessary port upgrade. The rationale is that even though congestion occurs on one side, it creates the impression of poor performance for all that interconnect with that network. It is an operational externality.

[441]In some RIRs actors votes, for instance when approving an activity plan or the budget, may be weighted based on the volume of number resources held.

discriminatory.[442] IX rights bundles differ most significantly in terms of management rights. More precisely, members have management rights and customers do not. Rule changes have been largely debates over the exercise of management rights that manifest in the archetypes in Section 6.2. Although not explicitly invoked in discussions, the logic of constitutional norms such as mutuality, neutrality, and non-compete shaped the rationale behind the exercise of these management rights.

6.4.1 Constitutional Rules

The associational membership IX regime, as a CRI, is based on three fundamental norms: mutuality, neutrality, and non-compete. These three norms are not mutually exclusive. Moreover, these provide the conceptual basis for IX operations, governance, and provide a common basis for reasoning about strategic decisions of IX management, IX participants, and others in the IX ecosystem that contribute to the outcomes described below. Each of these norms is described along with some of the specific aspects of IX organization and technical configuration, providing canonical instances that exemplify the range of options.

6.4.1.1 Mutuality

Mutuality is the basis for associational membership IX management. As noted before, IX staff and participants have reiterated that "the collective membership is the single stakeholder in the IX [firm]." Both neutrality and non-compete can be derived from mutuality. In the context of the IX community as an operational epistemic community, mutuality is a *normative* rationale for collective decision making. Mutuality can manifest in a variety of ways contributing to how a consensus process is actually enacted. As a norm, it is the premise for collective choice decision processes. The majority of this work describes consensus as a process, requiring organizational constructs reinforcing neutrality in such a way that one particular class of actor cannot capture the process. While the conceptual distinction between the norm and the process is useful, the two are rather entangled in practice.

Consider the norm and the process in the ideal-form of mutuality, rooted in small, ad hoc, volunteer IXes. In this model, Coasian bargaining as a means of decision making plays out in informally structured arenas. Much like the small communities invoked in early Coasian parables,[443] IX participants in these arenas have often had repeated engagements with other participants or may have a sufficient social network to ascertain reputation and credibility. The foundations of mutual accountability is rooted in the inevitable repeated engagements. In contrast,

[442]There may be corner conditions in which an IX cannot provision a particular port capacity. For instance, all ports of that size may be in use. For instance, it may be the case that a small actor wishes to upgrade a 1G link to a 2G link, but all of the 1G ports have been appropriated. There may be plenty of 10G and 100G ports available, though.

[443]The ranger and the cattle farmer comes to mind; the confectioner and the physician is the author's personal favorite; see R. H. Coase (1960).

modern forms have delegated decision-making processes, in particular technical operations, to a firm and an executive board to oversee that firm.

The collective membership quote above can be explained in terms of common resource management rights. The IX platform is the jointly provisioned common resource. Further elaborating the focus on "single stakeholder," explanations in interviews and fieldwork stresses that the IX does what its collective membership tells it to.[444] In the small volunteer IXes, this decision making process could easily occur around a large table. Like the RIRs, management rights of the IX are held collectively by the membership, hence the "single" stakeholder. As IXes grew, the decision process became more cumbersome.[445]

The ideal form of mutuality has been framed as a form of pure democracy: *all* decisions are made by the collective. The mechanics of mutuality in action has two broad dimensions: scope of management rights and mode of decision making. Over the course of IX development, the scope of management rights exercised by the collective has contracted and the mode of decision making has changed, to varying degrees, shifting the focus to holding the IX firm accountable. In this discussion and its manifestation in collective choice decisions in Section 6.4.2, structures implementing mutuality as an organizational mode reveal a range of mechanisms for eliciting decision-making direction from the constituency without the intransigence of unstructured discussions amongst potentially adversarial communities or the thrashing of mob rule.

Scope can be described in terms of principal agent delegation. In the ideal form, the collective manages all aspects of the common resource.[446] These decisions range from strategic deployment and infrastructure management decisions, finances, and handling daily operational tasks such as physical connectivity in the colocation facility. More commonly, a non-incorporated mode of delegation is in place. Volunteer IXes often share management responsibility amongst leadership and distinguished participants willing to manage the equipment that comprises these common fabric. Collectives identify more distinct roles, such as having a common board positions and operational roles fulfilled by volunteers that have demonstrated a sufficient credible commitment to be entrusted with those roles.[447] This management con-

[444]Given the overlap between the RIR and IX communities, it should not be surprising that this is also the management mantra of the RIR firm. As noted earlier, the difference is onto which set of rules, constitutional or operational, the will of the constituency modifies.

[445]Multiple interviewees have indicated early consensus processes became cumbersome as engineers belabored minutiae of operational and technical rules. One actor described the resolution of this process as the constituency developing trust in the leadership and later the IX as a firm. Another actor in a leadership position in an early IX indicated development of the firm was a strategic move to confer the firm with the authority to implement (exercise) management rights decided by the membership.

[446]Recall from Section 2.1 and 3.3 discussions of collective resource maintenance, in particular joint provisioning. Dam construction discussed is a collective effort of local participants in irrigation systems. In Japanese forest commons, both maintenance and harvest are collective, coordinated activities. In other commons, tasks may be delegated to specialized participants in the community. Ostrom's canonical case is the delegation of irrigation management, distinguishing between producers and those that sustain the provision of resources.

[447]Instances of these are the SIX in Seattle, LONAP in the London metro-region, and IX Leeds.

struct is an instance of joint provisioning comprising producers both developing and exercising management rights.

A common misperception is that for *all* associational membership IXes, the membership *is* the IX. In the case of small volunteer IXes, ad hoc volunteer management seems close to the ideal form of mutuality, but this is not wholly incorrect. In the ideal form, the membership is the primary organizational body that takes all decisions, from *a*) technical operational decisions such as IP delegations to participants; *b*) decision processes related to membership decisions; and *c*) to strategic decisions on expansion (new nodes) in the metro region or more broadly.[448] Contracting the scope of decisions taken up directly by the collective shifts the exercise of management rights from the collective principal to a more conventional model of principal-agent delegation. The principal retains the rights to change management rights, but delegates authority to exercise to the firm. A third and fourth body is added to the IX as a whole: the firm as an agent and the board as that agent's immediate principal. The collective membership has delegated the authority to exercise management rights created by the collective. The board is created as an elected representative of the principal to monitor and enforce limitations placed on the firm executing those management rights. As will be developed in the discussion of neutrality and non-compete, mutual, i.e. collective choice decision making kicks in when the strategy of the firm pushes against the current operationalization of one of the constitutional norms.

The other dimension in which mutuality differs is in how decisions are made. Membership IXes employ a mix of consensus and voting, depending on the IX. Decision-making processes are mutual in the sense that the membership contributes to evaluating the merit of changes in strategic direction. Some commercial IXes also take substantive direction from their customers. DE-CIX is commercial, but holds customer meetings to discuss service quality and potential demand for new services. DE-CIX has an elected advisory board for its platforms in Germany and is developing an advisory board for its North America platforms. Netnod is also a commercial IX but it has a board appointed by various external actors, regular participant meetings, and a leadership firmly entrenched in the community and other consensus based institutions in the operational epistemic community. In a number of interviews, participants and staff of associational membership IXes have likened Netnod to at commercial IX that behaves most like an associational membership IX. In effect, both DE-CIX and Netnod have leveraged elements of mutuality to elicit demand without the firm being directly beholden to a membership.

Moving farther along the commercial spectrum are large American IXes such as Equinix. These exhibit little if any elements of mutuality. Rather, these operate much the same way commercial actors operate, responding to market demand for services. While commercial IXes like Equinix do not directly compete with their customers, they do frequently provide their own facilities that compete with other

[448] In general, if an IX is large enough to be expanding beyond its metro-region it has typically made the transition to delegating operations to a dedicated firm, the canonical cases being the AMS-IX and the LINX.

co-location facilities. Moreover, many commercial IXes comprise a number of network infrastructure services and may even host other IXes (viz. Swiss-IX in Equinix Zurich and the LINX node in Equinix London-4).

Mutuality reinforces members' collective ownership and investment in the IX. Mutuality is the foundation of the IX as an honest broker/third party. This is especially salient to imposing uniform operational rules, and by proxy, stable rights. For instance, mutuality ensures that all network actors are treated equally on the fabric. Equal treatment is the foundation of neutrality; its various operationalizations are elaborated in Section 6.4.1.2. Succinctly, neutrality assures that no actor is privileged over another with respect to their use of and privileges guaranteed on the fabric. For instance, in the case of purchasing port capacity, prices for port sizes remain the same for all actors, regardless of the size of the actor or volume purchased.

Mutuality does not emerge out of the ether. In cases of developing an IX, the initial work is planting the seeds of mutuality: carrier neutrality and non-compete seem to emerge more easily. In a number of cases, developing a sense of mutuality requires community building, bringing together nominal competitors and developing the trust necessary to make the initial mutual investment. In the case of early IXes, this was the work of entrepreneurial engineers solving a common problem (tromboning and the attendant latency). In modern cases, conveners more resemble Mattli and Woods' policy entrepreneurs, many of whom will tell you that IX deployment is 80% social, 20% technical.[449] Modern entrepreneurs may be local or, in some cases, a mix of local actors that have reached out to external actors with experience developing IXes.

Mutuality is the basis for a variety of key IX attributes. Modern, representative forms of mutuality reify:

- distinct separation of strategic, tactical, and operational decisions, moving from constituency decision-making to the discretion of the organization/firm
- coupled with non-compete, mutuality helps preserve the domain-specific scope of the IX
- mutuality is the normative premise of negotiation within the compromise space, resulting in the balance of features and services provided by a given IX
- it is the basis for providing information available to the managers of the fabric (as third parties) to network actors that can reduce the uncertainty in resolving conflicts

The last point, provision of information, warrants some elaboration. One instance of pure information sharing has to do with the traffic from distributed denial of service attacks that may traverse the platform. The IX can reduce uncertainty by providing all of the actors affected by a DDOS with information regarding the source and targets that the firm can observer from its vantage point as the common switching

[449]A number of these cite the 80/20 rule, then go on to say the proportion is more like 90% social, 10% technical.

fabric, but, unless that DDOS affects the platform itself, the IX leaves resolution to those affected. This is a specific instance of neutrality, discussed in the next section.

6.4.1.2 Neutrality

Balancing a diverse constituency and non-discriminatory practices are key elements of neutrality. In terms of infrastructure management developed in that applications of Frischmann's model of infrastructure in Section 3.1, neutrality is the normative manifestation of non-discriminatory financing and appropriation pricing. In simpler terms, the spirit of neutrality is to ensure that no single actor or interest group is privileged by the IX firm. An early manifestation of neutrality is carrier neutrality. Simply put, *carrier neutrality* means an IXes' nodes should not be hosted at carrier facilities. Rather, IXes should be hosted at third-party facilities that do not have an interest in influencing IX operations.[450] The former articulation of neutrality is a general ideal form, the latter a specific instance of neutrality, but still an ideal form. Both have been adapted in different regions and as the IX regime has developed.

In terms of rights, neutrality ensures non-discriminatory access and appropriation rights. Manifestations of neutrality can be observed in the rules structuring the non-discriminatory delegation of rights. Borrowing Lessig's notion of an "architecture of control," types of neutrality in the organization contribute to ensure both organizational modes and the IX platform topology are "architectures of neutrality" rather than being bent to serve particular interests.[451] Types of neutrality can be roughly partitioned into facilities neutrality (platform infrastructure) and organizational neutrality. As these are discussed, it is important to note that not every type of neutrality needs to be present at a given IX. Subsets of the types of neutrality may be sufficient to offset potential deficiencies in others. The criteria for evaluating neutrality is whether the subset of structural neutrality is sufficient to provision non-discriminatory access to resource rights and those absent from the subset do not threaten non-discriminatory access.

As per above, *carrier neutrality* is one of the early forms of neutrality. Earlier lessons from the four US NAPs post NSF divestment have been cited in interviews as an influences. A frequently cited instance was that, for a period, participants at SprintNAP were required to purchase a circuit from Sprint to interconnect there. For Sprint's competitors, this meant reliance on a competitor for interconnection critical to their business. As elaborated in the discussion of OIX-2, access to equipment, for both the IX firm and the participant, is a factor in assuring service and connectivity.

[450] Early on, IXes were just another colocation facility tenant. It was not necessarily clear they would contribute to the value proposition of the colocation facility. Now IXes are sought after tenants, especially in Europe where IXes are prolific. To preserve neutrality in general, IXes now have a notion of data center neutrality, sometimes referred to as colocation neutrality, discussed below.

[451] See Lessig (1999) in general. Note that a premise of this work is a synthesis of the non-discriminatory infrastructure described by Frischmann and an argument that technical artifacts *do* have politics *in* context, a variant of Winner's argument in (1980). Here, public, private, and social goods atop the Internet, using the provisioned mode of non-discriminatory communication as an input, do have distinct, particularistic politics.

Limiting access to equipment could adversely affect participants performance.

Carriers have a vested interest in preserving traditional means of connecting with external actors, acting as the single interconnection point for transit and transport services. Carrier-neutrality means that IX facilities, namely the real estate in which switching equipment is housed, is not controlled by a carrier.

Carrier-neutrality has a number of constructive consequences:

- network actors have equal access to the facilities, mediated by the facilities owners themselves and IX management
- markets for transit and transport services
- competition amongst colocation facilities
- access to market-priced private-peering

Taken together, this ensures that carriers with a vested interest in preserving conventional means of acquiring access to the rest of the world/Internet through transit and transport arrangements could not unduly affect access to equipment or impose additional contractual and technical requirements on competing arrangements Carrier neutrality shifts control of facilities providing an alternative to conventional transit and transport relations to a third party, namely co-location facilities. IXes and co-location facilities have a mutually beneficial relationship—co-location facilities provide facilities for interconnection and a popular IX will bring business to the co-location facility.

To illustrate, consider an IX C hosted at a carrier A. Further, A may have created the IX to encourage interconnection between itself and the participants. A may benefit by selling its own services over the IX, even offering discounts on interconnection services to those purchasing A's core services.[452] B is a competitor of A and joins IX C to gain access to participants. In other words, B joins to gain access to a market. Given B is a competitor of A, if B needs service or access to its equipment, it is not in A's interest to support the operations of a competitor. The solution is to host IX equipment and the equipment of participants at a third party hosting facility.

As IXes grew, they became sticky (as per discussion in Section 6.1.2). By proxy, the colocation facilities they were located in also became sticky. Locating nodes of an IX in the physically diverse locations of a single colocation firm creates the potential for that firm to influence the operations of the IX. If the colocation facility also offers its own IX, an argument similar to the carrier-neutrality argument can be made. Even if the colocation facility does not have a competing IX, it could potentially increase the costs for the IX, knowing the transaction costs of relocating are prohibitive.

The solution to these problems is data center neutrality. *Data center neutrality* means multinode platforms place their equipment in data centers from different owners. For instance the AMS-IX, LINX, France-IX, and NetNod use a diverse set of data centers. Data center neutrality has a number of other effects. Data center neutrality is a form of redundancy for the IX. Poor performance of the data center firm

[452]This is often the rationale for ISPs to create IXes. This was noted in particular in discussions of the multiple small, poorly managed IXes in Paris before France-IX was founded.

may lead to deficiencies such as servicing cross-connects and access to equipment. Other deficiencies include security, redundant power supply provisioning, and cooling systems.[453] Diversification of data center provisioning creates both redundancy and competition.

As per the discussion of diversity, data centers now solicit IXes to locate nodes in their facilities. For instance, one IX provider described the process of selecting a colocation facility for the location of a new node. One condition for locating a node at a colocation facility is whether there are a sufficient number of existing participants or potential participants[454] at the colocation facility to cover the costs of locating a node there. This typically means the existing and/or potential participants will be provisioning exchange ports at that facility. If this criterion holds, creating the node benefits the IX, participants, and the colocation facility. In this case, the threshold for the number of participants is typically such that their membership fees and/or added value to the IX will "pay off" the costs of node deployment.

Alternately existing and potential participants may have presence at the colocation facility, but it is not clear these are sufficient to cover the costs of node deployment. In that case the IX and the colocation facility agree that if the colocation facility does not attract sufficient participants, the colocation facility will cover the difference in node deployment costs. This creates competition amongst colocation facilities, especially given the IX will not give one colocation firm all of its business. That said, some colocation facilities host multiple IXes. Having multiple IX nodes contributes to stickiness as well.

Another element of facilities neutrality is connectivity neutrality. Multinode platforms and platforms with resellers both incorporate transport as part of the platform infrastructure. Like relying exclusively on a single colocation facility firm, IXes also diversify their connectivity options. *Connectivity neutrality* means that no single connectivity provider is privileged by the IX, giving it leverage to influence IX operations. Where possible, membership-based IXes tend to provision diverse transport for purposes of redundancy and non-compete. On the surface, IXes providing any form of transport violates non-compete. The solution, elaborated as a relaxation of non-compete in Section 6.4.1.3, is to ensure any participant with sufficient transport facilities can potentially provide inter-node connectivity or reseller transport.

In addition to colocation facilities and connectivity, IX facilities also comprise switching equipment. Supplier neutrality assures investment in a single equipment supplier does not adversely affect IX integrity. Two forms of supplier neutrality have been identified: initial sources of IX equipment and platform redundancy. In early IX development, volunteer non-profit IXes have received equipment donations from a variety of sources. Although not necessarily a direct means of control, influence may be expected. Solutions to initial equipment donations follow other neutrality solutions: diversification of sources.

The professionalization of IXes came with demand for increased resilience and redundancy. Simple hardware redundancy implies equipment available for failover.

[453] Recall that all of these are specified in the OIX-2 standard.

[454] Potential participants are those that have expressed a credible interest in participating at the IX.

A common redundancy strategy is parallel fabrics. The LINX, NetNod, and PTT.br are all large IXes that have implemented redundancy by maintaining two LANs, i.e. two independent switching fabrics. Rather than run redundant fabrics with equipment from the same supplier, these IXes use equipment from different suppliers for each fabric. For instance, the LINX runs a Juniper and an Extreme network.

The second class of neutrality is administrative neutrality. Administrative neutrality assures that the IX's governing bodies are (1) representative and (2) independent of outside influences. In terms of political economy, administrative neutrality strives to avoid "regulatory" capture.[455] Representativeness means that a single interest does not dominate the board of the IX. Consider the board of an associational membership IX. For instance, the current LINX Council[456] comprises an employee of a hosting provider (NetConnex), an operations analyst (Sebastian Lahtinen), a network architect for the incumbent access and transport network (BT), a CTO of a colocation and IaaS (Infrastructure as a Service) provider (Markley Group), a director at a business and educations network provider (Exa Networks), and a strategy manager for a large video delivery firm (Netflix). In this case, a diverse set Internet infrastructure industry interests are represented on the board.

AMS-IX has a similarly diverse group: representative from a tier-1 (Level3), representative from a large search company and CDN (Google), a DNS services provider (Verisign), a national (non-Dutch) incumbent telecom (Turk Telekom), and a local Dutch business ISP (BIT). The AMS-IX board selection process is an interesting case at the intersection of representative mutuality, mutuality, and a form of credible commitment to joint management. When there is a vacancy on the board, the board interviews potential candidates and suggests a candidate for ratification by the membership. The argument behind this process is that this helps maintain a balanced board. One board member indicated that nomination also ensures potential board members all get along professionally, giving rise to a functional board.[457] For instance, when the last CDN employee left the board, that board member was replaced with another (Akamai left, Google came in). Some actors have critiqued the board nomination process. If the membership is not pleased with the nominated party they may choose not to ratify the nominated party and elect their own, independent of board influence.

France-IX is also an interesting case of administrative neutrality. France-IX was created to rationalize the French interconnection market. A number of the dysfunctional IXes were supported by local ISPs and incumbents.[458] The founding board of France-IX was specifically formed from well-known community members, two from outside France, two from inside France. These were Akamai, Google, Jaguar,

[455]Regulatory in this context, like most of this dissertation, refers to the nominal definition of regulation, creating order amongst social and economic endeavors. It does not necessarily mean *government* regulation.

[456]The LINX council is the equivalent of an elected board, its members elected from LINX members.

[457]This interview topic was the consequence of discussion of a rather dysfunctional board member that often created strife with the other four board members, more so than in the normal course of constructive conflict.

[458]See Section 6.5.2, in particular Table 6.3.

and Neo Telecoms. Further, there was the perception that local incumbents would attempt a board takeover. The France-IX articles explicitly built a mechanism into the initial elections to avoid this threat and staggered the introduction of elected board members rather than replacing the entire board at once. The strategy was to give France-IX time to gain critical mass, to become sufficiently "sticky" for a broad enough constituency to sustain itself against such a takeover attempt.

Operator, or participant, neutrality is a tacit objective. Operator neutrality assures that no single class of operator dominates the constituency. As a market, a successful IX is a market—both sides of the market, suppliers and consumers, need to be present. Although telecommunications operators, especially the large transit providers, are not often considered IX supporters, the IX should also welcome these participants and their contribution to IX activities. Early on, some IXes limited membership based on network investment, did not allow content providers, in effect did not allow actors that were not tier-2 ISPs. Since then the criteria has been substantively relaxed to requiring at a minimum a public ASN.[459]

Operator neutrality is the demographic manifestation of mutuality. Operator neutrality ensures sufficient representation of the diverse set of infrastructure industry interests are present and contribute to platform development. The result is an equally diverse market of interconnection options. Operator neutrality is considered "tacit" because it is so fundamental to the market; excluding classes of actors, such as content, only yields an incomplete market. Operator neutrality also demonstrates shades of non-compete. In the next section, membership criteria is described in terms of balancing critical mass while not poaching participants' customers.

6.4.1.3 Non-Compete

Following the discussion of operator neutrality, an IX's constituency[460] ideally comprises a diverse set of network actors, each of which provisions its own services, taking IX connectivity as a (non-discriminatory) factor of production. The spirit of the non-compete norm is to ensure that IX provisioned services do not compete with the services of its members. The ideal (degenerate) form of non-compete asserts that the IX will not adversely affect the revenue of its participants. As an aspiration, the ideal form conveys the spirit of the non-compete norm but it is impossible to implement in practice.

Consider associational membership IX regime's original objective: reduce transit costs for traffic exchanged amongst geographically proximate participants. In a number of cases the incumbent transit providers (frequently a regional tier-2 provider) were founding members of the IX.[461] Forming a local IX will inevitably

[459]There is certainly variation. The SIX solicits feedback on the e-mail list when evaluating a network's membership application. Some IXes "quarantine" new members to validate hygiene rules are followed. See Section 6.4.3 for a discussion of network hygiene rules.

[460]The constituency may be either members or customers, or in some cases, a mix of the two. This should be a mode of network actor::IX relation.

[461]For instance, BT was a founding member of the LINX. KPN was a founding member of the AMS-IX.

shift traffic from local transit relations to interconnection bundles on the IX. The benefits for participants include reduced tromboning, improved latency, and redundancy. Selecting for IX formation implies that, collectively, these benefits outweigh the costs to those that lose some transit business. It is also arguable that some of that business was replaced by transport to the IX.

Tacit in the early trade-off is the fundamental criteria for relaxing non-compete. In general, the trade-off is whether the collective gain is greater than the marginal loss by some actors. Non-compete has been relaxed when *a*) the motivating change improves the value of the IX as a whole and *b*) if the potentially competing service may be provisioned by the participants with whom it is nominally competing. Non-compete issues resolved by participant provisioning include connectivity between nodes in a metro region, reseller connectivity, and remote node connectivity for platforms with regional scope. For instance, early internode connectivity of LINX nodes within the London metro area was contested by Colt, a local fiber provider. Colt ultimately conceded. The contract for inter-node connectivity is up for tender; anyone can compete for that business. Moreover, internode transport is *exclusively* for platform traffic. The LINX does not offer a "general transport service." As such, the LINX does not compete with its members that provide more generalized point-to-point transport. This is generally the case for many of the associational membership IXes evaluated in this study. Arguably, the LINX, and other IXes that also have inter-node transport within their respective metro regions, contribute to demand for stable volumes of traffic for transport providers.[462]

A more controversial non-compete issue is internode or inter-platform connectivity beyond the metro-region. A number of IXes connect nodes over longer distances. An initially controversial inter-platform connection was between AMS-IX's platform in the Amsterdam metro and their platform in Hong Kong. While logically connected by a common transport provider, the contract is available to anyone that meets contractual agreements for interconnection. Like internode transport, inter-platform transport only supports interconnection and traffic exchanged amongst participants on one platform with participants at another. Inter-platform connectivity is not a general transport service offered in competition to pure-play transport providers. Under this interpretation, the IX is not directly competing with transport and transit carriers for long-haul data carriage.

6.4.1.4 Implications of Constitutional Norms

Non-compete has a variety of additional implications when linked with mutuality and neutrality. Non-compete, taken together with neutrality and mutual decision making, positions the IX as an honest broker/third party. In terms of larger NRS issues, the platform fits the non-discriminatory characteristics of infrastructure writ large established in Section 3.1. In its capacity as a third party, the IX can act as non-normative mediator between conflicting actors. For instance, a commonly cited instance of neutrality is the case of a DDOS across the fabric. The IX will

[462]Section 6.5.4 discusses the trade-offs faced by transport providers in the process of relaxing non-compete.

not interfere unless the DDOS effects the integrity of the fabric itself. That said, it will use its purview of the attack to share information about the sources and destinations of DDOS flows to those affected, facilitating coordination necessary for remediation.

Mutuality and non-compete act in concert to limit the functional scope of the IX. For instance, non-compete also covers a number of other services such as providing news servers, hosting,[463] and other services that may be provided by its membership. With some exceptions, the IX does not provide any application layer services that are not related to monitoring IX participation. Exceptions are network time services and DNS Anycast. Early on, news services were perceived to compete with those provided by ISPs, but time servers and DNS were generic network services. In effect, non-compete coupled with neutrality limits scope creep as well as serving to preserve neutrality.

While non-compete and mutuality have limited the functional scope, the geographic scope of the infrastructure, in particular the deployment of platforms by large "international" IXes, has grown. Counter-intuitively, the geographic scope of infrastructure has grown in part *because* of neutrality and non-compete. While there was contention over inter-node connectivity early on, transport providers have shifted from seeing IXes as competitive substitutes to seeing IXes as complementors in the larger infrastructure provisioning value network. For instance, each IX with a reseller program lists transport providers that offer connectivity to the IX. Among these, Atrato and IX-Reach are both prominent. IX-Reach, as the name implies, specializes in aggregate transport services to facilitate remote actors reaching IXes and the development of sophisticated platform bundles. As noted earlier, CABASE's carrier-member model is another instance.

Non-compete is largely discussed as a relation between the IX and its participants. A form of non-compete also exists between associational membership IXes. A simple interpretation would raise the specter of collusion between the IXes; this is not the case. In fact, there is healthy competition amongst the large IXes in terms of quality and resilience, but poaching of customers is not a productive endeavor. Non-compete amongst associational membership as a product of broader mutuality in the interconnection market. Consider Figure 6-3. The three largest IXes, AMS-IX,

[463]Hosting is an interesting instance of non-compete. In mature interconnection environments, IXes are hosted at colocation facilities. The IX does not have its own facilities in which to "host" servers. In contrast, in less developed regions, such as South America, Africa, and parts of Asia, IXes may be hosted in any facility the IX deems has sufficient power, cooling, and security. This may manifest in the IX leasing property. Under a conventional strategy of horizontal expansion and collateral benefits, capital investment in a facility for hosting an IX could easily support server hosting. That said, such hosting would compete with would-be hosting providers, especially in nascent infrastructure markets. Neutrality would also be affected by such hosting. If the IX becomes dependent on this revenue stream, it is less likely to give up that revenue stream in the interest of developing a colocation market in its region. That said, some IXes in this scenario walk this line. For instance, normally a CDN would place a cache in a colocation facility hosting an IX node. If the hosting facility is managed by the IX, the IX collects the hosting fees the CDN would normally pay to a colocation facility. Under a cost-recovery model, facilities fees may be quite low, attracting others that wish to host "on the IX" and further delaying the development of a local colocation facility.

LINX, and DE-CIX have similar patterns of participant overlap. Further, geographically proximate IXes have similar patterns of participants with one another and other IXes.

The patterns of membership with other IXes is driven by the distribution of diverse members that participate in a number of IXes. Participants diversity score is considered an indicator positively correlated with the value of effective IX management and performance to that participant.[464] When IXes compete with one another in a region, the fees of members common to those competing are essentially funding what some members consider unproductive development efforts. As discussed earlier, a key value of the IX is its role rationalizing the local interconnection market. Introducing additional IXes to a market does create competition, but it can also fragment a functioning market.

Consider two stylized instances of invasion, based on instances discussed in interviews.[465] Consider an existing IX S is operating successfully in a particular region and an ambitious IX A "invades" by deploying a node, either connected back to A's home region or as a disconnected platform in S's region. For the intersection of S and A's members, these common members are supporting a conflict between two organizations they fund through membership fees and port capacity. In the best case, a large metro-region gains a degree of diversity. In the intermediary case, the competition incents a potential complacent incumbent IX to improve its services.[466] In the worst case, the market will fragment. One outcome will be reducing interconnection potential (lower platform diversity) on both platforms. For those that with a vested interest in access to the complete market, this requires building in to *both* IX where before that actor needed only build in to, maintain, and monitor one. Both of these are costly to participants.

The second competition scenario is an attempt to *repair* an already fragmented market. The France-IX narrative, developed more fully in Section 6.5.2, is such an instance. Before France-IX, Paris had nine IXes, considered by the community to be an extremely fragmented market.[467] As an exercise of mutuality in the broader market, France-IX founders performed a survey to determine demand for a professionally run, mutual IX. In particular, the founders asked if potential participants would be willing to pay for such an IX: they indicated yes. France-IX displaced or took over a number of the dysfunctional IXes. Now France-IX has more than 250 participants. Although technically an invasion of a market with existing IXes, the France-IX effort is generally considered a positive development.

In the RIR and anti-abuse communities operational rules are typically the object

[464]It should be stressed that this is a broad indicator, on the granularity of "getting the sign right" and should not be considered as a deterministic predictor.

[465]This discussion focuses on mutuality across IXes but are considered in more depth in Section 6.5.2 and 6.5.1.

[466]For instance, when the AMS-IX deployed to Hong Kong, the existing IX announced service improvements.

[467]In early fieldwork, numerous actors independently encouraged the investigation of the France-IX story as the constructive development of an IX to replace dysfunctional IXes and rationalize a dysfunctional market.

and product of collective choice decision making. In the associational membership IX regime, provider-specific interpretations and applications of constitutional norms are the object of collective choice processes. The next section describes the range of collective choice processes in associational membership IXes.

6.4.2 Collective Choice Rules

Early on the object of collective choice rules were rather broad, serving as a vehicle to explore the mix of both organizational *and* technical IX management options. As per earlier discussions, the maturation process has delegated much of the technical to the IX firm. In the modern interconnection infrastructure, the object of collective choice rules is to *a*) maintain the bounds of management rights to be exercised by the firm, and *b*) to adapt implementations of constitutional norms to meet the demands of the interconnection market. This is particularly the case for medium and large IXes in Europe. In contrast to the scope of resource policy in the RIPE region, while the geographic scope of European IXes has grown, the scope the rules considered and frequency of updates by IXes collective choice process has contracted. This is especially interested because, as noted before, the individual constituents are members of both communities.

Mutuality is the normative basis of IXes collective choice process. Mutuality is the basis for the collective membership to *a*) structure and parameterize management rights and *b*) confer the authority onto the firm to implement those rights. Recall the common generalization of IX resource management: "the *collective* membership is the single stakeholder in the IX firm." This assertion of mutuality has been reiterated multiple times in describing IX management.

Collective choice rules in firm-based IXes often affect the scope of management rights. The nuance of management rights typically focus on the *bounds* of the platform, not the particular operational rules establishing particularistic appropriation and exclusion rights (or bundles). Particular appropriation rights, such as technical mechanics for participating on the IX, rules regarding capacity and port congestion, and others are largely well established; customer rights bundles are an obvious exception. In terms of the IX as a common resource system, the IX is the resource producer and

In the RIR system, management rights can be said to confer rights at three levels: *a*) technical operations of the resource management facility, in the case of the RIR, the registry; *b*) resource policy that shapes, parameterizes, and structures delegation practices; *c*) the strategic direction of the RIR firm. Rights delegated in the IX are *access* to interconnection options and exchange capacity. The parameters and structure of these bundles are much more static than the resource policy analog in the RIRs.

To better understand the object of collective choice decisions, consider the character of the two types of resource unit in play. Interconnection options are a joint product of the number of ports (not necessarily capacity) and simple platform diversity (number of participants utilizing ports). Interconnection options are not

scarce.[468] Exchange capacity *is* scarce under a static analysis. Under a dynamic analysis, under poor IX management capacity can be scarce, but it is renewable. A function of IX management is to monitor participant demand and utilization in order to incrementally increase capacity as necessary. IX capacity is far from finite in the same way IP addresses are.

Neutrality ensures that all actors are afforded non-discriminatory access (non-discriminatory pricing) and entry to the interconnection platform. While the types of neutrality have been refined as IXes have matured, the effects on the structure and accessibility of the basic appropriation bundle for participants[469] has not. Generalizing the three levels at which RIR management rights are applied, in the case of IXes, the second, operational rules, has remained largely static. What remains as the subject of collective choice decisions is the first, the technical operations of the resource facility (the platform) and the third, the strategic direction of the firm managing the common resource; the former will be addressed first.

Amongst IXes with a well-developed firm structure, operational management rights have been delegated to the firm. Technical operational decisions are the purview of the staff. In contrast, a number of IXes reported that early technical decisions were points of contention amongst the membership. That said, as those were resolved, participants both recognized the solution did not need to be debated each time and tacitly delegated authority to the staff. Leadership in multiple IX providers have noted that they have exclusive technical decision making power over IX operations.

As participants began to rely on the IX as a professional service on which their value proposition depended, they chose to delegate rather than exercise nuanced mutual decision making. This was an investment in creating a loci of provider-specific knowledge rather than relying on dissemination within a volunteer group. The result is a technical staff that can quickly respond to failures, maintaining a professional level of service. More importantly for participants depending on the IX, not only can staff respond, they are paid to respond and are held accountable when they do not. The formalization of this accountability (making accountability more durable) providers participants with greater guarantees when considering participation in an IX.

What remains as an object of collective decision making is the strategic direction of the firm. This typically manifests in refinements of or relaxations of constitutional norms. The range of each constitutional norms were presented briefly in the previous section; nuance in particular instances are presented in Section 6.5. Of the three levels of management rights, collective decisions largely shape the interpretation and implementation of constitutional norms. Within the scope of the strategic remit of the IX, the IX firm is delegated the discretion to define operational rules.

Of the three levels of management rights, the first two have substantive technical

[468]Recall the discussion that while they technically consume exchange capacity, this is relatively marginal compared to subsequent traffic exchanged.

[469]This statement is explicitly limited to accessibility for established participants. Topological changes documented in Section 6.2 *have* changed the geographic scope of accessibility by lowering barriers to participation.

components. The latter is strategic. The first, operational decisions, are technical aspects of managing the resource itself, rooted in the knowledge of the operational epistemic community. The second, resource policy, is a process of resource system management policy making informed by technical dynamics. The third, strategic direction, determines the scope and remit of resource managers. Strategic direction in the RIR deals with the functional scope of the RIR. In the IX, strategic direction is concerned with navigating the functional and geographic scope of the IX.

Collective decision making in the IXes vary in their mix of eliciting operational knowledge and voting. Eliciting operational knowledge and preferences on potential resolutions follows the general community engagement mechanisms. In some cases, consensus processes are used for decision making. In other cases, community engagement mechanisms resemble and have some of the characteristics of consensus processes, but decisions are affirmed by a majoritarian vote. For instance, CABASE uses a consensus process to decide general resolutions, debating these changes until a consensus is reached.[470] In the LINX, Article 18 of the Articles of Association describes the outcomes of what constitutes a valid conclusion to a decision in the LINX polling process:

> unless a poll is demanded a declaration by the chairman that a resolution has been carried, or carried unanimously, or by a particular majority, or lost, or not carried by a particular majority, and an entry to that effect in the minute book of the Company, shall be conclusive evidence of the fact without proof of the number or proportion of the votes recorded in favour of or against that resolution. (LINX, 2015c, Article 18)

As a part of the legal articles of, this describes the outcome and the actors, in this case the Council Chairman, the decision on consensus. In community vernacular, this designates "who calls consensus."[471]

In terms of the decision making process itself, the LINX engages in what is refers to as a consultation.[472] The notion of a consultation and how it is used is telling of the process. Consider the description of the LINX's consultation process:

> Regarding consultation prior to voting, although in general what we mean by this is to consult for a period (typically three months, to allow for successive LINX quarterly meetings to be used) ... We should also emphasise that consultation can take place in many ways—not limited to just quarterly member meetings in London. Quarterly member meetings

[470]This is based on an interview with Hernan Seone.

[471]The notion of who calls consensus can be traced back to early drafts of Resnick (2014) and the final document. More generally, this is an instance of what Hart would refer to as a rule of recognition. Rules of recognition are essentially those rules that indicate where authoritative statements of rules may be found, such as in a nation state's constitution and documented statutes. Here, the LINX articles highlight the seat of authoritative decisions on consensus to be the Council Chairman, an elected representative of the community.

[472]Other fora refer to the process of collecting information, engaging in consensus mechanisms, amongst its constituencies, as a consultation. For instance, ARIN refers to its Public Policy Meetings as "consultations."

> act as a focus, and some definite timetable discipline, but we are getting more sophisticated about using other mechanisms[.] (LINX, 2015a)

The phases of consensus are not as crisply delineated in the IX communities as they are in the RIRs. Rather, submitting an issue for an IXes general meeting docket may come from any member.

In the LINX, the issue is presented at one member meeting (as above) but not decided at that meeting. In the interim, the Council, membership, and firm discuss the issue via the communication mechanisms described in Chapter 4, in particular via member e-mail lists. LINX staff have indicated that the initial reaction to the consultation and the interim discussion decide whether an issue goes to a poll or not. If the issue has sufficient support, as gauged by the Council, the issue, as revised based on that input, is presented at the subsequent meeting.

Other IXes do not make explicit, formal mention of consensus processes, but interviews and documentation describe informal community engagement. For instance, in the AMS-IX Board Code of Conduct:

> Be an ambassador: Keep in mind that you are always recognized as AMS-IX Board Member, even when you are 'wearing a different hat' (AMS-IX, 2013b)

Taken by itself, this is a simple statement on representation. Interviews and private conversations with both staff and AMS-IX board members confirm that community engagement is a means to elicit preferences that inform how to handle potential changes to strategic direction or the remit of the AMS-IX.

For instance, various forms of the AMS-IX reseller program and the development of remote nodes required community approval. The informal component of community approval comprised board members presenting the community with draft resolutions and soliciting comments that were incorporated into resolutions. This comment solicitation process has the character and spirit of a consensus process in that it elicits input and direction from participants. For early proposals, this solicitation process, by both the staff and the board, is a form of problem elicitation. The quote above regarding being an ambassador is an instance of this. The Board Code of Conduct also goes on to encourage additional elicitation roles and basic mechanisms like using the AMS-IX business card when dealing with AMS-IX customers. In additional to firm management duties, the AMS-IX, and other IX boards, are in a continual process of problem identification and elicitation.

Informal consensus mechanisms play the role of problem identification, active consensus, and passive consensus. The official mechanism binding a rule is decided by a vote. For instance, in terms of achieving consensus, the AMS-IX reseller programs and remote node programs were both successes. In contrast, an early effort at developing a commercial services program was not. The AMS-IX staff suggested a services company that would offer additional services atop the AMS-IX fabric. This company was rejected by the membership in discussion before a vote occurred.

Of the three CRIs covered in this dissertation, the consensus process in the associational membership IXes is the most informal. While the specific mechanics (how

and when resolutions are presented) and nomenclature of consensus differs, the phases of problem identification, active consensus, passive consensus, and evaluation are present. Individual IXes have perhaps the smallest sets of active members (they also have on average the smallest number of total participants) of the organizations in the CRIs. While the geographic scope and volume of traffic on the Internet has grown, IX constituencies and consensus processes have largely retained the "close-knit, yet loosely organized" character of jointly managed resource systems. In smaller and some medium-sized IXes,[473] interviews indicated that discussions do take place at member meetings, but they are much more informal than those in the larger IXes such as the AMS-IX, LINX, France-IX, or CABASE.

6.4.3 Operational Rules

The specific operational rules of the IX are generally specified, monitored, and enforced by the IX firm. The bounds of what constitutes operational rules are set by the membership collective specifying management rights the firm is authorized to exercise. These are in contrast to tactical decisions that may require participant consultation or a strategic change requiring a change in the constitutional remit,. Operational rules of the IX are largely technical rules that ensure smooth participation and the technical integrity of the switching fabric. There are a number of common operational requirements across IXes: resources necessary for participation, network hygiene, rules regarding capacity, route server participation, and whether multilateral peering is required or not.

6.4.3.1 Technical Participation Requirements

Technical participation requirements can be partitioned into NRS resource requirements and private resource requirements. NRS resource requirements are simple: a public AS number for exchanging (provisioning) routes on the fabric. Some exchanges also require upstream connectivity independent of the IX. Recall the discussion in Section 2.1 that number resources are one half of the NRS resources necessary to participate in the control plane—the other half is connectivity and capacity. A public ASN delegation is necessary; an IP address delegation is not.

Consider the scenario where an IX participant p does not have an IP address delegation. Such participants typically have an existing upstream provider t (an LIR or the customer of an LIR), typically transit, that has delegated p some block of numbers. Participant p may then exercise interconnection options on the IX platform, provisioning routes to the blocks delegated to p by t.[474] Following the discussion

[473] For instance, LONAP, IX Leeds, and the SIX.

[474] If both p and t are participating on the IX, a third participant o may have an interconnection relationship with both p and t. In terms of BGP selection on prefix length and AS-path length, the route provisioned by p will certainly be shorter and the prefix will likely be longer—o will select the path provisioned by p over a longer path and shorter prefix AS-path provisioned by o. Further, it is likely that the cost of exchanging traffic is lower than exchanging traffic with t (given o is also likely a transit customer of t). This scenario is also a canonical negation of the degenerate non-compete norm: traffic exchange between p and o, over the IX platform, has reduced t's revenue from both p

of the IX serving to develop strategic interconnection bundles, IX participation may be the first step to developing non-transit redundancy and more direct routes to actors with whom the participant exchanges significant traffic. This may also be the precursor to acquiring a portable address delegation. Recall that the threshold for an initial delegation was historically lower in some regions for actors that are multihomed than those that are single homed. Participation in an exchange may help fulfill multihoming requirements, allowing a network to acquire a portable delegation sooner.

The private investment portion of technical participation requirements entails connectivity to the IX and routing equipment. These requirements differ for direct participants and remote participants. Direct participants "build into" the IX, provisioning transport and locating (hosting) equipment at the facility hosting an IX node. If colocating with the node is exclusively for IX participation, this may also mean procuring a router dedicated to IX participation. For would-be participants in the IX's metro-region, if they are already in the same facility this may not be a substantial investment, requiring only a cross-connect to the IX node.

For those outside the region, long distance transport and a dedicated router may have relatively substantive uncertainty. Reseller programs reduce the costs and associated risks of accessing interconnection options. Remote participants need only provision connectivity to the relatively local PoP of a reseller. Remote participants may already have connectivity at that PoP for their transit. This reduces transport and equipment costs. Moreover, the remote participant may initially provision a small amount of capacity, further lowering the costs of experimenting with the value of IX participation. If the remote participant garners value from exercising interconnection options, it can easily coordinate with the reseller to increase capacity. Moreover, given resellers are typically transport providers, if the remote participant garners sufficient value to become a direct participant, it has a pre-existing relationship with a provider that can provide transport to the IX.

Each of these participation processes are coordinated by the IX firm. When a participant establishes connectivity, either by building in or via a reseller, the IX firm delegates that participant an IP address to use for exercising interconnection options with other participants. The IX must also coordinate VLAN or MPLS identifiers (depending on implementation) with the reseller and participant for remote participants. In terms of the basic participation requirements (ASN and connectivity) as well as the more specific requirements in the following sections, the IX plays a neutral coordinating role. Refining the notion of having been delegated authority to exercise management rights, these rights may be more accurately described as rights to establish non-discriminatory coordination procedures.

---

and o. It may even be the case that o also receives its address delegation from t. Scaling this to a regional transit provider and a regional IX, it is not surprising that early large networks had mixed feelings about IX participation. The IX could lower the transit provider's costs, but also threatened to diminish revenue from the exchange of local traffic.

6.4.3.2 "On-Net" Traffic Over the IX Platform

One operation rule regarding traffic is whether traffic should or should not be exchanged between interfaces of the same member, i.e. whether the common fabric can be used as part of a member's "private" network. This latter case, exchanging traffic between interfaces of the same member, is an interesting case of non-compete manifest in the operational rules of the IX. In a strict interpretation of non-compete, traffic should not be exchanged between interfaces, especially between nodes within a platform. If it were not for the IX, an actor that wished to move traffic from one part of the region covered by the IX platform to the other would have to either build their own infrastructure, contract dark fiber, or contract transport. Relaxing non-compete for internode connectivity balancing lowering barriers to creating a local market against potential losses of typically metro-transport. Traffic traversing the common platform but between interfaces of the same interface is not an "exchange" of traffic.

Not all networks limit exchange of traffic between an organization's interfaces. LONAP explicitly notes this is an option:

> Members are permitted to pass traffic between their own ports and can even request private VLANs between their own ports or to other members, such as for DSL aggregation. We would remind prospective members that the LONAP exchange is not a replacement for an inter-site link and as a membership organisation our services are provided on a 'best efforts' basis. As such, relying on solely a LONAP connection for mission critical requirements would not be advised, despite our excellent reputation for service and stability. (LONAP, 2014a)

An alternate interpretation of the trade-off between constitutional norms explains this choice. The exchange platform is mutually (jointly) provisioned by the community through membership fees and port capacity fees. In particular, investment and provisioning of exchange capacity tracks utilization. Assuming participants heed the warning above, essentially a liability claim regarding using the fabric for "mission critical," a strict form of neutrality would override non-compete, allowing participants to use provisioned port and exchange capacity for whatever purpose does not harm the technical integrity of the fabric.

Two trends may contribute to preferring neutrality to non-compete in terms of competing with transport and dark fiber. The first trend is that IXes are expanding their customer base, as evidenced by remote peering and remote node models. The second trend is that some IXes are beginning to explore the enterprise customer market. The latter trend finds IXes developing services for enterprises that need to exchange large amounts of data, such as the pharmaceutical and financial industries, with other enterprise actors in their respective value networks. What differentiates these services from conventional interconnection is that the IX provides a simpler interface for selecting interconnection partners, essentially abstracting away the complexity of building and tearing down BPG sessions. In metro regions with such industry clusters, this kind of service may face similar resistance

from transport, dark fiber, or even colocation providers that would have otherwise provided bilateral connectivity for these actors.

6.4.3.3 Network Hygiene

Not all networks connecting to the exchange have the same degree of experience establishing connectivity with other L2 networks or with interconnection via BGP. IXes typically include what is commonly referred to as "network hygiene" rules. Network hygiene rules ensures that only the L2 and L3 traffic types necessary to participate on the IX, necessary to establish L2 connectivity, exchange routes, and exchange traffic, are allowed. In terms of basic connectivity, many IXes stress technical elements such as *a*) ethernet interface settings should be specified, not auto-sensing; *b*) ethernet frame types; *c*) one MAC address per port; *d*) unicast with very limited exceptions; *e*) no link-local protocols, in particular spanning tree; *f*) no export of peering LAN addresses; *g*) all route exchanges should be via BGP4; *h*) ASNs should not be private ASNs; *i*) routes should point to the advertising router;

Network hygiene rules are in place to ensure the technical integrity of the common fabric. For instance, spanning tree is the mechanism for eliminating L2 routing loops in a switching domain. Given the L2 domain of the common fabric is managed by the IX, no other actors should be invoking the spanning tree process. If actors do this, it would easily lead to instability and misconfigurations. One MAC address per port ensures only a single router is connected per interface, preserving the distinction that the boundary of responsibility is the patch panel, not some configuration of routers on the participant side.

Of particular interest relative to the discussion of origin rights in the previous chapter is the limitation on exporting peering LAN addresses. As developed in Section 2.2, route provisioning often assumes, and wants, routes to be advertised to as broad an audience as possible.[475] In contrast, IP addresses used on the IX are *exclusively* for the exchange of routes and traffic between participants. As such, those addresses are not destinations or sources of traffic in the same sense that Internet end users are. As part of the IXes bilateral relationship with each participant, those that have rules expressly prohibiting advertisement of IX addresses are effectively placing limitations on subsequent provisioning. In other words, they are withholding the largely tacit right for peers to provision routes appropriated in an interconnection relation. The rationale is that advertising infrastructure number resources opens the IX to attack, such as the attack on the LINX during the Cyberbunker-Spamhaus incident.[476]

The last of these operational rules, traffic should not be exchanged over the fabric between interfaces of the same member, speaks directly to the difference between a common resource and a private resource. Internet exchanges are often referred to in the community vernacular as public exchange points. Peering over these exchanges is similarly referred to as public peering versus private peering,

[475] Within the limits of interconnection economics discussed in Section 2.3.

[476] For a popular report of the incident, see Markoff and Perlroth (2013). For a discussion of number resources, see Gilmore (2014).

in which two private actors jointly provision either a cross-connect or transport to be used exclusively between the two. This work modifies the notion of a public fabric to more accurately, in terms of political economy frameworks, identify IX fabrics as common resources. The premise of the rule in point is that a single actor should not be able to leverage a resource provisioned for the exchange of traffic between market participants for private use. This is especially salient in terms of non-compete in multinode platforms. If actors could exchange traffic between their interfaces at two different colocation facilities the IX would be taking business away from transport providers those actors would have had to contracted with to achieve the same transport.

6.4.3.4 Capacity Rules

The common fabric is provisioned to ensure a congestion-free interconnection environment. Ideally, IXes have the capacity to handle the scenario in which all participants utilize their maximum port capacity. In private conversations, some actors have indicated there are some small- and medium-sized IXes for which port capacities provisioned by the participants exceeds the capacity of the switching fabric. This is not in and of itself a failure as long as these IX firms monitor actual usage and upgrade accordingly. That said, as an instance of standards enjoining this practice, OIX-1 requires the IX have sufficient capacity to meet demand if ports see traffic at full capacity (Open-IX, 2014b).

In most IXes, if a participant exceeds their port capacity, it simply congests. An exception is the LINX. Under Services Definitions and Fees Schedule, the LINX indicates:

> If the Average Measured Traffic on any port exceeds 80% or more of the port capacity, this charge–equal to the standard monthly port fee–will be levied in addition to all other fees. (LINX, 2014b)[477]

Actors that appear to be consistently approaching congestion are sanctioned. The fee is explicitly set such that consistent violation the 80% threshold results in a fee equivalent to a port upgrade that would avoid the fee. A subsequent result is that, with the exception of extraordinary traffic events, the LINX attempts to ensure that traffic exchanged over the IX does not congest. Although the general ethos of the LINX is vehemently for self-governance, this is an instance of regulation to ensure the quality and integrity of the common resource.

A contrast is the AMS-IX. In the course of discussing the degree to which the AMS-IX helps inexperienced networks, the discussion turned to the responsibilities

[477]The LINX goes on to define Averaged Measured Traffic:

> The Average Measured Traffic for a port is calculated by dividing the number of bits passed through the port by the number of seconds in the measurement period (a month). Traffic in and out of the port (based only on intervals where measurement data is available and valid) is measured separately to produce two average values, and the AMT is the higher of these two averages, measured in Mbits/second or fractions thereof. (LINX, 2014b)

of the IX versus the participant. One staff member made a point to indicate that the AMS-IX does not impose constraints on the traffic management practices of its participants. Paraphrasing, AMS-IX participants may use their capacity as they see fit. Participants can use none of it or consistently utilize 99% or even congest. Like the discussion of DDOS attacks across the fabric, as long as the behavior of the participant does not adversely affect the common resource, capacity, and the implications of over-utilization, port capacity management is the responsibility of those involved in the traffic exchange. Recall the scope of route provisioning is bilateral. Similarly, the AMS-IX rationale and the rationale of other IXes that allow port congestion is that this is an issue in the bilateral scope of the interconnection relation.

6.4.3.5 Route Server Participation and Multilateral Interconnection

The definition of an interconnection option is the right, but not the obligation, to establish an interconnection relation with any other IX platform participant. In the mechanics described in Chapter 2, interconnection is defined as a bilateral relationship. For actors that are willing to interconnect with anyone on a settlement free basis, a route server is a means for individual actors to "automatically" interconnect with other actors willing to engage in settlement free interconnection. Route servers are a proxy. Participants peer with the route server, provide an interconnection policy (in this case, willingness to peer with others on as settlement free basis), and the route server facilitates automatic bilateral peering between all matching pairs. The effect is multilateral peering between participants on the route server.

Early in the history of IXes, many of the interconnection relations on the IX were informal settlement free interconnection relations.[478] Route servers are a convenient way to reduce the transaction costs of IX participation. Rather than engaging the process of e-mailing another peering manager and manually establishing interconnection, a single BGP session can establish settlement free peering with a broad set of actors. In terms of the value of participating at a new IX and diversity (discussed in Section 6.1.2), route server participation is an indicator of guaranteed interconnection relations on that platform. For new participants, the more participants on the route server, the more attractive the IX is as measured by guaranteed, immediate, low transaction cost traffic exchange.

This use of route servers effectively implements "multilateral" peering. The actual routes provisioned have the same technical character as those provisioned bilaterally, but the granularity of control is not necessarily standardized. Each IX route server's parameters may be, and are frequently, different. For instance, communities may be used to selectively provision routes for appropriation by some actors, but not others. Consider the configuration information for the LONAP:

> LONAP permit[s] participants on the route-servers to filter their announcements such that they are not offered to certain other peers on the route servers. This is useful if you wish to prevent your prefixes

[478] Woodcock and Adhikari (2011) indicates this remains the case based on a survey by PCH.

from reaching your transit customers via the route-servers, or you wish to deny peering to some networks as a matter of policy. Today, the filtering logic is expressed with the use of BGP Communities:

8550:8550	Send prefixes to all other route-server participants
8550:ASN	Send prefix to only route-server participant with specific ASN
0:ASN	Do not send prefix to route-server participant with specific ASN
0:8550	Do not echo prefix to any other route-server participants.

(LONAP, 2014b, list structure modified from original HTML formatting)

A relatively unpopular operational rule in the IX community is forced multilateral interconnection. Forced multilateral interconnection typically means that every participant must interconnect with every other participant on the platform. The bundle of routes that must be provisioned comprises the prefixes delegated to the participant. For an IX attempting to develop critical mass, this may be very appealing, especially for would-be participants that currently pay to exchange traffic with large access networks participating on the IX. Depending on the routes from the would-be participants to the access networks, this may reduce the costs for both or deprive the access network of revenue.

Two general cases are possible: beneficial settlement-free or potential violation of non-compete. In the case that the would-be participant and the access network exchange traffic over transit, both are paying for transit and both save if exchanging traffic settlement free. In contrast, consider if the would-be participant is a customer of the access network, either as a transit provider or a customer of the access network. In this latter case, forced multilateral reduces the revenue of the access network. Reduced revenue violates the (degenerate) spirit of non-compete. In effect, the IX is offering forced multilateral as an enticement to the clear detriment of some of its existing participants. It is important to note that forced multilateral typically only requires provision of routes corresponding to prefixes delegate to a participant, not all of the prefixes reachable in the participant's stock. As such, participants may still engage in settlement-based contracts such as partial or full transit. Although forced multilateral is unpopular, it is implemented. For instance, it is implemented at CABASE. CABASE has seen attrition by large access networks on grounds of costs of maintaining capacity to support settlement-free access.

6.5 IX Issues

IXes have faced a variety of issues exploring variants of firm-based, yet mutually-managed organizations. Amongst these are *a*) early demutualization and commercialization efforts (Section 6.5.1), *b*) efforts at market rationalization (Section 6.5.2), *c*) adaptation and adoption of service level agreements (Section 6.5.3), and *d*) non-compete issues with multinode and transport (Section 6.5.4). As alluded to throughout the IX discussions, these issues track changes in the IX as class of interconnection facility in terms of *a*) archetypal topologies, *b*) corresponding changes in management rights, in particular classes of participants, and *c*) the im-

plications of these changes for joint management of the IX as an organizational structure continuously adapting to participant demand. The issues in the RIR system are adaptations of processes, rules, and rights bundles to match demands for resources and industry structure, but the RIR has not changed the fundamental structure of numbers or routes. In contrast, each IX performs step-wise modification of the platform to replenish interconnection options and switching capacity. While operational changes are within the remit of the firm, strategic changes often alter the organizational structure, topology, or both, of the IX. The following describes these, distinguishing structural, organizational, and rights configuration changes.

6.5.1 Demutualization and Commercialization

Mutuality is an ownership model that empowers resource appropriators to directly influence how resources are produced, provisioned and distributed. Revisiting the mantra of mutuality, the single stakeholder in the firm is the membership *collective*. This model empowers the collective to directs firm strategy to serve the interests of the collective. Mutuality is a contrast to stakeholders that are investors expecting a fiduciary return rather than participants that depend on the resource. This is a fundamental difference between stakeholders deciding strategy and pricing based on a profit margin and a non-profit mutual organization.

Recall from Section 3.3 the discussion of credible administration: recall Ostrom argues that external administrators may be less credibly committed to the integrity of the resource when they must balance that resource system's needs in a portfolio of other political interests. In terms of resource rights management, demutualization is a story of a commercial management as the administrative mode comprising a broad portfolio of interests. It is a story of the challenge to what is framed in common resource literature as credible resource management. Demutualization is a story of rent-seeking investors seeking to displace joint provisioning. Arguably, a commercial model could create more profit for the firm and conventional stakeholders, in this case real estate investors.[479] That said, these actors likely do not have the commitment to the market a mutually managed (and accountable) firm has to the interconnection market. The LINX demutualization narrative is an instance of actors that sought to commercialize the London interconnection market and, knowingly or not, threatened to fragment that market.

The late 1990's and early 2000's were the peak of the dot-com bubble. There was substantive investor interest in interconnection platforms. This created two issues for IXes. First, there was a question, at least in the LINX, regarding how investment would keep pace with growth. Second, investors saw a business opportunity in

[479]This is not intended to paint all data center and colocation real estate developers with a single "red" brush, condemning them with a reputation of rent seeking opportunists. A number of data center real estate developers are keenly aware of the interconnection market and coordinate with IXes to sustain that market. The LINX demutualization narrative illustrates how actors unfamiliar with the market but operating on simple notions of "more competition is universally good" can threaten the stability of a functioning infrastructure that is successfully facilitating a growing market.

colocation and interconnection platforms. At the time, the dot-com bubble was also giving rise to investment in colocation facilities.

An early mutual LINX, with a nascent firm structure and much of the operational decision making still in the hands of the collective, found itself in an increasingly artificially competitive market for interconnection platforms.[480] At the time, it was not clear that the dot-com bubble was in fact a bubble. The dot-com bubble attracted investment in data centers and commercial interest in the IX as an interconnection model. Within the LINX, there was genuine concern that the IX as a firm and an association, under a cost-neutral funding model, may not be able to raise the capital necessary to sustain growth. In effect, it would not be able to compete with privately funded interconnection platforms. Interviews with LINX leadership indicate that some emerging colocation facilities requested the LINX place a node in their new colocation facilities or that colocation facility would create its own IX to compete with the LINX. It is unclear whether these commercial actors realized this, but in effect, they were threatening to fragment the local IX market in an attempt to capture it.

Within the LINX the concern was over funding continued growth. One actor involved at the time indicated that there were questions about the ability of a non-profit to raise capital. They go on to say that non-profits are more limited than commercial entities. Taken together with dot-com speculation and commercial investment in colocation facilities by realty groups that were not exclusively dependent on and thus less invested in the integrity of the interconnection market, this created a volatile environment for the LINX. It was in this context that colocation facilities threatened, "build a node in our facility or we will set up a competing IX."

At the same time, the LINX was going through a process of professionalization. In the community at the time, many, but not all, IXes across Europe were ad hoc. There was little professional accountability, with little information about the guarantees. Many in the community have imbued IXes with a broad notion of "best effort" service.[481] Part of this was a variant of the ideal form of mutuality, in which a firm structure was considered unnecessary, even superfluous, and where volunteer effort, and contributions, are considered sufficient for IX operations.

This ideal notion of mutuality is but one implementation. Here it is interpreted as a techno-deterministic form of reductionism. Such an interpretation assumes the only function of the IX is to move traffic and improvements in IX performance and stability will exclusively and naturally follow technical improvements. Under this framing, an ideal, technically committed community of volunteers can provide sufficient IX operations, but it leaves little room for error. In more realistic settings, interviews reported high variability in responses to service requests. Multiple

[480]As interconnection platforms have been presented thus far, a competitive market for interconnection platforms is an oxymoron. Like the NRS, the interconnection market is a single global system, a single global market. While interconnection platforms have been framed as local markets, they are in fact loci of interconnection, the control plane equivalent of well-ordered bazaars for the exchange of routes and traffic.

[481]This notion of best effort service will be contrasted with stronger guarantees in the discussion of service level agreements, SLAs, in Section 6.5.3.

participants and members of IX leadership active during that time period recall complaints that dealing with the ad hoc volunteer organizations was difficult and that some participants, typically large networks, especially US networks, preferred more standardized, professionalized, "one-stop" shopping for IX services.[482]

Some in the LINX believed a management model closer to the ideal form of mutuality was holding the organization back. A pejorative view of the ideal mutual form was micro-management, in this case by an active cohort in the membership collective. For instance, early LINX membership criteria required demonstrating external transit. Multiple actors from the time period reported long, drawn-out discussions of precisely how to do this. In most large modern IXes, technical issues of this granularity have since been delegated to the staff. At the time, in the face of looming commercial competition, some in the LINX preferred a model in which more decision making authority, both technical and strategic, was delegated to the leadership and the board.

These pressures culminated in what has since been referred to by the IX community as the LINX demutualization event. One option was for a venture firm to buy out the IX membership and make the exchange a commercial firm. The bid was too low, to the point of being insulting to the membership. Commercialization also ran against the grain of existing best practices in running an exchange: neutral, mutual, non-compete. Demutualization ultimately failed and a leadership change followed. A group of actors in the LINX leadership, most notable among them being Keith Mitchell, the CEO at the time,[483] chose to form their own IX, PacketExchange.

While PacketExchange did not succeed as expected, it is an interesting instance of pushing the characteristics of the IX model. PacketExchange attempted to develop a connected regional platform. It did not eschew the notion of an SLA. Rather, the SLA was a natural extension of the accountability to general quality of service engendered in professionalization. Many of these were characteristics the community critiqued PacketExchange for. These same characteristics have since been implemented by at least one of the three largest IXes and by many smaller professionalized IXes.

Stepping back from the narrative, recall both the notion of institutional experimentation and the fact that this was occurring when IXes were still relatively new. The demutualization event was an inflection point for the LINX. The LINX community did not want to transition to a completely commercial entity. The LINX demutualization event did arguably accelerate the professionalization process. The demutualization event is also a lens into the type of organizational innovation within a mutual IX. Innovation can and does happen—as noted above, many of the commercial characteristics of PacketExchange have been embraced by IXes in Europe and abroad. That said, following the earlier arguments regarding credible management, mutuality is a check on a potentially risky change imposed by management

[482]Again foreshadowing the discussion of SLAs, the work of Gereffi, Humphrey, and Sturgeon (2005) is used to explain SLAs in the context of the disintermediation of specific assets to more general assets in Figure 6-1.

[483]Mitchell has indicated that once it was clear there was a conflict of interest being both the CEO and a supporter of the demutualization event, he stepped down as CEO of LINX.

of a valuable general asset, critical to participants value proposition.

Mutuality is a check on unfettered commercial expansion that threatens to expose a common resource that the infrastructure industry increasingly depends on to short-term profit seeking volatility and/or external influences on strategy that, absent mutuality as a check, may exacerbate non-compete issues. Recall the discussion of credible collective management in Section 3.3. Ostrom argues that actors with a deep vested interest in a resource, in the classic case those that depend on a natural resource such as water or a fishery for the entirety of their livelihood, are much more credible managers of that resource and the integrity of that resource than external managers. Her critique of external managers is directed to government administrators that must balance common resource interests against other interests in their administrative purview. Here the critique is extended to management of firms that hold stewardship over common resources, here applied to IXes. The IX may be but one of multiple assets in the portfolio of a a set of financial principals, leading to short term profit management decisions rather than longer term infrastructure investment decisions.

Further, recall the threat by commercial investors to build their own IXes in London. Simplistic applications of "capitalistic" competition strategies can lead to fragmentation of the market they are wishing to capture. Further, what such external administrators do not recognize is that they are interfering in the infrastructure that facilitates the interconnection market they are trying to garner value from. These actors are operating on a winner-takes-all (or most) model of network effects. Rather, network effects are at play, for instance in the stickiness of an IX node, but, as will be developed in the next section on market rationalization, "competing" with a functional IX platform is effectively introducing unnecessary transaction costs for market participants.

On the surface the LINX demutualization event is just that, a narrative of demutualization and the IXes that emerged. A closer analysis shows the role of mutuality in the interconnection market and the common resources that facilitate that market. The demutualization event is a rare surge at an organizational inflection point.[484] This analysis highlights elements of competitive infrastructure development. Further, it highlights the evolution from the ideal form to forms of mutuality with more commercial elements. Finding a balance that preserves the credible management of the mutuality based model with the agility of commercial models is an ongoing challenge amongst associational membership IXes.

6.5.2 Market Rationalization

In contrast to the environment of the LINX demutualization event, the French interconnection market of the late 2000's comprised a dysfunctional collection of small, poorly managed IXes that fragmented the interconnection market in France. In a number of private conversations, actors have indicated that, before 2009 when

[484]Other IXes have demutualized. For instance, the DE-CIX was initially a mutual IX but early on decided to transition to a commercial model.

IX	Short Name	Owner	Members	PDT	Incept
France-IX	France-IX	France-IX	185	187G	Nov 2010
Equinix Paris	Equinix-Paris	Equinix	187		2008
Paris NAP	PaNAP	T-Mobile	153	848M	Jul 2005
Free-IX	Free-IX	Free	106	2.89G	2000
Service for French Internet Exchange	SFIX	RENATER	95	39.7G	1995
PARIX	PARIX	France Telecom	46	14G	2001
Paris Operators for Universal Internet eXchange	POUIX	Gitoyen	33		
French National Internet Exchange IPv6	FNIX6	Novso	20		1 Nov 2002
MAE-Paris	MAE-Paris	MAE	18		
Gigabit European Internet eXchange	GEIX	ManyOnes SAR	8		
Mix Internet Exchange and Transit	MIXT	MIXT	7		

Table 6.3: IXes in France at the time of France-IX's founding. France-IX line indicates current participation.

France-IX was created, if there was any option to avoid IX interconnection in Paris, they would. Multiple actors indicated "if it was feasible to route around Paris, we would." At its peak, there were 9 small, many "free," IXes in the Paris region; see Table 6.3. These were a combination of local groups and subsidiaries of local providers. The adverse effects of a poorly tended IX and the dysfunction that arises when the community tolerates a diverse set of IXes is a common trope amongst network and IX operators. The result was that it was expensive and problematic to interconnect in Paris. Fragmented IX constituencies, poor service on the part of most of the Paris IXes, and the costs of monitoring these all contributed to unnecessarily high transaction costs.

Returning to a fundamental value-proposition of an IX, the objective is to *reduce* the transaction costs of interconnecting to a diverse set of actors. It is important to highlight that the value comes from having a diverse set of actors in a common platform, i.e. seamless *access* to a single diverse set from a common set of nodes. This is very different from having a diverse set of PoPs with a subset of the market participating at each. For an industry sector with already low margins and IX prices converging with transit, having to POP into multiple low-yield IXes increases

transaction costs relative to PoPing into one or two or purchasing transit. The redundancy benefits of multiple IXes does not hold either. Given the poor service quality and lack of overlap, diversity did not yield redundancy benefits.

The history of solving the problem in Paris was often just build another IX, but each followed the same pattern as before: make it free, try to capture network effects. Until 2008 and 2009, this contributed to fragmentation rather than helping it. A number of early interviews stressed that building an IX is not a panacea and will neither "magically"create an interconnection market nor will it rationalize a broken market. In this case, the community was faced with the problem of how to rationalize a long fragmented market. How does one introduce an IX that will not further compound fragmentation, one that will have the "gravity" to achieve critical mass quickly? The solution was a combination of credible policy entrepreneurs within the community that *a*) solicited community perceptions of interconnection to gauge demand for such an organization and *b*) had access to the resources to build it. Amongst the IX cases in this dissertation, France-IX is a relatively early instance of "modern" IX design and initial provisioning. To be absolutely clear, "modern" binds to the active design and provisioning that skips much of the mutuality building in early IX development. Rather, the survey confirmed a latent community—the work lay in design and buttressing initial credible commitment.

France-IX is the product of policy entrepreneurs both within the French market and those from international corporations with an interest in a more functional interconnection market in France. A process of community engagement to gauge demand gave rise to a regional IX[485] that rationalized the Paris interconnection market and may rationalize France. In an interview with one of the architects of France-IX's constitutions, Kaufmann indicated he did a side-by-side comparison of the LINX and AMS-IX articles of association. Lessons and mechanisms were drawn from both, and the experience of the founders, to develop the constitution of France-IX.

In terms of both governance and architecture, France-IX is arguably the product of modern IX design, building on the experience of the big three European IXes. In terms of governance, amongst the associational membership models, France-IX has greater freedom to innovate. The founders intentionally created a "forgiveness over permission" model of membership monitoring and enforcement of strategic decisions. This is in contrast to the conventional model of updating the IX mandate *before* experimenting with a new medium-term tactic or longer-term strategy.[486] This relaxes the traditional mechanics of mutuality but not the spirit of joint decision making and accountability. Architecturally, France-IX is one of the first IXes in Europe to interconnect independent IXes within a metro-region—the IX started by creating nodes in three different data centers and by connecting with SFINX. Since then France-IX has grown into an influential European IX that, like others, is growing both in its home country (Paris, Marseilles, Lyon) as well as coordinat-

[485] France-IX is arguably growing into a global IX.

[486] One may argue this is a hybrid of traditional associational membership and commercial IXes. In this sense, France-IX is a hybrid, somewhere between AMS-IX and DE-CIX in terms of commercial character.

ing to facilitate IX development as internationally through collaboration in French speaking African countries.

Community experience indicates a metro-area can support one or two IXes. Redundancy in a metro-area is typically provided by distributed nodes of one or two IXes, balancing data center neutrality, proximity to potential participants (especially in large metro-areas such as London, Amsterdam, and Sao Paolo), and most importantly providing a common interconnection fabric.

Even if all the IXes in Table 6.3 were well-maintained, fragmentation is still problematic. This was not the case. Many of these were what the community referred to as "me too" IXes. Such IXes were neither designed as stable cost-neutral IXes nor have self-sustaining revenue streams within their parent organization. IX services were offered as a free service.[487] The primary reason for the pre-France-IX network operators offering an IX was an attempt to improve their respective interconnection position in the interconnection market. When upgrades were necessary, the owner of the IX either did not have money for the upgrade, did not have the incentive, and/or had lost interest.[488]

In many cases, even if IX participants wanted to upgrade, it was not possible. Neither larger port capacity nor additional ports were available. The result was that actors invested in building in but the investment became a loss when the IX failed and/or stagnated. Despite a *nominally competitive* market of IXes,[489] supply did not meet demand for the desired connectivity bundles. More precisely, supply of a small set of high-yield, common points of interconnection was not being supplied.

Service level further compounded these problems. Before 2008, Paris IXes did not have consistent customer service and frequently failed to respond to participant correspondence. For instance, many did not have a 24/7 NOC. E-mails regarding service upgrades or failures received no response for weeks on end. In terms of governance, these IXes were not accountable to their participants through membership *or* paid customer contracts. This alone undermined the transition of the IX ecosystem from a relatively convenient, ad hoc, best-effort service to a professionalized service with performance guarantees[490] that could be relied on as an input to a participant's value-proposition. In comparison to the larger IXes in the region (LINX, AMS-IX, DE-CIX, Netnod), this was well below the standard of professional service.

By 2008, the market was fragmented amongst 10 participants. Typically, a minimum benefit of an IX is the ability to offload national or regional traffic. Consider IX

[487]Free service is not a bad model in and of itself if there are other revenue streams. For instance, both the SIX and PTTMetro provide service for free. That said, PTTMetro is cross-subsidized by NIC.br and the SIX relies on donations. The SIX is currently facing the challenge of how to maintain this model and cross the threshold into being a more commercially structured IX.

[488]There are a number of challenges with providing free IX services. Upgrades can be problematic. Dependence on cross-subsidy has neutrality implications.

[489]The market here is scoped to the metro-region. In the larger European market, demand was met at IXes such as the AMS-IX, LINX, and DE-CIX.

[490]Performance guarantees can come in many forms. SLAs are still a somewhat contentious issue in the IX world. That said, the industry seems to be developing SLAs appropriate to the best-effort model without baggage of traditional telecommunications industry SLAs.

Manchester, IX Leeds, NaMeX, and other regional IXes. The result is that local traffic traverses the IX, and national and international traffic goes either over transport to a larger IX or, in the worst case, via transit.[491] At the other end of the spectrum, connecting to AMS-IX, LINX, or DE-CIX gives participants access to metropolitan, regional, and international participants (and subsequently international traffic) as well as the options to offload traffic for the Netherlands, UK, or Germany, respectively.

This was not the case in Paris. In contrast to other states with a mature single IX or mature pair of IXes, an actor interconnecting in Paris needed to connect to a number of the Paris IXes *just* to access a significant portion of the French market. Moreover, SFR and Free, two incumbent ISPs provisioning their own IXes, stopped interconnecting on their own IXes, further degrading the value. Per IX, the result was less traffic, low service levels, high risk—precisely the things one *does not* want in an IX. In some markets, this kind of outcome perpetuates the idea IXes are unreliable. In aggregate, the result was higher operational and transaction costs for less traffic.

Collectively, the Paris IX community did meet the minimal requirement of offloading national traffic: local networks were interconnected in Paris, just not in one or two IXes. Collectively the Paris IXes behaved like a regional IX: most of the traffic exchanged was local traffic. Once divided amongst three or four IXes, this becomes a small amount of traffic per IX. Fragmentation further compounded the balance of traffic and the transaction costs for maintaining connectivity to multiple IXes. Absent a common interconnection point to get just local traffic, international actors were reluctant to join. This further compounded the problem: international traffic that would have traversed a healthy IX in Paris was routed through Amsterdam or London. The result was a rather inefficient IX ecosystem.

The last entrant before France-IX was Equinix Paris in 2008. Equinix entered the market, initially providing its IX service for free, cross-subsidizing the cost from its data center business. In contrast to the other entrants, Equinix was a step up. Equinix's model of using the IX to make data centers sticky meant it invested in providing professional services and a 24/7 NOC. In terms of neutrality, Equinix is carrier-neutral but not data center neutral. Like all elements of neutrality[492], one seldom finds an IX that meets all and it is typically the case that one balances the other. It is important to reiterate that critique on the dimension neutrality, here data center neutrality, is not a wholesale condemnation, simply a point of differentiation.

Paraphrasing comments in interviews, when the community discussed pre-2008 Paris, the story was the same: "Paris is a broken market, don't go there." The France-IX story is not a story of demand for regulation in the commons, but certainly not demand for *government* regulation.[493] A number of well-known community

[491]Transit is *typically* the worst case, but may be preferable if other options, such as the IX, are unreliable.

[492]See section 6.4.1.2 for the general list.

[493]Recall regulation does not imply state intervention, but rather introducing institutions that can impose order on economic and social interactions. In this case, an IX based on the associational membership regime.

members emerged as policy entrepreneurs to build a more effective IX, one based on well-known elements of the IX regime. At GPF 4.0 in 2008 representatives from Google and Jaguar presented an idea of rationalizing the Paris market based on a working group known as Pheon-IX.[494] By the following EPF (European Peering Forum) representatives from Neo Telecoms and Akamai had joined the effort to create what would eventually become France-IX.

France-IX development had two strategic goals: make this instance of IX development a community effort, and, building on that, consolidate the interconnection market. At EPF 4 the founders developed a survey that was deployed to gauge interest in the proposed IX.[495] The France-IX "demand" survey asked a number of questions, paraphrased here:

1. Are you happy with the state of the interconnection market in Paris?
2. Would you be interested in a carrier neutral IX?
3. Would you be interested in a data center neutral IX?
4. Would you *pay for a professionally* run IX?

As implied by the emphasis above, the last question is perhaps what set the then prospective France-IX apart. The survey was sent to approximately 200 peering coordinators around the world: approximately 60 to 70 percent of those that responded said yes, they would *pay* for a professionally run IX.

In contrast to the canonical story of *gradual* community coordination in service of IX development, this is a substantive deviation. Few of the IXes in Paris were community based. In contrast, France-IX's founders comprised well-known members of the network operator community. Multiple of the founders had experience with other IXes in Europe, both as participants and in leadership positions: Christian Kaufmann of Akamai was serving as chair of the AMS-IX board and Maurice Dean had served on the AMS-IX board.[496] France-IX is an instance of what is referred to as the modern IX development paradigm. In the late 1990's and early 2000's, IXes were not considered critical infrastructure and often "learned" by trial and error. Contemporary IX deployments benefit from community knowledge and experience. France-IX being one of the prominent instances, contemporary IXes are designed by experienced operators that recognize the value of professionalized services in an infrastructure value network in which IXes have moved from convenient but non-essential cost-savings centers rooted in best-effort operation paradigm to platforms network actors depend on to support their value-proposition.

Two consolidation strategies were considered. The first was to consolidate by absorbing the membership of the other IXes. The second was to consolidate by interconnecting existing IXes within a common logical fabric, much like metro-region

[494]Based on France-IX's description at (France-IX, 2015b).

[495]Another instance of applying a survey to gauge demand for IX deployment is the Open-IX initiative. A survey was deployed to gauge interest in and demand for a "European-style" neutral IX model in the United States. One interpretation of this is that the effort is a pushback against the dominant IX in the American market, Equinix.

[496]As a point of administrative nuance, Kaufmann stepped down to avoid a conflict of interest serving on the boards of two IXes in such close proximity.

IXes comprised of multiple, connected nodes. Ultimately France-IX is a product of both of these strategies. SFINX was the first IX to be connected. Members of some former Paris IXes were absorbed. Since then, France-IX members can interconnect at partner IXes SFINX, LU-CIX, LyonIX, Top-IX, and TouIX (France-IX, 2015a).

Given the mandate from the community, the next step was to build the IX. Funding sponsors were identified, as well as sponsors for hardware (Google), Interxion joined the founding members, Neo Telecoms provided fiber connectivity, and Akamai provided money and loaned one of its network architects' labor. Development was two-pronged: (*a*) the governance and organizational model and (*b*) building the technical platform. Note this parallels the partitioning of sources of neutrality discussed earlier. The governance streams and the technical streams were conceptually different, but did have some overlap. For instance, data center neutrality wants PoPs in multiple data centers, thus requiring a platform and node interconnection that will support that requirement. Another instance is IX-interconnection. A common L2 domain is generally undesirable not only for technical reasons but also for administrative reasons: while it should be convenient to peer seamlessly, connecting two IXes governed by different bodies should remain in two different administrative and broadcast domains. The simplest and most compelling reason in terms of integrity is that a failure in one domain should not cascade to others.

In December 2009 France-IX came online with Telecity, Interxion, and Telehouse as data centers. The Free IX was absorbed. SFINX was interconnected. POUIX did not join and is since no longer operational. The last visible listing via the Way Back Machine shows 30 members in March of 2012.[497] After absorbing Free-IX and interconnecting SFINX, participation in the France-IX ecosystem was commensurate with Equinix Paris. As of this writing, France-IX and Equinix Paris are the two largest, professionally run IXes in Paris.

6.5.3 SLAs

The adoption of SLAs is one indicator of professionalization of the IX in response to participants' reliance on IX services. The notion of a best effort service paradigm has been alluded to earlier. Best effort service is a product of two factors: projecting the best effort character of IP packet delivery and the volunteer character of early associational membership IXes. In the operator community, SLAs often invoke the controls imposed by strict SLAs in the telecommunications industry. These include quality of service specifications that are not a technical fit for Internet service delivery. That said, the notion of an SLA is broader than this industry specific instance. Here service captures both the firm's *a*) capabilities and resources as well as *b*) key variables of value network coordination such as complexity of transactions, ability to codify transactions.[498]

Some associational membership IXes still refer to the best effort model. Recall the discussion of traffic between interfaces contracted by the same firm. The quote

[497] See `http://web.archive.org/web/20120328202820/http://www.pouix.net/members-en`.

[498] Capabilities, complexity of transactions, and ability to codify transactions are key variables in the typology of global value chain governance offered by Gereffi et al. (2005).

from LONAP highlighted this practice was sanctioned in the LONAP. It also qualified the use of platform connectivity for critical uses, citing that while the LONAP has a great service record, it is still a best effort service. LONAP is an instance of a medium-sized IX, in a market with a larger IX (the LINX), that has relatively recently moved from a volunteer effort to hiring its first full-time employee. As such it had adopted a firm structure that signals to those looking for stronger service guarantees that the LONAP does have dedicated staff that can and will respond to service requests in a timely manner.

Other IXes also provide service level guarantees and remuneration in the case of a failure. For instance, Netnod's requirement that participants utilize both peering LANs is a commitment to a form of redundancy that increases mean time to failure (MTTF, a common service level variable). Netnod also establishes thresholds for down-time, reimbursing customers if down-time exceeds eight consecutive hours (Netnod, 2014, Section 7, Appendix 2). The AMS-IX has also introduced an SLA (AMS-IX, 2014). AMS-IX's SLA was initially a product of its GRX and IPX VLANs for exchanging mobile roaming data. The participants on these VLANs are mobile carriers that are more accustomed to SLAs. Provisioning an SLA was a requirement of the mobile carriers before participating on the IPX exchange.

In Netnod, SLAs are universal, the same criteria applies to all members. In the AMS-IX, all participants also receive the same service, but not all of the participants engage in the SLA contract. That said, technically there is no difference between the service provisioned for SLA and non-SLA customers. AMS-IX leadership indicated that they have their platform to well-above the standard set in the SLA. The SLA serves as an assurance to the participants, a signal that the IX firm is committed to a particular service level.

In the event of a failure, AMS-IX does not reimburse every participant affected. Rather, the participant must monitor the resources it has provisioned for failures that violate the SLA. From a skeptical perspective, this places the burden on the participant (principal) to monitor the firm (the agent). From the perspective of a firm in an IX regime that has historically embraced the best effort paradigm, requiring participant monitoring is a form of ensuring the service level contracted is genuinely of value to the customer. Not all applications are sensitive to the service thresholds offered and thus may not even show up on the participant's radar. As a firm provisioned by the membership and accountable to that membership for its expenditures, the expectation that a participant monitor its resources is both *a*) characteristic of the common resource regulatory paradigm and *b*) an obligation on the participants not to impose costs on the firm that do not benefit all participants.

6.5.4 Non-Compete and Topology: Multinode and Connected Platforms

A variety of multinode platforms and coordination models have been developed within the associational membership IX regime. Recall multinode platform topologies from Section 6.2. A multinode platform provides a common logical switching

fabric across IX nodes hosted at diverse colocation facilities, historically in a single metro-region. To implement this topology, the IX must provision transport between IX nodes. The IX has a variety of options for provisioning transport: build dedicated transport or contract various forms of transport.[499] The degenerate form of non-compete would preclude multinode platforms. Under a strict interpretation, *any* transport provisioned by the IX to facilitate traffic exchange between participants competes with the services provided by independent transport providers participating in the IX. Further, it may also diminish the revenue of transit providers whose customers participate on the IX platform.

Multinode became common, and some would argue necessary, for a number of reasons. The simplest and earliest reason was basic space and power constraints at colocation facilities.[500] Another rationale for multinode is the development of industry clusters; this is especially the case in the recent deployment of small- and medium-sized IXes in cities with downstream industries that increasingly demand greater bandwidth and lower latency.[501] Yet another reason is colocation pricing. As space becomes a premium, prices inevitably go up. Adding a new node at a colocation facility that has space and is looking to benefit from the stickiness of the IX typically yields cost savings for participants building into that facility. In effect, strategic node placement can reduce the costs of ix_{p-p}, ix_{co}, and ix_{xc},[502] ultimately reducing the cost of accessing the local or regional market of interconnection options. Multinode also facilitates colocation neutrality. As a trade-off, here colocation neutrality and transport neutrality seem to be preferred over non-compete issues with local transport providers. As discussed in Section 6.4.1.2, colocation neutrality is also perceived to create a more competitive, diverse colocation facilities market.[503]

Ultimately, the non-compete norm was relaxed, allowing for mutlinode platforms. The multinode debate is not a simple trade-off. The larger constitutional issue is whether the value of a multinode platform offsets the potential loss for transport providers. A second issue is whether this loss can be minimized, even eliminated. Multinode provides a number of benefits for non-transport participants: *a)* more nodes in different parts of the metro-region reduces the costs of building in; *b)* individual nodes can switch traffic locally rather than tromboning within the

[499] A fuller discussion of transport options was presented in Section 2.3.1.

[500] The LINX encountered space constraints in Telehouse early on. Colocation in NYC is notoriously space constrained. AMS-IX started in two locations in the Science Park area of Amsterdam, but grew out of those facilities. CABASE expanded to multinode in part because of space constraints in the Buenos Aires node.

[501] In this sense, the IX is considered a market catalyst, facilitating the local interconnection market supporting these objectives. As noted before, the IX is not a panacea, there is no guarantee that if you build it, they will come. Rather, IX providers require a credible commitment on the part of local participants. See discussion of remote platforms in next section.

[502] In the case of the latter, cross-connects, some European colocation facilities have offered "the first cross-connect to the IX" free in order to attract customers. Typically this is a smaller cross-connect with the expectation that participant will grow and purchase more.

[503] This is a fundamental argument in the development of Open-IX. Open-IX drives competition by attempting to level the playing field via standards. In Europe, IXes have sufficient gravity, or stickiness, to demand higher quality services.

metro-region,[504] reducing potential points of failure in the path; *c*) multiple nodes create the opportunity for redundant connectivity to an IX; *d*) colocation facilities have limited physical capacity; *e*) more nodes, yielding more participants improves the value (diversity) of the IX. Colt, a fiber provider in London, pushed back against multinode connectivity when it was presented to the LINX.

The solution has been to minimize the impact of internode transport on transport providers that offer there services in the broader market. IXes *do not* offer such a general transport service. The only traffic that flows over internode transport is traffic exchanged between participants. Also recall that on some IXes participants cannot send traffic between their own interfaces on the IX, circumventing the need to provision transport between discrete networks managed by the same organization. Limiting flows to internode traffic minimizes, but does not eliminate the competition. To further minimize the impact, IXes do not necessarily own their transport. Rather, they may provision transport on the market, typically from participants. Provisioning internode from participants offsets the perception of the IX as a substitute, replacing it with a more complementary relation in the infrastructure value network.

The efforts to minimize competition is a way to minimize the IX as an substitute service. The growth of IXes has increased demand for transport to the IX and the volume of traffic between IXes. Reseller programs are an instance of mutually beneficial development. Reseller programs facilitate adding additional participants to the IX. Reseller programs also create additional potential business for transport providers. The result is that the IX becomes a complementer to the transport market rather than a substitute.

CABASE offers an illustrative instance of the complementarity between IXes and transport. CABASE is an multinode platform whose geographic reach spans the nation state of Argentina. Topologically, CABASE is a hub and spoke configuration. The hub is the central switching node, located in Buenos Aires. All other nodes are remote nodes. As a platform provisioning a common fabric, transport is necessary to transmit traffic from remote nodes to the central node or to another remote node. CABASE nodes have distinguished members, carrier members, that have infrastructure sufficient to carry traffic from remote nodes back to Buenos Aires to be switched at the central node. Each carrier member charges each member the per volume rate that would hold for the total traffic level for the remote node. This is very attractive for small members that would have to pay a higher per volume rate if purchasing transport to Buenos Aires on their own.

[504]This implementation of a node harkens back to the early nomenclature of an IXP as an IX *point* rather than a provider with greater geographic reach.

Chapter 7

Anti-Abuse

GREATER THAN 90% of e-mail *transmitted* on the Internet is spam.[505] Fortunately, not all of that e-mail reaches the user's inbox. Even assuming all the routes these messages traverse are legitimately provisioned and appropriated, this does not mean that that the *traffic itself* is legitimate. In terms of resource rights, *access* conferred does not mean that resources along the way from a sending party to a receiving party—in particular the delivery infrastructure of receiving networks and the resources of end hosts—is open for appropriation by anyone with access.[506] In the vernacular of the anti-abuse community, traffic, such as e-mail, is legitimate if the receiving party has consented to communicating with the sending party. Just because the routing system provisions a path, and that path may be legitimate from the perspective of bilateral route exchanges at each point of interconnection, *does not* mean all possible traffic flows along that path are consensual. To enforce resource rights in the messaging value network, the anti-abuse community has developed a set of anti-abuse norms around the notion of consent and reputation-based coordination mechanisms to enforce those norms.

In general, resource access is defined to be non-subtractive, but, in the routing system, it comes with the implicit assurance of best-effort delivery to the prefix advertised. Some abusive actors frame this kind of access as carte blanche to send unsolicited messages to any users for whom they have identifiers, such as e-mail addresses.[507] A premise of consensual messaging is that the communication is mutually beneficial; it has personal and/business value.[508] Abusive messaging not necessarily mutual beneficial and by definition not consensual. Rather, most

[505]According to M^{3}AAWG (2014d, p. 2), 90.1% of mail in the first quarter of 2014 was classified as abusive based on how it was handled; 90.2% in the second quarter.

[506]Returning to the rights developed in Chapter 3, this is the fundamental difference between access and appropriation.

[507]E-mail has been the historical abuse vector. New asynchronous messaging mechanisms have emerged, such as SMS messaging, instant messaging, and social network based messaging systems. Each is suffering from many of the problems related to abusive messaging that have been observed in e-mail. As stated earlier, identifiers grant access, but, absent consent, one does not have rights to appropriate resources necessary for delivery by an agent of the recipient or the recipient's processing resources.

[508]The value garnered from consensual messaging is not necessarily symmetric, though.

spam recipients consider unsolicited traffic at best a nuisance, at worse a severe curtailment of the value of having e-mail. Aggregating individual recipients' diminished utility to their respective (access) network providers (a type of receiver in the anti-abuse vernacular), spam is an abuse externality that diminishes the value of a downstream uses facilitated by number resource assignments.

As the Internet grew into a platform for online sales, the messaging value network became an attractive vehicle for marketing. Abuse externalities range from simple, naïve externalities rooted in local optimization to those that seek to intentionally extract value by virtue of negative externalities. These externalities grew in frequency and magnitude as more and more value networks in the global economy integrated Internet facilitated messaging into their workflows and as tools supporting their revenue streams. Some classes of composite negative externalities, have even developed into value streams in and of themselves. For instance, malware delivered in spam may add infected machines to a botnet that is, in turn, on offer in the black market for a variety of DDOS-for-profit schemes.

In the face of these negative externalities, interest groups (constituencies) seeking to protect their own value proposition in an adversarial, multi-sided messaging value network began to emerge. Receivers, those handling the bulk of incoming e-mail traffic and, in particular, the costs of abuse externalities, comprise one constituency. Receivers developed norms around consent as the primary criteria for distinguishing legitimate from illegitimate messaging, typically requiring opt-in. Senders are at the other end of the messaging value chain. Sender range from legitimate marketing firms that conform to anti-abuse norms to malicious actors whose entire revenue stream is based on extracting value from simple and composite negative externalities.

Anti-abuse norms and the attendant best practices have evolved to identify, mitigate, and, ideally, eliminate abuse externalities and the sources of those externalities. The anti-abuse regime is enforced by various reputation mechanisms. Reputation, tempered by graduated sanctions, is effective as both a positive and negative selective incentive for compliance with anti-abuse norms. Reputation brokers, comprising reputation monitors and reputation aggregators, emerged as a MVN constituency dedicated to collecting, evaluating, and distributing reputation indicators. One class of reputation aggregator comprises IP blocking lists (IPBLs), that bind reputation indicators to blocks of addresses. Receivers use combinations of messaging indicators collected within their indicators (local messaging indicators) and externally appropriated messaging indicators (including reputation indicators, collected from various vantage points across the Internet) to selectively filter traffic from actors with a reputation for abuse. The community authorizes credible reputation aggregators by continuing to use reputation data provisioned. Less credible aggregators are simply ignored and, left unused, have no effect (and no revenue stream).

Abuse externalities can, and have historically, led to substantive tension between senders (and their hosting providers) and receivers. Unchecked, adversarial appropriation and application of reputation indicators can (and in some cases has)

led to unnecessary escalation[509] and economic losses on both sides. Fortunately, a number of communities, most notably in this chapter, M^{3}AAWG, have emerged to mediate engagement between these constituencies. M^{3}AAWG in particular serves as an arena for *a*) coordinating the mechanisms used to confer reputation and *b*) generating and sharing operational knowledge, i.e. creating a knowledge commons, necessary to navigate what is referred to in this chapter as the legitimate sending parameter space. The MVN as a system is framed as a jointly provisioned, commonly managed resource in which reputation is bound to number resources in the underlying routing infrastructure. Like the RIR system and the IX regime, many of the facilities require, and are built upon, information jointly provisioned from diverse vantage points in the Internet infrastructure topology.[510]

The anti-abuse regime is CRI that, like IXes broader role in the NRS, can enhance the value garnered from rights held by actors in the NRS. In contrast to both the RIRs and IXes, as an enforcement regime, anti-abuse is more widely known for diminishing the value that can be garnered. Norms in the anti-abuse regime are enforced indirectly by the decentralized application of reputation indicators. Reputation conferred in the anti-abuse community can either enhance or diminish the value of downstream number rights, in particular sending rights. In general, reputation is bound to numbers to *a*) protect valuable, legitimate downstream number use rights of actors experiencing extensive negative externalities and *b*) diminish the rights of actors producing, and often profiting from, those externalities. In the latter case, appropriation and application of reputation indicators has the effect of a negative selective incentive, revoking downstream use rights identified as the source of externalities.[511]

Also in contrast to the RIRs and IXes as CRIs, norms created by the community are not operationalized into resource policies that directly shape management rights or facilities. Rather, the products of consensus processes are messaging industry norms (BCPs). BCPs explain the structure and dynamics of how downstream rights are affected by the application of messaging and reputation indicators, in particular those that are used to identify and mitigate abuse externalities. BCPs are produced in arenas provisioned by organizations such as M^{3}AAWG and the APWG (Anti-Phishing Working Group), but are not directly binding on all MVN participants. Each faces a different set of monitoring and enforcement incentives. For instance, while M^{3}AAWG members are regularly evaluated as to whether they continue to adhere to ant-abuse norms, BCPs are not applied instrumentally in these evaluations. As another instance, when reputation aggregators overstep their bounds, for instance engaging in aggressive listing escalation, extorting actors for delisting, or exhibiting high-false positive rates, they suffer in the market for credible, actionable

[509] Escalation is the unproductive side of graduated sanctions.

[510] More precisely, the vantage points are elements in the Internet topology that are dedicated to some portion of the MVN: sending and receiving mail transfer agents, end user hosts, reputation aggregators and monitors' infrastructure, and routes appropriated for the exchange of traffic amongst these actors.

[511] Section 5.2.3 discusses this negative selective incentive in the context of RIR number resource delegation.

reputation indicators. As such, while BCP advice is not directly binding, those that follow BCP advice have sufficient network effects to discipline reputation aggregators and senders.

The remainder of this chapter develops the role of the anti-abuse community in the NRS. Section 7.1 provides an overview of the anti-abuse community, its values, and origins combating spam. A cornerstone of anti-abuse is assuring enforcement of community norms. As yet another contrast to RIRs and IXes, where resource structure was presented first, followed by the constituencies that appropriate resources, this chapter first presents constituency dynamics as a means to understand why the resource structures for jointly provisioning and distributing reputation indicators have been developed. Section 7.2 describes the constituencies in the messaging value network, their relationships, and the role of number resources as the least fungible identifiers to which reputation indicators may be bound. Section 7.3 describes how resource aggregators jointly provision and distribute reputation indicators.

Section 7.4 resumes the established format for CRI study chapters, describing constitutional, collective choice, and operational rules at play in the anti-abuse regime. Constitutional rules (Section 7.4.1) are rooted in the notion of consent, explicit facilitation of consensus processes, and non-binding operational rules. Collective choice rules are presented in Section 7.4.2, standing up M^{3}AAWG's focus on problem identification and subsequent BCP development. Section 7.4.3 describes and explains BCPs describing how to collect, monitor, interpret, and navigate indicators in the legitimate sender parameter space. Broadly speaking, operational rules help network actors negotiate the parameters dictating what constitutes abusive behavior and how to manage reputation when an actor is imbued with a negative reputation.

Section 7.5 addresses some of the contemporary issues facing the anti-abuse regime. Section 7.5 opens with a discussion of whether efforts at instilling anti-abuse norms are still effective or whether the community is increasingly treating indicators instrumentally, finding the anti-abuse community devolving into a single-minded focus on deliverability. The problem of instilling norms versus instrumentality is a fundamental challenge and is addressed throughout the chapter. In terms of regime mechanics, this is a dilemma of authority versus coercion. The provision of reputation indicators is a source of power: do actors in the MVN comply with anti-abuse norms because they are credibly committed to those norms as authoritative or do actors see reputation as a coercive tool, complying instrumentally, and only as much as necessary to avoid value-diminishing sanction?

In terms of NRS stability, a focus on enforcement creates tension with other institutions in the NRS, in particular the RIR system. As briefly noted earlier, anti-abuse enforcement often diminishes rights *tacitly conferred by* number rights delegations, but not explicitly guaranteed by, the RIR system.[512] Conversely, the anti-abuse com-

[512]Recall the only rights guaranteed by the RIR system, i.e. rights the RIR system is willing to enforce, is uniqueness and host enumeration. Recall RFC 2050 notes that delegation does not guarantee the delegated numbers are routable. In a similar vein, the RIRs do not make any guarantees on the integrity of downstream use rights tacitly conferred when the basic rights bundle is delegated. Such guarantees will be even more diffuse when the only means of number delegation is the

munity is critical of RIRs' unwillingness to deny delegations to "known" abusive actors. This tension is discussed in Section 7.5.2.1. One factor in this tension is early instances of sanction escalation. Graduated sanction is a well-known mechanisms in the common resource management literature, often a form discretion to adapt sanctions based on extenuating circumstances, here whether an actor is unaware of the norms or a repeat offender. As the chapter is developed, graduated sanction is presented as a mechanisms for instilling anti-abuse norms early rather than severe sanctions, which may yield satisficing via instrumental responses. An adjacent issue to revocation and escalation is the interaction between reputation and transfers markets (Section 7.5.2.2).

7.1 Implications of Reputation in the NRS

The simple definition of abuse is nonconsensual end-to-end traffic. The following places abuse and reputation in the larger context of the NRS, unpacking the notion of abuse as a family of security externalities, relating these to the basic mechanics of number reputation, and providing an overview of common indicators. Taken together, this framing highlights how abuse affects *a*) value derived from number resources and routes and *b*) how anti-abuse communities use reputation monitoring and enforcement to limit the costs of abuse externalities.

7.1.1 Defining Abuse

Consider an early articulation of anti-abuse principles offered by one of the first BLs, MAPS:

> At MAPS we believe that all information exchange on the Internet should be consensual, and unless you choose to receive email from a third party, you should not have to accept it. The RBL is our way of assisting email and network providers with identifying and refusing email from known senders of unsolicited email. By subscribing to the MAPS RBL Service, these providers can reduce the impact of spam on their own network and focus their resources on providing their customers with better support and services. (MAPS, 2004, p. 1)

Ensuring a network's customers can garner contracted value, and that the contracted value is not diminished by non-consensual traffic, is important to individual networks. E-mail is but one form of abusive messaging traffic.

Moving beyond simple consent, abusive traffic is an externality on the receiving host, typically conferring value to the sender. Under the modern incarnation of abuse, (D)DOS attacks, scans to identify hosts vulnerable to exploit, traffic to activate those exploits, development of botnets, and other security externalities are all forms of abuse. Abuse externalities diminish value in a number of ways. A number

transfers market.

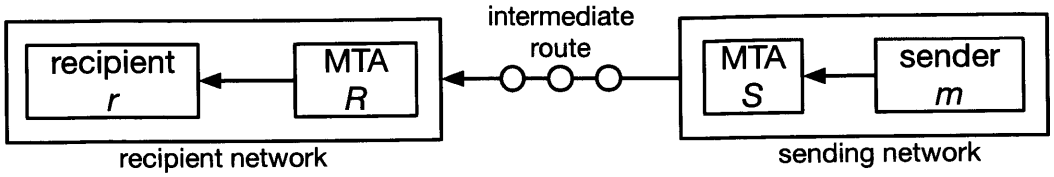

Figure 7-1: Simple messaging value chain (MVC).

of resources are appropriated in the course of end-to-end communication. The cost of abusive traffic is distributed over this chain (or in the broader case, network) of resources.

Consider a simple illustration of the messaging value chain (MVC) in Figure 7-1. A sender m may be either an individual or, in the case of most anti-abuse scenarios, a firm sending bulk (typically marketing) messages to a large number of recipients. Mail transfer agent (MTA) S is the host sending messages on behalf of m. The path between S and R is a route appropriated by the network actor hosting S. R is the receiving MTA, or, as will be referred to in this chapter, the delivery agent. Finally, r is the receiving host, typically an end user. For the time, the simple abuse externalities considered are borne by the receiver as costs at R and, if abusive traffic makes it past R, by the end user at r. Costs include filtering abusive messages at R, investing in feedback monitoring tools, and the diminished value of e-mail for r when she has to sift through spam to find genuinely valuable messages.

In developed regions, connectivity is cheap and e-mail is not the bulk of traffic volume. This has not always been the case. Moreover, it is not necessarily the case in developing regions where connectivity is still relatively expensive. Early on, when e-mail was a significant portion of traffic exchanged, spam could impose a significant costs. (D)DOS attacks by their nature diminish network availability, and subsequently value. Other exploits, such as malware and botnets, are intended to illegitimately appropriate the computational resources of r and traffic resources of the network hosting r, appropriating these into the abusive (and illegitimate) value networks of abusive actors. In these latter, remaining undetected may mean balancing abusive activity that diminishes observable resources with leveraging those resources to garner value for the abuser.

Common to all of these forms of abuse is the difference between consensual engagement and illegitimate appropriation. Legitimate end-to-end sending differs from the contractual agreements to exchange traffic at the AS-level in interconnection relations; the distinction also highlights which actors are incented to mitigate abuse externalities. Recall from the mechanics chapter that administrative domains are differentiated by AS number, typically by distinct routing policies, essentially a tuple comprising *1*) an interconnection bundle and *2*) policies that dictate the conditions under which those routing resources will be appropriated. Interconnection relations are largely bilateral contracts addressing aggregate flows between

networks, typically in terms of sets of addresses aggregated into set(s) of prefixes. Networks, especially those between R and S in Figure 7-1 garner value from traffic volume. These networks have neither the incentive nor the information necessary to determine whether messaging traffic is or is not abusive. Externalities and selective incentives typically bind to m, S, R, and r.

With the exception of large infrastructures and host deployments, such as is the case with firms such as Google and Akamai, most network actors have a limited purview into anything near the full range of abuse types and origins. Like information sharing in the NOGs and other operational epistemic arenas, formal and informal mechanisms for sharing information have been developed. Some of these information sharing arenas are open, such as the APWG. Others are closed groups that have been referred to by some as "fight clubs." As per the quote from the movie, "the first rule of Fight Club is that you don't talk about Fight Club."[513] As may be obvious, these actors are breaking that first cardinal rule.

Fight clubs are one instance of security information sharing communities, SISCs. Information sharing in the anti-abuse communities shares some qualities with the NOGs, but differs significantly in scope. Like in the NOGs, information sharing within SISCs is a means to reduce information asymmetries. Unlike the NOGs, the scope of information sharing and requisite trust relationships differ substantively; this will be discussed more fully in Section 7.2. Dissemination of the characteristics of abuse externalities can be widespread, such as through semi-public or commercial security information exchanges (SIEs). Dissemination may also be limited to a small, sub-industry-specific group, such as those dealing with financial institutions.

Anti-abuse communities initially formed to collate and distribute information collected from various vantage points. Later, as will be developed in the sections on constituencies (Section 7.2) and IPBL resource structure (Section 7.3), elements of these information sharing regimes have been made more durable. SISCs vary in their strategies for making their norms durable. In some cases, community size is limited to maintain stability. In other cases, SIEs are developed to make norms more durable by formalizing and automating information sharing processes amongst vetted participants. One of the largest closed working groups is M^{3}AAWG, a professionally run anti-abuse working group whose conference, like its NOG counterparts, serves as the convening arena for function-specific groups that adhere to general institutional norms established in the anti-abuse community.

Although the quote from MAPS above refers to the RBLs[514], the spirit of the statement generalizes to reducing abuse or eliminating the magnitude of abuse externalities writ large. Anti-abuse information can benefit the firm, but requires investment in learning how to interpret this information and how to process information appropriated. Within a given SISC, experienced members offset learning costs for new members. Based on the fact that industry actors do make these investments indicates that, at least for some, the cost of application is lower than the

[513]This is a quote from the 1999 movie *Fight Club*. See the IMDB entry (*Fight Club*, 1999) for a description.

[514]RBLs are an early instances of what became IPBLs discussed here.

magnitude of abuse externalities. Community efforts at making this knowledge more durable also strive to reduce learning, appropriation, and deployment costs.

The community's articulation of what constitutes abuse is a normative statement that distinguishes consensual traffic as acceptable and non-consensual traffic as abusive. In terms of the operational epistemic community, and the anti-abuse community as a CRI, this notion of abuse and the right to enforce that norm is a foundational constitutional norm. These norms operate within the larger NRS rights structures, but are not as aligned as the partition of number rights and routing norms in the RIRs and IXes, respectively. Number resource management in the RIRs and route provisioning within the operator community are interdependent, both necessary for the function of the control plane. In the larger NRS, RIRs have *historically* eschewed enforcement. This is a combination of cultural norms of operational sovereignty and audience costs when enforcement threatens that sovereignty.

Historically, LEAs have not had the operational knowledge or capacity to effectively investigate and act on online behaviors. A common characteristic of commonly managed resources is endogenous, jointly provisioned enforcement mechanisms. Ostrom describes this in terms of monitoring and enforcement (E. Ostrom, 1990, p. 94–99), which will be developed in this chapter. In many of the historical common (pool) resource systems, communities existed within a single jurisdiction and comprised nested structures in which resource management authority was delegated. CRIs in the NRS have neither a common jurisdiction, nor do they operate within a common hierarchy of authority that can resolve conflicts between them. In the absence of other NRS-wide enforcement mechanisms, anti-abuse has assumed particularistic management and exclusion rights to mitigate abuse externalities. As will be discussed throughout this chapter, this has created tension with actors in the RIR system that do not hew to anti-abuse norms.

7.1.2 Binding Reputation Indicators to Numbers

To understand anti-abuse's role executing particularistic management and exclusion rights and the conflicts that ensue, anti-abuse mechanisms are placed in the larger NRS context. Nominally, Internet connectivity means that any host uniquely identified by an IP address has an unfettered route to any other host uniquely identified by an IP address. Connectivity guarantees S has a route to R in Figure 7-1. The stock of routes appropriated to create a full routing table provides *access* in the form of known paths to any host on the Internet; R can reach any number of receivers such as R. Recall that access, more accurately, entry, described in Section 3.4.1, is the "claim to be *present* within... a domain," (Blomquist, 2012, emphasis added here, loc. 9428). The appropriation of bandwidth in the data plane facilitates traffic delivery from S to R.

Anti-abuse mechanisms, in particular filtering based on reputation, can affect both appropriation processes. Following the difference between spam transmitted and spam delivered in the introduction, a canonical anti-abuse strategy is to limit the *effects* of abusive end-to-end traffic flows (externalities). Binding reputation

indicators to IP addresses or address blocks facilitates a form of collective enforcement. Given this reputation information, any network provider R may choose to reject traffic from any address or collection of addresses, say for instance, S. Acting on reputation indicators is a form of "regulating use patterns"in the data plane; it is the application of management rights. Recall from Section 2.2 that the control plane "is used to direct, measure, and repair the data plane," (D. D. Clark et al., 2003). R uses local and distributed reputation indicators to decide whether traffic from S is should be trusted. In effect, reputation information is used to direct, or shape particular types of traffic in order to limit the externalities potentially imposed on R by S.

Figure 7-1 is a single messaging value chain between a sender $m \in S$ and recipients $r_1 \ldots r_n \in R$. More realistic abuse externalities manifest in diverse segments of the MVN. Consider an example scenario based on SimpleNet. AS 24 in Figure 2-1 is a network comprising abusive hosts. For the immediate discussion, assume it is known absolutely that AS 24 is a source of abuse. Credibility and false positives will be discussed shortly, then later in Section 7.2. Consider further that AS 24 is engaging in a novel, previously unobserved form of abuse. The degenerate case is the absence of reputation information. Under the degenerate case, only networks comprising hosts targeted by AS 24 will be affected by this novel mode of abuse; for this example, hosts in AS 25 and AS 22 are being targeted. Assuming those networks have the resources to monitor messaging indicators and identify abuse, absent information sharing, only those targeted networks will be aware of this type of abuse.

Consider additional subsequent abuse from hosts in other networks, say AS 27. If the malicious hosts in AS 27 target hosts in AS 22 and AS 26, AS 22 will have reputation information on *both* abuse sources, but AS 25 and AS 26 will only have information on AS 24 and AS 27, respectively. Subsequent abusive activity by hosts in AS 24 and AS 27 may target hosts in other networks, even though the abuse is now known to existing targets. In the degenerate case, there are no means to share information about abuse externalities.[515]

Repositories for documenting and sharing *a*) categories of security externalities, *b*) observations of those externalities in the Internet, *c*) indicators that provide evidence of these externalities, and *d*) attribution information, in particular observed sources, related to these externalities are referred to here as security information repositories, or SIRs. Security information exchanges (SIEs) are instances of SIRS. SIRs are typically maintained by a SISC. Informal SIRs may be little more than closed e-mail archives, but more pragmatic SIRs may be implemented in structured text, spreadsheets or databases. More formal still, RBLs and SIEs such as maintained by ISC[516] and Stopbadware (stopbadware, 2015) comprise *a*) the technical resource facilities for implementing SIRs and *b*) organizations for maintaining these, both of which are embedded in the anti-abuse regime. As may be obvious,

[515]The degenerate case is the textbook economic transaction: no additional information, no cooperative information sharing.

[516]An instance is the ISC's passive DNS system, see (Behjat, 2010)

SIRs are yet another jointly provisioned management facility.

(R)BLs (Realtime Blocking Lists) are function-specific instance of SIRs that serve as reputation aggregators specializing in the distribution of reputation indicators. RBLs were developed in the mid-1990's to mitigate spam externalities as commercial interests appropriated increasing volumes of Internet resources for marketing messaging. RBLs were developed to share information amongst networks targeted by abusive actors. The general idea behind an RBL is to serve as a repository for network reputation information. Given such a repository, and participation in that repository by a significant subset of network actors, abuse vector information can be monitored, disseminated, and acted upon at a significantly lower cost than individual networks could acquire the same information.

Reconsider the earlier abuse scenario. Abuse from AS 24; hosts in AS 22 are targets, then later AS 26. If AS 22 submits abuse characteristics and observed sources of abuse to the repository and AS 26 monitors that repository, AS 26 will be able to appropriate that information in order to mitigate security externalities originating in AS 24. Similarly, consider if hosts in AS 26 are targeted by hosts in AS 27 before AS 22. AS 26 reports abuse from hosts in AS 27 to a repository. Assuming AS 22 appropriates this information before the malicious hosts in AS 27 turn their attention in that direction, AS 22 can mitigate externalities imposed by malicious hosts in AS 27.

Thus far, it has been assumed that all of the information in SIRs, in particular BLs, is *credible*, i.e. it is an accurate and precise representation of actor's behavior. Perfect information is a useful ideal comparator, but unrealistic. Rather, SIRs comprise institutional mechanisms that attempt to ensure credibility by validating information from multiple sources. Here joint provisioning coupled with precise attribution helps reduce false positives that damage the reputation of the reputation aggregator.

RBL providers, and reputation aggregators in general, are participants in the larger anti-abuse community that uses and contributes to the development of abuse-monitoring tools; a subset of these tools are discussed in Section 7.4.3. Mechanisms such as spam traps are distributed at various vantage points to collect evidence of abusive activity; see Section 7.4.3.1 for detail. Based on sources identified, reputation aggregators assemble lists of prefixes that regularly produce abusive traffic. Some reputation aggregators also provide information regarding why the number resource is listed. This attribution information contributes to evaluating the credibility of reputation indicators provisioned.

Reputation aggregators have the *effect* of imbuing nominally non-rival number resources perceived to harbor abusive hosts with negative reputation. In effect, they attribute negative reputation, but it only binds when used by a critical mass of other actors in the MVN. *Credible* reputation aggregators are used by a broad array of network actors, often a critical mass, to filter abusive traffic. Use of a particular reputation aggregator's indicators by a critical mass of networks, typically access networks in their role as a receiver such as R in Figure 7-1, creates network effects that, by virtue of application in e-mail and/or more general traffic filters, diminishes the utility of the listed addresses to the holder of those address rights. Diminished

utility translates to *a*) diminished value to abusive actors and *b*) diminished overall magnitude of externalities created by listed actors. In terms of downstream number resource rights write large, listing by credible reputation aggregators diminishes the *range* of uses for the listed prefix.

Returning to spam transmitted versus spam delivered, consider where reputation indicators are applied. Reputation indicators are appropriated by networks such as R in Figure 7-1. In effect, filtering at R denies m access to delivery services. In the case of spam, diminished use results in lowering the deliverability indicators often monitored by senders. Relaxing the scenario to presume m is not malicious, but potentially a naïve new sender that does not yet know good sending practices, the objective of the anti-abuse regime is to teach m good sending practices. Part of these practices is to monitor various messaging indicators, described in the next section, to identify deliverability problems, interpreting these indicators as signals that its messaging strategy may be violating one or more anti-abuse norms.

Another way to frame this dynamic in terms of the NRS is that IP reputation creates a form of rivalry. Recall from Chapter 3 that number resources are nominally non-rival. In general, a resource is rival *in a particular use* when appropriation by one actor reduces the potential value that may be appropriated by another—one person can eat an apple, two people can split an apple, but both cannot appropriate the full value of consuming the apple. Rivalry also occurs when one use diminishes the value of another use—loud motorboats along a river diminish the riparian rights of those living along the river. Rivalry created from "competing" uses, a result of fragmented resource rights, often manifests as externalities.

Participants in the RIR community and the anti-abuse community have some overlap, but there is also conflict. Organizations that find themselves imbued with reputation are members of the RIR community. Some of these have representatives that participate in collective choice processes in fora such as M^{3}AAWG. Others are not aware of anti-abuse practices. Others still are regularly listed by reputation aggregators. A number of RIR community members contest the legitimacy of reputation aggregators, and IPBLs in particular as one of the most visible instances of reputation aggregators. In terms of rights and obligations, there is conflict between the *a*) operational rules developed by the anti-abuse community limiting abusive traffic and *b*) the absence of enforcement mechanisms in the RIR communities. As with all externalities problems, the problem is ultimately about who bears the cost of either remediating abuse externalities, the costs of selective incentives, and endogenizing the costs of legitimate (consensual) sending practices.

Before moving on to a discussion of constituencies and the relationships amongst these constituencies, the next section introduces some of the more frequently used messaging indicators used by the anti-abuse communities to identify sources of abuse and develop corresponding reputation indicators.

7.1.3 Messaging Indicators

The tension between transparency and security is distinctly manifest in anti-abuse norms compliance. Anti-abuse decision processes rely on monitoring and effec-

tively interpreting a number of well-known messaging indicators in the context of common anti-abuse strategies. BCPs, such as those described in Section 7.4.3, describe many of these indicators in the context of their application, for instance as a means to ensure good sending practices (Section 7.4.3.4) and how indicators can help diagnose the causes for listing by an IPBL (Section 7.4.3.5). The challenge is ensuring sufficient information is available to facilitate good-faith compliance by credible participants in the messaging ecosystem but limiting gaming by satisficers and malicious actors.

Historically the challenge was *a*) provisioning sufficient information *b*) jointly from diverse sources and *c*) making this process and the distribution mechanism durable. The contemporary challenge is, simply put, that the community may have done too good a job at making indicators and their application durable, but may have unintentionally diminished the stability of consent-based norms. One fundamental problem facing the community as of this writing is whether the spirit of anti-abuse norms, rooted in notions of consent, are being displaced by application of indicators to instrumentally drive deliverability as the ultimate outcome rather than interpreting indicators as a signal to re-evaluate the bases of a sending campaign based on anti-abuse principles. This problem is referenced throughout the chapter and addressed in its own right in Section 7.5.1.

Returning to the simpler problem of describing common indicators, across the BCPs and other industry anti-abuse documentation, a few key indicators for evaluating mailing lists are presented consistently:

Inbox Placement Rate is likely the most intuitive: IPR is the proportion of messages sent that actually make it to recipients inboxes.

Bounce Rates account for some of the undelivered messages that lower IPR. Various types of bounce rates indicate that the MTA did not recognize the address and, subsequently, is delivering a message indicating the user does not exist. As discussed in Section 7.4.3.4, bounces should be a signal to remove that address from an e-mail list.

Complaint Rates are another intuitive indicator, essentially representing actors that have hit the "this is spam" or "this is junk" in their e-mail client upon viewing an e-mail.

Sending Volume Thresholds may be set based on actual number of connections the MTA can handle or a sending rate, over a particular sampling period, the receiving organization has correlated with abusive messaging.

Spamtrap Hits represent the number of spamtraps a list has sent messages to. As the name implies, these are e-mail addresses that are not affiliated with a real user, but are seed online and in systems where abusive actors frequently harvest e-mail addresses. Given there is no real actor behind these addresses, consent is not possible and thus nearly all messages sent to Spamtraps, with limited exceptions, is non-consensual.

Message Content is generally difficult, but particularly problematic elements can be identified, such as large images and malicious URLs.

Each indicator is a general form of a family of indicators. For instance, BCPs and RFCs document a family of bounce rate and unknown user types and how they may be interpreted to inform sending practices. Similarly, there is a family of spamtraps. The character of each class provides different kinds of information about actors that encounter these traps. The role of the BCP process as a knowledge commons is to document these families of indicators and share best common practices for deployment, data collection, and, perhaps most importantly, interpretation.

Each actor deploying and monitoring these indicators sets the thresholds they will react to and the attendant response themselves, based on their value proposition. Experience managing abuse and knowledge shared amongst members of the messaging operational epistemic community informs these thresholds and responses. For example, closed groups of receivers may share information regarding what they consider to be effective sending volume thresholds. Such a group may even coordinate, deciding on best practices and establish ranges of thresholds. Ultimately, though, these coordination activities are informative, not standardized and enforced by the group or some external entity. The authority to set thresholds lies exclusively with the network actor, in this example, with the receiver. This is another exercise of operational sovereignty by the receiver.

Thresholds in the indicator space are not single points in that space. That said, while threshold levels *are* set individually, coordination through information sharing and best practices result in a clustering effect. Indicator thresholds writ large are more accurately represented as distributions. The centroid of such a distribution reflects the consensus of the community. Variance represents different value propositions and contexts. Skew may reflect an ongoing shift in the trend. Assuming that the values of the indicators do represent anti-abuse norms, i.e. they have not veered into instrumental delivery metrics, skew may represent a shift in norms related to that particular indicator. Rationalization of a threshold indicator from an incoherent (seemingly random) distribution of values to a coherent cluster is an indicator of coordination and, in the case of the operational epistemic communities considered here, an outcome of an effective consensus process.

7.2 Messaging Value Network Constituencies and Dynamics

The messaging ecosystem has been framed, like the Internet infrastructure itself, as a two-sided market. In the simple case, messaging comprises two very general, and very obvious, classes of actors: senders and receivers. Determining how well-behaved these actors are in terms of the externalities and selective incentives they can impose on one another is not as simple. Anti-abuse norms describe what should be done in terms of consent. The messaging indicators described in the previous section provide an introduction to how those norms are monitored. The dynamics

presented in this sections describe the how the monitoring and enforcement elements, the core of anti-abuse efforts, affect *a*) the integrity of the MVN and *b*) the mechanics of the incentive structures amongst MVN participants, monitors, and enforcement processes.

To limit the effects created by abusive traffic, vendors have emerged to offer solutions that mitigate or completely eliminate the propagation of abuse externalities. One largely technical solution has been spam filters. While appealing for their automation, abusive senders have consistently found means to circumvent spam filters. The result has been an cycle of ongoing escalation. While this may be good for the employees of spam filtering and anti-virus firms, it does not get to the root cause of the problem, value garnered from imposing abuse externalities.

Other vendors have chosen to leverage rights in routing infrastructure. Reputation aggregators, in particular IPBLs, provision reputation indicators that appropriated by access networks (receivers) to limit abuse externalities. Reputation is not perfect—there is not a single, canonical source of reputation information.. Rather, feedback loops in the ecosystem of professional senders, reputation monitors, reputation aggregators, and receivers continuously adapt how reputation is provisioned, conferred, and/or rescinded. A variety of feedback loops, discussed in the remainder of this section, exist between the simple actors illustrated in Figure 7-1.

A simple parable, that of a chain letter is used to make the roles in Figure 7-1 more concrete. This parable highlights that, even in the simplest abuse scenario, feedback loops can be observed. Further, this parable illustrates the three fundamental anti-abuse strategies for resolving externalities: dialogue, technical filters, and coarser grain blocking. As will be developed, feedback loops exist not only between MVN participants, but also between the strategies of dialog and coarser-grain blocking. The ideal situation is dialog, but often blocking is necessary as a selective incentive to get various kinds of abusive actors to the table to credibly develop solutions.

7.2.1 Messaging Dynamics and Consent Amongst Simple Senders and Receivers

Receivers may be individuals, such as m in Figure 7-1. Receivers may be firms for whom individual receivers are a primary customer or employee base. Such receivers may either provision MTA infrastructure (such as R in Figure 7-1) or they may outsource MTA management. In the case of individual receivers, these are individuals that receive (and send) e-mail, SMS messages, and IMs. Receiving firms include access networks, online e-mail providers, and enterprises.

Delivery requires appropriating resources at R and m; these are private resources maintained by a receiver firm and it's customers, respectively. To gain access to messaging resources, senders must acquire receivers' unique identifiers, such as e-mail addresses, mobile phone numbers for SMS messages, or user identification information for IM systems or social networks. Legitimate appropriation (use of) identifiers requires consent. In terms of rights, consent confers a sender

with rights to appropriate resources in R and r. The community has defined a number of criteria for how strong an assertion of consent is: whether that consent has been documented, the precise types of messages consented to, and whether consent has been recently affirmed (or renewed).

The simplest way to acquire a messaging identifier is for the receiver to share that identifier with a (potential) sender, such as two people exchanging e-mail addresses at a party, business function, or conference. Tacit in this exchange is the consent for one actor to send the other particular kinds of messages (appropriating resources in R and r). E-mail addresses may be collected by firms for a variety of reasons, among them to communicate regarding services offered, such as confirming the order of a product. This latter is referred to as *transactional*. Other early reasons for soliciting messaging identifiers was for e-mail lists, such as newsletters. Ideally, *interested* parties, those that believe they will garner value from receiving these messages, would share their messaging identifier with actors maintaining these kinds of lists.

Each of these scenarios above is consensual. Note that this provides access—appropriation is the act of sending and receiving messages along the MVC in Figure 7-1. Consider how the assumptions underlying each of these may be subverted, shifting the value of messages from consensual messages with value to the recipient to unwanted messages with no value. Worse yet, in bulk, unwanted messages may diminish the utility of messaging service by increasing the burden of finding valuable messages amongst the detritus.

Although less prevalent now, early Internet users experienced the annoyance of chain letters. These were often sent by friends with whom recipients had willingly (consensually) shared their messaging identifier. Depending on how much a person sent chain letters and how many of their friends sent them, sorting these and deleting them could be a mild annoyance. A number of strategies are available: *a*) send a message to the sender indicating the recipient is not interested in chain letters; *b*) creating filters to delete messages that look like chain letter; or *c*) failing other solutions, blocking all messages from that particular sender. As it turns out, these strategies represent the general strategies for remediating abuse externalities writ large.

None of these are perfect solutions on their, but they do illustrate the fundamental notion of recipients' consent and the nuance involved in managing that consent. When exchanging e-mail addresses, the recipient did not precisely and explicitly specify their preferences regarding the types of and range of content they consider acceptable. The first strategy is an attempt to more precisely specify recipient preferences; it is an attempt to clarify the rights the recipient conferred onto the sender. The second two strategies are attempts at enforcing those rights.

Consider the first strategy. The first attempts to stem the flow at the source. The expectation is that the sender will use good judgment, but this requires sender expend resources to determine which recipients want a given chain letter and which do not. When this fails, the individual receiver attempts to reach out to the sender, attempting to establish a dialog necessary to transmit finer-grained messaging preferences. The receiver expends resources to remediate, but is providing the infor-

mation necessary for the sender to endogenize subsequent costs based on a more precise specification of receiver preferences. Some senders will respond, others will simply ignore the receiver and continue to pursue their local value proposition.

Failure to resolve the externality via dialogue, enforcement efforts focus on limiting access to r's inbox or limiting delivery at R. The second strategy is an investment in spam filters, resources that automatically detect and filter undesirable messages. Power users may develop simple filters, such as filtering specific lists or based on keywords sent from particular users. More sophisticated commercial spam filters require a learning period, and even then still experience false positives. End user filtering is a largely technical solution and is not considered further than it is an expenditure of resources at r and one form of escalating rights violation sanctions.

The third solution assumes all content from the offending sender is non-consensual and simply blocks these messages. This latter is, at both the infrastructure level and the messaging (application) level is referred to by the community as "blackholing." As the verb implies, the unwanted traffic is irretrievably destroyed. Wholesale blackholing of a particular sender can have collateral damage for both the sender and the recipient. In this simple scenario, it is very likely legitimate (consensual) messages the recipient would have garnered value from will be discarded (blackholed). For instance, the sender may contact the recipient (through other channels, such as a phone call) to inquire why their e-mails did not receive a reply after a number of those that would typically elicit a response went unanswered. At this point the recipient may respond that the cost of dealing with the chain letters outweighed the benefits of consensual communication and thus blackholing created the net benefit. More likely, the recipient may respond that earlier requests to cease sending unsolicited messages met no response and this escalation was necessary to establish a dialog.

Although rather simple, this scenario illustrates the core notions of opt-in, consent, thresholds maintained by receiving firms, graduated sanctions, and dialog negotiating messaging rights.[517] Before elaborating these further, the notion of open access is presented, followed by an articulation of abuse by a firm. It is rather stylish to refer to the Internet as "open." Moreover, open access is often confused with commonly managed resources. Open access means that any actor with entry rights also has unfettered appropriation rights. As established in the contrast between access at the routing level, this is nominally the case but, in the face of diminished value from such an open regime, exclusion mechanisms have been introduced. Consent is fundamentally about conferring prices appropriation rights to senders. It is in this sense, with the introduction of exclusion (and management) rights and mechanisms, that makes the MVN, built a top a nominally open Internet communication platform, a *common* resource system rather than an open system.

Under an open access model of messaging, any actor possessing the messaging identifier of another may send that actor messages. Open access has become a rather popular term in Internet governance and regulatory arenas. In this work,

[517]This particular metaphor and the generalization to anti-abuse strategies, has been validated with a number of actors in anti-abuse leadership positions.

open access precise definitions and implications. Under an open access model, any actor with knowledge of messaging identifiers has access rights and rights to send messages to those identifiers (appropriation). Any actor with connectivity may appropriate messaging resources, at any frequency, in any volume, regardless of the consent of the recipient. Actors whose value proposition relies on abuse often make open access arguments. Some even go as far to argue that anti-abuse efforts are a form of censorship, for instance the Stophaus Project (Stophaus, 2015).

Most *individuals* do not have a social network large enough to make their personal address book sufficiently valuable to risk being blocked. In the most innocuous case, unsolicited interpersonal messaging is a case of poor judgment. In the less innocuous, but non-malicious case, perpetuating the chain letter may be perceived to garner some value ("it only works if you send it to five other people!") greater than the sender values the cost imposed on the receiver. Both of these cases frame abuse externalities with the character of operational externalities: they are either borne of naïve operational practices or non-malicious local optimization typically in ignorance of the externalities generated or their impacts.

The rationale behind both is a form of cost optimization. Poor judgment is a degenerate case: the sender does not know it is imposing a cost.[518] This will be referred to as a *naïve abuse externality*. The latter case finds a sender that recognizes it is imposing a cost but, absent swift and certain sanction,[519] the sender may not realize the externalities generated and continue with its local value proposition optimization. This case will be referred to as *operational abuse*.

The dynamics between simple senders and receivers can be extended to a recipient interacting with a sending firm. Consider the example of the firm that collects messaging identifiers from its customers in order to communicate regarding service delivery. Today, many solicitations of messaging identifiers come with an immediate disclaimer that *a*) the messaging identifier will not be shared with other firms, *b*) will not be used to send unsolicited messages, and *c*) will only be used for the purpose outlined at the time of sharing those identifiers. Experienced users are more reluctant to share their primary communication identifiers without these assurances, a form of specification of use (appropriation).

This was not always the case. Early on, a number of firms quickly recognized that, in the course of their day-to-day business, they acquired large corpi of messaging identifiers. Marketing departments realized electronic messaging was a low-cost means of advertising. Further, other firms that did not have immediate access to those corpi were willing to pay for those corpi, again for the purpose of advertising. In terms of NRS resource rights, one could argue that early marketing departments and firms did not distinguish between *a*) access and consensual appropriation and *b*) non-consensual appropriation via unsolicited marketing messages. The early extremes help understand contemporary messaging boundaries.

[518]This may be likened to the inexperienced network operator creating operational externalities in the routing system in Section 2.2.

[519]Across index crimes it has been empirically shown that the celerity and certainty of a sanction are greater predictors of deterrence than the conventional wisdom around the magnitude of the sanction.

Like the interpersonal exchange of messaging identifiers, sharing messaging identifiers with the firm was under-specified. That said, it was not nearly as under-specified as the interpersonal exchange. A reasonable assumption on the part of a customer is that the firm would only send messages pertaining to immediate services contracted. A nïave assumption on the part of the customer is that sending messages outside the services contracted would be costly. Experience has taught many users otherwise. Under the nominally open character of Internet messaging, sending marketing messages is apparently not only low-cost, but is perceived to be a valuable marketing tool. Experience also taught users that firms that do not specify precisely what they will use identifiers for may assume carte blanche, including selling or sharing those identifiers, a form of alienation. Rampant alienation of messaging appropriation rights has led to extensive access and appropriation by actors far outside the scope of the original relationship. As a result, left unchecked, such practices would lead to an overwhelming volume of abuse externalities—again recall the chapter's opening statistic, $> 90\%$ of e-mail traffic *transmitted* is spam.

The simple case of interpersonal messaging identifier exchanges are only nominally externalities. The recipient has an existing relationship with the sender. Following the strict definition offered by Coase, an externality occurs when a *transaction* between a and b has an effect on actor c which is not party to that transaction. In the chain letter example, a sender a is fulfilling a "contract" with the actor that send a the letter, sender b. Recipient c is not privy to this contract until a sends the letter to c. At this point, a has included c in an unsolicited transaction. It is abuse as defined in Section 7.1 but is either poor judgment or violation of use specification a between a and b, but only nominally an externality.

Consider instead a scenario that will be referred to as an extractive abuse externality, or simply extractive abuse. Assume that a has never directly exchanged messaging identifiers with or directly received identifiers from b. Actor a has no existing relationship with b and, subsequently, no basis for assumed or explicit consent. Actor a is contracted by c to send marketing messages to a list of message identifiers I. Actor b's message identifier is in I ($b \in I$). In this case, marketing messages from a to b are non-consensual appropriations by a, that are costly to b, in order for a to fulfill its contract with c. The origin of the list may be from c, or a may have acquired the list in order to win and/or fulfill the contract with c. In either case, in contrast to the scenario in which a has a pre-existing relationship with b that may be construed to limit sending behavior, a has no pre-existing relationship with b. Rather, a is not only abusing its access to b's messaging identifiers (and all of those identifiers $i \in I$), a is also abusing the open, low-cost character of Internet-based messaging.

Although many in the anti-abuse community consider extractive abuse egregious, the case above is neither malicious, nor does it appropriate resources explicitly for subsequent abuse. *Composite extractive* abuse surreptitiously exploits vulnerabilities that are used to *further appropriate* resources from r or R for use in later abuse campaigns. Following the scenario above, a's surreptitiously, typically maliciously, appropriate b's resources to impose further externalities on others in service of a's contract with c or future contracts, typically with other abusive or

malicious actors. In the case above, abuse is an operational externality intended to, as noted above, minimize costs. In the composite case, costs are reduced (externalized) for the contracting principle (c) and the contracting agent (a) at the cost of actors in I that have no interest in the advertisements sent by a on behalf of c. The fact that these campaigns make money indicates that some set of recipients do consume these advertisements, although it is unclear whether this is because of the "compelling" content of the advertisement or simple curiosity. In the case of abuse as a security externality, illegitimately appropriated resources are often used to further abusive activities, such as abusive messaging from botnets or DDoS attacks for hire.

7.2.2 Making Reputation Durable

The strategies described in the previous section can be generalized to:

1. coordination mechanisms establishing precise specifications of what constitutes consensual messaging,
2. filtering messages that "look" like abuse,
3. limiting appropriation by blackholing traffic from abusive actors.

Fora such as M^{3}AAWG and APWG are the arenas in which actors engage in the first strategy, establishing precise specifications of what constitutes consensual messaging. These specifications include various consent models and how messaging indicators can be used to monitor the practices of senders from the perspective of a number of units of analysis: *a*) per e-mail campaign segment, *b*) per e-mail campaign, *c*) per e-mail list, *d*) per sender, *e*) per professional sender's shared infrastructure, and *f*) per address block. These levels of attribution allow receivers to use message indicators to precisely signal what they consider abusive. Understanding how these levels of attribution are signaled and the corresponding message indicators facilitate: *a*) monitors' efforts to signal their clients of poor sending practices, *b*) aggregators' efforts to reduce false positives when binding reputation indicators, *c*) and senders' ability to accurately target remediation efforts.[520] The levels of attribution and interpretations are products of the anti-abuse operational epistemic community, documented in BCPs.

As noted earlier, the second and third strategy are means to sanction norm-violations. The second strategy, filtering, is a largely technical approach based on statistical machine learning and, while still widespread, is outside the scope of this work. The third strategy, credible sanction, is facilitated by the indicators produced in the course of the first polity. Messaging indicators are used to monitor and evaluate positive and negative reputation; most of the indicator thresholds are framed in terms of the thresholds at which an actor may accrue negative reputation. Well-defined and well-understood indicators make reputation more durable. In turn, durable reputation indicators facilitate limiting appropriation.

[520] Again, the ability to accurately target remediation efforts hints as the balance between remediation based on first principles versus instrumental remediation based only on avoiding message indicator thresholds.

Given the granularity at which abuse can be localized, consider the problem of attribution and the spectrum of abuse externalities. While the granularity of abuse can be identified, binding reputation to the appropriate actors, a form of attribution, is more difficult. Broadly speaking, attribution problems are confounded by cheap pseudonyms. *Cheap pseudonyms* are identifiers that may be easily replaced at little or no cost.[521] Cheap pseudonyms substantively lower the barriers to effective anonymity and substantively increase the costs of holding actors accountable for behaviors costly to others; here, cheap pseudonyms make eliminating the source of externalities difficult. Unfortunately messaging identifiers, such as e-mail addresses and instant messenger handles, are classic cases of cheap pseudonyms.[522] Abusers make extensive use of cheap pseudonyms to avoid filters based on messaging identifiers.

Recall the scenario from Section 7.1 in which abuse originates from a /28 comprising 16 hosts. A subset of a particular AS's address block may be delegated to an abusive actor. The quality of attribution depends on the scope of this abusive actors' messaging. Attribution data may be limited to a small number of actors that have been targeted in attacks such as spear-phishing. In the case of simple extractive abuse, targets may be widespread, yielding a large, diverse sample.

Ideally, abuse would be attributed to the organization or individuals profiting from abuse. In lieu of this ideal, in the MVN, IPv4 addresses are the most costly and least fungible identifiers the community can consistently bind reputation to. This is a trade-off between an ideal enforcement strategy that binds to the root of the problem (the source of externality) versus an enforcement strategy that binds proximate to the last known identifier, but not directly to, the source.[523] In other cases, reputation is bound to actors that support, or are complicit in, abusive activities, but are not necessarily the architects of that activity; see Section 7.3.1.3 for elaboration. Binding reputation to addresses makes reputation more durable and more actionable. More commonly, reputation is bound to a *block* of host addresses; the precise mechanics of proximate attribution and the potential for collateral damage will be elaborated in Section 7.2.2.1.

Recall the number resource delegation hierarchy discussed in Section 5.2.1 and illustrated in Figures 5-4 and 5-5. Further recall that $L2 - *$ delegations require resource justification based on reported utilization. For LIRs, white-washing, simply

[521] See Friedman and Resnick (2001), Feldman and Chuang (2005), and Feldman, Papdimitriou, Chuang, and Stoica (2006) for entries into the literature on cheap pseudonyms and identity white-washing in general.

[522] There have been efforts to make these less fungible. Domains are less fungible, but still relatively cheap. That said, domains may become the least fungible identifier available for attribution purposes once IPv6 reaches critical mass.

[523] The language here is unfortunately overloaded. The source of the externality is the firm that benefits from abuse externalities. In terms of the routing system, an address is a form of resource identifier and, when discussion routes, the origin address is often referred to as the source of the traffic. Given the routing system is about neutral transport, most analyses are not concerned with the actual provenance of the data or the implications of that data unless, as with other discussions of neutrality in the RIR system and IX regime, that traffic damages the integrity of the facilities necessary for system function.

throwing away tainted identifiers (addresses) and acquiring clean, new, (cheap) identifiers, is *not* a low cost task. In contrast, networks prefer to avoid replacing identifiers, i.e. renumbering in the RIR vernacular. For the LIR, renumbering requires coordination with the RIR, typically with some justification, and the operational cost of reconfiguring network devices.

That said, as one traverses the sub-LIR delegation hierarchy in Figure 5-5 down to end users, identifiers do become more fungible. LIRs delegate at least the minimum appropriation bundle for a subset of IP addresses when they provide connectivity to a customer. In practice, the criteria for justifying this delegation is not as onerous as a direct delegation from an RIR. It is also, as per discussion in Section 5.2, typically scoped to the length of the service contract with the LIR. The result is that end users have *a*) more potential sources of number delegations and connectivity and *b*) lower accountability for how those resources are used than LIRs. Addresses that are expensive for the LIR may be relatively cheap pseudonyms for end users.

LIRs, as the rights holder with direct accountability to the RIR, *arguably* have the most vested in ensuring number resources they have been delegated do not accrue negative reputation. While a rational assertion based on networks' reliance on number resources for control plane and, subsequently, Internet infrastructure, participation, a number of factors affect how much effort LIRs put into ensuring the integrity of delegated number resources. Actors exhibit a range of commitment to ensuring the integrity of number resources, often contingent on their specific value proposition. For this discussion, multiple scenarios will be discussed: *a*) credible commitment to resource integrity, *b*) actors complicit in facilitating or perpetuating abuse externalities, and *c*) hijacking of resources held by others for the purpose of abuse.

Consider the options available to end users. End users, especially those buffered by an intermediary that has been sub-delegated number rights (i.e. an end user conferred number rights in an $L5$ delegation in Figure 5-5), are much less vested in the reputation of assigned number rights. Such end users typically have a multiple choices for upstream connectivity. If the actions of an end user assignee result in those numbers accruing negative reputation, that abusive actor can simply acquire connectivity (and number assignments) from another provider. Although numbers as identifiers are not as cheap as messaging identifiers, end user assignments are much more fungible than LIR assignments. Given end user assignments are cheap to abusers but expensive for LIRs, they are susceptible to a form of moral hazard. The degree to which LIRs, playing the role of the insurer in the classic moral hazard narrative,[524] are susceptible is a function of LIR monitoring and enforcement of anti-abuse norms through its rights to revoke number rights.

[524]See Moss (2004, pp. 36–39) for a brief history of the roots of the terms adverse selection and moral hazard in the fire insurance industry of the late 19^{th} century. Moss (2004, p. 38) reports that characterizations of moral hazard can be traced back to the 11^{th} century amongst the Charitable Brotherhood of Valenciennes.

7.2.2.1 Credibly Committed LIRs

Fungibility of pseudonyms and moral hazard are functions of LIR commitment to anti-abuse norms and the ability to authoritatively impose those norms on clients.[525] Consider an LIR C credibly committed to either limiting abusive behavior by its clients or, at the minimum, ensuring its resources remain pristine for the sake of its value proposition. In this case, C invests in monitoring the status of the number resources that have been delegated to C. In particular C monitors those numbers that have been further sub-delegated to actors that *a*) are not directly contractually committed to the RIR, *b*) do not face the same requirements for acquiring number resources, *c*) for whom numbers are more fungible. When C identifies abusive customers it is incented to contact those customers to further incent them to remedy the situation.

The set of C's downstream clients is designated D. Consider C's options once an abusive actor $a \in D$ has been identified. C has delegated a a resource set r_a from C's larger resource set r_C. For this example, r_a is a /22, comprising (1024) IPv4 addresses. C and a have, at minimal, a contractual relationship. From the perspective of RIR resource delegation, C is ultimately responsible for maintaining accurate registry information for a, although it may have delegated this task to a.[526]

Recall from Chapter 2 that number resources, along with upstream interconnection relations, are the means to participating in the control plane and Internet connectivity in general. From the perspective of the anti-abuse community, C provides a, at minimum, the resource set r, and frequently, data plane capacity necessary to establish Internet connectivity. Taken together with the notion of abuse writ large, C has greater responsibility in number reputation stewardship than registry accuracy. In terms of rights and obligations, C's right to the messaging resources accessible via connectivity to receivers (such as R in Figure 7-1) is predicated on the obligation to endogenize abuse externalities. Under this framing, C *should* act to limit a's abusive activity. Networks provisioning delivery resources have a range of options to enforce that obligation, limiting a's abusive activity. Some of these may be established at the outset of the relationship, such as in a contract, but this requires a direct relationship between the two actors. Other enforcement options are the reputation-based sanctions discussed thus far.. The dynamics of how these play out are discussed first in the case of a credible network C, transitioning to a less credible network E.

Ideally, C explicitly limits abusive behavior in its contract with a, allowing C to credibly threaten to sever the contract if a does not limit its abusive activity. For C, a's behavior degrades the potential value of C's number resources. The value of a resource rights bundle is denoted $v(r_a, b)$ where r_a is the resource set delegated to

[525] This statement implicitly sanctions anti-abuse norms as, in and of themselves, legitimate. Recall that there are actors, that believe that best effort end-to-end connectivity is a sanction for (unfettered) open access.

[526] See Section 5.3 for discussion of the joint provision of registry data and delegation of registry access and utilization rights.

a by C[527] and b is the corresponding rights bundle. The basic appropriation bundle is denoted $\mathcal{B}$, the complete bundle of rights to downstream uses will be denoted $\mathcal{D}$. The total potential value of a resource r_a, v_{r_a} is thus

$$v_{r_{a\in C}} = v(r_{a\in C}, \mathcal{B}) + v(r_{a\in C}, \mathcal{D}) \quad (7.1)$$

$$v_{r_{a\in C}} = v(r_{a\in C}, \{\mathcal{B}, \mathcal{D}\}) \quad (7.2)$$

When a's behavior lands those resources on an IPBL, some number of downstream uses of $r_{a\in C}$ are limited. For example, consider reputation indicators that limits e-mail sent from hosts identified by $r_{a\in C}$. Here, this reputation indicator's semantics is that a is a frequent source of abusive e-mail. The rights bundle will be denoted e. The downstream rights that contribute to C's value proposition are denoted $V \in \mathcal{D}$; note that $e \in V$ and the remainder of the downstream rights in V are denoted $v_1 \ldots v_n$. The value of r_a under blocking will be denoted v'_{r_a}:

$$v_{r_a} = v(r_a, \mathcal{B}) + v(r_a, \{v_1 \ldots v_n, e\}) \quad (7.3)$$

$$v'_{r_a} = v(r_a, \mathcal{B}) + v(r_a, \{v_1 \ldots v_n, e\}) - v(r_a, e) \quad (7.4)$$

$$= v(r_a, \mathcal{B}) + v(r_a, \{V - e\}) \quad (7.5)$$

$$= v(r_a, \mathcal{B}) + v(r_a, V') \quad (7.6)$$

The result of blocking is not only diminished downstream rights, but diminished resource value $v'_{r_a} < v_{r_a}$, derived above and as expected from the notion of rights as a claim to value developed in Section 3.1. Note that $v(r_a, \mathcal{B})$ represents the value of the connectivity appropriated by a from C. Further note that the semantics of $\{V - e\}$ is a simplification that reflects the pejorative perception of some operations in the RIR community: a single reputation indicator completely revokes all e-mail rights. In other words, it is a strong, absolute, and from their perspective, illegitimate coercive power.[528] More realistically, e is not completely revoked, but, as per the language used thus far, diminished by that subset of actors that act on reputation indicators binding negative reputation to r_a; a more complete specification is provided in terms of local reputation functions in Section 7.2.4.1.

Consider C comprises customers $\{c_1 \ldots c_n, a\}$, each of which derives value from downstream uses in V. The value of C is:

$$v_C = \sum_{1\ldots n}^{i} v_{c_i} + v_a \quad (7.7)$$

$$v'_C = \sum_{1\ldots n}^{i} v_{c_i} + v'_a \quad (7.8)$$

Diminishing the value of r_a diminishes the value of r_C, incenting C to encourage a to limit abuse.

[527] For explicitness, this may also be presented $r_{a\in C}$.

[528] The term *coercive* is used here as an intentional contrast to the current frame, that anti-abuse norms are legitimate and *authoritative*. By proxy, enforcement of those norms is legitimate.

Consider an ideal situation, where a is unintentionally abusive. Network a monitors high impact reputation indicators and remediates abusive activities of its hosts to minimize the impact of abuse listing. In this scenario, a eliminates abuse on its own volition. For instance, consider the case where a is a hosting provider with a large number of non-abusive clients. Many reputation aggregators do not list individual hosts, but rather one or more prefixes. Moreover, absent remediation, reputation aggregators may increase the size of the prefixes in a that are listed as a means to attract the attention of a *or* C. For a credible hosting provider, the impact on a's value proposition is the potential loss of clients as they move to providers that can offer all of the services based in V. As a result, a's value proposition is its own incentive to preserve the integrity of assigned resources.

Now consider that b is a network that either does not have resources to track abuse or is simply unwilling to expend the costs necessary to endogenize abuse externalities. As with v'_C above, network C's resource value is diminished by b. Assume C has included a limitation on abuse in its contracts with downstream actors. Network C can incent b to remediate its abusive hosts. At this point b has two general choices. Network b can remediate, but it may lose those clients, thus limiting its own revenue stream. Network b may be able to find another provider, perhaps LIR E or F that may not be as credible.

In the former case, like operational externalities described in Section 2.2.1, C endogenizes the costs of the abuse externalities of its downstream clients. C must balance keeping clients and maintaining the value (usability) of its resource set (r_C). If C simply disconnects abusive networks, it effectively gives customers to E or F. If C retains abusive networks like b, the value of its resource set will be diminished. Consider the impact of b's latter choice, to renumber into E or F. If b changes providers, the value of $r_{b \in C}$ remains diminished, at least for a time. The result is that these resources are less valuable to both C and its potential clients, thus diminishing the marketability of C's services to all but abusive actors.

Further consider b's latter option, taking advantage of the fungibility of resources to renumber its abusive and non-abusive hosts with resources from other LIRs such as E and F. Network b rescinds its resources delegated by C and establishes a relationship with F. Network F delegates a "clean" set of resources to b, $r_{b \in F}$. In this case, "clean" means the full set of downstream rights $\mathcal{D}$, including all of $V \in \mathcal{D}$ is undiminished. As network b continues its abusive activity, V is again diminished to V', limiting the value of $r_{b \in F}$ for both b and F.

7.2.2.2 Breaking Windows

At this point in the scenario, b has now diminished the rights set of $r_{b \in C}$ and $r_{b \in F}$. To attract non-abusive clients, both C and F must engage in remediation efforts. At the minimum this means ensuring subsequent utilization of these resource sets is non-abuse, potentially requiring investment in increased monitoring and client vetting.[529]. Credible reputation aggregators will delist these resource sets once suf-

[529]The community is aware of this problem and has documented community expertise on the topic, see Section 7.4.3.2

ficient time has passed with no observed abusive behavior. For C, this is costly, especially if C has developed a reputation for clean hosting addresses but at a relatively premium cost.[530]

Further consider F's value proposition. F has low margins and a loss of revenue may be much more damaging than for C, which if wise, has invested some of its premium cost in remediation efforts that minimize the costs of the occasional abusive client.[531] In the short term, it may be less costly for F to simply ignore b's abusive behavior and continue to collect revenue. Depending on the extent of b's abuse and the reputation aggregator, graduated sanctions by the anti-abuse community may only affect r_b or may affect a broader prefix, and in turn F's other customers. Consider the listing of a broader prefix and a network $r_{k \in F}$ in a resource block adjacent to $r_{b \in F}$. Network k is a non-abusive actor. Denial of downstream uses will likely lead k to complain to F. Complaints, and threats to find another provider, to F may be sufficient to incent F toward C's position.

On the other hand, if F cannot remedy the situation, or cannot remedy the situation in a timely fashion, k is faced with the choice of compensating for diminished rights or finding an alternate provider. Consider the case where network k takes its business to a more consistent actor such as C. Here fungibility of downstream delegations is a remedy abuse externalities, not a cheap pseudonym for an extractive abuser. Here, F is left with a dirty block resulting from b's activities *and* the non-abusive actor k whose behavior could have contributed to redeeming that reputation has now changed providers. F may be able to delegate those resources to non-abusive actors, but, with a growing reputation for dirty blocks, F's market for clients may be increasingly limited to abusive actors.

The result is that F itself is increasingly imbued with the reputation of its clients. As an increasing proportion of r_F is imbued with negative reputation, $v(r_F, D)$ is diminished to some $v'_F < v(r_F, D)$. Network F was initially presented as complicit, but not necessarily encouraging, abuse. Shifting these static point analyses to a dynamic analysis, a progressive sequence of abusive clients may drive $v'_F \ll v(r_F, D)$. As v^a_F diminishes, F is eventually considered just as abusive as its clients.

The progression of F from an LIR comprising largely non-abusive actors to one that comprises largely abusive actors can be likened to the development of a bad neighborhood. In particular, it is an instance of the broken windows theory:

> Consider a building with a few broken windows. If the windows are not repaired, the tendency is for vandals to break a few more windows. Eventually, they may even break into the building, and if it's unoccupied, perhaps become squatters or light fires inside. (Russell, 2013)[532]

The progression of F is a canonical instance of broken windows. It is unclear,

[530]This premium cost accounts for monitoring and enforcement that ensures clients are not susceptible to diminished downstream rights due to dirty resources or the abusive behavior of actors in adjacent IP ranges.

[531]See Section 7.4.3.5 for a discussion of the value of remediation capacity.

[532]The original is (Kelling & Wilson, 1982). The quote here is an accurate representation of the broken windows theory, but is mis-attributed to the original Kelling and Wilson article.

though, how easy it is to fix broken windows.

At some point, F becomes more like network E. E does not necessarily engage in abusive behaviors itself, but is not only complicit, but is reliant on abusive actors for business, serves as a harbor for abusive actors, and substantively lowers the barriers to abusive actors gaining connectivity. In the case of network E, a large proportion of, if not all of, E's resources r_E have diminished rights. E may have a combination of long term clients and high turnover clients, actors that jump from resource block to resource block in attempts to mask their identity, and subsequently, to mask their abusive activity. In the worst case, E actively facilitates abuse, serving as a minimal barrier entry to Internet connectivity.

Consider a malicious, professional abusive network, an actor known for promulgating spam as their primary value proposition. Further consider there are a number of providers such as E. Like E, collectively these are not abusive, but they are complicit, providing a fungible space of low-value tinyednoteCheck tainted versus dirty. addresses in which abusive actors may move around in as an attempt to mask their identity. Although low-value, the cost of these pseudonyms is likely amongst the lowest in the market. Further, consider that, historically, those that have been successful at supporting abusive actors can demonstrate need, thus acquiring more space from the RIRs.

Acquiring additional space is not the same as the space being fungible. Historically, this new space has been previously undelegated, thus clean. Abuse supporting networks may be able to delegate this space for a premium, but the value drops once a credible reputation aggregator lists that space. The result is that the abuse supporting networks damage the integrity of the number stock by reducing the potential value of that stock. Amongst this class of abusive network, the local stock of dirty resources grows.

The juxtaposition of an institution delegating resources but not policing *downstream* use and anti-abuse actors rescinding tacit downstream rights ($D \rightarrow D'$) has created tension between the RIRs and the anti-abuse community. Recall that the criteria for number delegation is based on demonstrated utilization, *not* precisely what those downstream uses are.[533] RIRs' mandates do not cover how resources are used, simply that they are used to route traffic on the Internet. RIR staff and participants often reiterate that they do not want to be the "Internet police." Whether that traffic is abusive or not is outside the scope of the RIRs' remit. Aside from quarantining addresses when they are recovered, the RIRs do not engage in efforts to stem abuse themselves.

7.2.2.3 Serial Extractive Abusers

In the previous section, customers of E were portrayed as extractive, if not composite extractive, abusers. These actors derive most, if not all of their revenue from abusive traffic. In contrast to non-malicious, operational abuse, these actors do not

[533]For instance, a recent e-mail on the RIPE Anti-Abuse WG list is a recent assertion that, even though abusive downstream behaviors are not desirable, they may be legal, or at least not illegal, in some countries; see the discussion following (Doering, 2014).

have a legitimate revenue stream to anchor them to a particular resource set and, more importantly, to the reputation of that resource set. Rather, the most egregious of these move from one cheap resource set to the next, sending messages for a period, then moving on to the next. These actors are always looking for means to subvert filters and blocking lists.

Spamhaus provides one list of serial abusers, the Register of Known Spam Operations (ROKSO) database (Spamhaus, 2014e). As per Spamhaus (2014e), the ROKSO "collates information and evidence on known professional spam operations that have been terminated by a minimum of 3 Internet Service Providers for spam offenses," (Spamhaus, 2014e). As per the template of malicious actors described above:

> [t]he majority of the spammers on the ROKSO List operate illegally and move from network to network and country to country seeking out Internet Service Providers with poor security or known for not enforcing of anti-spam policies. (Spamhaus, 2014e)

The ROKSO is an instance of Spamhaus and volunteer investigators identifying and documenting well-known malicious actors. It is also an instance of collaboration between anti-abuse and law enforcement—Spamhaus provides additional information to "qualified" law enforcement to facilitate investigations.

7.2.3 Professional Senders and Delivery Strategies

Despite the beliefs of some hardcore anti-spam idealogues, not all bulk mail is abuse. Returning to the consensual basis of Internet traffic, legitimate senders are those that have a well-specified, consensual relationship with targeted recipients. Legitimate senders are actors whose revenue stream is rooted in sending messages on behalf of firms or in service of a legitimate value proposition, such as online sales, typically as a means of low-cost marketing. Following both the chain letter framing and the application to the anti-abuse constituencies thus far, the problem in those scenarios was what precisely constituted *well-specified*. As an arena, the anti-abuse community is a compromise space shaped by constitutional norms and in which the legitimate sender parameter space is specified. Perhaps more importantly, the community continuously re-evaluates, in other words adapts, what constitutes legitimate relationships between classes of senders and receivers.

Participants in the anti-abuse community, here in particular M^3AAWG, contribute to maintain this compromise space as a knowledge commons. Receivers serve as both monitors and distributed enforcement agents. Reputation aggregators coordinate with receivers to identify abusive actors and exchange reputation indicators with cooperating agents. Legitimate senders derive value from understanding both how to engage in the compromise space and the legitimate sender parameter space. In particular, credibly legitimate senders actively participate in this process to further reinforce the norms that facilitate efficacious delivery rates. Once again, the challenge is to foster credible commitment to anti-abuse norms, not simply satisficing by instrumentally adapting to messaging and reputation indicators.

Less credible, yet nominally legitimate senders free ride. These are not pure free riders. Nominally legitimate senders range from non-contributing compliance to instrumental satisficing. At the compliance end, nominally legitimate senders comply with sender norms, but do not contribute to the development of those norms. Those in the non-contributing compliance category do not degrade the MVN but the do not contribute to the knowledge commons, either.[534] At the other end, satisficing, senders continuously skirt the boundary between legitimate and illegitimate sending. Satisficers treat messaging and reputation indicators instrumentally rather than as signals to re-evaluate their sending practices. Actors that comply but do not actively contribute may not be ideal, but they do demonstrate that a viable revenue stream can be generated under requirements to endogenize abuse externalities.[535]

7.2.3.1 Warming up to Legitimate Sending

Practices for adapting and navigating the legitimate sending parameter space are the operational rules of the anti-abuse community. Understanding how to navigate is especially critical for legitimate sending. These operational rules attempt to balance *a*) endogenizing the costs of legitimate messaging, *b*) creating a credible commitment to consensual sending norms, *c*) developing new indicators and tools that effectively signal deviations from good sending practices, and *d*) promoting these practices by sharing knowledge necessary to effectively interpret messaging and reputation indicators (signals).

The previous section described the degradation of IP resources delegated to actors moving along of spectrum from complicit to extractive abuse. Operational rules militate against this decline in resource value by specifying precisely how to demonstrate and engage in legitimate sending. Effectively operating as a legitimate sender comprises *a*) the costs of maintaining an accurate list of consensual receivers and *b*) monitoring feedback from receiving firms in their role as reputation monitors

[534]Free riding in management facilities provision in common (pool) resources is subtly different than free riding in a knowledge commons such as depicted here. In CPR cases in the literature, such as irrigation provisioning, part of equitable joint provisioning requires all participants contribute to the construction of dams and system maintenance. In a knowledge commons, work contributions are not symmetric. Novice participants may best serve the knowledge commons by simply learning the current norms and practices from experienced actors. As such, what may appear to be free-riding may, in simple terms, be novice actors staying out of the way of the experts. As their experience accumulates, they will likely have contributions. Unsurprisingly, this is a characteristic of IETF consensus processes and is also evident in the consensus processes documented here. The converse of this is when newer experts, steeped in contemporary issues, must push back against, and in some case displace, more experienced actors clinging to older norms or best practices that may no longer be effective in the current environment.

[535]In some of the smaller, informal SISC, typically "fight clubs," if a participant is no longer actively contributing on a regular basis, that participant may be asked to politely withdraw until such time they can either resume active contribution or they are in a position to resume active contribution. These SISCs are more selective in their membership, often inviting actors in only after they have demonstrated anti-abuse investigation and remediation skills in the wild. In contrast to the issue of free riding in Footnote 534, these participants are actively selected for their specialized knowledge and ability to actively contribute quickly in the event of a security incident. As such, lack of contribution does detract from the mission of the SISC.

offering messaging indicators and from reputation aggregators such as IPBLs to ensure ongoing compliance with legitimate sending practices. Boundaries of the space of legitimate sending parameters are not discrete. Rather, these boundaries are more accurately described as loosely coordinated distributions of parameters.

The provenance of these distributions is the set of monitors and enforcers in the MVN. Monitors in this space comprise receiving firms and IPBLs. These actors loosely coordinate their thresholds for what constitutes abuse, but do not all agree. Receivers also maintain statistics on observed rates and deviations from average rates from known IPs. Some actors publish these thresholds, such as sender rates described in Section 7.1.3. Others do not publish their thresholds for messaging indicators. If these thresholds were published by all, especially those for unknown IPs, it would be a recipe for snowshoe spammers[536] distributing their illegitimate messaging, using thresholds as a weighting scheme to fly just under the radar of rate-based message blocking.

Despite these thresholds, the dynamics of these rate distributions are known and respected by legitimate senders. While receiving networks do not advertise these thresholds, they recognize that new legitimate senders may emerge and may need to navigate these distributions. Consider the simplest case, a new legitimate sender Q with a large, well-maintained messaging list. Further assume that all of these recipients have opted-in and are genuinely interested in receiving marketing and transactional e-mails from Q. Q maintains its own sending management processes and has a resource set r_q dedicated to is sending activity. Finally, consider that r_q is newly delegated—no actor participating in the control plane has any experience with resource set r_q or its sending rates.

Given this perfect information about Q, it is clear that S *is* a legitimate sender and its messages *should* be delivered. In terms of the anti-abuse community, r_q is "cold" resource set. Immediately sending from r_q at what would be Q's ideal rate (within other sending parameters of course), would trip receivers thresholds. One strategy recommended by credible third party senders[537] and the anti-abuse community is to "warm up" the sending resource set, in this case r_q. "Warming up" the resource set is a metaphor for familiarizing recipients with sending resource set r_q.

Warming up, or familiarization, is achieved by slowly increasing the sending rate to acclimate recipient firms. As noted earlier, credible, sophisticated receivers keep track of the sending rate of known senders above a particular threshold, considering blocking, or diminishing delivery rates in the event of sudden deviations, typically sudden spikes in sending. Q's comprises clients of multiple receiving firms. It is

[536]The definition of snowshow spamming from Spamhaus is:

> Like a snowshoe spreads the load of a traveler across a wide area of snow, snowshoe spamming is a technique used by spammers to spread spam output across many IPs and domains, in order to dilute reputation metrics and evade filters. (Spamhaus, 2015b)

They fly under the radar because no one IP or contiguous range sends sufficient spam to trigger threshold-based indicators.

[537]For instance, see SendGrid's white paper (SendGrid, 2014).

this diversity of threshold choices across these firms that create the distribution of thresholds. When recommending warming up r_q, a consistent rate, across this variety of recipients, is recommended.

A number of factors warrant a consistent slow-start. A consistent rate is limited by the lower valued range of the distribution of thresholds. A slow-start strategy for warming up a sending resource set r_q allows receivers to use the growth period to evaluate the character of messages and the legitimacy of the sender based on a variety of factors *in addition to* sending rate. For instance, consider if Q is less than ideal. Receivers may monitor bounce rates during the growth rate: a high rate of multiple bounces from the same set of addresses may be a signal of poor list hygiene. As established at the outset of this scenario, Q is a credible sender, so the same messaging identifier does not bounce twice and Q does not hit any of the receivers' (or reputation aggregators') spamtraps. In effect, familiarization by way of positive signals from messaging indicators establish sender legitimacy. In the event of poor indicator performance, recipient firms have the opportunity to signal the sender as a justification for increasing the sender-specific threshold rate.

Warming up an IP address is one instance of navigating the dynamic parameters of the legitimate sending parameter space; see Section 7.4.3 for elaborations in the discussion of BCPs. Warming up also illustrates the investment in monitoring and adaptation necessary to maintaining legitimate sending practices and the potential for accruing positive reputation. Further, warming up is an instance of endogenizing the costs of legitimate sending.

Not all senders have sufficient knowledge or investment in the MVN to monitor and adapt to these dynamic parameters, though. For instance, consider an online retailer L. L sees messaging as a low-cost marketing mechanism that supplements other marketing mechanisms in support of online sales. Further, L may not be familiar with messaging norms. The combination of these factors is a recipe for naïve operational abuse.

Endogenizing good messaging practices may not be feasible for actors such as R. Operational capabilities and attendant costs to may not be available to actors such as R. On an *individual* basis, monitoring and adaption necessary for good sending may be too costly to yield a return on investment. A firm dedicated to collating and interpreting messaging indicators may be able to take advantage of economies of scope and scale to reduce the costs of monitoring and adaptation. The next section describes firms that have emerged as professional senders that also serve as reputation monitoring agents for their clients.

7.2.3.2 Promulgating Credible Commitment to Legitimate Sending

For individual senders whose primary value proposition is not directly related to MVN dynamics or messaging infrastructure, the cost of legitimate practices, in particular the learning curve and measurement costs, may be quite high. Actors in the anti-abuse community have developed a variety of commercial and non-profit services for *a*) promoting legitimate sending practices, *b*) abuse monitoring and enforcement practices, and *c*) reputation monitoring. Thus far, senders are actors

that benefit from originating messages that are consumed by, and ideally acted on by, recipients. For instance, acting on a message may be an impression or a sale of goods directly or indirectly from those messages. Instances of what are referred to here as originating senders are marketers such as online and terrestrial retailers. For small non-infrastructure actors, referred to as origin senders, seeking commercial growth, the perception that e-mail is an immediate low-cost, high ROI tool is attractive, but comes with the risk of operational abuse.

Rather than each origin sender *re*discovering how to navigate the dynamics of the legitimate sender parameter space, professional senders have emerged to provide origin senders with services that reduce these costs. Professional senders have invested in relationships in the MVN, tools to process messaging indicators, and services for making this information accessible and actionable by origin senders focused on their primary value proposition. Professional senders' value proposition is a deep understanding of and services facilitating navigation of abuse monitoring and enforcement strategies and tactics employed by both IPBLs and senders. Origin senders minimizing the costs of legitimate sending services by outsourcing service provision to professional senders.

Professional senders benefits and expertise grows with their client base. As the number of clients increases, network effects accrue to both the professional sender and its clients. The professional sender garners more experience with legitimate *and* illegitimate sending dynamics from origin senders that must remediate operational abuse. It also garners knowledge from observing reactions of monitors when legitimate senders encounter new thresholds or veer towards those as result of changes in practices. Credible ESPs[538] are repositories of knowledge on how to interpret indicators and changes to sending practices that will yield increased deliverability and delivery ROI. If not obvious, this is different from instrumental application of indicators to increase raw deliverability rates.

Veering toward the boundaries of legitimate messaging practices may be unintentional or intentional. Unintentional instances may be a change in messaging that does not account for messaging best practices or changes in thresholds by monitors. Instances of intentionally veering toward illegitimate practices may be more aggressive marketing campaigns. Interviews have reported that there are many actors that satisfice relative to legitimate sending practices. Satisficing senders use professional senders services to hover as close to illegitimate sending practices as possible without sufficient reputational damage to cause local revenue loss. Satisficing is compliance in exclusively to avoid sanction. Satisficing as a strategy implies the firm either has not identified how to garner value from legitimate sending or believes the changes attendant with credible commitment will yield less revenue than satisficing. The difference in motivation follows the distinction between coercion and legitimate authority.

Professional senders can be a part of a constructive, reinforcing feedback loop between reputation aggregators, monitors, and origin senders. By promoting and

[538]ESPs are e-mail service providers, specialized firms dedicated to hosting e-mail servers and deliverability services for e-mail campaigns and marketing.

facilitating legitimate sending practices, professional senders help origin senders maximize their marketing objectives. Such a reinforcing feedback loop relies on a number of factors, the first of which is professional senders' good reputation and their investment in maintaining that reputation. Reconsider the notion of "warming up" and the sending resource set. Newly delegated resources do not have a bad reputation, downstream rights have not be abused and are not diminished. Newly delegated resources do not have a good, or *enhanced*, reputation, either. These resources will not be accorded the same treatment if, as per the warming up scenario above, a new sender materializes and begins sending legitimate messaging full bore. Rather, the process of familiarization described above is a dynamic process of imbuing that resource set with a reputation. It enhances the value of downstream rights, in particular sending rights, denoted s.

Professional senders' value proposition is that they have invested in good sending practices and, as a consequence, have enhanced their sending resource set, and by proxy, their sending rights within the messaging value network. Consider various sending rights using the nomenclature developed earlier:

1. the sending rights of a previously undelegated resource are denoted s,
2. enhanced sending rights held by a legitimate professional sender is denoted s^p,
3. the sending rights of a operationally abusive, say a nïave origin sender, s^o,
4. the sending rights of a satisficer s^s, and
5. the sending rights of an extractive actor s^a.

The result is a value ordering of:

$$v(r, s^p) > \quad v(r, s) > \quad v(r, s^o) > \quad v(r, s^s) > \quad v(r, s^a) \tag{7.9}$$

$$v(r, s^p) \gg \quad v(r, s) \gg \quad v(r, s^o) \gg \quad v(r, s^s) \ggg \quad v(r, s^a) \tag{7.10}$$

Professional senders' value is rooted in a good reputation that helps avoid sending thresholds and listing by reputation aggregators. As a result, professional senders can credibly offer higher delivery rates and the *potential* for higher click-through rates.[539]

Large professional senders have number delegations directly from the RIR; credible professional senders are an instance of LIR C in earlier scenarios. Professional senders' resource sets fungibility is limited by both their reputation investment and needs-based requirements of the RIR. At a minimum, professional senders comply with legitimate sending practices because the integrity of their resource set is critical to their value proposition. Ideally, given the potential knowledge base, professional senders are active participants in the anti-abuse community, promulgating legitimate sending practices and contributing to the development of these sending practices in fora like M^{3}AAWG. Participation in these broader knowledge commons

[539]This latter, higher click-through and actual viewership, is more closely linked to the character of the message and traditional marketing dynamics: content of the message, demographic segmenting, uses of consumer history, formatting and design of messages, etc. That said, sender reputation is key to getting those messages into the inbox.

allows professional senders to more closely *a*) monitor the changing dynamics of messaging abuse and the attendant indicators, *b*) supplement their own knowledge base, and *c*) learn about monitors and enforcers that may not be in the professional senders collaboration network. This information alone provides inputs into professional senders processes for understanding and ensuring deliverability. Yet again, the balance of instrumentality and norms is evident.

Participation also creates a feedback loops between senders and the monitor-aggregator-enforcer community. False positives by monitors can be discussed in closed fora, allowing senders and monitors to revisit the established common image of what constitutes legitimate and illegitimate sending practices. In some cases, this results in a discussion of observable indicators and their interpretation. In terms of the three strategies for remediating abusive messaging, this is a form of more precisely specifying what constitutes legitimate messaging and how to identify that legitimate messaging. It is also an instance of the feedback loop between strategy three, reputation attribution, and strategy one, dialog.

From the perspective of promulgating legitimate sending practices, a market of professional senders promotes the integrity of both the MVN and underlying resource sets. Consider the incentives for professional senders to ensure all of their clients engage in legitimate sending practices. Recall blocking is not always at the individual IP address level. In the case where a single address delegated to a professional sender is dedicated to a single client, if that client misbehaves, it will likely land that address and adjacent addresses on a blocking list. Not only is the abusive actor affected, but other clients of the professional sender as well. This damages the credibility of the professional sender and, by proxy, the professional sender's value proposition.

The result is that credibly committed professional senders ideally attempt to enhance the reputation of their resource set as a source of competitive advantage. These professional senders monitor their clients behavior and even proactively limit sending that, in the professional sender's experience, will be considered illegitimate or is in the penumbra of legitimate and illegitimate. Two models of revenue generation have been discussed in interviews. The first is rooted in the assumption that consensual marketing campaigns will effectively identify actors genuinely interested and thus, while there are fewer targets, each target has a substantively higher probability of engaging. The second is rooted in the satisficing approach. These actors optimize along the boundaries of legitimate sending practices. Satisficers recognize that compliance is necessary, but still fundamentally operate on casting as wide a net as possible. While satisficers are technically engaging in legitimate sending practices, monitors and enforcers recognize these actors are not credibly committed to these practices. This tension has led to a significant subset of monitors and enforcers that recognize the active incentive as coercion (rather than credible compliance), and, subsequently, that these actors are simply coerced abusers.

Given this characterization of the legitimate sending space, what constitutes a legitimate sender, whether credible or satisficing, is an investment in continuously updating their priors on the parameter distribution. Like other common resources

discussed in this dissertation, the legitimate sending space is a consequence of compromise amongst nominal adversaries engaging to maintain a common resource Here, a common set of indicators and parameters are continuously developed, negotiated, and applied to make consensual sending norms durable. These are made enforceable in part by projecting the implications of credible commitment, nominal compliance, and deviation onto rights bundles in the underlying number resources. Venues for managing the legitimate sending space created by this ecosystem are communities such as M^{3}AAWG and the APWG.

7.2.4 Distributed Reputation Amongst MVN Participants

No single actor has the information or capability to authoritatively imbue a particular prefix with a reputation value. As alluded to earlier, a sub-constituency in the network operator communities do often focus their attention on reputation aggregators that have seen widespread use, such as Spamhaus. Reputation attribution, the impact on downstream services, and, consequently, the impact on number resources' value is the product of collective action. Reputation indicators are provisioned by a variety of actors. Receivers are a class of reputation monitors that have direct access to messaging indicators produced by their clients; receivers can make decisions solely on local data but often incorporate other sources to enrich and validate their local view. IPBLs are a class of organizations that combine reputation monitoring (for instance deploying spamtraps), data collected from diverse partners (such as receivers), and investigations into instances of extractive abuse to (jointly) provision reputation indicators.

When reputation indicator appropriators use indicators in their filters, that action serves as a sanction, diminishing the value of the resource. With the exception of a few large receivers, only the application in filters by a large set of receivers (network effects) will, in aggregate, diminish value significantly. The following builds on the resource value nomenclature to precisely illustrate mechanisms by which diverse reputation indicator appropriators use of indicators bind reputation to the subject of the indicator.

7.2.4.1 Local and Aggregate Reputation Images

Messaging and reputation indicators are produced and consumed by a variety of actors in the MVN. Senders, professional senders, receivers, vendors, and IPBLs all consume indicators in their roles as reputation monitors and/or reputation aggregators. Each consumer may generate indicators in the course of its own business practices as well as consuming indicators produced by others. Professional senders, receivers, vendors, and IPBLs are instances of actors that may also provision indicators based on information these actors have access to from their unique vantage point and customer base. In turn, each individual consumer of reputation indicators combines indicators from different sources to create its own image of the reputation of a senders.

Consider the application of messaging and reputation indicators, such as by R

in Figure 7-1, to mitigate abuse. *Local* information may include *a*) the outcomes of various spam identification filters, *b*) sending volume thresholds conditioned on *c*) previous history with known or unknown senders, and *d*) bounce rates to revisit some of the indicators described in Section 7.1.3. Reputation attributed to particular actors based on this information will be referred to as the *local reputation image*, produced by a *local reputation function*. An *private reputation function* and the attendant private reputation image, is parameterized *exclusively* by local (private) information. For large actors, private reputation functions may be sufficient; contemporary instances are GMail and Yahoo!, two of the largest receivers in the MVN. For medium and small actors, private reputation functions may be insufficient.

Recall the information sharing scenario in Section 7.1.1, in which different networks experienced different forms of abuse from different actors. Consider a set of locales that provision private reputation images, enumerated $p_1 \ldots p_n \in P$. In the basic case, the effects on sending rights bundle s is the projection of private information onto s. One approximation of an actors' reputation is the aggregate effects of private reputation images on the sending rights of a particular resource set. The sending rights of a resource r under the private reputation image of p_i is denoted s_{p_i}. The aggregate effect, an aggregate reputation image, transitions from (approximately) s_P to diminished sending rights s_P^o or even s_P^e.

Sending rights sets resulting from the projection of local reputation images are denoted by

$$\overrightarrow{s_P} \equiv \{s_{p_1}^a, s_{p_2}, s_{p_3}^e \ldots s_{p_n}^o\} \tag{7.11}$$

Refining the value nomenclature used thus far, the value of a particular resource set under a set of private images is *defined* as simply:

$$v_P(r, s) \equiv v(r, \overrightarrow{s_P}) \tag{7.12}$$

The refinement in Equation 7.12 highlights that value is always conditioned by the reputation of some set of actors, typically receivers, in the MVN.

Expanding this to make the aggregate effect of private reputation images more explicit, we can precisely specify the approximate value of a bundle of sending rights as:

$$v_P(r, s) \approx \frac{\sum_{p_i}^{P} [v(r, s_{p_i})]}{\sum_{p_i}^{P} [v(r, s_{p_i}^p)]} \tag{7.13}$$

This is an approximate reputation and value effect, it is more useful as a concise specification of what contributes to that effect.

Equation 7.13 requires some explanation relative to the earlier, simpler nomenclature. The earlier nomenclature $v(r, \mathcal{B})$ and $v(r, \mathcal{D})$ was presented in Section 3.1

as an intuitive means to partition infrastructure rights ($\mathcal{B}$) from downstream rights ($\mathcal{D}$). Here the nomenclature is refined to more accurately represent of the maximum value that may be garnered from rights bundle s, assuming the exercise of those rights is consensual.

For a resource block c exercising rights s to send to receivers in $r_1 \ldots r_n \in p_i$, the function v is defined as:

$$v(c, s_{p_i}) \equiv \sum_{i=1}^{n} v(c, s_{r_i}^{\cdot}) \tag{7.14}$$

In Equation 7.14 each element in the summation represent the value of exercising sending rights with that particular participant. Ideally, under ideal consensual sending (s^p), the value of v for a consensual receiver will be positive and all others (those that have not consented and thus should not be on the list) is 0. An example for $r_1 \ldots r_7$ is,

$$v(c, s_{p_i}) = \sum \{+, +, 0, 0, 0, 0, 0\} \tag{7.15}$$

where of the seven recipients, two garner value and five are either uninterested or have untapped (latent) demand.

The maximum value possible *under consent* be the situation in which all actors in p_i that would consent (have demand), had consented and were on the list sent by c. This is unlikely. The nonconsensual version of this formulation attempts to tap into this ideal set by sending to all recipient in p_i regardless of interest, presuming they will garner value from proportion of recipients genuinely interested. The externality manifests in the set of actors that are not interested, and for whom messages from c are a cost. Following Equation 7.15, an instance of the externality is,

$$v(c, s_{p_i}) = \sum \{+, +, -, -, -, -, +\} \tag{7.16}$$

where the sender gets one more impression, but at the cost to the receiver of four negative. Depending on the history of the sender, this kind of behavior would be characterized as at least naïve operational abuse, s^o.

The ideal maximum value for overall sending is

$$v_*(c, s) \equiv v(c, \overrightarrow{s_*^p}) \tag{7.17}$$

This is the value that all credible senders arguably aspire to. It is the situation in which every recipient on their list has opted-in continues to value messages from the sender. Conversely, this means that there are no recipients that experience costs due to unsolicited messages from the sender. In other words, there are no negatives ($-$) in the recipient vector such as in Equation 7.16.

Returning to Equation 7.13, the expression $v(c, s_{p_i}^{\cdot})$ represents the value of the resource set c when exercising (sending) rights bundles s with respect to network

p_i's filtering choices. Ideally, in the case of a credible sender with good reputation

$$v(c, s_{p_i}^{\cdot}) = v(c, s_{p_i}^{p}) \tag{7.18}$$

If the sender is having problems, it is the more likely case that

$$v(c, s_{p_i}^{\cdot}) < v(c, s_{p_i}^{p}) \tag{7.19}$$

for a number of $p_i \in P$.

The loss created by poor sending practices, as imposed by the private reputation images in P, can then be specified as:

$$l_{+,P}(c, s) \approx v_{+}(c, s) - v_{P}(c, s) \tag{7.20}$$

This is the loss relative to what could have been garnered from good sending practices. Note the use of $v_{+}(c, s)$ rather than $v_{*}(c, s)$ As noted earlier, $v_{*}(c, s)$ is perfect sending practices and a list comprising all actors that would be interested and garner value from those sending practices; this is aspirational, but likely unattainable in practice. The notation $v_{+}(c, s)$ will be used to denote sending practices equally as good as the ideal, but recognizing that there may still be latent demand. As may be obvious,

$$v_{+}(c, s) < \qquad v_{*}(c, s) \tag{7.21}$$

$$l_{+,P}(c, s) < \qquad l_{*,P}(c, s) \tag{7.22}$$

and the difference is what drives marketing senders toward either better recruitment mechanisms, satisficing with how lists are constructed, or creating externalities (such as the simple externality in Equation 7.16), all in an attempt to capture the difference.

The function l facilitates specifying the relative effects of different reputation images. For instance, $l_{+,P}(c, s) < l_{+,G}(c, s)$ where P is a set of small private actors and G is Google's private reputation image. This expression *asserts* that the sending rights diminished by the group of small actors is less than the sending rights diminished by Google's large sample. A contemporary question in the anti-abuse community is whose reputation image has the greatest effect on sending practices? Generalizing, reputation images comprise:

individual private reputation images created by small and medium sized actors such as p_i above;

uncoordinated private images such as v_P;

aggregate reputation images created by reputation aggregators as a means to consolidate and coordinate information collected by private actors and the aggregator itself (denoted in general as v_{RA});

large reputation images generated by large private actors (denoted $v_{LP}(r, s)$ and

$l_{LP}(r, s)$) such as Google ($v_G(r, s)$ and $l_G(r, s)$) or Yahoo! ($v_Y(r, s)$ and $l_Y(r, s)$) whose sample sizes and diversity may rival that of the reputation aggregators;

mixed heterogeneous reputation images is the image common in the wild, where private actors combine local data with reputation indicators appropriated from reputation aggregators to add diversity to their own image.

Private reputation functions, such as v parameterized by p_i, have limitations for small actors, but actors such as Google, with some of the largest recipient bases, may be able to act independently. Historically, reputation aggregators have been the third party consolidators of private information. As noted in early discussion of IPBLs, they are susceptible to market forces—if they are too aggressive in their sanctions or have to many false positives, they are deprioritized and fall to the wayside. Their role as a reputation aggregator is their primary value proposition and thus incents effective consolidation and accountability.[540] For instance, in a number of interviews, actors attribute substantive weight to the reputation indicators[541] provisioned by Spamhaus.

Other actors, most notably professional senders, have indicated that large recipients, such as Google and Yahoo!, represent large swaths of their market. Both of these actors utilize large private reputation images. A question to be addressed in Section 6.5 is, if it is the case that

$$l_{LP}(c, s) < l_A(c, s) \tag{7.23}$$

there may be a shift in authoritative reputation indicators to a small cohort of large private actors. As may be obvious, this is a dimension of the larger to anti-abuse norms versus indicator satisficing issue that has presented frequently at numerous points in the anti-abuse discussion throughout this chapter.

7.2.4.2 Monitoring Reputation

While origin senders can monitor messaging indicators, the background and/or capability development necessary to make those indicators actionable is a substantive investment. Professional senders reduce the costs of infrastructure and monitoring the diverse set of reputation indicators, helping origin senders navigate indicators produced by the images described in the previous sections. Reputation monitors and aggregators integrate the indicators discussed into more actionable reputation scores, such as Return Path's Sender Score.[542] Reputation aggregators have extensive relationships with monitoring and enforcing agents. These relationships supplement aggregate indicators observable by the sender, such as changes in delivery

[540]This is more evident in the shift from strong sanction to what is presented later as graduated sanction in service of promulgating good sending practices.

[541]In Spamhaus' vernacular, this is blocking list advice, invoking the non-binding character of anti-abuse operational rules. See discussion of resource structure in Section 7.3.

[542]A single valued indicator such as Return Path's Sender Score is a flat, unit-less score that arguably tracks $l_{+,RP}(c, s)$.

rates, bounce rates, etc, with perspectives from within receivers that partner with reputation aggregators. Within the messaging ecosystem, a number of the professional senders use reputation aggregators to ensure the integrity of their sending resource sets, and, subsequently, to protect their value proposition.

In their role as a monitor, reputation aggregators can also play a mediation role between senders and other aggregators, such as IPBLs.[543] For instance, satisficing senders may have larger lists and more recipients than credible senders, but they face a higher risk of complaints.[544] In particular, because these actors live in the penumbra of legitimate sending practices, the "Report as Spam" button provisioned by mailbox providers such as Yahoo! and Google Mail is an ever present threat. A high complaint rate in p_k will definitely affect the reputation image based on local information, i.e. it will diminish the local value of that resource to $v(r, s^a_{p_k}) < v(r, s_{p_k})$. More aggressive sending practices also increases the probability of hitting spamtraps. If a satisficer hits a spamtrap, they may very well find the IP address used to send that message, or even a block containing that sending IP address, on the blocking list maintaining that trap.[545] In terms of diminished value, given the widespread appropriation of BL reputation data in mixed reputation functions, a spamtrap hit could have a substantive effect on l.

Professional senders, in their role as a Reputation monitor, can help satisficers move away from the penumbra and back toward core legitimate sending parameter space via good practices. Assuming satisficers are willing to allow the reputation monitor to mediate their sending practices, reputation monitors can act as intermediaries to both avoid damaging practices and accelerate repairing a damaged reputation. This latter repair function should be qualified. In the role of reputation monitor, professional senders stake their own reputation as a credible enforcement agent each time they act to repair the reputation of a sender. Unsurprisingly, reputation monitors cannot (and will not attempt) to repair the reputation of a malicious abusers. They can help repair the reputation of marginal operational abusers, some satisficers, and the occasional credible sender that acquires a less credible subsidiary or campaign.

A canonical case is facilitating the repair of a credible sender that acquires a less credible subsidiary or has a client that, under new management, abuses the repu-

[543]IPBLs are what might be called a pure aggregator: they do not provide the consultation services that a commercial monitor, such as Return Path, provides, but they do produce and distribute aggregate reputation indicators. As will be discussed in Section 7.3.2, this boundary is not hard and fast, especially amongst contemporary IPBLs that are becoming more communicative with those attempting to remediate origin senders that have strayed into abusive practices.

[544]Reconsider the recipient vectors in Equations 7.15 and 7.16. The satisficer is after the potential gain from latent demand, the transition of 0_7 in Equation 7.15 to $+_7$ in 7.16. Depending on how aggressive the sender is, it may impose the externalities denoted by $-_3 \ldots -_7$. Satisficing is an attempt to garner latent demand and minimize externalities by instrumentally optimizing on messaging and reputation indicators rather than treating indicators as a signal to consider a more holistic approach perceived to be more expensive.

[545]The development, deployment, and maintenance of spamtraps is an art in the messaging industry. It is also an instance of operational rules. Operational rules in the anti-abuse community are discussed in general in Section 7.4.3; Spamtraps are discussed in Section 7.4.3.1.

tation it has developed. In these cases, the reputation monitor can provide information on the credibility and legitimacy of the "good" campaigns and demonstrate that the abuse originated from a particular "bad" campaign. To effectively repair reputation, the professional sender would have to commit to more closely monitoring the offending campaign. Making this credible commitment to the reputation monitor, that monitor *may* then stake its reputation on that commitment to appeal to the blocking list on behalf of the professional sender. In the ideal case, the professional sender corrects the problem by either remediating the sending practices of the abusive campaign, suspending the abuse campaign, or even severing the relationship with that client. In the abuse case, the professional sender may suspend temporarily, but does not credibly commit to a complete remediation, allows the abusive sender to continue sometime later, and finds itself again with diminished reputation. In this case the reputation monitor recognizes this and, at the least is unwilling to further lend its credibility to the professional sender. In the worst case the reputation monitor may sever its ties with the professional sender.

Another instance is a new sender inadvertently engaging in operational abuse. Mechanically, it is similar to the satisficer but is different in terms of norms promulgation. In the case of the satisficer, the satisficer is well-aware of norms and good sending practices. By definition, the satisficer is incented to comply in fear of coercion rather than endogenizing legitimate sending norms as authoritative bounds on revenue seeking behaviors. The satisficer has made an informed choice.

In contrast, a relatively uninitiated operational abuser is not yet either a satisficer or a credibly committed sender. As implied by "uninitiated," this new actor has not been exposed to either legitimate sending norms or sanctions for violating those norms. The first sanction, listing on a blocking list or some other form of diminished reputation, is an opportunity to illustrate anti-abuse norms, provide insights into effective tools for compliance, and, perhaps most importantly, demonstrate that compliance *does* yield corrective benefits. A characteristic element of enforcement in common resource management regimes is graduated sanction and discretion. As discussed in Section 7.4.3, a number scenarios highlight how discretion can be used to limit both attrition and further rule breaking. Also consider empirical evidence from deterrence theory: multiple studies have shown that the celerity and certainty of sanction has a greater effect on deterrence than the severity of the sanction. Taking these together, immediate sanction demonstrates an actor's behavior *is* being monitored and that the community does have the means to enforce it. Relaxing the severity of the sanction relative to that imposed on a consistent operational abuser or an extractive abuser sends the signal that the anti-abuse regime is not simply persecuting senders without consideration of their revenue streams, genuine demand for marketing messaging, or, most importantly here, their willingness to remediate.

Finally, consider the scenario in which a credible sender with a long history of good sending practices finds itself on a blocking list. For instance, consider a firm H that provides hosting services akin to those provide by Amazon or Rackspace. These actors offer commodity virtual machines to any actor, without the vetting and monitoring performed by professional sending firm. Further consider that the

hosting space offered by these firms may have multiple customers per IP address.[546] One abusive actor can have adverse effects on a number of (legitimate) non-abusive actors. Such hosting providers have a good reputation because they do have policing efforts in place and they respond to abuse reports promptly.

A prompt response is precisely the effect intended by reputation-based sanction. For those credibly committed to anti-abuse norms, correcting abusive messaging is simply their duty in the messaging commons. Satisficers comply to avoid sanction. Despite the perception by some actors in the sending and network operators communities, reputation aggregators, in particular BLs, prefer prompt remediation—they do not benefit from networks remaining blocked. For instance, multiple interviews amongst anti-abuse leadership, both those in the reputation monitoring community and the blocking community, have decried blocking lists with poor delisting practices or that extort listed actors by requiring payments for delisting. The instance of these recounted have since ceased operation. Prompt, verifiable remediation is beneficial for the blocking list: it achieves *a*) the larger goals of the anti-abuse community (remediating abusive practices) and *b*) improves the reputation of the BL as a credible and legitimate reputation aggregator.

As noted earlier, across interviews of anti-abuse community members, it has been stressed that experienced actors do not rely on a single reputation indicator, such as a single IPBL or a simple local indicator such as bounce or send rate. Rather, each appropriator of reputation information applies their own mixed heterogeneous reputation function, using and weighting different messaging and reputation indicators in different ways. While these weighting functions are different, they do tend to weight the content of IPBLs, both commercial and non-profit, heavily. For instance, when an actor g lands on a reputable blocking list such as Spamhaus, it is almost guaranteed that the value of g's resources will diminish. Depending on how blocking list reputation is used and the severity of the abuse, a reputable blocking list can push a resource to nearly unusable for a given set of rights. The next section describes anti-abuse resource structures, focusing on IPBLs as jointly provisioned resources that facilitate imbuing number resources with reputation.

7.3 Resource Structures Provisioning Reputation

In the previous section, discussion focused on understanding the constituencies in the MVN. Sender, receivers, and vendors relationships were described in terms of their position in the MVN and how reputation serves as a mechanism for coordinating the remediation of externalities. The discussion culminates in Section 7.2.4.1's nomenclature for describing how different actors *consume* reputation indicators, combine these indicators, and the comparative outcomes in terms of various parameterizations of reputation functions. Section 7.2.4.1 treated actors generated reputation indicators largely as black, or at best, dark grey, boxes. This section opens those boxes to present how reputation aggregators, with a focus on blocking

[546] In both the hosting and professional sending space this is referred to as *shared infrastructure*.

lists, decide to bind (or attribute) good or bad reputation to a particular prefix. The following sections describe some of the behaviors BLs have offered as rationales for listing, the BL nomination processes and mechanisms, delisting processes and mechanisms, and a description of Spamhaus' various blocking lists as an instance of different types of reputation indicators (advice).

7.3.1 Why Did I Get Block Listed?

One of the contributes of the anti-abuse community is the knowledge commons necessary to understand why resources are "blocked," i.e. why they are attributed with negative reputation. The following describes a number of the more well-known rationales, ranging from principals such as abuse origination and complicity to mechanics such as relays.

7.3.1.1 Abuse Origination

As may be obvious from earlier discussions, the canonical reason an actor c is placed on a blocking list is because the BL operator has what it considers sufficiently strong evidence that k is engaging in abusive behavior. For instance, the original MAPS RBL Introduction indicates:

> The original focus of the MAPS RBL when it began operations in mid-1996 was on identifying the sources of dedicated, professional spammers. (MAPS, 2004, p. 3)

This introduction references dedicated, professional spammers. The generalized modern incarnation of originating spam is originating abusive behavior. In the typology of abusive actors identified here, "dedicated, professional spammers" are simple and composite extractive abusers that leverage abuse externalities in service of, and typically as their exclusive, revenue stream.

At the other end of the spectrum, operational abuse is not an intentional exploitation of abuse externalities. Recall naïve operational abuse is likely rooted in actors ignorance of good sending practices and norms. Graduated sanction is an opportunity to educate actors that are trying to enhance an otherwise legitimate revenue stream.[547] In that case, IPBL listing due to operational abuse is a means to send a signal that norms exist and c is violating those norms. Framed as such, answering "Why was I listed?" is more appropriately framed less as "because you are originating abusive behavior ¡full stop¿" and more constructively "because you need to learn the rules, the behavior listed here violates those rules and here is why." The typology of abuse externalities developed in Section 7.1 differentiates based on knowledge, internal economics, and revenue streams of the abusers.

[547]For the purposes here, a legitimate revenue stream does not derive the bulk of its volume from externalities. In this sense, simple extractive abuse can be legitimate, but composite extractive, such as using spam as a botnet infection vector, is very likely not legitimate.

Mechanically, abuse externalities function in fundamentally the same ways. Following the argument around graduated sanction and deterrence, distinguishing between types of origination based on the motivation provides insight into which sanctions, if any at all are most effective at remediating the problem. As such, when considering why one is listed, the context is often more important than the actual act.

7.3.1.2 Unconfirmed Identifiers and Double Opt-In

Double, or confirmed, opt-in to mailing lists is a widely cited mechanism for assuring legitimate sending practices. For instance, double opt-in has is recommended by M^{3}AAWG (2011a) in its Sender Best Practices (discussed more extensively in Section 7.4.3) and has been transposed into the whitepaper marketing material of professional senders such as SendGrid (SendGrid, 2015), Constant Contact (Constant Contact, 2015) and reputation monitors such as Return Path (Return Path, 2015). Opt-in requires that users submit their messaging identifiers based on interest in a list rather being solicited without prior exposure to the list or its supposed content. Opt-out is the situation where, through whatever identifier collection mechanism, an actor's messaging identifiers are added to a list without their knowledge or consent—upon receipt of unsolicited messaging, the onus is on the recipient to opt-out. In the latter case, sender's following the opt-out strategy *a*) do not have consent, violating the fundamental framing of abuse, and *b*) place the burden of filtering the list on the recipient. Opt-in explicitly requires consent, and, assuming the description of the list is credible, i.e. it is not deceptive or misleading, opt-in is a signal of interest in the topic or the product the list is focused on. MAPS (2004) makes the link between opt-in and opt-out explicit by indicating that "[t]he opt-out approach violates our fundamental principle: *all communications must be consensual*," (2004, p. 4, emphasis in the original).

Single opt-in can be abused, though. In particular, single opt-in is not a strong, verifiable form of consent. Consider a malicious, or maybe even simply a mischievous, actor c that has access to the messaging identifiers of some set of recipients R. Actor c can "opt" the recipients in R into a wide variety of lists without their knowledge or consent.[548] To avoid this form of abuse, the notion of double opt-in has been developed. When an actor d opts into a list, say via an online form, the list manager must confirm the actor submitting the form is in fact the owner of the submitted messaging identifier i. A confirmation message is sent to i containing some mechanism such as a preference panel or a confirmation link. Such as confirmation assures that the owner of i did in fact opt-in. This also serves as evidence of reaffirming consent.

BLs may list actors that they have identified as using unconfirmed messaging identifiers in their lists. BLs can determine this via their own spamtraps and by working with mailbox providers to create spamtraps used to test and validate list confirmation strategies. A BL, potentially in conjunction with a mailbox provider,

[548]Under this specification, opt-in becomes opt-out when actor c is the owner of the list R is unknowingly opted into.

creates a set of addresses U that will be submitted to mailing lists. If the list does not use confirmed opt-in at all, identifiers in u will begin receiving messages; some strict BLs may consider this abusive. If the list does use confirmed opt-in, the BL will simply ignore the confirmation(s). In the case where the list does use confirmed opt-in, if the BL receives messages regardless of confirmation, this constitutes abuse. As may be obvious, listing due to confirmation technique is an attempt to strengthen operationalizations of consent. In effect, this is an attempt to make consent durable.

7.3.1.3 Complicit Intermediaries

Recall from Section 2.2.3 that externalities often include complicit intermediaries. In the case of route hijacking, actors that do not filter routes appropriated to confirm the origin is in fact the legitimate holder of prefix rights *are not* the source of the hijacking. That said, they are *complicit in perpetuating* the hijacking. This notion of complicity assumes that these actors are not actively engaged in the abusive activity. Nevertheless, by avoiding costly monitoring (endogenizing costs) these actors contribute to externalities that diminish the integrity of the control plane. In the case of messaging abuse, a number of behaviors may be considered non-malicious, but nonetheless complicit in perpetuating abuse.

Two early forms of complicit activity are open relays and open proxies. In the case of an open relay, professional spammers can use the relay to send spam from an otherwise legitimate mail server. Open relays were initially used by system administrators for redundancy purposes: if a user's local outgoing mail server was not available, others could be used to send outgoing mail. In the early days of spam campaigns, abuse was uncommon, but innovative abusers recognized the abuse potential of open relays. Bianchi (2013) describes Cyber Promotions, noted as "the first to start spamming Internet users on a massive scale." Initially Cyber Promotions used their own servers, but their IPs were later blocked. Cyber Promotions recognized they could use open relays to send spam. According to Bianchi (2013) other spammers followed suit and "open relay hijacking" became widespread in the early 2000's.

One remedy to this problem was to create blocking lists specifically for open relays. Historical instances listed by Bianchi (2013) include the MAPS RSS, SURBL, ORBS, and RSL; Bianchi (2013) also notes that more general blocklists also include open relay lists, such as the AHBL, DSBL, Five-Ten-SG, NJABL, and SORBS. Abuse of open relays peaked in 2001, then began to decline. A combination of BL listing and software developers change default configurations remediated the problem. In the case of the former, when legitimate mail bounced, system administrators investigated, determined they were on a BL, closed the relay, got delisted, and legitimate operations continued. Bianchi (2013) reports that by 2005/2006, the open relay problem seemed to have been solved.

7.3.1.4 Resources Supporting Abuse

Complicit intermediaries are, in some cases, described as unwitting victims turned accomplices.[549] In contrast, some Internet resource holders actively support abusive actors. These supporting actors do not originate abusive behavior themselves, but actively and intentionally contribute to an environment that facilitates abusive behavior. As such, simply listing the *immediate* source of abusive messaging does not eliminate all of the resources *supporting* abusive operations. Rescinding the rights of supporting services is another mechanism for limiting abuse, here by limiting abusive actors' operational capability.

MAPS (2004, p. 6) lists a number of support services: *a*) hosting web pages supported by spam; *b*) providing resources such as DNS, advertisements, hit counters, and backend processing to sites promoted by spam; *c*) credit card processing for sites supported by spam. Recall the discussion of credible versus satisficing LIRs. Those examples focused on actors leveraging hosting services to originate abuse, i.e. the first criteria for listing discussed in Section 7.3.1.1. Such hosting providers faced potential attrition and loss of revenue when enforcing legitimate sending practices.

Spam supporting firms face similar problems. The support services above are common support services for a wide variety of legitimate and illegitimate online commercial activity. For many of these hosting providers, they follow the appropriation ethos of the larger network operator community: hosting and infrastructure providers offer tools, they are not there to police how those tools are used. Further, as with the issue of the marginal LIR F, many of these services are low-margin product offerings—these providers rely on economies of scope and scale to remain in business. Taken together, some of these operators will argue that it is not their role to police how their clients use nominally non-abusive services. In effect, they take the position that their services do not in and of themselves have normative implications and thus the provider should not be sanctioned for how they are used.

Adding supporting firms to the BL is akin to sanctioning the source of the spam *and* the supporting (abuse) value network. Ideally, this sanction would operate on the same principle as graduated sanction as have those related to naïve operational abuse activity. The ethos of neutral service provision, in its degenerate form, may preclude discrimination, many providers to preclude spam and other abusive activities in their AUPs. In the case of actors that have only a few abusive clients, those may not warrant the collateral damage for their legitimate clients. Listing is an opportunity to correct the behavior with a credible signal. In other cases, some supporting services may have tailored themselves to support spammers as a niche market. In these cases, blocking those actors will have a low risk of collateral damage for legitimate senders and may even have the collateral benefits of limiting the activities of other abusive actors and other (potentially yet unknown) illicit actors appropriating those services.

[549] See discussion of spam relays in (MAPS, 2004).

7.3.2 Joint Provisioning of Reputation Indicators

Each IPBL is single source of composite reputation indicators for the stock of IPv4 numbers. Like maintaining RIR registry data, IPBL data is also jointly provisioned. As developed in Part I, very few single actors have sufficient purview across the private networks that comprise the Internet to speak accurately and precisely to the character of all other networks. IPBLs enlist the help of actors that *a*) have purview into private networks that experience losses from abusive activity and *b*) benefit from IPBL resource provisioning, in part because those actors must cope with the repercussions of abuse. The challenge for IPBLs is to provide reputation indicators in near real time while maintaining their reputation as a credible source of reputation information. A number of "nomination" processes have been developed, in part to balance this trade-off.

7.3.2.1 BL Nomination Processes

Discussion of whether to place a resource on the IPBL is a joint decision. The term nomination process was taken from early MAPS documentation but has been generalized here to refer to the joint evaluation of a resource a reputation aggregator is considering attributing with reputation. Common Internet infrastructure management functions, here the development of reputation indicators, their distribution, and their enforcement, *requires* joint provisioning. Like RIR registry data, this is because the necessary information is only available from particular private networks. Unlike RIR registry data, binding a reputation indicator to a resource is the product of consensus amongst actors with different sources of evidence of abuse. Different IPBLs have different decision mechanisms.

MAPS (2004) describes its RBL nomination process.[550] The first step in MAPS' process is investigation: *a*) determining if the e-mail is in fact spam, *b*) IP origin of the e-mails, *c*) actors responsible and how they may be contacted. Each of these is a potentially subjective assessment. Whether a message is actually spam may be a point of contention, especially when distinguishing between consensual marketing messages and non-consensual. Origin is less subjective but often requires multiple perspectives to confirm (cross-validate) a common source. Tracing origin is part of determining which actors are responsible for the IP addresses. Not surprisingly, one of those sources is the RIR.

Despite being an early influence on the anti-abuse community, in contrast to some of the modern blocking lists, MAPS' notification process included an element of forewarning. MAPS attempted to notify those responsible for the IPs currently in the nomination process that they were being considered for listing. MAPS provided reasons for listing, supplementing this reasoning with samples of spam attributed to the nominee. Reporting nominations was an attempt to

> ensure that a 'good-faith' effort to notify the responsible parties has been

[550] It is important to note that this nomination process describes the original incarnation of the MAPS RBL. Since then it has changed hands multiple times and is not a commercial service offered by Trend Micro.

> made before a listing goes live, and that they have a chance to respond to the list. (MAPS, 2004, p. 9)

The 'good-faith' effort is interpreted here as an opportunity to establish a dialog with the actor in question. A response is an opportunity for the BL to evaluate where this actor is in the spectrum of naïve operational abuse to composite extractive abusers. Consensus amongst those that present the evidence of abuse, its interpretation, and whether the nomination goes forward is, ideally, a credible assessment of reputation. It is also an opportunity for discretion and graduated sanction.

MAPS and modern BLs provision a number of different blocking lists, each of which documents one or more observed behaviors. An open relay can be unambiguously tested by attempting to use that relay to send messages. MAPS, along with some modern BLs, automatically and preemptively list open relays as network elements that facilitate abuse. The deterministic character of the threat, open or closed, also warrants altering the deslisting process. One protocol is for the network administrator to contact the BL indicating the relay has been closed. A credible BL will retest the relay as soon as possible and, upon confirming the relay is closed, delist it.

Another instance of the joint provisioning of a special purpose blocking list is a list of customer, or end user, IP addresses that should not have certain downstream rights.[551] Typically these are dynamically assigned addresses. In some cases, the access networks managing dynamically assigned addresses submit those address blocks for listing on their own volition. As a use of BLs, this is a preemptive listing that is intended to further reinforce limitations on downstream uses imposed by those delegating number resources. By voluntarily listing these resources, on a specialized list, with the specific semantics that these are dynamic end user addresses, the network actor adds another mechanism to the various means of enforcing limited downstream uses of dynamically assigned numbers.

BL listing is often construed as the BL manager punishing actors that deviate from that BL's operationalization of anti-abuse norms. The mechanics of distribution undermine some of those negative perceptions. The instance of DULs (dynamic user lists, i.e. dynamically assigned number ranges) illustrates that BL listings are not always a punishment. In common case, distributed enforcement is harnessed as a defense mechanism. In the case of dynamic addresses listed on a DUL, distributed enforcement is harnessed by the actor holding the rights for those addresses in order to further enforce the limitation *the holder* has imposed on its own consumers and network.[552] This highlights BLs can, and do, serve as a more generalized rights enforcement mechanism.

[551] Instances are running an e-mail server, proxy, relay, or other services that, if misconfigured, could be put to abusive purposes. The idea is that many of these services have historically had very open, and thus very appropriable and subsequently abusable, configurations. The average end user may not have the expertise to secure such services.

[552] Using the resource value nomenclature, if d is the set of dynamic end user addresses and q is the set of rights (services) blocked by a networks AUP,

$$v_{+,DUL}(d, \mathcal{D} - q) > v_{+,DUL}(d, \mathcal{D}) \tag{7.24}$$

Return Path is a reputation monitor that helps senders navigate the reputation indicator space. Return Path's value proposition is to map out the messaging indicator space described in Section 7.1.3 and offer a service that helps actors stay in the legitimate sending space bounded by those indicators. To achieve this goal, ReturnPath partners with a wide number of receivers to collect individual reputation indicators. In effect, Return Path uses its network of receivers and mailbox providers, what it refers to as the Trusted Cooperative Network, to sample and estimate the distributions of reputation indicators. It then uses these to create what it refers to as the Sender Score, a single valued indicator ranging from 0 to 100.

Like other indicators, pinning decisions to single valued thresholds of the Sender Score does not guarantee improved outcomes. The Sender Score is a starting point for improving an actor's overall reputation. Return Path conveys this in the following description:

> Any Sender Score below 100 means your sender reputation can be improved. Pushing your Sender Score to 100, however, isn't necessarily the best way to optimize your email program results. Email senders with a Sender Score below 70 typically experience aggressive email filtering applied to every email coming from the IP address attached to the Sender Score. Email senders maintaining a Sender Score above 70 typically see filtering criteria applied to individual emails and email campaigns instead of entire IP addresses. (Return Path, 2014a)

The sender score is another, albeit aggregate, indicator in a larger, and continuously changing, legitimate sender parameter space. Return Path goes on to stress that, as developed conceptually in the previous section (7.2) "[e]ach email provider uses your Sender Score to filter emails in different ways," (Return Path, 2014a). The result is that a high score, such as a 90, may improve delivery rates in one ISP but have no effect in another. Return Path advises that network actors should also keep appraised of other indicators as a means to continuously monitor the nuance of their reputation, using this nuance to as feedback to improve sending practices and, subsequently, delivery. It is important to note that while this is credible advice for any aggregate indicator, it is also self-serving: Return Path's business is to make these indicators available to senders for a fee.

Return Path is one of a number of *commercial* actors whose value proposition is the joint provisioning of reputation information. Cloudmark, discussed below, uses a collective of trusted evaluators to train message filtering services. Historically, provisioners of reputation information were volunteer and/or non-profit organizations. Of the major blocking lists referenced by the community, Spamhaus and SURBL are the two most frequently discussed as credible non-profit BLs. Return Path and Cloudmark are commercial reputation monitors and aggregators. Reputation appropriators can and do utilize a diverse set of reputation indicators from diverse providers as inputs into their local reputation function.

asserting the costs of handling potential abuse of those rights diminishes the value of dynamic end user addresses.

Cloudmark does not confer reputation on number resources, but the reputation mechanism is worth discussing as an instance of the family of consensus processes in the messaging industry. Like the actors that nominate BL listings, Cloudmark's Global Threat Network (GTN) service relies on a network of trusted evaluators to bind reputation. In the case of the GTN, message fingerprints are imbued with reputation. The trust network comprises actors in a broad set of receiving networks. An evaluator is a distinguished recipient in receiving networks that can report messages they consider spam.

Consider when an evaluator receives a message and marks it as spam. If this is the first observation as spam amongst all evaluators, the fingerprint of the message is placed in the Nomination Server (a form of registry). As more evaluators observe the message, evaluations will be elicited. Some will mark it as spam and some may not. At a particular threshold, the system decides *a*) whether a message is, with acceptable assurance, spam, or *b*) with acceptable assurance it is not spam or in some cases *c*) if it is tied. Once the message is imbued with reputation with acceptable assurance, abusive messages (spam) are automatically placed in the user's spam or bulk folder rather than landing in the inbox. In terms of reducing externalities and receiver resource consumption, this system continues to use delivery resources (MTA agent R in Figure 7-1) but reduces the burden (externality) on end receivers (r).

On the surface, the Trust Network looks like a simple voting system. Evaluators' votes *do not* have uniform weight.[553] Credibility weighting is updated with each assessment. If the evaluator agrees with the overall assessment, their credibility increases. If the evaluator is in the minority of an assessment, that evaluator's credibility is diminished. Credibility grows slowly, but is diminished rather quickly; it is hard to reach the highest echelons of trust but quite easy to fall from these high echelons.

Cloudmark's trust network developed in much the same way that the trust networks within the larger anti-abuse community developed. The community began with a core set of trusted actors. In terms of this work, that trusted set is the subset of the operational community that is considered to embody a sufficiently diverse set of views that, taken together, can credibly assess whether a message is spam (more generally abusive). At first, accruing credibility is a form of growing the set of like-minded evaluators—one cannot accrue credibility unless one's evaluations match those of the initial trusted set.[554] The ideal growth situation is that as new receivers join the threat network, evaluators that hew to existing norms, manifest

[553]Recall from Section 3.2.5 that a key distinction between consensus processes and majoritarian voting is that contributions in consensus processes are not uniformly weighted, fungible votes that can be easily exchanged.

[554]There are of course corner conditions. Consider a scenario in which there is a small number of trusted evaluators and a large number of new evaluators. Further, consider the trusted evaluators have a very broad notion of spam and the new evaluators have a very narrow notion of what constitutes spam. It is possible that, with a sufficiently high number of new evaluators, they could overwhelm the trusted evaluators. If this happens a sufficient number of times, the trusted evaluators are displaced with a more diffuse set of actors with, on average lower credibility but collectively their image of what constitutes (or does not constitute) abuse.

in trusted spam evaluations, are distinguished from those with deviant views.

As highlighted in Section 7.2.3.1, abuse parameters are dynamic, not static. These parameters change, in part, as *a*) perceptions of what constitutes abuse, *b*) how indicators interpretations change as signals of good or bad practices, and *c*) how good sending practices themselves change and evolve. One manifestation of that change is the shift in credibility of evaluators in the trust network. For instance, early on, a number of actors considered *any* unsolicited marketing messages as spam. That has since been relaxed to recognize introduction messages and transactional messages. In the course of recognizing these as non-abusive, some trusted evaluators that continued with the old norm lost credibility and those that hew to the new norm garner additional credibility. In some cases, the evaluators gradually change the norms through their assessment processes. In other cases, the change is a combination of sufficient subset of these actors changing and others being left behind as their credibility diminishes. Ultimately, weighted assessment is a mechanism to ensure that the assessment function tracks the actors considered the core credible assessors in the messaging operational epistemic community.

7.3.2.2 Delisting Resources

Mechanically, delisting resources is the process of removing that resource from a BL. Delisting is not a reputation panacea, it will not automatically restore reputation. From the perspective of the BL, delisting can signal a number of possible outcomes. The general rationale for delisting non-voluntary listings is that the evidence of ongoing abuse has changed or a credible actor has interceded on behalf of the listed resource holder as part of a larger remediation process.

Consider elaborations of each delisting scenario and sources of evidence for delisting. In the case of diminished abuse activity, when no abusive activity has been observed from a non-voluntarily listed resource within a set period of time, credible BLs remove the resource from the list. There are instances where a resource may be a voluntarily listed resource with respect to one list and a non-voluntarily listed resource with respect to another. Consider a consumer access network J that has submitted the blocks of resources it dynamically blocks to a DUL. J has recently become a victim of a zero-day malware that contributes its exploited hosts to number of botnet CC's. At this point, a subset of J's dynamic hosts are listed on general BLs for abuse and all of them are listed on the DUL. Once J remediates *and* no abuse is observed, its hosts are removed from the credible general BLs but remain on the DUL.

Consider J has a history of credible participation in the anti-abuse community: J has consistently responded to reports of abuse by its clients and has acted promptly to remediate those clients or block (endogenize) the abusive activity. In this case, it may be sufficient for J to contact the BL, or BLs, listing that network and request delisting on the good-faith assertion it is taking care of the problem. Not all network actors have sufficient credibility with the BLs, if the BL is even willing to delist without demonstrable evidence of diminished abuse. As discussed earlier, this is where reputation monitors and professional senders can act as reputable intermediaries.

Both have a value proposition that relies on their credibility.

In the case of reputation monitors, their value is in their credibility as brokers of individual and aggregate reputation indicators derived from data collected from trusted sources. For professional senders, their value proposition is rooted in their ability to achieve high inbox delivery rates as a result of their credibility and reputation for facilitating good sending practices. This is the strategy of the credible professional sender committed to anti-abuse norms. These actors are effective intermediaries because they are credibly committed to, by virtue of their value proposition, helping actors maximize the revenue from sending by engaging in legitimate sending practices. When a professional sender intercedes for an actor, especially in its role as a reputation monitor, it is a risk to the professional sender's credibility and, subsequently, their own value proposition. The professional senders' close working relationship with origin senders and their investment in origin senders' credibility allows them to operate as assessors of the commitment of senders to remediate abusive behaviors.

BLs remediation processes incorporate combinations of timing out and explicit delisting requests. In terms of timing out, modern BLs considered credible[555] delist resources when abusive activity is not observed for some period of time. A number of historic BLs have been chastised for listing, then never delisting resources; moreover, that behavior was one of the factors offered for why they were *a*) not considered credible and subsequently *b*) unused and/or no longer operational. The latter means of delisting, dialogue, may or may not require intermediation. The objective of dialogue is to adjudicate credible commitment to delisting conditions as an indicator those actors understand the anti-abuse norms that led to their listing. This does not mean that the actor is credibly committed, merely that they understand how to satisfice to messaging and reputation indicators. Reputation brokers and (legitimate) professional senders can serve as known, credible intermediaries that have sufficient access to the nuance of senders' behavior.

At the end of the discussion of listing due to netblock inheritance, MAPS (2004) indicates that

> [w]hen it is brought to our attention that an IP address is no longer under the control of a spammer, we will work with the new user to remove the address as quickly as possible. Loss of connectivity hurts us all. Spam hurts us all even more. (2004, p. 8)

Listing and delisting criteria both rely on effective indicators of reputation, both those observable by the BL and those offered by credible third parties. MAPS as well as other credible BLs distinguish between preemptive and non-preemptive listings. Preemptive listings occur when there is a high certainty of abuse or potential abuse

[555]The sample of community members' criteria that constitute what is "considered" a BL credible comprise formal interviewees and private conversations during fieldwork. A subset of actors, in both the anti-abuse community and the broader network operator community wholesale vilify BLs. These actors provide insight into how BLs effected revenue derived from downstream uses, but their particularistic focus on revenue did not distinguish amongst the variety of BL listing and delisting practices. Some, but not all of these actors were frequently listed on BLs.

from a particular resource. For instance, the listing of open relays is preemptive. Another instance is the preemptive listing of resources that fall under the control of known professional, malicious abusers such as the actors listed on the Spamhaus ROKSO.[556]

In contrast, non-preemptive listing attempts to contact the resource manager before listing a resource on the BL. MAPS implemented this policy in what it referred to as the notification phase of its resource listing process. As noted in the previous section, this is a manifestation of graduated sanction, frequently observed in common resource management regimes. In terms of delisting, this blurs the line between listing and delisting, but again highlights the power of *sending the signal* that the community is capable of prompt and certain sanction for those violating anti-abuse norms. In effect, the specter of sanction may be sufficient to create a deterrent effect.

In the MAPS BL process, if a resource was listed, it remained listed until MAPS was notified there has been a change in behavior or stewardship of that resource. Upon notification, the first step is to ensure the actor requesting delisting is an authoritative agent of the steward, not an end user or other party. It is further qualified that this authoritative agent should be "an abuse representative or server administrator," (MAPS, 2004, p. 8). This qualification ensures that the corresponding agent is not just a representative, but is in a role that has the operational capacity to *actually remediate* the offending behavior rather than provide empty assurances.

The next step in the MAPS process requests an explanation for why the notification sent when the listing was originally considered did not receive a response. The explanation facilitates gauging a number of characteristics of the listed actor. For instance, if a notification was sent to the abuse address for the network but was not acted on, that is evidence of *a*) at best an overextended organization; *b*) worse a satisficing actor that only deals with abuse when sanctioned; or worse yet *c*) an actor who completely ignores abuse notifications.[557] More optimistically, in the case of nïave operational abuse, the delisting process is an opportunity to educate the actor on the value of monitoring abuse messages and developing better relationships with the anti-abuse community, both to develop better practices and to reduce the transaction costs of future engagements.

The next step in the MAPS remediation process is for the network actor to explain what has been done to remediate the abuse and what will be done to ensure the problem will not happen again. Open relays are an easy instance: the agent can simply ask for the BL to retest. The more difficult scenario is that of the credi-

[556]The ROKSO is the Register of Known Spam Operators. One threshold criteria for ROKSO listing is that an abusive actor that has been denied service by three or more ISPs for abusive behavior. See (Spamhaus, 2014e) for details.

[557]A potential exception is abuse notifications sent to the upstream of the abusive actor rather than the abuse address of the abusive party. This is only loosely an exception. An actor committed to anti-abuse norms should be parsing their abuse address for reports regarding its downstreams and forwarding those to the downstream actor's abuse address. If the upstream knows the downstream is a good remediator, it may contact the BL to let it know that it has forwarded the report to the downstream and subsequent reports should go to the downstream.

ble LIRs discussed in Section 7.2.2.1. The credible LIR C can emphasize its history of quick remediation. No LIR can perfectly assure its clients will be not be abusive. LIRS can develop a reputation for *a)* monitoring BLs and *b)* a willingness to quickly and effectively enforcing anti-abuse norms via its AUP. In the extreme case, or in the case of a repeat offender, the LIR can demonstrate to the BL that it has terminated its relationship with the abusive actor(s). In these cases C must develop its own norms and processes for deciding when to extend its reputation and when to assert stronger sanctions when signals that there is monitoring in place fail.

More realistically, immediate termination is not likely. As developed earlier, it is not in the LIR's interest to terminate paying customers just because they appear on a BL. The compromise often sees the LIR faced with convincing the BL that it knows which of its clients are responsible and that it is working to correct the problem. Consider the situation in which the LIR resource assigned to the abusive actor, and listed on the BL, is exclusively assigned to the abusive actor. In this situation, the BL listing only affects the abusive actor. Under this circumstance, the LIR may notify the BL it is working on remediation and ask for delisting when successful. Here, BL Listing does not affect the LIR's other clients.

It is often that case that the listing creates collateral damage. In the case of a resource block, say a /28, 16 IP addresses, a number of different senders may be affected by a single listing. It may be the case that only the senders behind one of these addresses, or the senders behind a discontiguous subset $a \in /28$ are abusive. The other senders, those in $/28 - a$, may be perfectly legitimate senders. Under this scenario, the LIR may ask the BL to either scope the listing to only the abusive client or, especially if the listed block is shared amongst abusive and non-abusive clients, the LIR may ask for a temporary delisting while it remedies the situation.

Alternately, shared infrastructure often assign multiple clients to a single number resource (address). Under this configuration, the single resource is a channel for collateral damage created by a single abusive actor. Even the finest grain BL listing, a single address, has the potential to create collateral damage for legitimate senders.

Consider the /28 above. Each address supports 6 senders, the block supports a total of 96 senders. Depending on the granularity listed by the BL, only a few abusive actors are sufficient to impact all the senders behind this block. Even a low proportion of illegitimate senders in a block can create collateral damage for all the senders. For instance, consider if the boundary is 4 address blocks, 4 abusive actors, one in each block of 4 addresses. The result is collateral damage for the remaining 92 senders.

From the perspective of the BL, these delisting dynamics are intended to serve as selective incentives enforcing good sending practices. Collateral damage creates pressure on the actors that provide the number resources and connectivity necessary to engage in abusive sending. It also creates ill-will amongst those that suffer collateral damage. In what is considered here to be the constructive scenario, those suffering collateral damage will pressure the LIR to remediate the abusive senders by either suspending services for abusers or terminating their contract.

Collateral damage does not always serve as a constructive selective incentive. Both abusive senders and legitimate senders suffering collateral benefits may see

blocking as an illegitimate revocation of services by the BL as an unrelated, third party actor. These actors see the BL as the sole source of service revocation. Following earlier discussion, revocation is effected by the aggregate of local reputation images created by BL appropriators. That said, delisting, removing abusive actors and those suffering collateral damage from the parameter set of these local images, is under the control of the BL. Despite enforcement of listings by listed actors' peers in the control plane, namely access networks appropriating BL indicators to mitigate abuse externalities, tension focuses on the denial to delist.

In the constructive scenario, the selective incentive results in the LIR eliminating or endogenizing abuse. Three general scenarios can be distilled from the discussion thus far, considering whether delisting may contribute to remediation: *a*) quick remediation by a credible LIR, *b*) frequent recidivism by a satisficing LIR, and *c*) abuse facilitation. The former, quick remediation, is the ideal scenario for dialog. The remediator is known and there is an expectation that dialog will lead to credible action—there is reason to believe the LIR is not satisficing. Scenario (b), intermediation on behalf of the frequent recidivist, requires taking a risk on whether dialog can remediate a potentially low-margin LIR that cannot afford to lose customers. It is unclear at what point an LIR can remediate its broader reputation, whether there is sufficient demand from legitimate senders to recover from the lost revenue for abusive senders. For the professional abusive sender, delisting is a very unlikely option. Rather, for known professional abusive senders, preemptive listing is an attempt to *a*) eliminate possible avenues of number resource appropriation and connectivity by *b*) deterring others from providing operational support for fear of collateral damage.

Delisting from the Spamhaus SBL is an instance of the general trends in delisting described thus far: depending on the existing relationship, delisting may involve combinations of *a*) dialog with known, trusted actors; *b*) validation that abuse has been remediated; *c*) longer investigative procedures in the case of known, professional abusers on nominally legitimate services (Spamhaus, 2014f). The first of these is the easiest case, falling into the category of a credible LIR whose business is largely with legitimate downstream network actors, some of whom may be senders. As noted above, this credible LIR has established a track record for promptly handling actors that engage in abuse. Spamhaus' delisting procedure indicates that

> [w]here we have a proven working relationship with any Internet Service Provider, the SBL team implicitly trusts the Internet Service Provider's Abuse Manager and will normally remove listings on the Abuse Manager's word that the reason for listing has been corrected or terminated. (Spamhaus, 2014f, footnote 1)

This is not a blank cheque to the LIR, though. If evidence of abuse continues, the resource will be relisted and the LIR will lose that credibility.

Consider the process both the credible, known LIR and the credible, yet unknown LIR must consider when remediating abuse. The trust relationship above presumes that the credible LIR has reviewed the SBL and acted on the information available. An SBL listing contains evidence of abuse used by the SBL team to justify

listing. That information is provided in the publicly available SBL listing itself so the responsible LIR can use that evidence to identify the source of abuse, remediate that source, and credibly report to Spamhaus that they have in fact resolved the issue. Typically a good track record means LIRs have well-defined procedures for handling abuse. For instance, the LIR may have a mechanism for automatically transposing SBL listing information into an e-mail along with references to pertinent portions of the LIR AUP, incenting the abusive sender to remediate or face suspension.

As per the previous section, listing on Return Path's Reputation Network Blacklist is based on a predictive model parameterized by data from Return Path's Trusted Cooperative Network. Simply put, Return Path places faith in its model: if a sender corrects their deviations from legitimate sending, the model will recognize this change and the sender will be automatically removed from the BL (Return Path, 2014b). Following the discussion of credible actors (our credible LIR C) with clients on BLs, Return Path (2014b) indicates that an actor can request a temporary removal while the cause of the listing is investigated. A number of BLs offer automatic or what is referred to as temporary delisting. The BLs provide a web interface for requesting a delisting, on the presumption that the abuse has been or will be remediated.

The period of time from the temporary delisting to the point where a resource may be relisted will be referred to as the temporary delisting window. The duration of the temporary delisting window depends on a number of factors: *a*) whether the BL explicitly tracks temporary delistings, *b*) how frequently reputation data informing the BL reputation function is collected and/or processed, *c*) severity of the abuse from the temporarily delisted actor(s). For instance, some BLs set an explicit "time out" on temporary delistings before triggering a re-evaluation. Others, such as Return Path, already have a high sampling rate, using a combination of pre-existing resource reputation (such as previous track record for delisting) and observable indicators to determine if a sender's abusive behavior has ceased, diminished, remained the same, or increased.

The former two cases are signals of abuse remediation. Eliminating abuse certainly does not warrant relisting. Diminished abuse may be sufficient for delisting; it may also foster a dialogue regarding how to further diminish abuse. Abuse remaining the same indicates either a failure to remediate, but not necessarily why. In the naïve operational case, a credible LIR may have encountered a problem it does not have experience with. For an LIR with a positive record for remediation, the diplomatic approach of the BL is to contact the listed actor to determine that status of abuse remediation.

In terms of MAPS' listing procedure discussed earlier, this is akin to contact before initial listing. The final scenario is the case of an actor that does not attempt to remediate or is gaming the temporary delisting window to avoid diminished revenue. Depending on the BL, it may try to establish a dialogue before relisting or may simply relist the resource. Once relisted, the steward of the resource is once again faced with diminished value as a result of listing.

Given these scenarios, the temporary removal strategy is best applied by the known, credible LIR with well defined processes for remediating the wayward

client. These actors have well defined processes for remediation and, based on their experience remediating others, are not risking their broader reputation by requesting the temporary removal. In the ideal case of the credible, prompt LIR, the problem is quickly resolved and the resource is not relisted—remediation "works" and ongoing monitoring by Return Path does not identify further abusive activity. This is a testament to both the remediation processes of the credible LIR and Return Path's monitoring processes and knowledge of the legitimate sending parameter space.

Temporary removal from the BL, in this case Return Path, may have a negative impact if the requesting agent cannot remediate abusive behavior promptly. Return Path indicates that, in the case of a neutral resource,[558] one that is either not in a block with a strong reputation for prompt remediation and is not in a block with a reputation for recidivism, this actor should remediate *before* using the delisting service. This avoids the risk of further damaging the listed actor's reputation by failure to remediate within the temporary delisting window. This is the recommended strategy for Return Path's blocking list.

Pollard (2013) argues this strategy for Spamhaus listings. In general, those with neutral resource valuation are unlikely to have substantive experience remediating abuse. As such, on their own, absent the guidance of an experienced LIR, even a well-intentioned, legitimate actor may fall outside the temporary delisting window Relisting when an actor is in the course of credible remediation creates (or, for some, reinforces) the perception that BLs are aggressive and unwilling to coordinate with good-faith remediation efforts.

7.3.3 BL Service Provision

Blocking list feeds may be implemented with a variety of data streaming or database query tools. A common implementation is the use of name servers. These are referred to as DNSBLs. In particular, blocking lists are implemented in the style of WHOIS servers. DNSBLs have similar query parameters to name servers: resources such as number ranges or domains. DNSBLs also benefit form the resilience of Anycast deployment.

For simple individual queries, users can point their favorite `whois` client at a BL name server and query a particular IP address or range for reputation indicators. The following describes the name server model in terms of BLs provided by Spamhaus and the application of Anycast to make BL reputation indicator distribution durable in the face of attacks by abusive actors.

In the RIR model, WHOIS servers map IP addresses to information about the organizations to whom those addresses are delegated. In the BL model, BL servers map IP addresses to reputation indicators. Recall there are multiple types of blocking lists, for instance those that provide abuse indicators and DULs offered by stewards of dynamic blocks of end users. In many cases, a BL provider maintains a

[558] Here a neutral resource is the term use in this work, not a term used by Return Path. The notion of a neutral resource refers to $v(r, s)$ such that $s^{good} < s < s^{a}$.

number of specific types of DNSBLs. For instance, MAPS maintained their primary blocklist, the RBL as well as the DUL (dynamic user list), the RSS (relay spam stopper, listing open relays), the OPS (open proxy stopper, listing open proxies), and the NML (non-confirmed mail list for mailing lists that did not confirm recipients on the list).

As an illustration of the variety of classes of blocking lists, Spamhaus currently maintains:

SBL "Spamhaus Block List Advisory is a database of IP addresses from which Spamhaus does not recommend the acceptance of electronic mail." (Spamhaus, 2014g)

XBL "Spamhaus Exploits Block List is a realtime database of IP addresses of hijacked PCs infected by illegal 3rd party exploits, including open proxies (HTTP, socks, AnalogX, wingate, etc.), worms/viruses with built-in spam engines, and other types of trojan-horse exploits." (Spamhaus, 2014c)

PBL Policy Block List "is a DNSBL database of end-user IP address ranges which should not be delivering unauthenticated SMTP email to any Internet mail server except those provided for specifically by an ISP for that customer's use. The PBL helps networks enforce their Acceptable Use Policy for dynamic and non-MTA customer IP ranges." (Spamhaus, 2014d)

DBL "Spamhaus [Domain Block List] is a realtime database of domains (typically web site domains) found in spam messages. Mail server software capable of scanning e-mail messages body contents for URIs can use the DBL to identify, classify or reject spam containing DBL-listed domains." (Spamhaus, 2014b)

DROP "DROP (Don't Route Or Peer) and EDROP are advisory 'drop all traffic' lists, consisting of netblocks that are 'hijacked' or leased by professional spam or cyber-crime operations (used for dissemination of malware, trojan downloaders, botnet controllers). The DROP and EDROP lists are a tiny subset of the SBL, designed for use by firewalls and routing equipment to filter out the malicious traffic from these netblocks." (Spamhaus, 2014i)

ZEN The Spamhaus Zen list combines the data from the SBL, PBL, and (E)DROP lists into a single BL. (Spamhaus, 2014j)

Different lists facilitate mixing and matching of the indicators. There are also some distinguished subsets. For instance the Botnet Controller List (BCL) "is a specialized subset of the [SBL] ... used by cybercriminals to control infected computers (bots)," (Spamhaus, 2014h).

Another specialized Spamhaus list is the BGPf list. BGPf is a feed comprising the BCL, DROP, and EDROP, distributed via BGP. The BGPf FAQ indicates it differs from the SBL, XBL, and PBL in that those are largely for use with SMTP filters.[559] In

[559]Spamhaus also provides `rsync` based text feeds of all of the lists so appropriators can use the data sets in custom applications, not just those that have built in compatibility with Spamhaus' services.

terms of downstream rights, those lists advise limiting or denying the sending rights, rights bundle s in previous examples, of networks listed. The BGPf is used to filter some, if not all of the traffic from listed resource ranges. BGPf lists actors that are hosting malware that may leverage a variety of attack vectors across the spectrum of traffic types. For ranges that are wholly controlled by malicious actors, wholesale blocking, may be warranted. In terms of resource rights, when BGPf appropriator n uses the BGPf feed for wholesale blocking, it denies each resource holder on that list all downstream rights with respect to the n's hosts. From n's perspective, BGPf is a mechanism for protecting its network from malicious attackers. Not surprisingly, attackers see this as a denial of targets. Some actors classified as abusive have gone as far to argue that any blacklisting is denial of rights to free speech.

7.4 Anti-Abuse Rules

A number of actors in other communities, especially those listed on BLs, have claimed the anti-abuse community is a group of vigilantes pursuing their own interests. When initially faced with the dynamic character of the legitimate sending space (developed in Section 7.1), the lack of bright white lines can unfortunately contribute to that perception. That said, as described in terms of constituencies and structure in the previous sections, the anti-abuse community *does* *a*) have well-defined norms for what constitutes abuse, *b*) document standards for how abuse indicators are *disseminated*, *c*) contribute to value network dynamics for credibly demonstrating legitimacy, *d*) have feedback loops for signaling trends amongst particularistic constituencies without fomenting gaming of "bright white line" thresholds.

The foundations of consent as the basis for anti-abuse norms and practices was established in Section 7.1. Equally important is the ethos that these rules are not applied as binding by some contractual relationship with one or more principals such as in the RIR and IX systems. Rather, these norms rely on mechanisms reinforcing a voluntary, yet credible, commitment to legitimate sending practices. The following sections develop a discussion of the constitutional norms of the anti-abuse community (7.4.1), the variant of the consensus process used to implement collective choice (7.4.2), and a sample of the best common practices that have been produced by that process (7.4.3).

7.4.1 Constitutional Rules

The anti-abuse regime collectively creates a messaging and reputation indicator environment in which abuse externalities is remediated. Unlike the RIRs and IXes, anti-abuse as a whole or as sub-communities does not create facilities such as the registries or IX fabrics. In those cases, constitutional rules bound the management of those facilities. Rather, the anti-abuse regime's constitutional rules are about maintaining the consent-based principles of anti-abuse: what is and is not abuse and how to mediate engagement amongst potentially contentious MVN participants.

7.4.1.1 User Consent as the Basis of Abuse Redux

The definition of abuse, presented in Section 7.1 is rooted in consenting traffic between hosts, typically the users behind those hosts. Simply put, any traffic exchanged between hosts that cannot be traced to a strong expectation of consent, such as a friend passing along an e-mail address, or explicit, well-specified consent by the recipient of that traffic, is abusive. Reconsider the abuse indicators laid out in Section 7.1.3, forming the foundation of the legitimate sending parameter space and form the basis of reputation images discussed in Section 7.2.4.1. Many of these indicators highlight that messages are unwanted (complaint rates), contain malicious content, and may be a drain on the receiving network and its users' resources (sending volume thresholds). Abuse is not only non-consensual, but, as illustrated by the various references back to Figure 7-1, it is also an illegitimate appropriation of receiving network and end users' resources in particular messaging value chains within the MVN.

This notion of abuse is a recurring theme in the best practices of a variety of actors. Reputation managers and BLs reiterate that legitimate, non-abusive sending practices will keep a sender off the BLs and in the good graces of receivers. In terms of resource value, this avoids the kind of value degradation discussed in Section 7.2.4.1. Legitimate professional senders vested in promoting good sending practices argue that non-abusive sending practices improves the per message ROI—endogenizing sending costs reduces list sizes but improves the quality of the remaining list. Legitimate senders argue they experience the benefits of good sending practices in revenue and inbox placement rates. Reviewing whitepapers and recommendations of these legitimate actors, the consent-based notion of abuse, as a principle, does seem to have been embraced and promulgated as a constitutional norm. .

While there has been historic tension between the anti-abuse community and the network operators community, these two communities share a common belief that their private networks are operationally sovereign. Sovereign here does not mean above the laws of the jurisdiction in which these networks are incorporated or deployed. Rather, both communities hold strong beliefs that internal operations and resource use is the sole choice of the network owner and operator. It is oft stated in the network operator community, "My network, my rules." As discussed in terms of abuse, any network can choose what traffic they will accept, typically at the border, but in principle at any point in their network. This does not mean that any network will simply discard traffic capriciously or that every network is equally incented to police consent. The scope of this norm, in other words the boundary, has been historically limited to actors at the ends of a route.

Consider an AS path in Figure 2-1, say, `42.123.0.0/16: 25 23 27 I`, used by `AS 24` to send traffic, including e-mail, to hosts in `42.123.0.0/16`. In the current system, concern with consent has historically bound, unsurprisingly, to those whose value proposition is affected: `AS 27` as a receiver that wants to protect its users from abusive messaging and `AS 24` which wants to generate revenue from sending

messages to hosts in 42.123.0.0/16.[560] Intermediate networks do not necessarily have the incentive to consider whether traffic is consensual. From their perspective, especially those whose value proposition is transport or transit, more traffic is more revenue. Moreover, following neutrality norms, transport and transit is "dumb middle network" and does not care about downstream effects, only whether it generates more traffic or not.[561] For instance, consider a bilateral interconnection relation over a direct transport link (for instance, a relationship such as in the top of Figure 6-1). The transport provider presumes any issues regarding types of traffic will be resolved between the endpoints.

IXes' neutrality norm makes the same presumption. Recall the instance of a DDOS between two IX participants from Section 6.4.1.2. The IX will only interceded if that DDOS affects the quality of service on the IX node or the IX platform itself. In other words, the IX will only intercede if the DDOS affects others on the platform. If the IX began interceding based on traffic categories or perceived behaviors, it would be engaging in discriminatory traffic management. This also holds for the bilateral transport relationship.

Transit providers, on the other hand, typically indicate that sending bulk and/or unsolicited commercial messages, often referred to as spam in quotes, is prohibited. Some networks even reference activities that will land an upstream on a BL, for instance, Hurricane Electric's AUP states that:

> No Hurricane customer shall ... [d]o anything that threatens the integrity of Hurricane's network or the utilization thereof by other persons ...
>
> No customer shall do anything that could get any portion of Hurricane's IP space (or address space announced by Hurricane on behalf of Customer) put on blacklists such the SBL (Spamhaus Block List) as maintained by Spamhaus (http://www.spamhaus.org/) or other similar organizations, or perform activities that would cause portions of the Internet to block mail or refuse to route traffic to any portion of Hurricane's IP space (or address space announced by Hurricane on behalf of Customer). (Hurricane Electric, 2014)

This is a reflection of efforts by the community encouraging network access providers to not provide supporting services. That said, even though a network may have anti-abuse or anti-spam messaging in their AUP, it is not a guarantee that actor is strictly enforcing the AUP, if at all.

Enforcement is, in part, a function of how the LIR values its number resources and, in particular, rights related to messaging. As established earlier, endogeniz-

[560]There is an exception for lists such as the BGPf (Spamhaus, 2014a), which presumes all traffic is bad, and will be addressed shortly.

[561]This is under normal conditions of uncongested traffic flow. There are certainly cases, such as a DDOS, when an upstream may be able to deliver traffic but the downstream indicates the bulk is attack traffic, and to filter traffic from the source(s) of the attack. That said, even in this case, the upstream is being asked by the downstream to help protect the rights of the downstream to ensure the integrity of the downstream resource.

ing abuse externalities is costly and requires coordinating amongst a diverse set of actors. The transit provider may have a large subset of downstreams that can provide reputation information, but not the breadth of a reputation manager such as Return Path. Moreover, mixing transit with reputation management would create a conflict of interest between rating other networks' reputation indicators and those of its downstreams.

The transit provider must also balance its own neutrality norms and enforcing anti-abuse norms that require inspecting traffic content. Technically transit providers can monitor messages traversing its network and apply some filtering based on content. One problem is that transit providers, especially intermediaries removed from immediate contractual relations with senders or receivers, may not have the incentives to respect the feedback of senders or receivers. The diverse set of transit providers, without incentives to participate in feedback loops or simply having insufficient information, may have more divergent filtering practices. It is also problematic that intermediaries are not accountable to those sending or receiving. Taken together, intermediaries have neither the incentive nor the information necessary to act as effective filtering agents.

Once traffic reaches its destination, in this case, AS 27, it has reached an actor that has both incentive to enforce (protect resources) and access to local and distributed aggregate reputation information sufficient to make consistent local decisions. For conventional anti-abuse norms concerning specific sending rights, neutrality and end-to-end issues do not seem to conflict with inspection of messaging data. In the case of AS 27 as a mailbox provider, AS 27 is not inspecting the contents of all traffic, just that traffic that reaches an endpoint in its domain, namely the receiving mail server. While consent lies with the users, strictly speaking, an e-mail transmission constitutes three end-to-end connections: sender to local e-mail server, sending e-mail server to receiving e-mail server, and receiving e-mail server to recipient. Filtering at the second hop does not interfere with end-to-end connectivity. Further, DPI is not necessary given the filters applied are typically part of the receiving SMTP server.

Thus far, the scope of consent-based norms of abuse have been largely limited to the edges. Reconsider the BGPf list provided by Spamhaus. Like other BLs, the BGPf list is jointly provisioned based on reputation indicators derived from consent-based anti-abuse norms. That said, when BGPf is appropriated and utilized by an upstream of some set of networks that upstream's filtering will limit available networks. In the case of a transit provider f appropriating and using the BGPf, the range of numbers reachable from f contract its "transit" offering to less than "the rest of the Internet."

Recall that consent-based abuse norms root consent in that actions of end-users. Under this provenance of abuse, indicators are user-centric. Note many of the indicators in Section 7.1.3 are aggregate rates of effects on end-users: delivery rate, bounce rates, complaint rates. In contrast, other modes of traffic can create externalities for either networks or governments, but are not necessarily abuse in the sense developed here. For instance, one way to frame the Pakistan-YouTube conflict is an attempt to legitimately limit what the government considered abusive

traffic. The traffic creates an externality because it creates exogenous costs for Pakistani law enforcement. From the perspective of Pakistani law enforcement, this protects law-abiding Pakistani citizens.

Consider the scenarios under which user consent-based notions of abuse are aligned with state sanction. If the state is considered a genuine agent of their constituency, one that has been delegated the authority to decide what communication will be consensual and what will not be, the notion of consent-based abuse holds. This is the ideal scenario for both a representative government and the scenario in which the government is a perfect agent of the body of users. That said, short of issues like child pornography, the intersection of what users in a particular state consider abusive is likely to be very narrow.

At the other end of the spectrum, an alternative to the government as a perfect agent is anti-abuse practices that privilege user preferences over government preferences. In this case, the focus on the user as the source of consent privileges the directly expressed preferences of the user over the decision making authority delegated to the state. Rather than disagreeing with which actors have been imbued with reputation, in this case, governments wish to appropriate the operational capability, authority, and credibility to imbue reputation as a way to pursue their own interests. In this case, absent coordination with the anti-abuse community, reputation indicators from legitimate anti-abuse community members will remain authoritative independent of state authority. Actors appropriating and using these indicators will be, in the terms of transnationalism, forming interest groups across national boundaries.[562] Further following the transnationalism model, the interests of the anti-abuse community are not necessarily aligned with that of the state. The result is the potential to fulfill the prediction of conventional IR, that any authority that is not state authority is competition.

Fortunately, the existing anti-abuse community does not fall at either of these "ideal form" extremes. Anti-abuse actors have developed relationships with agents of the state. In particular, anti-abuse actors such as Spamhaus actively coordinate with LEAs, in particular providing support in investigations of fraudulent activity. In these situations, notions of what constitutes abuse and illegal activity are aligned. Returning to the notion of a common image of integrity, the anti-abuse community and LEAs share a common image of legitimate behavior. In addition to this image, operational capabilities of the anti-abuse community complement the authority of LEAs.

7.4.1.2 Nonbinding Best Practices

Appropriators in the RIR system and the associational membership IX regime have contractual relationships with facilities managers. These contracts form one basis for binding appropriators, in their roles as resource users, to operational rules. In those contexts, operational rules generally dictate how *a*) facilities manage the stock of resources, such as numbers by the RIRs and interconnection options and

[562] See Nye, Jr. and Keohane (1971) for general discussion of transnationalism.

platform capacity in the IXes;, *b*) the dynamics and limitations on legitimate appropriation; and *c*) how monitoring and enforcement of that facility is performed. Operational rules are created by the constituency of the facility as the aggregate owner of that facility and the resource stock managed thereby.

In contrast, the norms and rules in the anti-abuse community establish non-binding best practices intended to foster legitimate messaging practices. In effect, the messaging industry has *a*) identified harms emerging from unfettered use of message sending and transport structures and *b*) introduced a distributed, non-binding set of operational procedures intended to rationalize and stabilize the message sending ecosystem (market). Tacit in the discussion of constituency dynamics and the legitimate sending parameter space is that these best practices, a subset of which are discussed in detail in Section 7.4.3 are *non-binding*. In many cases, reputation indicators are considered to be *advisories*, such as the labeling of the BLs provisioned by Spamhaus.

Reconsider the dynamic character of the legitimate sending parameter space. Sharing explicit, single-valued thresholds would nominally provide guarantees regarding what behaviors would result in delivery and what behaviors would not. These guarantees could certainly be used by credible senders and reduce the costs of continuously monitoring. Abusers could also make use of this to game the intent of those thresholds. Gaming the intent effectively follows the black letter of the thresholds while skirting the spirit of anti-abuse norms. A particularly pernicious instance is a snowshoe spammer. Simplifying the problem to rate limiting thresholds, publicly published, single-valued thresholds could be used to calculate how to distribute malicious messages to particular domains without tripping those particular alarms.[563]

Further reconsider the operational sovereignty of networks with respect to managing the traffic each accepts, here in particular messaging traffic. These actors may, and often do, set their thresholds and parameter values independently of one another. Imposing a single valued measure would not be appropriate for the diversity of networks' value proposition or the number of receiving hosts. More importantly for consistency though, non-binding norms provide qualitative bounds rooted in the spirit of anti-abuse principles that adapt to changing conditions based on general feedback loops described in Sections 7.2 and 7.3. Part of the dynamic character of the legitimate sending parameter space that makes it robust is a function of the non-binding character that allows continuous adaptation. Intrinsic enforcement is a characteristic of how reputation is bound in the anti-abuse communities is a substitute for the "artificially" binding character of contracts as coordination enforcement functions.

The non-binding character of best practices does not mean that the effects of *how* reputation indicators are used is not binding. Non-binding best practices are in and of themselves advice. As per the discussion of reputation images, when appropriators act on that advice, the attendant reputation binds. When a large number of

[563]Of course, this does not mean they will not land on internal and distributed BLs for other reasons, such as poor list hygiene, content, and other feedback loops.

appropriators act on that advice (or a large receiver) it subsequently affects value that can be garnered. In this sense, M^{3}AAWG and other organizations that create such non-binding best practices are offering a common image of what constitutes legitimate sending practices that sustain the integrity of the messaging system, but do not introduce the rigidity of strictly binding operational rules and the potential to invite locking in rules (capture) by requiring potentially costly re-evaluation in order to adapt. The result of building systems based on this non-binding advise creates what has been referred to here as the legitimate sending parameter space, a key part of which *does* bind reputation to number resources as identifiers. In this sense, the knowledge commons itself is the mechanism for developing a set of reputation indicators. The common normative frame, the constitutional norms describe here, are what knit these together into a set of indicators that follow the spirit of a consent-based notion of abuse.

7.4.1.3 Participation Facilitation in M^{3}AAWG

M^{3}AAWG's investment in facilitation mechanisms is an instance of conflict resolution that has been elevated to the level of a constitutional norm. The implication is that an ethos of comity is diffused through its application in leadership training, ORT facilitator training, and working group facilitation strategies. Effective facilitation of engagement across participant constituencies is a necessary skill in each of the CRIs. Recall that Early on, M^{3}AAWG was faced with dysfunctional, adversarial engagement amongst actors in the community. Specific and intentional investment in facilitation skills is considered an investment in diffuse, continuous conflict resolution.

Reconsider two aspects of the MVN: *a*) the intrinsically adversarial positions of those attributing reputation, those animating that attribution through the appropriation of reputation indicators, those attributed with reputation as a result of the former two processes; and *b*) those same actors engaging in ORT discussions that are intended to engender *constructive* adversarial engagement as a means to elicit diverse expert perspectives. The goal of the facilitator is to transform these nominally adversarial relations into comity amongst participants. When a facilitator runs an ORT or a working group, the process of eliciting information and conflict mitigation is tightly interleaved. For instance, ensuring all actors are heard is a critical function. In some scenarios, the objective is to diffuse adversarial situations that emerge in the course of a discussion. In some cases, interviews have indicated that this can be accomplished by structuring the discussion in order to give both sides an opportunity to present their ideas, reminding the group as a whole that this is the goal. In more severe cases, it has been necessary to let folks get unproductive critiques out of their system, then saying, "Now that we've gotten that out of our way, we can get down to real work."

7.4.2 Collective Choice Rules

Like in other fora, the consensus process observed in M^{3}AAWG is an exercise in navigating a compromise space amongst potentially contentious industry actors. In the case of anti-abuse, navigating the compromise space means accommodating the value-propositions of senders and receivers in the messaging industry. A number of firms that specialize in both populating and navigating the legitimate sending parameter space emerged: *a*) anti-virus and spam filter vendors; *b*) firms providing reputation aggregator services provision and distribute reputation indicators; *c*) reputation monitors specialize in third party tools for keeping track of the various indicators that may be appropriated to navigate the legitimate sending parameter space; *d*) professional sending organizations are technically senders, but *credible* professional senders have invested in commoditizing the complexities of good sending practices, reputation monitoring, and managing a sending infrastructure that monitors and helps enforce good sending practices. M^{3}AAWG's model of eliciting industry challenges, clarifying the issue space, and developing a best common practices document is the vehicle by which these actors both navigate and, perhaps more importantly, *adapt*, the legitimate sending parameter space. Following the notion of a non-binding set of BCPs, the consensus process described here is the means by which indicator interpretations and navigation processes are modified as a means to adapt while hewing to common image of anti-abuse.

As per interviews with M^{3}AAWG leadership, the M^{3}AAWG collective choice rules are a descendant of the IETF consensus process. Like in other CRIs, the collective choice process is a descendant of the IETF process but it is not a carbon copy. Rather, the consensus process has been adapted to the particularities of anti-abuse as an epistemic domain. While participant facilitation was discussed as a constitutional norm in Section 7.4.1.3, it is a critical element of the collective choice ethos in the M^{3}AAWG community. In this variant of the consensus process, problem identification receives substantively more focus, both in how the consensus process is structured and as a part of developing the anti-abuse knowledge commons. The following subsections trace the elements of consensus established in Section 3.2.5 as they are implemented in M^{3}AAWG's process of developing best common practices.

7.4.2.1 Problem Identification at the Open Round Tables

M^{3}AAWG's Open Round Table (ORT) sessions are one possible entry point into the process of developing best practices.[564] The objective of the ORT is to identify salient challenges faced by one or more of M^{3}AAWG's constituencies and convene M^{3}AAWG participants into constructive dialogue exploring those challenges. There are three general steps in problem identification and development: determine which problems will be discussed at a given meeting, select six of those for group discussions, and perform the actual discussions. The first step is a combina-

[564]Other entry points are discussed in Section 7.4.2.2. The ORT is distinguished here because it is a process for problem identification. The other entry points presume a problem has been identified and documented.

tion of previous meeting feedback and eliciting suggestions on an e-mail list before an upcoming M^{3}AAWG meeting. A combination of suggestions and ongoing topics are placed on a list. At the beginning of the ORT session the group selects from the topics lists.

The ORT sessions span two days. In terms of structure and organization, the ORT session for each day is broken into three distinct time slots. All six topics run concurrently; each topic convenes a small working group in each of the three (sequential) time slots. Each of the three "time slots" are approximately 30-45 minutes. Once the topics have been selected, participants select three different topics, one for each time slot. Alternately, if a participant is particularly interested in one topic, they may choose to remain in that topic for two, or even all three, of the time slots. The objective is to allow participants to engage in multiple topics if they so choose to collect input from a diverse set of participants across each of the three time slots.

ORT sessions are managed, or facilitated by, a volunteer M^{3}AAWG member, typically a member with an interest, experience, and/or expertise in the topic area. Members may volunteer to facilitate. The group or M^{3}AAWG leadership may volunteer known effective facilitators if no one steps forward for a topic that has substantive demand. While many members of the community have participated in the facilitator training program, as with any group, there are a number of actors that are known effective facilitators and that are known to be credibly committed to the ORT process in general.

The role of the facilitator in the ORT is to lend their expertise, but also to ensure all perspectives are heard, to order the discussion, and to serve as an active scribe (on a flip chart). Part of the facilitation process is to use the flip chart notes and facilitator skills to structure the discussions and ensure as many of the ideas generated are developed and explored within the allotted time. Depending on the style and expertise of the facilitator, as well as the topic-expertise of the participants, the facilitator may develop a baseline discussion in the first time slot, then build on the first to push the second further, and again for the third.

The two OTR sessions have distinct functions: the first is idea generation, the second is critique and evaluation for subsequent work. The first session is intended to explore the problem space. This is driven by a discussion that is intended to clarify the problem itself, characterizing scenarios under which the problem occurs, what factors contribute to the problem, what trends in the problem have been observed, and, perhaps most importantly potential solutions to the problem. Following the general rules of brainstorming, the objective is to get the full diversity of characterizations and ideas on page (flip chart) without immediately launching in to a critique.

In topic areas comprising a single constituency, say senders, there may be less contention than in a mixed group, say, senders and vendors. A key element of the idea generation phase is to elicit the full diversity of ideas related to the topic, especially conflicting perspectives, but limit the discussion to getting these perspectives on page. A role of the expert facilitator is to help participants explore these ideas by asking questions that delve further into session utterances. In this role, the fa-

cilitator, especially the expert facilitator that may have their own opinions, is to encourage and to elicit elaboration, but avoid slipping into critique. In summary, the first session is about painting as complete a landscape of the topic as possible, flagging areas of conflict for the next day's discussion.

The second day of OTR is dedicated to sifting through the previous days ideas and distilling a potential plan of action for moving forward on the topic. Recall the topic is a challenge facing the community—the second OTR session is an attempt to catalyze action around transforming the ideas offered into action items that contribute to the anti-abuse knowledge commons. In this series of discussions, groups walk through the ideas generated the previous day, critiquing these ideas and playing out different scenarios, criteria, requirements, solutions, and postulated outcomes against one another. These are certainly not problems that can be solved in two approximately two and a half hour sessions. Many of these topics will require substantively more work to produce even a presentation for general consumption, much less a document that may be sanctioned by M^{3}AAWG as a BCP. The intent is to make a first pass at the ideas that seem to have merit, even a collection of contending ideas. Part of the critique of these ideas is an evaluation of what would be necessary going forward to further develop the topic.

While part of the second session is about structuring the ideas, as may be implied above, another objective is to catalyze interest in carrying the topic forward for further development. For instance, second session actors may identify compatriots in other firms that they would like to collaborate with to develop the topic further. These actors may work for different firms but have similar problems. These actors may be in different constituencies, but recognize a potential to work together as representatives of their constituencies and those constituencies' value propositions. In either case, the ORT sessions are one way for potential collaborators to meet one another and establish a productive relationship. It is also an opportunity to better understand the challenges that will be involved in a commitment to further work.

Taking these together, at the end of the OTR sessions, a number of participants in a topic area have contributed to ideas and a potential plan forward on a given topic. Recalling the notion of prestige in (operational) epistemic communities, participants in the OTR may see an opportunity for prestigious contribution to the anti-abuse community. Moreover, documentation of the OTR is geared to sustain this. OTR facilitators are required to fill out forms describing the progress of the session, potential steps forward, and to elicit champions for future deliverables. At the end of the OTR sessions, facilitators provide quick summaries of the topic, key ideas discussed, and plans for subsequent development.

7.4.2.2 Drafting Best Practices

One of M^{3}AAWG's primary internal and external deliverables is the development of best practices documents. Best practices are one codification of the anti-abuse knowledge commons generated and sustained by the anti-abuse community. The process of drafting best practices serves a number of purposes. The first, and obvious, is the mechanics of structuring and documenting best practices developed in

the M³AAWG community. Mechanics include common document structure, style, editing.[565]

The second element of the BCP development process is to document the various entry points into the that process. As an model of developing consensus, the ORT sessions discussed in the previous section (7.4.2.1) serve as one entry point into the BCP development process. Other entry points include submitting proposals, via the Deliverable Proposal Form, to committee chairs. As per M³AAWG (2014c, p. 3), "[a] document can be proposed from any area within M³AAWG, i.e., from a Member, a Committee, through a M³AAWG Round Table discussion, or suggested by the M³AAWG Board or a Senior Technical Advisor." This is a formal process that is intended to transmit a problem statement for evaluation by the appropriate committee or directly to the board.

The formal process is documented in the Deliverable Proposal Form (DPF) that is filled out by document Champions. Champions[566] "help coordinate and drive the overall document process from concept to publication," (M³AAWG, 2014c, p. 1). A DPF requires two Champions that, along with potential Writers and Contributors, develop (and document in the DPF) the scope, purpose, and audience of the document (2014c, p. 1). Writers coordinate with the Champions to contribute the actual writing of the content, either for the entire document or specific sections; they may or may not be Champions themselves (M³AAWG, 2014c, p. 1). Contributors are M³AAWG members that offer advise, information, commentary, or critiques of the content, but have not committed to writing content themselves. (2014c, p. 3) These roles are documented in the DPF. Submitting a DPF facilitates two functions: evaluation of the problem itself and credible assessment of commitment to the proposed work. In the case of any DPF, the committee chairs, potentially in consultation with Senior Technical Advisors, evaluate the problem and the proposal for developing a BCP.

Once the DPF has been submitted and approved, the actual process begins. In terms of the general consensus process, the M³AAWG BCP development process does not have an explicit active and passive phase. Recall from Section 5.6.2 that the RIR PDP processes have distinct, time-bounded phases of discussion that map to active and passive consensus.[567] The BCP development process does not have explicitly time-bounded phases. That said, it does require monthly progress reports from the Champions to the M³AAWG Program Manager and Committee Chairs.

Returning to active and passive consensus, the BCP development process does have elements of these phases. Resource policies in the RIRs often go through multi-

[565]These mechanical guidelines will not be elaborated here, but it is noted they are documented in (M³AAWG, 2014c) and available to M³AAWG members.

[566]The development process describes a number of distinguished roles: Champions, Writers, Editors, Senior Technical Advisors (STAs), and Contributors. These roles are left capitalized in this section to signify they are distinguished roles in the development process and should be considered as such. Each will be defined as used.

[567]Further recall that, at the discretion of the policy shepherds, a phase of the development process can be extended to make time for changes elicited in the most recent phase or may be moved to a previous phase for rework.

ple passes through the active consensus phase, but the ideal-form is a single pass.[568] That said, as with any consensus process, depending on the complexity of the topic and issues elicited during feedback, the process may require a number of iterations. The consensus development phase comprises multiple cycles of contribution and committee review, writing, and editing. Each of these is discussed in turn, along with mechanisms documented to manage the process and keep it on track.

Contributions, by formally defined Contributors, as well as participants in the managing committee and interested M[3]AAWG members, can occur through a number of the mechanisms: face-to-face meetings, e-mail, teleconferences, essentially any mechanism for coordinate document management. In the BCP development process, contributions occur during committee sessions at the M[3]AAWG meetings. Early in the development process, the document may be presented by the Champions soliciting Writers and Contributors and early contributions as part of a larger agenda for that session. As content is developed by the Champions and Writers, the beginnings of a document form. Once there is sufficient content and structure, the Champions may request a dedicated session at the next M[3]AAWG meeting for committee based writing. To elicit contributions from the broader member base, a draft of the document and discussion agenda may be distributed on an e-mail list dedicated to this particular BCP development process. In this session, Champions and/or Writers will walk participants through the existing content and structure, eliciting comments and critiques.

Early in the document development process, this may be information elicitation, collecting the experience of the operational epistemic community. For instance, Section 7.4.3.1 summarizes a BCP on Spamtrap operations. In one of the later sessions observed in the development process of this BCP, the session comprised a who's who of members with substantive experience developing, deploying, and monitoring spamtraps. These participants comprised representatives of large mailbox providers, receivers, BL investigators, and reputation manager representatives. In other words, the session comprised actors with a vested interest in spamtrap operations and those that have substantive experience with spamtrap operations. In this particular session, the agenda had distinct issues and points of contention to handle in an already well-developed document. Points focused on soliciting specific experiences, providing counterpoints to a particular deployment strategy, or discussing what types of spam might be observed with a particular strategy.[569]

This instance illustrates how the BCP consensus development process comprises elements of both active and passive consensus processes. It is active in the sense that some new information is still being elicited from the group and potentially debated as it is introduced. This elicitation is akin to the active modification of a policy in the RIR PDP process or a draft in the IETF. It is passive in the sense that

[568]This is a point of interpretation: multiple active in a row could be one continuous.

[569]As an important aside, descriptions of observed sessions cannot go into detail on particular strategies, issues discussed, or who discussed them. A cardinal rule of these closed fora is to ensure a free flow of information, with the assurance that various views uttered in the course of discussion will not be repeated outside the M[3]AAWG community. In order to serve first as a participant observer and later an academic member in this community, the author agreed to these rules.

elements of the process that have reached consensus earlier may have been written up based on contribution from the last face-to-face session or the last teleconference comprising Champions, Writers, and Contributors. When distributed to the session, participants have the opportunity to review these newly written sections to ensure they accurately reflect their contributions. Like passive consensus in the RIRs, this is also an opportunity for participants to contribute to the current state of the BCP they have not yet reviewed.

These elements of active and passive consensus are interleaved at a finer-grained level than in the RIR process. That said, there are distinct decision points that help navigate which portions of the BCP are still in contention and which have achieved active consensus. While the differentiated constituencies discussed in Section 7.2 have learned to coordinate in the M^{3}AAWG as a common arena, contention over issues such as opt-in, block listing, and other mechanisms that affect the reputation and revenue of actors in the messaging industry remain. The general mode of operation for deciding on the content of a particular BCP is a loose form of consensus. In the course of contributions and discussion, parties often come to a consensus, typically an agreed upon compromise.

Disputes and disagreements do arise, though. In some of these cases, a more formal consensus process is applied first and, failing that, a vote. The BCP development process provides a high-level description of what is referred to as the "Approval" process:

> A Committee or SIG approves documents by consensus under the leadership of the Chairs. Consensus is usually achieved by addressing the comments and the proposed changes received from committee members in meeting sessions and using the committee mailing.
>
> If there is not a clear consensus for the document or a section of the document, then the Chairs can call for a vote during a committee call, at a meeting session or by email. Chairs also can call for a vote as a last check for agreement and consensus. (M^{3}AAWG, 2014c, p. 5)

Like other fora, the Chairs, here the equivalent of policy shepherds, have discretion in deciding consensus. The BCP development process indicates this Approval process can be invoked for "positions, paragraphs or content," (M^{3}AAWG, 2014c, p. 6).

Comparing this process to the RIR process and the IX process, the BCP development process shares commonalities and differences with both. In the RIR PDP, when a member is describing his/her contention with a particular policy, the policy shepherd or chair of the session often asks questions to clarify and verify precisely what portions of the policy the member is taking issue with. The first question is typically a scoping question, narrowing the issue to, say, a particular choice of threshold, as an instance.[570] Often a follow-up question is whether this is the only

[570]For a specific instance, in a recent ARIN policy proposal discussion, thresholds at which needs criteria would not be evaluated for a delegation were the topic of debate. Some members had contention with particular thresholds (a /18 versus a /26) while others objected to the policy as a

point of contention for a given member in that policy. This is often asked when that member is in a minority that is holding up consensus. Such a strategy is intended to narrow in on a potential compromise that will let the remainder of the policy move forward. That said, these are taken into account, the policy is modified, and consensus of the *whole* is re-evaluated.

In the case of BCP development, the finer grained Approval process allows the group to make decisions on individual points of contention as a means to move the process forward. In contrast to the PDP processes, the finer grained Approval process is available as a mechanism to resolve conflicts *in situ*, as they occur, at the discretion of the Chairs. The general mode of loose consensus in the course of developing the BCP document is akin to the general problem identification, discussion, and consensus processes that occur in IX communities. Recall from Section 6.4.2 that many of these communities present a potential operational or constitutional rule change to the community in order to foster discussion, then later invoke a formal consensus or voting process to decide on that change. The BCP can be seen as a variant of that model, potentially punctuated by localized Approval processes necessary to move the discussion forward.

After some number of cycles, the BCP is ready for a complete review before being sent to the M^{3}AAWG Board for approval. The first step is a document level Approval process. As with the Approval process in general, the decision regarding whether the document has reached consensus is at the discretion of the hosting Committee Chairs. Once the Committee Chairs have decided that the document has reached consensus, it will be sent to Executive Director to send on to the Board for review. If the document is not approved by the Committee Chairs, it is returned to the BCP development team members, along with feedback, and a revision plan is put in place.

7.4.2.3 Speaking in a Common Voice

Although a mechanical point, the process of editing and speaking in a common voice warrants some discussion. M^{3}AAWG, like its constituents, is an organization that lives and dies on its reputation. For M^{3}AAWG, part of that reputation is a consistently articulated tone of its message to the outside world. One vector by which that tone is expressed is the BCPs sanctioned by the community and the Board, collectively as M^{3}AAWG. The mechanics of committee writing, editing for M^{3}AAWG style, and review by the Board all have the explicit goal of creating a common voice. A common voice contributes to maintaining this consistency.

In parallel with the discussion of content development documented by M^{3}AAWG (2014c) and described in the previous section, M^{3}AAWG (2014c) also documents the means to coordinating the articulation of that content. While the team of Champions, Writers, and Contributors add content to the document, a combination of a lead Writer and an Editor create consistency. Champions' coordinate conference calls for this team. Within the team, a Champion or a Writer is assigned the role of

whole. The objective of eliciting this where it was not clear was to determine of the major contention was with the existence of a threshold or the actual value.

the Lead Writer. The Lead Writer maintains the draft. When the document is being developed by Committee or on Committee calls, the Lead is the only actor that can modify the document (M^{3}AAWG, 2014c, p. 4).

The role of the Editor is to help organize the document. The only point at which the Editor's role is explicitly invoked is when content development is complete, but before being sent to the Committee Chairs. Depending on the how many draft iterations, the Editor role may be invoked to help organize the current content, highlighting potential points of contention. Once consensus has been achieved, the next step is to shift editing to a common voice. A key point of the common voice is that the document is, like the recommendations in general, advisory. Ensuring the advisory character assures the non-binding character.

7.4.3 Operational Rules

To understand how the knowledge commons around messaging indicator dynamics is codified into operational rules, a sample of M^{3}AAWG BCPs are presented. One theme of this chapter has been navigation of messaging indicators. BCPs serve as *authoritative* recommendations for navigating, sustaining, maintaining, and altering the legitimate sender parameter space. The sample of BCPs summarized here focus on recommendations for operational development of messaging indicators, how to navigate indicators as the bounds of the legitimate sender parameter space, or a combination of both. While being as explicit as possible with this advice, these BCPs must also walk a fine line not to offer mechanisms that facilitate subverting anti-abuse principles. In some cases they offer advice on how to avoid the appearance of attempted subversion or gaming.

Operational rules discussions start with three articulations of messaging indicators dynamics: *a*) spamtrap operations (Section 7.4.3.1), *b*) ESP vetting (Section 7.4.3.2), *c*) feedback mechanisms (Section 7.4.3.3) These describe the operational characteristics of tools for *a*) discriminating between abusive and non-abusive behavior, *b*) operational strategies for deploying and managing these, *c*) dynamics of these tools and indicators that represent tacit knowledge developed in the field, and *d*) pitfalls of these tools and indicators. None of these offer strict thresholds for indicators. Rather, these describe the *range* and *dynamics* of observable types of responses, vantage points from which they are observable, and how experienced practitioners have interpreted them.

The second group of BCPs discussed are referred to as navigational, offering: *a*) overall documentation of good sending practices (Section 7.4.3.4) and *b*) strategies for remediation for getting listed on a BL (Section 7.4.3.5). These BCPs advise senders how to negotiate the landscape (parameter space) created by the anti-abuse community. Note this landscape is shaped in part by following the first three BCPs discussed. The former, sending practices, maps concepts from consent-based anti-abuse norms to messaging indicators and how these serve as guideposts to stay on the "legitimate interior" of the sending space. The latter, remediation strategies, provide advise for those that find themselves on the illegitimate exterior of the parameter space. In particular, the BCP described in Section 7.4.3.5 describes how to

diagnose the sending practices that gave rise to blacklisting, standard practices for remediation as both ex ante risk mitigation advise and ex post advise on remediation when necessary.

7.4.3.1 Spamtrap Design and Operations

M^{3}AAWG (2013) is a canonical distillation of knowledge developed in a particular epistemic sub-domain: spamtrap design, deployment, and operations. Spamtrap operators provide goals, types, dynamics, and pitfalls of spamtrap development and operations. Spamtrap design and operations is perhaps the flagship instance of tacit knowledge: although the concept is rooted in the notion of honeypots, successful spamtrap operations are the product of experience and the exchange of ideas with other experienced spamtrap operators and analysts. From the perspective of the anti-abuse constituencies, in particular BLs and receivers, providing this BCP improves the knowledge commons by increasing the number and sophistication of actors incented to monitor abuse. In terms of numbers, adding additional monitors improves the sample of abusive activities that can be observed. New observers, or contributors, provide additional vantage points, enriching the decision models and the diversity of vantage points contributing to aggregate indicators such as Return Path's Sender Score.

Sophistication of the actors is also a contribution. Like any data analysis project, the quality of the sample is paramount. M^{3}AAWG (2013) provides a common language for describing the types of spamtraps and the kinds of information that can be gleaned from those spamtraps. In academic terms, it is discussion of spamtrap methodology, validity, and operational mechanics. Such documentation helps both experienced and novice spamtrap operators improve the quality of their deployments and the data generated. Aggregators and BL operators consuming this data also benefit from having richer data sets, have to spend less time correcting misconceptions and teaching the "basics," focusing more of their time developing targeted spamtrap deployments.[571]

M^{3}AAWG (2013) starts by linking spamtraps to conventional concepts in security research:

> Computer security researchers have long made use of "honeypots," servers and/or networks designed as traps to detect, deflect or in some way counter and research the abusive use of information systems. To an outsider, a honeypot generally looks like an ordinary service such as a Web server, mail server or network server, but it has additional instrumentation for close monitoring. (2013, p. 1)

Spamtraps are a means to instrument the MVN as a messaging commons.[572] Spamtraps are e-mail addresses, or domains of addresses, that are not currently in use

[571]This parallels the knowledge commons components of organizations like Euro-IX, training programs in RIRs and NOGs in developing regions, and the education series in NANOG.

[572]In the anti-abuse community, the messaging ecosystem typically refers to the actors discussed in Section 7.2 and the attendant resource dynamics. The messaging commons is a synonym, but used to stress that this ecosystem as a whole is fundamentally resource that increasingly relies on

by end users and that are distributed in such a way as to facilitate sampling of abusive messaging behaviors and responses related to address acquisition, opt-in, list hygiene, bounce notifications, and other indicators of how abusive actors engage in the messaging commons. Recall from Section 3.2 that a characteristics of operational epistemic communities managing a common resource[573] is the development of tacit, then explicit, knowledge of how that commons functions. Here, spamtraps are *mechanisms* for leveraging deep understanding of how the MVN operates as a commons to subsequently generate a number of indicators of abusive behavior. Again, this is a particular means of instrumenting the MVN as a commons.

Further following the characteristics of commons, no singular actor has the resources, or access, sufficient to observe all of the points at which rule violations may occur. Commons monitoring mechanisms often require contributions from participants. For instance, recall the discussions of time-based allocation of water in rural irrigation systems: adjacent farmers were often present to monitor the start and completion of their time slot at each transition. In terms of resource rights and incentives, the farmers are incented to monitor the time slot in which they can exercise appropriation rights and, given how the system is structured, it is a relatively low (transaction) cost for them to monitor the appropriation of the actor before them and after them in sequence. In simpler terms, systems such as the irrigation system with "adjacent" appropriators are interdependent in a way that facilitates and incents decentralized, yet uniform monitoring of resource appropriation. Receivers are akin to the farmer observing legitimate, and potentially illegitimate appropriation, in the course of their participation in the common MVN.

Disseminating spamtrap operations knowledge improves the monitoring capability of actors "adjacent" to legitimate and illegitimate resource appropriation is occurring. In the case of farmers, enforcement is immediate. In the case of the spamtrap operators, behavior may be analyzed for a variety of local goals.[574] Indirect enforcement is implemented by application in reputation functions and aggregate reputation functions, the latter via aggregate analysis by reputation monitors and BLs that deploy spamtraps and, in some cases, share spamtrap data. For these latter to make use of the data, common, or standard methodologies are necessary. Samples of these from (M^{3}AAWG, 2013) are described to illustrate.

One of the most immediate contributions by M^{3}AAWG (2013) is the characterization of spamtrap addresses:

Pristine spamtraps "have never existed as a legitimate destination for email but are still receiving email. This occurs due to misspellings of addresses and parsing errors of automatic address harvesters; for example, truncating parts

reputation and the dissemination of messaging indicators informing that reputation, to function effectively, or as intended.

[573]More precisely Chapter 3 is rooted in common pool resources, but as noted there, many of the dynamics of common resource management institutions applicable to broader commons.

[574]From (M^{3}AAWG, 2013, p. 2), these include *a*) refining local spam filters, *b*) creating reputation lists based on various heuristics, *c*) monitoring clients' bulk mail lists, *d*) capturing malicious payloads, *e*) identifying phishing campaigns, *f*) identifying and detecting malicious URLs and domains, and *g*) data leakage.

of email addresses, collecting Message-IDs as addresses, etc."

Seeded spamtraps "have been deliberately 'offered' by listings on websites (especially via tools like 'wpoison'), by deliberately infecting machines with lots of preloaded email addresses, and by other methods."

Repurposed spamtraps were "[o]riginally valid e-mail addresses that are now being used as trap addresses."

Existing spamtraps "use[] instrumentation from production mail servers delivering e-mail to real users."

These definitions offer both a description of the type and provenance of trap addresses, but also indications of what types of information can be gleaned from "hits" on these addresses. For instance, pristine addresses are variants of legitimate addresses that can help identify dictionary attacks[575] and, as noted in the definition itself, "parsing errors of automatic address harvesters," (2013, p. 2). Note that seeded addresses are defined by how they are deployed or had once been deployed, not the character of the construction of the string (address). Repurposed is defined in terms of the provenance, or the context of the address. Later in (M^{3}AAWG, 2013), strategies for using repurposed addresses, such as conditioning for 12-months to allow infrequent legitimate senders to receive and process bounces, are presented to avoid privacy issues.

Given these baseline definitions, the remainder of (M^{3}AAWG, 2013) describes various issues with addresses, nuances of analysis, security issues, elements of information sharing strategies, and finally common hints and pitfalls at effective spamtrap operations. Nuances of analysis, security issues, and hints and pitfalls are largely particularistic technical elements of spamtrap operations. The discussion of issues with addresses and information sharing are more geared toward how spamtrap data is used, implications for these instruments in a live environment (the messaging ecosystem) and the challenges of sharing without compromising the utility of spamtraps as an instrument. For instance, the section entitled "Issues with Addresses" describes some of the dynamics experienced when putting addresses in place: not all received mail is spam. Two instances are described to illustrate: bounces arriving at the trap when the trap address is forged to send spam[576] and

[575]Dictionary attacks create strings that may be useful for a particular attack by creating either random permutations from a particular set of tokens or common variants from a set of existing sequences. Both are common in attacking the passphrases of authentication systems. The latter is used to create addresses that look like users' names, especially to automatic address collectors but are in fact spamtraps.

[576]Any address can be forged to send abusive messaging. Given spamtraps are intended to land on abusive lists, it is not surprising that these may be forged to send abusive messages, especially if the actor sending the abusive message considers that address to be reputable. When this happens, legitimate servers may send bounces that get routed to the spamtrap. A naïve count-based implementation may result false positives (type I errors). The implementer should make this possibility clear to actors using this data.

reconditioning issues.[577] These are less about the spamtraps themselves and more about recognizing common but not necessarily obvious sources of type I errors that operators that have run spamtraps for long periods of time have encountered. While the potential for type I errors is known to anyone with rudimentary statistics knowledge, in the messaging ecosystem, reputation is paramount. Measuring type I errors as if all are homogeneous, with the same impact, is naïve. For instance, a type I error from reconditioning that lands an emergency alert system on a BL, then limiting the effectiveness that system can have, has substantive implications.

Developing an effective image of reputation in the messaging ecosystem requires sharing information. M^{3}AAWG (2013) lists a number of actors spamtrap operators may benefit from sharing data with and that may benefit from that data: *a*) other abuse researchers; *b*) LEAs; *c*) mitigation groups;[578] *d*) commercial senders, especially in the context of "traps generating reputation information"; *e*) service providers (2013, p. 5) This list is immediately followed by an assertion that:

> Important: It is important to keep in mind that there are entities on all sides of the email ecosystem who may wish to subvert your spamtrap data for their own use and gain. Take great care in deciding who you trust to receive your sensitive data, should you decide to share it at all. (M^{3}AAWG, 2013, p. 5)

Tacit in the list above is that these actors are credible, where credible in this context means they, *at a minimum*, follow the general norms espoused by consent-based notions of abuse. Most actors require more specific shared values to share spamtrap data; typically it requires a credible commitment to both anti-abuse norms *and* the promulgation of those norms.[579] The sharing discussion goes on to *a*) encourage NDAs, *b*) discuss issues of potential spamtrap identification, *c*) offer recommendations for removing data to avoid spamtrap identification "is directly counter to RFC 5965, which defines the Abuse Report Format (ARF) and encourages full disclosure"[580] (M^{3}AAWG, 2013, p. 6) and *d*) suggest limitations on use of data (like

[577] As per the definition, repurposed addresses may have been held by a valid user and may still receive intermittent legitimate messages. Reconditioning keeps these addresses from being reissued to end users and sends standard rejection messages (see RFC 5321 (Klensin, 2008)). It is presumed legitimate senders will only need to see such a rejection message once and remove the e-mail from their list. Actors that do not may be seen as abusive. After a period of time, the reconditioning period, it the address may be used as a spamtrap. That said, a number of legitimate classes of messages may be received, for instance emergency alert systems that may only be activated every few years. That said, good e-mail list management practices indicate that opt-in-confirmation messages be sent by both marketing-based lists and these emergency lists to confirm addresses are active.

[578] Mitigation groups are a subset of SISCs described in Section 7.1. Mitigation groups are largely communities focused on sharing information to support immediate mitigation efforts, they do not necessarily have resources dedicated to (or even the resources to dedicate) long term data collection.

[579] This is generally the case for most anti-abuse data that required resource expenditure or for which sharing puts the reputation of the source at risk.

[580] See Shafranovich, Levine, and Kucherawy (2010) for RFC 5965. ARF generally encourages full disclosure of abuse reporting information. Spamtrap data comprises similar evidence of abuse, in particular copies of abusive messages. It differs in that releasing that data without redaction of information identifying the spamtrap can diminish, or completely eliminate, the utility of the trap.

opening URIs).[581]

Data from spamtraps can be quite useful for remediation efforts. In the hands of a credible ESP, this information can be used to identify operational abusers and malicious abusers among its customers that may have slipped past the vetting process (discussed in the next section, 7.4.3.2). M^{3}AAWG (2013) actually references remediation best practices and describes both legitimate (credible) efforts and what at the best could be seen as satisficing, but is more likely the sign of a recidivist:[582]

> With such information, the customer can be identified and appropriate remediation steps can be taken. The best practice is to recommend that the sender identify how the addresses were collected, repair the problem, and/or educate the customer as appropriate so that the problem does not repeat. If the ISP just removes a few offending recipient email addresses, it does not address the root cause of the issue and it does not prevent it from being repeated. (M^{3}AAWG, 2013, p. 6)

In addition to distinguishing the credible from the abusive, this discussion is also further evidence of ensuring strategies align with the general tenets of legitimate sending rather than simple aversion to sanctions (instrumental satisficing). This last portion of this discussion speaks to ESPs responding to potential misbehavior of clients; the next Section discusses strategies for ESPs' vetting process, a means to minimize, or at least become more prepared for, the potential for misplaced investment in instrumental manipulation of indicators rather than credible commitment to anti-abuse norms.

7.4.3.2 ESP Vetting Practices

Part of the services provided by a credible ESP is to provide information to senders that help them make better sending decisions. Providing these services does not guarantee it will be put to effective use. ESPs have a range of credible senders, satisficers, and naïve operationally abusive actors as clients—the differences are degrees of legitimate sender compliance. The range of sender compliance is evidence that these actors' motivations and perceived value propositions differ despite common information. This argument was developed in Section 7.1 in terms of networks' clients and the reputation of the LIR responsible for number resources in use by senders. The range of sender compliance is not a simple or deterministic assessment. Rather, evaluating the potential for sender compliance requires substantive experience to ask the right questions (detailed below) as well as some knowledge

---

Once an address is a known trap, abusive actors will remove from lists and likely broadcast that the address is a trap.

[581]Opening malicious URIs can be damaging. For instance, malware URIs may infect the machine opening that URI or invoke ads whose revenue goes back to the abusive actor. Even some experienced anti-abuse participants have opened these, creating conflicts inadvertently implicating other in abuse.

[582]Assertion of satisficing and recidivism as behaviors is the interpretation of the author, it is not explicit in the cited text of (M^{3}AAWG, 2013).

of the potential sender's previous history. M^{3}AAWG (2011b) is a codification of that experience.

The messaging indicators discussed in most of this chapter have focused on identifying abusive behavior either in the course of that behavior occurring or after. The vetting procedures described in (M^{3}AAWG, 2011b) transpose the espoused criteria into questions ESPs should ask of potential new clients. These can also be used to supplement the monitoring of existing clients. Recall the reputational threats of abusive actors to a credible LIR from Section 7.1. The introduction to (M^{3}AAWG, 2011b) evokes the same sentiment:

> Email Service Providers (ESPs), who send large volumes of email on behalf of their clients, are at the mercy of their worst clients' worst practices. Common problems such as e-appending, poorly run affiliate programs and past data corruption can create delivery and reputational issues not only for an ESPs problem senders, but for all of the ESPs other clients as well.(2011b, p. 1)

ESPs clearly have an incentive to develop vetting practices.

One of the first sections of this BCP is simply and clearly entitled "Why Vet?" The simple answer is that vetting facilitates risk management through client portfolio management.

> When proper pre-send vetting is performed, ESPs can preempt damage caused by bad clients to both recipient domains and to their own sending reputation. (M^{3}AAWG, 2011b, p. 2)

The BCP also indicates that

> a variety of ESPs ... [came] together for healthy conversations about how to vet clients to avoid these issues. After much discussion, it became evident that not all ESPs use the same methods. (M^{3}AAWG, 2011b, p. 1)

From the perspective of the anti-abuse regime complex, common vetting practices create more consistent enforcement processes. Taken together these highlight the benefits sender vetting has for firms and the regime.

The vetting BCP is structured as a series of questions for potential senders, partitioned into pre-sending and post-send vetting techniques. The majority of the BCP is pre-sending questions and what the community considers the characteristics of acceptable responses. In effect, it is a template for interviewing clients. Categories within pre-send include *a*) corporate entity formation and history; *b*) infrastructure and process; *c*) sending history and patterns; and *d*) list, data collection, and management practices (M^{3}AAWG, 2011b, pp. 2–6).

Corporate history questions probe the origins of the company, the reputation of the principals, whether these principals have been involved in firms responsible for abusive behavior, and other indicators of generally disruptable corporate behavior. Evaluating corporate entity formation and history is a point of tension between the

anti-abuse community and the RIRs. RIRs confirm an actor requesting numbers is a legal entity, but are often critiqued for doing little more than a nominal check. Here, ESPs are LIRs whose downstream value-proposition has incented more in-depth initial corporate vetting than other LIRs may engage in. One interpretation is that this is a failure of the RIR system to engage in general vetting procedures. Another interpretation may generalize this to assert downstream incentives, such as the need for sender vetting, will give rise to vetting procedures in other classes of downstream rights that face rights-diminishing externalities. Downstream rights enforcement will be taken up more fully in Section 7.5.

The next step in pre-send vetting is to understand the applicant's existing infrastructure and processes. This is an assessment of how sophisticated a sending infrastructure the sender has developed, if and to whom the actor has outsourced sending infrastructure, and the reputation of that infrastructure attributable to the applicant. For instance, an immediate question is whether an applicant has worked with other ESPs and why they left. This intends to identify abusive actors that treat ESPs as a sources of fungible sending resources.[583] Infrastructure and process questions push this further by suggesting questions regarding previous IP addresses, domains used, and control over DNS records. Through community resources, namely public and private BLs, the ESP can develop a better understanding of the applicant's sender compliance.

In Section 7.1.3 a number of messaging indicators were presented. The discussion of sending history in (M^{3}AAWG, 2011b) is an application messaging indicators—inbox placement, types of messages, frequency, segmenting practices, list sharing, history with BLs—and the implications of these measures. For instance, (M^{3}AAWG, 2011b, p. 4) suggests querying for segmenting criteria, with the following template for evaluating responses:

> Generally speaking, campaigns sent to segmented, targeted lists tend to perform better in most respects than campaigns sent to a large generic list of addresses subscribed from different places. Accurate answers to this question can give the ESP some level of expectation regarding frequency and magnitude of reputation (and concomitant delivery) issues. (2011b, p. 4)

Here, the BCP moves beyond simply proselytizing particular indicators. The BCP explains what those metrics can tell the ESP about the risk they are taking on, potential threats to reputation, while also conveying general best practices such as effective segmenting.

[583] A more sophisticated technique is referred to as waterfalling. From the Spamhaus Glossary:

> A list owner is "waterfalling" when they run the same illicitly obtained address list through a series of ESPs, each time cleaning bounces, complainants and maybe non-respondants, and then hoping to move up to a cleaner ESP with better deliverability. The result still includes spammed addresses but fewer spam complaints to the ESP. (Spamhaus, 2015b)

In effect, fungibility is leveraged as a filter.

The last two points of the sending history inquiry focus on how BL remediation was performed and sending metrics for the last 3 months. (M^{3}AAWG, 2011b) suggests asking if an actor has been listed on a BL and, if so, how was it managed.

> Possible constructive responses to a block listing generally include a review of list acquisition or hygiene procedures and implementation or tightening of sender best practices. Any answer that hints at infrastructure changes to evade a block listing is a potential red flag. (M^{3}AAWG, 2011b, p. 4)

Here the BCP again distinguishes between activities one would attribute to a credible sender and those one would attribute to a satisficer or an actively abusive actor. Requesting historical messaging indicators is offered as a vector for garnering further information about the

> quality of their lists and... the likely frequency or magnitude of potential issues. Poor historical performance, for example, may indicate deficiencies in the manner in which the senders' lists are assembled and maintained. (M^{3}AAWG, 2011b, p. 4)

Again, the suggested analysis supports both anti-abuse strategies and the benefits for the ESPs' risk portfolio. Depending on the perspective (ESP or sending industry writ large), one outcome is a collateral benefit (positive externality) of the other.

Finally, (M^{3}AAWG, 2011b) drills down into how the lists themselves are managed. The questions here focus on opt-in implementation, feedback loops (discussed in the next section, 7.4.3.3), managing unsubscribe requests, bounce indicators, sources of list contents, affiliate marketing strategies. Each of these, again, re-affirms anti-abuse best practices. For instance, the opt-in discussion suggests "[s]enders should create and maintain an auditable trail" regarding consent to participate in an e-mail list. Poor feedback, such as complaints and labeling messages as spam conveyed via feedback loops (FBLs) may be another indicator of poor opt-in (permission, consent) practices. These questions probe the sender regarding how they handle the metrics themselves. At this level of granularity, the queries shift from past behavior to whether these actors are familiar with industry standard messaging indicators, technologies, and how they may be used.

From the perspective of the ESP, these indicators are services the ESP will be providing these same the prospective client. As noted before, the answers provide priors on sending compliance. The answers also indicate of how much work will be necessary to move these senders from poor practices to more compliant. The ideal case is, as described in Section 7.1, a shift from naïve operational abuse to a credible sender. Recall that M^{3}AAWG (2011b) started by indicating ESPs are "at the mercy of their worst clients' worst practices." Understanding one's risk portfolio is an effective confluence of ESP reputation management and incentives to improve the practices of actors engaging in the global messaging commons. This positive feedback makes credible ESPs valuable gateways to good practices and experts in constructive remediation.

Thus far the focus has been on evaluating how actors that are aware of these metrics put them to use, either in support of more general sending practices or to game those metrics. By definition the naïve operational abuser is abusive as a result of ignorance of the norms and local optimization rather than out of malice or an explicit abuse-based value proposition (such as extractive abuse). Many of these questions are just as much about evaluating whether potential clients are aware of these issues as they are about recognizing indicators of abusive behavior. Naïve operational abusers *are* characteristically different from satisficers or credible senders that simply want to contract ESP resources and service to improve their messaging indicators, in particular reputation and inbox placement. As discussed in Section 7.2, they present an opportunity to instill potentially neutral parties with good sending practices while also increasing those parties' return on investment in sending services.

Taken together, sender vetting is a body of recommendations that helps the ESP manage its reputation and negotiate sender compliance. Sender vetting also further promulgates attendant norms and, and ultimately, the anti-abuse regime itself. The vetting practices' question and response pairs described above are illustrative of how collateral benefits are key to driving anti-abuse compliance. Although this section has characterized vetting as part of the discussion of indicator dynamics, it can also be seen as recommendations for how the ESP and its clients can navigate good sending practices. The more general framing of good sending practices is documented the best sending practices document discussed in Section 7.4.3.4. The next section (7.4.3.3) describes feedback loops, FBLs, as source of complaint indicators.

7.4.3.3 Feedback Reporting Mechanisms

A number of e-mail clients have a mechanism, typically a button in the primary interface, for flagging a particular message as spam. When e-mail clients and receivers follow abuse reporting standards, invoking the abuse reporting function creates an abuse report. The abuse reporting function transposes certain elements of the message's content and metadata into a standard, machine-parsable abuse reporting format (ARF). The ARF standard is documented in RFC 6650 (Falk & Kucherawy, 2012).

FeedBack Loops (FBLs) such as implemented using ARF are one source of messaging indicators developed by the community. For instance, across a given receiver, a local complaint rate for a given list can be created. ARFs generated from the report spam button are generated based on the user's evaluation, they can be considered an accurate reflection of their preference set, what is consensual, and subsequently, what is or is not spam.[584] ARF data can be shared or the receiver can share aggre-

[584]There are of course exceptions. Users may mark messages with private information as spam or may mark messages from sender such as emergency services as spam. Both of these are false positives. As such, while user generated ARF reports *can* provide high-fidelity indicators, they do require some insignificant amount of cleanup. As noted in later discussion, this is especially the case if an actor chooses to share this data with other anti-abuse practitioners.

gated metrics such as complaint rates. Reputation aggregators make use of FBLs from different receivers to build a larger, more diverse sample for measuring complaint rates. FBLs are a key resource for managing the messaging ecosystem.

M^{3}AAWG (2014e) provides an overview of the purpose of the ARF and its benefits from the perspective of the receiver. As a recommendation, the objective of the documentation is to provide high level dynamics and references to more detailed technical documentation:

> This document will not attempt to restate the comprehensive information in RFCs 2142, 3912, 5965, 6449, or 6650, nor whats in the Regional Internet Registries policies on abuse points of contact. Rather, it is expected that the reader is, or will become, familiar with those documents in order to best understand the context for the recommendations in this document. (M^{3}AAWG, 2014e, p. 1)

Before describing more of (M^{3}AAWG, 2014e), the assertion above warrants additional discussion. The IETF is largely geared to protocol development. Operational epistemic communities build on IETF and other technical protocols, providing feedback based on industry and operations experience. In the document described here, as well as the other documents in the set of anti-abuse operational rules, the primary focus is to highlight the set of economic and industry relations that give rise to abusive behavior, those that deter abusive behavior, and those that have adopted legitimate sending practices.

In effect, operational rules augment RFC technical specifications with the context necessary to achieve their intended objectives. In the broader set of operational epistemic communities, operational rules highlight the economic and industry factors that impinge on "ideal" applications of protocols and standards. Reconsider the discussion of routing in Chapter 2. Industry practice follows the BGP protocol, but exercise discretion when faced with options afforded by distinctions between "SHOULD"and "MUST" as articulated in RFC 2119 (Bradner, 1997). Like the instance of route flap discussed in Section 2.2.2.1, operational rules in general augment technical specifications with rule structures that facilitate navigating the uncertainties in each of the function specific domain instances of infrastructure resource management. RFC writers are largely protocol designers. In contrast, the architects of operational rules have tacit knowledge of economic and industry dynamics necessary to maintain the broader commons.

M^{3}AAWG (2014e) is an instance of evaluating protocols on page vis-à-vis operational applications; in other words, they explore where and how well protocols *as designed* can adapt to the operating environment. The quote above provides the background technical references for nominal implementation. The remainder of the BCP describes the benefits of feedback loops, which actors should and should not be sending ARF reports, standard delivery channels, and how to receive and process ARF reports. In terms of who should and should not send, (M^{3}AAWG, 2014e) indicates that not all actors have the resources to implement ARF reporting. Further, some messages generated using ARF processes, such as "those focused on harass-

ment, copyright violations, notice of blocks on IPs and domains" are not suitable reports. In fact, ARF reporting of these may be considered abuse in and of itself.[585]

The discussion of where to send ARF reports highlights the necessity of relationships and other standard practices to make effective use of ARF reporting. M^{3}AAWG (2014e) indicates that most ARF reporting occurs based on bilateral agreement, such as "enrollment by one party in the other's feedback loop (FBL), in keeping with RFC 6440," (M^{3}AAWG, 2014e, p. 2). Typically this entails an ARF receiving address that is different from the standard abuse address. M^{3}AAWG (2014e) indicates "[a] separate mailbox for ARF reports both allows full control of traffic into the mailbox, since only known senders are permitted, and enables full automation of the mailbox processing, since all messages to that mailbox should be in ARF," (2014e, p. 2). This encourages implementations that demonstrate a credible commitment to ARF processing and signals this credibility to actors submitting ARF reports. Absent bilateral relationships, M^{3}AAWG (2014e) suggests making use of the abuse contact information available in the RIR's WHOIS registry or from commercial abuse contact databases.

7.4.3.4 Sending Practices as a Roadmap to the Legitimate Parameter Space

This section, on sender best practices, and the next (7.4.3.5), on strategies for getting off of a blocking list, are two lenses into credible sending practices. This section discusses the BCP that covers what has thus far been referred to as legitimate sending practices. In other words, it is a field guide to being a good sender, staying on the interior of the legitimate sending space. Section 7.4.3.5's discussion is also rooted in anti-abuse sending practices, but presents strategies for demonstrating remediation efforts and the implications of trying to game BL's delisting processes. In effect, it is both a field guide for remediation efforts and highlights points where the BLs engage in discretionary graduated response.

Consider the introduction to (M^{3}AAWG, 2011a):

> This BCP creates a greater transparency between senders of bulk mail and the receiving operators. This transparency helps distinguish legitimate mailers from spammers and the BCP also advocate[s] technologies and practices that help to make email a more secure and reliable communication channel. (2011a, p. 2)

Here, transparency refers to making the dynamics of the legitimate sending parameter space more clear without facilitating gaming. Making e-mail more secure and reliable is community vernacular vernacular for the general concept of consent-based integrity developed in Sections 7.1 and 7.4.1. Secure, accountable messaging and reliability depend on the integrity of common facilities protecting the messaging commons from abuse—these include jointly produced facilities such as BLs and knowledge commons such as the BCPs.

[585]This is the interpretation of the author, not directly documented in (M^{3}AAWG, 2014e).

(M^{3}AAWG, 2011a) is framed around two primary issues: *a*) enhancing sender accountability and reputation and *b*) managing delivery errors and list maintenance (2011a, p. 2) The first recommendation is to adopt e-mail authentication mechanisms. M^{3}AAWG (2011a) cites M^{3}AAWG (2008) as a survey of authentication mechanisms. The authentication discussion immediately transitions to number reputation, recommending the use of dedicated addresses. As established in earlier discussions, dedicated addresses are useful for accumulating reputation (both constructive and destructive) and limiting the collateral effects of destructive reputation. These dedicated ranges are expected to have valid reverse lookups, one piece of evidence used in "clearly identify[ing] the brand and Web site of the sender," (M^{3}AAWG, 2011a, p. 3).

Senders are encouraged to use as many authentication standards as possible. M^{3}AAWG is active in sharing information on the operational practices of many of the recommended authentication mechanisms. These mechanisms include SPF (Sender Policy Framework), based on linking IP addresses to domains, and DKIM (Domain Keys Identified Mail) which uses a digital signature that can be cryptographically validated by the recipient. In addition to authentication techniques, ensuring consistent, clear headers is recommended. Maintaining accurate domain registration information, accurate number registration information, consistent "From" name, and consistent identification of brand and campaign are among header consistency recommendations. Simply put, these are recommendations to help ensure what a novice sender may see as innocuous inconsistencies are not interpreted as intentional obfuscation of the sender's identity—each of these recommendations has a counterpart in malicious sending practices.

Even if not actually abusive, expecting receivers to either blindly accept inconsistently identified and/or unauthenticated messages increases their risk. The expectation that these actors will have mechanisms for differentiating innocuous from malicious inconsistencies imposes a cost (an externality). Some of this bulk mail, especially transactional messages such as shipping notifications and e-mail receipts, are valuable to end users. Inconsistency places the receiver in the position of having to sort these, a form of operational externality that does not involve abuse but is an artifact of poor sending practices. Under this framing, good sending practices can be seen as a list of practices that may not seem necessary, but which receivers have identified as operational externalities and potential indicators of abuse externalities. As such, good sending practices such as consistent headers may be seen as endogenizing costs in order to avoid more costly reputational damage resulting from poor sending practices.

Another recommended sending practice is transparent content. This includes consistency in attestations regarding delivery, consistency in how redirects are presented, avoiding sending large images or attachments, and being clear about tracking pixels. These are generally mechanisms for avoiding filters, avoiding annoyed users flagging the message as spam, and being clear that the sender will know when a user views the message. This recommendation is followed by recommending establishing FBLs with ISPs that provide them. In addition to reinforcing the constitutional norm of operational sovereignty through M^{3}AAWG (2011a)'s discussion of

AUPs (Acceptable Use Policies), it also states:

> Senders should actively monitor abuse-related complaints from individuals and ISPs with an understanding that *every ISP has the right* to set their own abuse and complaint thresholds. (M^{3}AAWG, 2011a, p. 6, emphasis added here)

As discussed in the previous section, FBL monitoring is one basis for deriving complaint rates. Senders monitoring FBLs themselves can stay ahead of potentially growing complaint rates, typically by ceasing a campaign with a growing complaint rate before it leads to ISP-level blocking, or worse yet reputation attribution by a reputation aggregator such as a BL. Again, endogenizing the costs of monitoring FBLs and potentially re-evaluating a campaign is arguably less costly than having that campaign blocked. It is almost certainly less costly than the collateral damage of having an address supporting multiple campaigns blocked due to poor practices on one. The cost is even greater for ESPs that may have (shared) sending infrastructure supporting multiple campaigns for multiple clients—as may be obvious, these sending practices are aligned with ESP vetting strategies to help avoid abusive senders to begin with.

The second focus of (M^{3}AAWG, 2011a) is a focus on how manage delivery errors and how these inform list maintenance. The latter, list maintenance, is also referred to list hygiene. One recommendation that bridges accountability and list hygiene is keeping track of list age, refraining from using inactive accounts, and keeping track of clients' list sources. List age speaks to, among other things, opt-in and whether actors have recently reaffirmed their consent. Inactive accounts are frequently converted to traps (repurposed spamtraps, see Section 7.4.3.1). Poor hygiene can lead to BL listing if senders hit these traps. Keeping track of clients' list sources is directed to ESPs and is a broad invocation of vetting practices.

The core of the discussion of messaging errors focuses on conveying operational experience interpreting RFC-standard delivery codes, Delivery Status Notifications (DSNs). The general recommendation is to develop operational capacity to process and evaluate these codes. The recommendations *a*) help differentiate between the causes of connection timeouts and common responses (blocking); *b*) offer advise on retrying during transient failures; *c*) highlight permanent failures that should warrant delisting the address and may indicate violation of the AUP. The discussion concludes with a reiteration that "[d]isruptions in the communication stream are bound to happen; it is the sender's responsibility to keep logs of what they send and what is sent back o them in order to anticipate message disruptions based on good analysis" (M^{3}AAWG, 2011a, p. 8) of messaging indicators. This is reinforced by another assertion of a network's sovereignty, hear a domain's sovereignty, over it's operational decisions:

> Keep in mind that a senderys domain is governed by their rules and the desires of their end users. Final authority based on what is allowed in and what is denied rests with the controller/operator of that domain. (M^{3}AAWG, 2011a, p. 8)

Like ESP vetting as a mechanism to reduce costs by understanding the risk portfolio, good sending practices reduce the risk of blocking and lost revenue. That said, as above, disruptions do happen and networks may find themselves blocked. The next section describes mechanisms for getting delisted.

7.4.3.5 Help, I'm on a Blocking List!

Following the field guide metaphor, (M^{3}AAWG, 2014a) is a step-by-step guide for *a*) determining if sending issues are in fact due to listing, *b*) assessing the impact of the listing, *c*) remediation, and *d*) communicating remediation efforts to the BL. (M^{3}AAWG, 2014a, p. 2) These steps amount to a root cause analysis, a cost-benefit analysis for remediation efforts, and is complemented by recommendations for how to avoid the appearance of gaming or exacerbating the listing. In the discussion of constituencies, the difference between experienced LIRs engaging in remediation and novice LIRs was offered. M^{3}AAWG (2014a) is a confirmation of that characterization. Consider the introduction to (M^{3}AAWG, 2014a):

> Nearly all email systems at some point have delivery issues because their sending IPs or domains are included on a blocklist ... these listings can trigger a panic reaction inside the blocked company. Therefore, understanding the established procedures defining how to triage and respond to the situation is important to ensure a timely and effective resolution. (2014a, p. 2)

This latter point, "established procedures" foreshadows a theme focused on investing in remediation operations as a risk mitigation option that runs parallel to the primary theme of delisting strategies.

Like other BCPs, (M^{3}AAWG, 2014a) is born up by constitutional norms. A tacit norm throughout these BCPs has been the responsibility of the sender to either perform their own indicator monitoring or contract a third party with that capability. In terms of common resource management, monitoring has been presented as an often intrinsic element of the appropriation workflow. Like the irrigation farmer metaphor presented earlier, the points at which monitoring can occur are not distant. In contrast, and following the discussion of endogenizing costs in previous sections, monitoring is more costly than simply observing the timing of irrigation appropriation. Here, monitoring requires operational expertise to collect and evaluate available data. Messaging and reputation indicators are relatively cheap to collect but require operational knowledge to interpret and translate into actionable information. With this expertise, actors can make good on the recommendations in the introduction to (M^{3}AAWG, 2014a): they can leverage early detection and understanding of messaging indicators to develop a "triage procedure" and a "disaster response plan" to reduce the impact of listing.

The first recommendation is to recognize and understand how "blocklists make[] decisions" and how to make "resolution paths clearer" (M^{3}AAWG, 2014a, p. 2). This understanding helps better verify that the sending failures are due to attribution by a blocklist and the justification for that attribution. M^{3}AAWG (2014a)

explains general blocklist categories and functions. A variety of list types exist, and each BL or BL organization has different policies and thresholds for listing. M^{3}AAWG (2014a) indicates that, like the endogenization arguments,

> listees cannot rely on the people running the blocklist to tell them exactly why they are listed. Listees are expected to troubleshoot their listing issue using generally available information without relying on the list's personnel for specific information. The lack of support from list operators makes understanding the how and why behind a listing critical for troubleshooting. (2014a, p. 4)

This does not imply all BLs do not provide information justifying blocking. Rather, it indicates actors should not rely on it as part of a well-developed remediation strategy. That said, many BLs offer elided information that is perceived to provide sufficient information to determine the cause of the listing. As may be obvious, "generally available information" refers to the common messaging indicators discussed throughout this chapter.

The BCP differentiates between internal lists maintained by a receiver and external lists, such as those managed by Spamhaus and Return Path. While there are many different lists, M^{3}AAWG (2014a) indicates that "the most widely used lists tend to have good, consistently applied policies and clear delisting criteria," (M^{3}AAWG, 2014a, p. 4). Conversely, lists that do not have these qualities are used with much less much less frequency.[586] Such poorly run lists pollute the messaging indicator space, creating delivery problems for both senders and receivers. As a result, with some exceptions, these lists are generally eschewed by the community.

Internal, or private, lists are maintained based on the local reputation function of a single receiver and may not be susceptible to the same pressures.

> Internal lists implement policies of the specific receiver. Sometimes, these policies may seem excessive and overly burdensome on senders. Receiving systems are in a position to decide the rules for sending to them, e.g., no receiving domain can be forced to accept email which they do not wish to accept. It should be assumed that the list is providing value to the receiver, although senders may not understand the underlying logic. (M^{3}AAWG, 2014a, p. 4)

The operational sovereignty of the receiver is further affirmed here. Further, while information about criteria is incomplete, contributing to the diverse distribution of messaging indicators for internal lists, M^{3}AAWG (2014a) asserts that one should "assume that the list is providing value" despite absence of a rationale or policy statement. The lack of information further confirms the value of keeping track of indicators rather than reactively responding to listings.

[586] In the course of interviews, actors listed 5 to 6 of the most credible BLs. They also described the characteristics of dysfunctional BLs. The most frequently cited deficiencies were poor updating and lack of delisting criteria. One of the most egregious cases was a list seeking payment for delisting, referred to by at least one interviewee as extortion.

Once the listing has been verified, the next step is to determine the impact. Recall BLs have different listing timeouts. The simple, local objective of a listing impact analysis is to determine whether the cost of the impact warrants the effort (cost) of demonstrating credible remediation to the BL maintainer. In some cases, a block may result from a temporary aberration in sending patterns. For instance, Return Path indicates its BL samples frequently and will recognize both abusive behavior *and* when it stops. Once it stops, the resource is automatically delisted.

If the resource blocked does not create significant impact (cost) and the abuse has verifiably ceased, it may be less costly to simply wait for listing timeout. Consider a more problematic case. An ESP has a sender that is intermittently listed on a BL with a short timeout. Moreover, the frequency of listing is not regular. If this listing affects other ESP clients, i.e. via shared infrastructure, it can be problematic. From the perspective of other clients the ESP has a stochastic failure mode. As such, while the simple view of waiting out a listing may be effective in some one-off scenarios, this situation warrants remediation as a vehicle to root cause analysis. The result is that the ESP is incented to invest in remediation strategies and capabilities that solve immediate problems and help identify problems that could be recurring, causing the seemingly stochastic failure mode.

The recommendations of the impact analysis are a combination of a root cause analysis and a cost analysis. Immediate questions for senders address where the network is listed, how many customers the block is affecting, whether the abuser been disconnected from other networks, whether the listing is an indication of a malware infection, whether there is a botnet command and control node (or nodes) on the network (M^{3}AAWG, 2014a, p. 7). Each of these hones in on the source and scope of the problem. ESPs have additional issues to consider. As noted before, ESPs should evaluate what proportion of their client base's recipients are affected. Conversely, it could be the case that the block only affects a small domain, but one with a strong relationship with sender and receiver, potentially warranting resolution (2014a, p. 7).

Following earlier BCPs, bounce analysis is a critical operational capability. In the case of listing impact analysis, bounce analysis attempts to identify the source of the block and what additional campaigns may be affected. Bounce analysis strategies read like a data analyst's handbook for bounce processing: *a*) identify bounces that have a high incidence rate; *b*) bounces *without* a citation to a specific BL *may be* temporarily correlated with bounces *with* a specific citation, but this *does not* mean they are from the same list; *c*) partitioning the data into domains and particular listings can improve the analysis. The objective is to sort the data in a way that highlights breadth and impact of a listing while avoiding false positives (M^{3}AAWG, 2014a, p. 8).

Bounce analysis can provide substantive direction, in the remediation process, but it does not always provide sufficient information for a sender to successfully remediate the problem. Earlier the BCP admonished relying on BLs providing the data necessary for remediation in lieu of monitoring messaging indicators and retaining good sending records. That admonishment did not mean BLs do not provide data, but rather that BLs are not guaranteed to provide information explaining a listing.

Endogenizing the cost and demonstrating good bounce monitoring is a signal of credible commitment and willingness to develop this kind of remediation capacity. Demonstrated remediation capability builds also reputation for being a good remediator with reputation aggregators. Many BLs do provide data, but not all of them provide the same data, nor do they necessarily make it public. Rather, given each BL uses potentially different data sources and criteria, each may offer different types of data and require different attestations that a credible remediation is in process before sharing that data.

As a codification of operational knowledge, M^{3}AAWG (2014a)'s listing of different BL data access models is the product of senders' and reputation aggregators' experience engaging with BLs in remediation efforts. Like MAPS, some blocklists have notification procedures intended to incent a dialogue on remediation *before* listing. M^{3}AAWG (2014a, p. 8) indicates that these notifications, typically sent to a network's abuse address if other contact information is not available, comprise actionable data. In other cases, such as Spamhaus, a public IP address lookup service (a web form) provides a link to a listing record containing elided information justifying the listing. Spamhaus' DNSBL also has well-documented return codes[587] that provide information why a particular address is listed. In other cases, the listed actor may need to contact the BL, demonstrate they are an actual representative of the listed block, that they have done their due diligence investigating potential reasons for the block on their side, and request additional information necessary for remediation. Here again, good monitoring and due diligence signals credible commitment to anti-abuse norms.

While the community generally eschews BLs with poor policy descriptions and delisting policies, some do exist and may impact delivery. This is often the case for regional blocking lists or lists used in commercial blocking software. M^{3}AAWG (2014a) recommends that the listed actor contact the receiving organization in an attempt to remediate the situation. In some cases, the receiver will make an exception. It is not in the receiver's interest to block legitimate messages to its customers. In other situations, receivers may not consider any exceptions (M^{3}AAWG, 2014a, p. 8).

Remediation is a necessary cost to BL delisting. Taking action to contact the BL if the block is time sensitive is an additional cost that may or may not be warranted. If the impact analysis indicates the cost is warranted, (M^{3}AAWG, 2014a) offers advice on how to engage with the BL. Like the diverse mechanisms for gaining access to BL listing data, the dynamics of engaging with BL operators is also rooted in experience that would be costly to acquire for new actors. Reiterating the asservation to do as detailed a root cause analysis as possible, the BCP indicates that, when engaging the BL personnel:

> If one does not know the exact cause, achieving a resolution can become much more difficult. (M^{3}AAWG, 2014a, p. 9)

The dialogue between the listed actor and the BL is as much about assessing the credibility remediation efforts as it is about sharing listing rationale. Framed as

[587] See (Spamhaus, 2015c).

externalities and endogenizing costs, a demonstrably detailed, yet concise, impact analysis is a signal of willingness to commit resources to credible remediation. This not only provides information the BL can use to help the actor understand what behavior was objectionable, but it also builds credibility with the BL staff itself. The BCP also highlights "one highly ineffective approach for getting delisted from spam trap-based blocklists is to ask for the spam trap address 'in order to remove it form our mailing lists'," (M^{3}AAWG, 2014a, p. 9). Spam traps are indicators of larger sending problems. The BCP indicates that requesting the "trap address may actually prolong the listing because it demonstrate to the operator a reluctance to *resolve the underlying list problem* and lack of understanding in how the blocklist operates," (M^{3}AAWG, 2014a, p. 9, emphasis added here).

Credible mediation signals are an important element of the messaging ecosystem in general. Messaging indicators create the boundaries of the legitimate sending parameter space, but are only correlative indicators of behavior. Recall that the range of sender compliance is distinguished in terms of motivations for abusive behavior and magnitude. Understanding actors' motivations and reputation is the foundation of graduated sanctions. This is the case in the messaging commons and is documented in the broader literature on rule enforcement in other common resource management regimes. In both the specific case and the general case, the discretion intrinsic in graduated sanction is an opportunity to help actors (re)integrate into the normative structures that the community has identified as supporting resource system integrity.

BL operators have such discretionary authority over BL listings. All such enforcement agents with such discretionary authority are faced with problems assessing the credibility of actors deviating from community norms. The case of BL operators is an extreme point because they encounter a broad spectrum of abusive actors ranging the across the entire sending compliance spectrum. Characterizing these requires additional information and history on the actors in question. The broad cost of this is that it is a high barrier for new appropriators in the messaging ecosystem. Trust and established reputation can be mistaken for collusion and cronyism. Communicating credible, observable results and outcomes of remediation efforts is one way to both build and sustain a relationship with individual BLs and enforcement agents in the BL community in general.

Step 4 of (M^{3}AAWG, 2014a) offers a recommendations for communicating remediation outcomes. Given the problems with gaming and satisficing, this step begins by noting that

> The listed party must keep in mind that if action is communicated, yet never actually taken, a transient blocklisting may turn into a permanent one. (M^{3}AAWG, 2014a, p. 10)

Some BLs have e-mail addresses staffed by actors trained to deal with delisting; others provide a web form. The former reaffirms the recommendation to provide a concise summary of remediation details. Web forms are a potential double-edged sword. In some cases, submitting a delisting request gives the submitter, and the resource, the benefit of the doubt and immediately delists the resource. That said,

many of these also monitor such delistings, immediately relist on evidence of unremediated abuse, and make further note that remediation signals for this resource are not credible. The result is that the immediate resolution may have longer term impacts on resource reputation, making future delistings more difficult.

These mechanisms can be used to establish a reputation for good remediation practices. This is especially the case for ESPs. An ESP that has invested in good messaging indicator monitoring, in particular *a*) list monitoring, *b*) bounce and reject monitoring, and *c*) feedback loop monitoring can quickly identify the resources, clients, campaigns, and even campaign segments responsible for a listing. Once remediated, the ESP may then use the delisting form. Credible blocklists will also make note of actors that use automatic delisting facilities and for whom additional abuse indicators are not observed during the timeout period.

Legitimate BLs and receivers recognize the range of sender compliance and operational capacity. There is some recognition, albeit tacit, that good sending practices are learned, not necessarily intuitive, especially for marketers coming from telephone and snail mail marketing traditions. Effective remediation and the use of signals such as automatic delisting is a mechanism for signaling that learning process. BCPs such as the ones discussed here on delisting, sender good practices, and ESP vetting all have the tacit objective of improving senders capabilities to comply with anti-abuse norms. Building a reputation for remediation may be seen as an extra cost. Rather, it should be seen as a risk mitigation strategy. In the event that a more persistent problem occurs that requires additional information from a BL, having a record of effective remediation can increase the probability the BL considers you a credible remediator rather than a satisficer or an abusive actor trying to game the system. Improving capabilities does require investment, but if demonstrated through delisting procedures, that investment can signal credible commitment, yielding the longer term benefits of being seen as a good, reliable remediator.

7.5 Anti-Abuse Issues in the NRS

In contrast to the largely endogenous issues faced by the RIR system and the IX regime, two of the three issues considered are exogenous, having to do with its reputation attribution and the implications of the interaction between reputation and a transfers market. The balance of norms and instrumentality has been referenced through the anti-abuse discussion. These threads are tied together in the next section to describe the value of principle-based indicators for maintaining a common authoritative image. This discussion also foreshadows the more general discussion of contingent social order the next chapter, Chapter 8. The discussion of revocation highlights operational conflicts between two common images of authority that are not intrinsically at odds with one another: legitimate route provision and consent based messaging. What is missing is a common decision making process to remediate those tensions, an issue that is taken up in Chapter 8.

7.5.1 Norms and Instrumentality

Reconsider the messaging indicators presented in Section 7.1.3. The discussion of BCP in Section 7.4.3 discussed the provenance and interpretation of these indicators extensively. The first set of BCP's were distinguished as messaging indicator dynamics, i.e. how to develop and interpret messaging and reputation indicators from the perspective of those creating and disseminating those indicators. The second set were navigational, targeting consumers of these indicators. Across the five BCP evaluated, as well as others reviewed for this work, the common theme of consent based messaging and the underlying principles was generally consistent.

For the BCP's reviewed in this work, messaging indicator dynamics can simplified to a few simple design questions. Of the data available, which provide indicators that a particular stream of traffic is abusive? What does this indicator tell receiver about consent? Perhaps the easiest is the complaint rate, a specific indicator that messaging is not consensual. The final question is the most important: how can this indicator, or an aggregate statistic for this indicator, be used to signal credible senders that they may have a problem with their sending practices that may ultimately result in negative reputation attribution?

The key to this final question *is not*, which indicators are the most effective indicators for attributing negative reputation? Instead, the ethos expressed in the BCPs was that an essential component is the role of indicators sending a signal to the sender to re-evaluate its sending practices. In many cases, BCPs offer both interpretations of the indicator in terms of immediate practices and in terms of consent. For instance, certain bounce rates may mean a list has old entries that are no longer live e-mail addresses. The broader implication is that the sender is not validating consent, not keeping track of which users on its list actually engage versus those that do not. In effect, there may be many others for whom these messages are unwanted, but their addresses are live. The recommendation is better list hygiene, or, in consent terms, ensuring those that are on the list still consent, are still interested.

In contrast, the more instrumental approach, often affiliated with abusive actors or those simply satisficing, is to remove the bounced e-mail addresses. In the short term this is a low-cost solution. Moreover, it is part of many of the BCP's recommendations: invest in or develop tools for processing messaging indicators such as feedback loops in order to get the best information necessary to avoid negative reputation attribution. Such an automated system can serve a satisficer or a credible sender. The credible sender will introduce rules that trigger an evaluation of sending practices if indicators pass a certain threshold or deviate from the norm for some statistically significant period of time. List re-evaluation may be costly and depending on the sophistication of the indicator thresholds, may be a false positive. Following the bounce example, the instrumental satisficer may simply remove the bounce and continue sending from the rest of the list. This may even keep the bounce rate below thresholds for negative reputation attribution. That said, the implication of the hygiene interpretation is that there many be more (or many more) on the list for whom these messages are abusive. The credible sender will remove

them, the instrumental satisficer will not.

There are a number of benefits of what will be referred to as the principled indicator approach. Recall from the chapter that credible signals are useful communication mechanisms, for instance graduated sanction as a means to establish a dialog with a naïve operational abuser. One the other end of the spectrum, there is room for error on the part of those provisioning indicators. A principle-based interpretation facilitates a dialog between the producer of the indicator and the consumer of the indicator, starting from a common authoritative image of legitimate sending practices. Under an instrumental approach, absent a common image, the discussion will be much more discordant, potentially devolving to escalation tactics rather than a compromise, or better yet, revised interpretation of the indicator that satisfies both actors' preferences. A number of actors working for reputation monitors have indicated that while the false positive is much more rare in the contemporary anti-abuse regime, when they do occur, a credible reputation monitor can convince a reputation aggregator a particular indicator interpretation is causing harm rather than sending an effective signal.

The second benefit is a longer term benefit. Recall the notion of a common image is a shared understanding of the mechanisms that animate a resource system. The principled indicator approach ensures operational decisions are rooted in the most recent interpretation of that common image. This use reinforces the common image as a logic of appropriateness,[588] a normative framework in which to evaluate phenomena that are known to the system, but also as a framework for evaluating *new* observations. A simple instance is the emergence of abusive behaviors on mobile platforms and instant messaging platforms; many of the indicators and lessons from e-mail abuse were recognized and easily transferrable. Further, the common image facilitated the use of conceptually similar tools in an operationally different environment. As will be developed more generally in Chapter 8, the common image facilitates adaptation amongst nominal competitors.

7.5.2 Reputation Attribution

The role and value of the common image in the previous section highlights how constitutional norms effectively structure bargaining amongst nominal competitors. The following issues are tensions from operational interpretations of common images that are fundamentally not at odds with one another. The next section grounds tension over revocation via reputation attribution in the difference between access and appropriation. Following that discussion, a potential negative interaction between reputation attribution and transfers is discussed.

[588]Logics of consequences and of appropriateness are discussed by Abbott and Snidal (2009, loc. 1573-1578). In a logic of consequences, enforcement imposes a costly sanction. In a logic of appropriateness, enforcement is a signaling mechanism to "communicate the commitment and concern of managers and stakeholder."

7.5.2.1 Revocation

In Section 5.2 RIR number rights revocation was presented. That discussion also foreshadowed this discussion, that the attribution of negative number reputation by the anti-abuse community similarly revoked downstream uses considered "inalienable." Setting the stage for the discussion in Chapter 8, while the common image of legitimate route provision and consent-based messaging are not fundamentally at odds, the reputation mechanism creations operational tensions. The tension is rooted in a subtle difference between access granted to a prefix when a route to that prefix is appropriated versus the right to use that route, to send traffic to hosts enumerated in that prefix.

Fundamentally, abusive messaging is the case when one actor sends unsolicited, non-consensual traffic to another. As noted in the discussion of the MVC, the networks in between, those on what is labeled the intermediate route, do not care either way: they are paid to move anonymous bits. That said, the recipient network does in fact care, it is a cost to deliver that traffic. The provision of a route occurs at the AS level, between network edges. The obligation of "upstreams" between the sending network along the intermediate route is to simply move bits. As ASes "on the interior" of a path, their provision of the route has committed them to carry bits from adjacencies they have provisioned. The recipient network has made no such commitment. As the terminating network, the only actors it is account to are its immediate clients, in this example, e-mail recipients.

In this sense, any route *appropriated* by an actor only confers access to the terminating network. It is just that, the provision of the route, the information that provides some assurance some set of networks have sufficient information to deliver packets destined for that network, *not* an unlimited provision of the receiving networks private (interior) resources. In the very early Internet, where everyone knew everyone else (academic networks), attribution was easy and one could, as in the simple chain letter parable, simply call up the responsible party and have a civil conversation about whether the traffic was consensual or not. As the Internet scaled, this norm has been morphed by *some* (not all) into the perception that the Internet is an open access system: any actor can appropriate the resources of any accessible network. Recipient networks' demand for reputation indicators and the investment in their local reputation functions is evidence that at least this subset of number resource holders believes routes provision access, not unfettered appropriation. Moreover, this investment, in both reputation functions and the social order created by the anti-abuse community, supports the framing of routes as access.

Returning to the perception of revocation, when a network, for instance a less than credible hosting network, is attributed with negative reputation, it perceives that it has been denied some set of its inalienable number resource rights. Under this framing that hosting network is an extractive abuser. Its revenue stream is based in part on that extractive behavior. Further, the hosting provider is likely not monitoring messaging indicators, as described in the previous section. Thus, it may not receive the initial "warning" signals from recipient networks and reputation aggregators. More likely, it may not realize it has been attributed with negative

reputation until reputation aggregators escalate, in one of the worst cases, only after a client leaves because critical downstream uses have been diminished. At this point, the hosting provider perceives the reputation aggregator as impinging on its revenue and the aggregator sees the hosting provider as an unresponsive extractive abuser.

A frequent counter argument by the hosting provider is that it is not responsible for its clients, it is merely providing a service. In effect, the hosting provider is arguing that it is akin to the first hop to the west of the sending network in Figure 7-1. Topologically, that may well be the case, but when the client is either hosting equipment in the hosting provider's facilities or it is single-homed with the hosting provider, those (now diminished) number resources are technically the responsibility of the hosting provider to which they are delegated. Finally, the root of the problem, the difference in the interpretation of fundamental network access and appropriation rights, means that the two contending actors have a fundamental conflict over the common image. As such, they do not benefit from the starting point shared by actors in the previous section. This situation is one factor contributing to the poor relationship between the anti-abuse community an the RIR community, in particular between anti-abuse and the RIPE community.

7.5.2.2 Reputation and Transfers

A key characteristic that makes IPv4 addresses effective identifiers in the anti-abuse space is that, relative to e-mail addresses and domain names, IPv4 addresses are the least fungible of the available identifiers. The notion of lowering the friction in transfer markets runs counter to this quality can have negative consequences for both the anti-abuse community and the RIR communities. Recall the relationship between reputation attribution and number asset value described in terms of binding reputation to indicators in Section 7.1.2 and the effects of reputation on number asset value in Section 7.2.2 and 7.2.2.1. The discussion in Section 7.2.2.1 highlights how abusive actors can "infect" a credible LIR C, potentially diminishing the value of its number assets to the point that LIR C can only attract extractive abusive actors, those that either would not be served by more credible actors. The infection rate becomes even more pronounced under the scenario where unfettered transfers are no longer accurately reflected in the registry.

Consider LIR E from the discussion of credible commitment, the LIR whose number asset value has been diminished by serving abusive actors. In a low friction transfers market, it may be cheaper for LIR E to simply sell-off the rights to its damaged assets to LIR G and purchase the rights for undiminished assets. The positive outcome would be that G may have very specific uses for these assets that are unaffected by use-specific reputation; in this case G just got a really good deal. Further, assuming these specific uses are not themselves abusive, these assets will eventually recover their "neutral reputation" value. The negative, and more likely, outcome is that another abusive actor, LIR B, may purchase E's damaged assets. In this case, those assets will retain their negative reputation, if not garner more.

Playing this dynamic out over multiple cycles, transfers lead to increasing num-

bers of assets being attributed with, and retaining, negative reputation. It is possible LIR E could attract credible senders, but it is unlikely. The current assets delegated to E would likely also garner negative reputation. LIR B may likely transfer its damaged assets elsewhere, again likely to another abusive actors.

To make the scenario even more grim, recall that the selective incentive for maintaining registry accuracy is demand for subsequent addresses. Further recall that IPv4 depletion will result in transfers being the only means to acquire number rights. A combination of abusive actors and little incentive for accurate registry maintenance opens the door to a market of unregistered numbers with little recourse for recovery (revocation) by the RIR. One member of RIR leadership referred to this as the "nightmare" scenario. In this scenario, a potentially growing set of number become highly fungible, low-value assets that circulate primarily amongst composite extractive abusive actors. Given the mechanics sketched here and in this chapter in general, a number of permutations are obvious, but some degree of overall asset value seems inevitable.

Recall the heeds-based criteria for allocation in the RIR system. Historically, a collateral benefit of needs-based evaluation was that it ensured numbers were much less fungible. Reconsider two scenarios for LIR E under historic needs-based delegation and a needs-based criteria for transfers. In both cases, to justify either delegation, LIR E must demonstrate to the RIR that it has a business-based need for more addresses.[589] For E to acquire more numbers, it must justify why its current delegation is insufficient. While E could certainly lie, a simple query to the RIR's favorite set of reputation aggregators could confirm E is requesting more space because it has damaged the existing space. Given the nature of E's business, it would have to fabricate a massive influx of customers that require additional physical asset investments it would have to use to justify a new delegation. Here, needs-based evaluation served to limit infection by abusive actors to existing delegations, not necessarily feeding these actors undiminished number resources.[590] Absent some check on fungibility, the more addresses may be available to abusive actors.

The argument above is not an apologia for needs-based evaluation criteria. Fungibility as it relates to reputation is a constructive collateral benefit, it is not the intended purpose. Needs-based was initially a means to conserve a finite resource. Returning the discussion of registry accuracy in Chapter 5 discussion of RIR participation, the lesson is that while the initial role of needs-based may soon become obsolete, it has a number of accuracy-related collateral benefits that are still necessary. While much of the debate in the RIR community has focused on whether needs-based is applicable, in the language developed here, does it contribute to the social order, the debate should be over how to introduce a mechanism that preserves collateral benefits such as audits and number fungibility. While the latter

[589]Recall a canonical instance of need is an access network's customer growth, requiring additional addresses to assign to new customers.

[590]This is the spirit of needs-based delegation, but may be contested by actors in the anti-abuse community, especially those critical of the lax criteria for *initial*, typically small, needs-based delegations. That said, this argument arguably holds for the scrutiny applied when requesting additional (subsequent) delegations.

has been a discussion amongst RIR leadership, the community seems focused on the former. RPKI is one natural solution, but as developed in Section 5.7.4 and in Part III, RPKI has its own issues to contend with.

Part III

Contingent Social Order in the NRS

Chapter 8

Authority in the NRS

NUMBER RESOURCE SYSTEM participants, like other appropriators of commonly managed resources, have heterogeneous, often divergent resource utilization patterns and preferences. Common resource management regimes mitigate this contention to preserve the integrity of the resource system. In Section 3.2.3 the notion of a common image was introduced in the context of describing operational epistemic communities. Two common images of integrity, routing and consent-based messaging, have been the focus of the discussion. These images are the current foundations of social order in the NRS writ large. Notions of social order and relational order are introduced in the next section (8.1) to supplement descriptions of Ostrom's principles in action with theoretical concepts that stand up *a*) consistency between NRS common images and CRI constitutional norms despite operational tensions and *b*) how the NRS operates as a complex of interdependent relational authorities. This language also sets up the discussion of authority in Chapter 9, in particular the challenge of demonstrating the credibility knowledge assessment derived from relational authority to more conventional formal-legalistic authorities in the global political arena.

Ostrom's eight design principles (1990, pp. 88–101) offer criteria for successful common resource systems. In addition to facilitating comparison amongst common resource systems, these principles can be applied to understand the structure and processes that animate these systems. Applied to the CRIs *and* the NRS as a whole, these principles are used to highlight contention and potential strategies for mitigating that contention. In the explanatory mode, Ostrom's principles provide the building blocks from the common resource paradigm, linking the adaptive capacity necessary in a contingent social order to monitoring, how that monitoring information feeds into rule-making processes, and the role discretionary graduated sanction.

Ostrom's principles describe the integrated set of functions frequently observed in successful common resource systems: *1*) well-defined boundaries, *2*) congruence, *3*) collective choice arrangements, *4*) monitoring, *5*) graduated sanction, *6*) conflict resolution, *7*) rights to organize, and *8*) nested enterprises. The analysis here shows how these principles contribute to the NRS social order, rooted in substantive-purposive authority. In particular, distinguishing this mode of authority

links operational epistemic communities, consensus, and adaptive capacity to the NRS social order. In their evaluative mode, these rules speak to *a*) opportunities for participation and the attendant effects; *b*) the role of sanctions and discretion in applying these sanctions; *c*) and the structure of institutions and organizations in this complex. This latter, the structure of the institutional complex, is especially salient to the NRS. Rather than a hierarchical nesting of institutions and organizations implied by Ostrom's eighth principle, the NRS, and Internet resource management in general, is most effectively characterized as a networked relational authority. Each CRI is framed as a relational authority that creates social order based on one of the two common images that characterize NRS management.

Presenting the NRS as a social order has the potential to invoke notions of competition with order created by the state system. Early on, the cyber-libertarian model of Internet governance[591] eschewed government engagement, arguing its focus was on purely technical aspects of infrastructure management. In the early to mid-1990's, when the Internet was emerging as a nascent economic engine, this was a recipe for durable stability—ICANN was thrust into the global political milieu while the NRS institutions quietly maintained the routing system under the hood. Those days are long gone. State-based institutions are increasingly aware of the importance of the routing infrastructure. The last section of this chapter discusses engagement with external actors, in particular government agents as the NRS "comes further out from underneath the hood."

8.1 Common Images and Relational Authority in the NRS

As will be discussed in the next section (8.2), individual NRS CRIs are stable, but must contend with developing new capabilities to effectively develop coordinated external engagement strategies necessary to navigate the broader global political arena. The character of authority, in particular the role of contingent social order, contributes to evaluating the stability of CRIs and the NRS as a whole. Chapter 9 builds on this nuance to explain the relationship between *a*) type of authority developed by these operational epistemic communities and *b*) the potential for credible knowledge assessment to serve as *durable* political capital. Three common images of social order have been discussed in this work. The first is consensus-based decision making, developed in both Chapter 3 and specific models of consensus in Part II. In Chapter 2, the integrity of the routing system was framed in terms of the conjunction of legitimate route exchanges and efforts to minimize operational externalities such as route flap and security externalities such as prefix hijacking. This image of *routing integrity* ensures that any network has *access*, via legitimately provisioned routes, to any other network enumerated by publicly routable numbers. Legitimate route exchange hinges on respecting the basic number appropriation bundle, in particular prefix origination rights, managed by the RIRs.

[591] See Mayer-Schonberger (2002) for a discussion.

Chapter 7's discussion of anti-abuse norms is rooted in the integrity of end-to-end communication—in effect, given access to paths to all other hosts on the Internet, does an actor create externalities for end hosts? Consent-based anti-abuse norms assert that the only legitimate host-to-host traffic is consensual. Under this framing, integrity means all traffic exchanged is consensual. In the ideal form the recipient must have, at some point prior to traffic exchange, consented to receiving traffic from the sender. The major focus of this image of integrity is the messaging value network (MVN), but it is easily extended to security externalities such as botnet-based DDOS attacks.[592] Routing is a Layer 2 process; the consensual character of host-to-host traffic exchange is a Layer 7+ process. Stated as such, consensual and non-consensual traffic may traverse legitimate or illegitimate paths from host to destination. Logically, these are not inevitably at odds with one another, but, as evidenced by the tension over attribution of reputation, they certainly can be. Later in this discussion these images will be framed as the broad social order provisioned by the NRS.

Contention over an image of system operation is not novel to the NRS; consider Ostrom's discussion of the pumping race in Los Angeles area groundwater basins:

> The solutions to the pumping race, however were not imposed on the participants by external authorities. Rather, the participants used public arenas *to impose constraints on themselves*. (1990, p. 110, emphasis in original)

Many common resource systems in the "conventional" literature,[593] exist within a particular national jurisdiction. That external jurisdiction can serve as a fail-safe for the common resource system. If operational rules disenfranchise a particular class of participants, they may appeal to state-jurisdiction. A failure-mode of common resource management is when a class of participants leverages state-based rules (typically rights) that allow that class of actor to garner more value at an overall cost to other participants and/or at a cost to the integrity of the resource system as a whole. Leveraging state-based rules is a variant of jurisdiction shopping common to international jurisdiction problems.[594] As per above, the ideal scenario is mutually agreed-upon constraints imposed by appropriators, on themselves, in service of a commonly shared image of the integrity of the NRS.

A key premise of these self-imposed constraints is what is referred to as a "common" image of system operation upon which participants first agree, then use as the

[592]In the language of abuse externalities developed in Chapter 7, a botnet-based DDOS attack is a composite extractive externality. Modern manifestations require significant coordination costs. These manifestations are typically economically, politically, and/or ideologically motivated. Although early compromises may have been driven by prestige (which is a value incentive in and of itself), most compromises are driven by the aforementioned motivations.

[593]The conventional literature refers to the perception that common resource institutions only work in small, inconsequential settings: historical Icelandic pasturelands, late 18^{th} and 19^{th} century Japanese forest management, other local resources managed by rural participants dependent on that resource.

[594]Downstream services on the Internet have been plagued by jurisdiction problems, documented by (Zittrain, 2005) and (Kohl, 2007).

basis for developing constraints expected to ensure the integrity of the system. Following the groundwater basin discussion, after eliciting analysis from the Division of Water Resources of the State of California Department of Public Works (1990, p. 111),

> [t]he parties then shared a single, authoritative "image" of the problem they faced. They also would confront a new "default condition" [E. Ostrom (1986)] if they could not agree on their own solution. (1990, p. 112)

Without offering the full analysis,[595] the authoritative image provided sufficient information on the character of the basin for participants to devise a redistribution scheme independent of the "default condition," litigation in which the redistribution outcome was less certain. In effect, the authoritative image was the basis for devising operational rules acceptable to the diverse set of parties involved.

Returning to the NRS, two broad images, routing integrity and consent-based messaging integrity, are the foundations of social order in the NRS. These two broad images speak to resource management outcomes, but not how those outcomes are implemented—these images are the constitutional norms of the NRS. The object of the two common images of the NRS are the resources themselves: routes comprised of number resources and how number resources are attributed with reputation. In each CRI, the constitutional rules and how they structure the operational rules create common images for *facilities management* in service of these two images of resource system integrity.

To make the distinction clear, the object of CRI rules mix *a*) resource management *facilities* participation and *b*) resource management facilities maintenance in service of *c*) broader NRS participation. Consider each of the CRIs. The RIRs' rules govern number delegation and the documentation of these delegations in the registry (facility), providing the identifiers necessary to participate in the Internet routing system. The object of associational membership IXes' rules are almost exclusively about managing and sustaining IX platforms as facilities that create a neutral, non-discriminatory platform for route and traffic exchange in local interconnection markets. Anti-abuse norms sustain the messaging value network by attributing numbers with reputation, with diffuse rules shaping the operation of facilities such as IPBLs. Rather, each has created its own social order whose foundations are articulated in its constitutional norms, adapted via collective choice rules, and implemented through operational rules. These rules (re)distribute unique identifiers, lower barriers to diverse route exchange opportunities, and create self-help enforcement opportunities. Each is a necessary component, but none are individually sufficient to operate a functional routing system.

In contrast to the LA groundwater basin, the NRS institutions do not have an external authority to which to appeal. In developing his notion of relational authority, one which will be applied here to explain authority at the CRI and NRS levels, Lake (2009) draws on Bull (1977):

[595] See Chapter 4 of E. Ostrom (1990).

> [A] social order is "a pattern of human activity that sustains elementary, primary, or universal goals of social life" [(Bull, 1977, p. 5)], including ... *an assurance that property will not be subject to challenges that are constant or without limit*, and an expectation that promises and agreements, once made, will be kept... *Social order possesses large externalities and, in some ways, approaches a public good.* As a result, individuals will typically seek to free ride on the efforts of others, purely voluntary efforts will produce less social order than desired, and the net outcome will be collectively suboptimal (see [Olson (1965)]). (Lake, 2009, p. 334, emphasis added here)

The NRS common images are the foundation of the social order Internet communication rests on. Consider the first point emphasized above, the resource rights (property) protected by each CRI: RIRs document and are increasingly enforcing number rights, IXes provide consistent (neutral, non-compete) access to interconnection options, anti-abuse norms protect receivers' resources from abuse externalities. CRIs provision facilities for sustaining their function-specific component of this social order.

The second emphasized point, social order approaching a public good, is common to discussions of relational and private authority. In their discussion of self-regulatory mechanisms in the global political arena, Mattli and Woods (2009a) refer to the common interest rather than the public interest, providing a foundation for the remaining discussions in (Mattli & Woods, 2009b). The private authorities discussed in (Mattli & Woods, 2009b) create similar large externalities, but cases are scoped to particularistic issues such as instances in the typology of industry self-regulatory mechanisms described by Abbott and Snidal (2009). Abbott and Snidal go on to indicate that "government is no longer the only game in town and may no longer be the most important game in town," (2009, loc. 2000–2002) In many cases these authorities are alternatives to government regulation and may be seen as competing. Governments may paint the NRS's social order with the same brush, seeing the NRS as competing for authority. While the NRS social order does create substantive downstream public, private, and social goods, NRS rules and policy implementing the NRS's social order are largely scoped to common interests rooted in maintaining a non-discriminatory infrastructure. Here and in the next chapter (9, the NRS's social order will be argued as complementary to, rather than competing with, government authority.

Authority in the NRS is not derived from the state, nor does it have the formal-legalistic character of state-based authority. Nor does it rest on the specter of formal-legalistic authority proselytized by "international law." Lake builds on social order and the social contract to argue for a notion of relational authority:

> Social contract theories contain within them an alternative conception of authority in which obligation does not follow from the office of the ruler but from a bargain between ruler and ruled.[] Relational authority is premised on an exchange between ruler and ruled in which A provides a social order of value to B sufficient to offset the loss of freedom

> incurred in his subordination to A, and B confers the right on A to exert the restraints on his behavior necessary to provide that social order. *In equilibrium, a ruler provides just enough social order to gain the compliance of the ruled to the taxes and constraints required to provide the social order. A gets a sufficient return on effort to make the provision of social order worthwhile, and B gets sufficient social order to offset his loss of freedom.* If A extracts too much or provides too little social order, B can withdraw his support and A's authority evaporates. In this way, relational authority is *contingent* on the actions of both the ruler and ruled. (Lake, 2009, p. 334, emphasis added here)

Of course, the NRS does not use the language of "ruler" and "ruled," these are generalizations from a particular political science paradigm's vernacular. Nevertheless, the general notion of a social contract and the provision of social order holds, and arguably has greater explanatory power than principal-agent theories.

Consider each CRI in terms of relational authority. RIRs and IXes hew closely to the equilibrium model of sustaining social order. RIRs provision a registry documenting delegation rights, collect membership fees to fund registry maintenance, and are beholden to members for operational rule development; members benefit from stable, well-documented number rights delegations. In the RIR system, the obligations necessary for routing system function, the basic appropriation bundle of origination and enumeration rights, are documented in a jointly provisioned registry. Registry services facilitate self-help remediations for a variety of operational and security externalities. Perhaps the most clear cut instance of equilibrium is in the associational IX regime. Participants confer management rights to the firm, essentially rights to enforce social order in exchange for a platform facilitating lower barriers to interconnection market access. Unlike the RIRs, which are regional monopolies, the European market offers a variety of IX options; Open-IX is a case of attempting using standards to catalyze such a market in the US.

The "ruler," here more appropriately the management collective,[596] does not need to explicitly provide a set of facilities.

> Alternatively, the ruler can produce order indirectly through what is commonly referred to as leadership. Rulers may not actually defend property rights themselves, for instance, *but they may facilitate the organization of property holders to defend their own rights.* As long as rulers are consequential for resolving such collective action problems, they can still be credited by the ruled with fulfilling their part of the social contract [(Blau, 1964, pp. 213–215)]. (Lake, 2009, p. 335, emphasis added here)

Both the NOGs and M^{3}AAWG serve this "alternative" leadership role. NOGs play this role by acting as a knowledge commons, promulgating operational knowledge.

[596] In the case of RIRs and the IXes, the management collective is in fact a firm that has been delegated a particular bundle of management rights. In contrast, NOGs and the anti-abuse regime are collectives that, while they do create firms, those firms are almost exclusively for the management of finances related to running regular conferences that constitute convening fora.

M^{3}AAWG plays this leadership role more explicitly facilitating engagement between contending constituencies and the development of BCPs. As developed in the analysis of BCPs in Section 7.4.3 and the non-binding character of these operational rules, these rules *facilitate* self-help.

Remaining consequential is particularly salient for the RIRs and M^{3}AAWG. In the case of the RIRs, a critical question is whether the RIRs will have a purpose once the IPv4 free pool is completely depleted and actors only need one IPv6 delegation. As will be discussed in terms of adaptation and congruence, the issues of RPKI and transfer rights are critical in congruence (Section 8.2.2), in essence, whether the RIRs can continue to contribute services of sufficient value to warrant sustaining the *current* social order. The question of whether M^{3}AAWG moving from an anti-abuse conference to a deliverability conference is also a change in social order, here a change regarding for whom the latter will be of consequence to in contrast to the former. In both cases, the CRI is facing a transition, ideally from one stable social order preserving the integrity of the routing system to another.

The context in which these CRIs operate is far from static. Similarly, the operational rules ordering these relational authorities are far from static. Ensuring the social order provisioned remains aligned with resource dynamics, industry structure, and incentives, a form of what Ostrom refers to as congruence (see Section 8.2.2), is an enduring problem in conventional state-based regulation. Consider Lake's discussion of maintaining social order:

> Without the desired social order, the ruled have no reason to subordinate themselves voluntarily to the commands of the ruler, and without the compliance of the ruled, the ruler lacks the *endogenous* means to produce the social order. This equilibrium becomes more robust as members of the community of subordinates are vested in *the existing social order or acquire assets that are themselves specific to the particular order obtained.* (Lake, 2009, p. 335, emphasis added here)

The most clear instance of dependence on an asset is scarce IPv4 addresses. RIR members have historically followed the rules in order to ensure access to additional resource delegations. The IX narrative in Section 6.1.1 and illustrated in Figure 6-1 tells a story of disintermediation, generalizing asset investment within the associational membership IX platform; the notion of stickiness (Section 6.1.2 tells a story of dependence on interconnection option topologies unique (specific) to the IX regime.

In the anti-abuse regime, receivers depend on the messaging indicators provisioned by dedicated reputation monitors and reputation aggregators, both of which are adjuncts to the ideal MVN, products of the anti-abuse regime that facilitate self-help remediation of abuse externalities. In effect, the value of the social order is an incentive[597] to "subordinate themselves voluntarily." Critical to the NRS and

[597]This incentive is different than a selective incentive. Selective incentives describe specific, targeted incentives. Here the incentives, the value of the social order, is at the level of constitutional norms. Selective incentives are the particular incentives found in operational rules such as the various applications of messaging and reputation indicators.

stressed by both Lake and Ostrom is the endogenous character of maintaining what Lake refers to as subordination and what Ostrom refers to as credible capacity for enforcement.

A static set of operational rules that fell out of congruence with industry structure would quickly fail Lake's criteria. Fortunately, social order also includes the constitutional rules and, perhaps more importantly, the secondary collective choice rules. Here, collective choice rules are the endogenous means of adapting.[598] As the character of industry demand changes, collective choice rules serve as an alignment mechanism, ensuring rules are congruent with industry demand, incentives, and structure (again see Section 8.2.2 for a specific discussion of congruence). Rulemaking is a key source of power in relational authorities. Collective choice rulemaking is where CRIs diverge from the distinctions of "rulers" and "ruled" in Lake's articulation.

> In some cases, subordinates may be able to exploit the ruler and extract a disproportionate share of the joint benefits created by order—the proverbial strong power of weak allies [(R. O. Keohane, 1971)]. But more frequently, the ruler will use her ability to set the rules to bias the social order toward her own interests. *The power to write rules has been long recognized as an awesome power and may be one of the most important benefits of ruling.* Indeed, at the extreme, the ruler may skew the rules to such an extent that subordinates are indifferent between remaining in the contract and reverting back to anarchy, although *most rulers likely value future gains sufficiently not to push subordinates toward such fragile, knife-edge equilibria that can be easily perturbed by exogenous shocks.* (Lake, 2009, p. 336, emphasis added here)

In Lake's articulation, the value of future gains is sufficient to "tie the hands of" rulers. In the CRIs, a combination of diverse, participant-elected boards and collective choice processes animated by the "ruled" serve to keep the firms, the "rulers," in check. Each of the CRIs employs a collective choice process congruent with the function-specific management facilities maintained by that CRI. Recall the distinction between the RIRs and the IXes: RIR collective choice operates on operational rules, IX collective choice operates on constitutional rules.

The ability for an operational epistemic community to, on its own endogenous inertia, continue to adapt to avoid the "anarchic state of affairs," is a key capability of CRIs; this capability, credible knowledge assessment as a mode of regulatory adaptation, is a source of political capital discussed in depth in Chapter 9. Following the quote above, the second emphasized point speaks to a common, counterintuitive failure mode in "adaptive" technical systems: overfitting. Consider the uncertainty in common resource management introduced in Section 3.2.1, the object

[598] Historically these have been endogenous. Recall barriers to participation rooted in technical knowledge. That said, there are instances of exogenous influences. IX and anti-abuse collective choice processes are limited to CRI members. In the RIR communities, any actor can participate in the policy development process. Recent exogenous actors have been LEAs and IPv4 brokers.

of knowledge commons discussed in Chapter 4, and the motivation for consensus-based adaptation in both collective choice discussions and the CRI issues comprising the last sections of Chapters 5–7. Optimization models impute precisely bounded operational conditions. Optimization metaphors rooted exclusively in static efficiencies are not good fit for either policy prescriptions in the CRIs or as explicans. Rather, the range of options afforded by constitutional and operational rules, explored via collective choice rules, provides the slack necessary for adapting monitoring mechanisms in real time and the discretion to apply graduated sanction (see Section 8.2.3).

Explaining NRS CRIs as relational authorities highlights the contingent character of authority in these communities. The contingent character of authority also protects relational authority by incenting congruence, ideally reducing the need to seek exogenous authority. More importantly for the remaining discussion, predominantly in the next chapter, relational authority is a contrast to conventional framings of authority. Lake contrasts relational authority with more conventional notions of formal-legalistic models. Lake draws on Flathman (1980)'s distinction between an actor *as* an authority versus an actor *in* authority. Flathman builds this distinction based on two modes of authority. The more common of the two modes, formal-legalistic, confers authority to a role. Authority does not necessarily inhere in the actor conferred that role.

Like principal-agent models, under the formal-legalistic model, authority can be conferred onto any actor and, a more important distinction for this work, that authority can be rescinded by the conferring principal. In the extreme, under a principal agent model authority is role that is easily conferrable and rescindable—authority is fungible under the formal-legalistic model. Even though such fungibility is clearly not the case, it severely limits the explanatory power for the institutions considered here. Flathman's notion of an actor *as* an authority an instance of substantive-purposive authority. Under this model, authority inheres in the actor as a consequence of that actor's unique capabilities for provisioning social order. As a conceptual foundation of relational authority, and in the instances of CRIs framed as relational authorities, the "ruler's" capability to create, maintain, and, most importantly, adapt, a social order is the source of that actor's legitimacy. In the next chapter, this is framed as not only a source of legitimacy, but a regulatory capability is held by CRIs and will be argued as the political capital on which non-subordinate comity with conventional state-based agents rests.

8.2 Evaluating CPR Design Principles

Ostrom provides eight design principles for CPRs, derived from empirical studies.[599] These principles distill the trends identified in Ostrom's narratives into inter-related principles that not only describe functioning CPRs, but also provider the basis for evaluating these. For each resource management regime, the discussion of con-

[599]For a summary, see Table 3.1 in (E. Ostrom, 1990) and pp. 91–101 for elaboration.

stituencies highlights both participants in that regime and federating agents that facilitate information sharing within that subset of the operational epistemic community.[600] While functionally distinct, these CRIs are tightly interdependent. An analysis through the lens of Ostrom's integrated principles facilitates an analysis of the NRS as a whole.[601]

Ostrom argues that "[a]ppropriation, provision, monitoring, enforcement, conflict resolution, and governance activities are organized in multiple layers of nested enterprises," (E. Ostrom, 1990, p. 101). Ostrom goes on to indicate that "[e]stablishing rules at one level, without rules at the other levels, will produce an incomplete system that may not endure over the long run," (E. Ostrom, 1990, p. 102).

8.2.1 Clearly Defined Boundaries and Collective Choice

A common trope in Internet governance is that the Internet governance and architecture is "open." Not only is this inaccurate from the perspective of resource management, but it is misleading with regard to how resources are provisioned, appropriated, and maintained. Well defined boundaries are the key differentiator between common property and open access resource systems. Boundaries are about effectively *delineating* access and exclusion, not determining who should have access rights or who should be excluded.

> So long as the boundaries of the resource and/or the specification of individuals who can use the resource remain uncertain, *no one knows what is being managed or for whom.* Without defining the boundaries of the CPR and closing it to "outsiders," local appropriators face the risk that any benefits they produce by their efforts will be reaped by others who have not contributed to those efforts. At the least, those who invest in the CPR may not receive as high a return as they expected. At the worst, the actions of others could destroy the resource itself. (E. Ostrom, 1990, p. 91)

A number of NRS issues are fundamentally boundary issues: prefix delegation, RPKI, and transfers are particularly salient issues.

---

[600] In the case of the RIRs, see the discuss of the NRO in Section 5.4.5. In the case of the IXes, see the discussion of IX associations in Section 6.3.2. In the case of anti-abuse organizations, see the discussion of M^{3}AAWG in Section 7.2.

[601] There is not currently an officially recognized federating agent amongst the NRS institutions. [602] Rather, members of these institutions engage at common arenas such as NOGs, RIR meetings, IGF meetings, ICANN meetings, and occasionally at special purpose dedicated meetings. For instance, while any of the RIR meetings would have sufficient representation to hold various NRO meetings, the hosting RIR would be at a disadvantage given it is extremely busy managing the meeting. To solve this problem, NRO meetings are now frequently held at ICANN meetings—representatives from the five RIRs will be present and none are preoccupied with meeting management. As an institutional complex, the NRS comprises a *mix* of overlapping community-based organizations contributing to overall NRS management. As highlighted before, one leader of the RIR community characterized the system as a "complex web of authority" rather than strictly hierarchical. In the studies, federating agents are one step up the hierarchy, but serve to coordinate rather than delegate authority.

Some common resource analysts believe that boundaries are all that is necessary for effective regulation. Ostrom elaborates in her transition to the discussion of congruence:

> Since the work of [Ciriacy-Wantrup and Bishop (1975)], the presence of boundaries concerning who is allowed to appropriate from the CPR has been used as the single defining characteristic of "common-property" institutions as contrasted to "open-access" institutions. The impression is sometimes given that this is all that is necessary to achieve successful regulation. Making this attribute one of seven, rather than a unique attribute, puts its importance in a more realistic perspective. Simply closing the boundaries is not enough. It is still possible for a limited number of appropriators to increase the quantity of resource units they harvest so that they either dissipate all potential rents or totally destroy the resource [C. W. Clark (1980)]. Consequently, in addition to closing the boundaries, some rules limiting appropriation and/or mandating provision are needed. (E. Ostrom, 1990, pp. 91–92)

Effective boundaries, and by proxy their documentation by a commonly acknowledged authority, is considered a minimum, necessary condition. As noted above and the contribution of the discussion of these principles, NRS management also has a distinctly adaptive quality that can be explained by a combination of Ostrom's principles and relational authority.

In conventional CPRs, boundaries may be "bright white lines," but still require coordination, monitoring, and enforcement. For instance, the bounds of the meadow in which sheep may graze has observable boundaries. Another instance of observable bounds is an irrigation system a community relies on to water its crops. The topology of the former is rather immutable; short of substantive landscaping it may be difficult and/or costly to alter a rock outcropping or clear forest to expand the meadow.[603] On the other hand, the irrigation system is man-made, a mutable, artificial construct that enhances existing water resources. Other instances are less observable, such as underground reservoirs providing a general purpose water source. The reservoir relies on indirect inferences based on operational knowledge to determine the scope of resource management.

In the NRS, all boundaries are quite mutable. Facilities' boundaries, described in the next section, are rather clearly defined, either as databases (registry or reputation information) or interconnection platforms. These are well-defined and well-managed given they are administered by CRI firms. Boundaries amongst NRS resources, namely numbers and routes, are a bit more complex. Numbers are documented in the registries, but a post-deployment world threatens the integrity of the facilities documenting those "bright white lines," and, by proxy, threatens the integrity of the NRS. As established in Section 2.1, no single actor has the purview to assemble an accurate, complete view of the global stock of routes. Section 8.2.1.3

[603] Altering the topology has historically been costly, especially for rural CPRs before technologies for substantive landscaping existed. Even with current technologies, such change would be costly.

describes the boundaries in the stock of routes in terms of interconnection markets and how IXes serve as a means to expand the boundaries, providing participants greater *access* to more valuable routes (resources).

NRS-level boundaries are distributed and not centrally managed, thus require more monitoring and more nuanced graduated sanction to ensure integrity. Boundaries in the routing system are more accurately characterized by how network topology and infrastructure economics parcelize number resource stocks and stocks of routes. Both are dynamic, but they operate at different clockspeeds.[604]

8.2.1.1 Facilities Boundaries

Boundaries of many of the facilities maintained by the CRIs are centrally managed and well-defined: registry databases, IX platforms, IPBIs are instances. Number registry access and utilization is limited to members with bulk access to offload costs of extensive use. Consider the boundaries of the registry in terms of entrants. The registry operates in the spirit of open-access, but is partially rival in terms of database load. As discussed earlier, actors invoking high frequency or large queries will be rate limited to ensure equitable access for all entrants. Rate limiting is an instance of monitoring the boundary and applying appropriation rules (discussed in the next section).

IX participation is limited by strictly controlled physical access to interconnection facilities. IPBLs are, like the registries, implemented as a database that is rate limited and accommodates bulk access for selected users. In contrast to NRS stocks, in particular the jointly provisioned stock of routes, management resources are centrally managed by a single firm rather than a federation or a market.

As databases, the registry and IPBL boundaries are not commonly contested.[605] At a technical level, IX boundaries are also well-defined, often placing the threshold of responsibility at the patch panel between participant routers and IX platform infrastructure.[606] Geographically the diameter of a platform has seen debate both within individual IXes and in the community. Section 6.5.4 of the IX chapter argues for a notion of non-compete between associational membership IX providers that limited counterproductive competition that would fragment an otherwise functional

[604]Evaluating the NRS in terms of boundaries further affirms the differences in the management of numbers and routes. The two types of NRS boundaries correspond to the two types of provisioning discussed in Section 2.1, protocol defined boundaries and operationally defined boundaries. The stock of routes is much more fluid. The stock of routes changes with the economic conditions (costs) of traffic delivery and congestion. The stock of routes and the resulting typology is shaped by the composition of contracting modes that dictate the legitimate provisioning and appropriation of routes.

[605]Contention over boundaries is very different than contention over use of content. Recall from Section 5.3 that use of registry content has seen discussion in terms of bulk access and implications for operations.

[606]In the case of remote participants, the boundary is a bit more nuanced. The reseller serves as an ideally transparent extension of this boundary, creating the effect of connectivity at the IX over longer distance transport. Thus, technical issues experienced by the remote participant are taken first to the reseller, then potential to the IX provider.

market. In the case of independently managed platforms, the technical boundaries are well defined while the attendant loci of market activities overlap.[607]

8.2.1.2 Numbers Boundaries

The mapping of prefix sets to origination rights give one set of boundaries. For numbers, this the "what is being managed and for whom" emphasized in the distinction between common resource management and open access. Delegation modes have been historically dominated by allocation. Boundaries in the number stock are dynamic in the sense of consistent delegation, but with the exception of mergers, there has been much less transfer activity relative to allocation activity. Allocations from the free pool largely added to the topology.

Consider delegation within the number resource stock. Number delegation creates boundaries, shaped by different operational rules in the RIR system.[608] In terms of principles, the range of parameters is a function of congruence mediated by collective choice (Section 8.2.2). Within the formal-legalistic framework, these boundaries are enforced by contract with the firm. RPKI purports to enhance rights enforcement, here boundary enforcement, through a system of cryptographic assertions of "legitimately documented boundary demarcations." RPKI is a double-edged sword, enhancing boundary documentation and enforcement to the point where it can be easily appropriated by external actors wielding formal-legalistic authority

[607] Recall the distinction in Section 6.2 between geographic scope of participation versus the geographic scope of the platform infrastructure. In the sense of the former, participation, the market is the market for routes, and is often global, especially for large IXes such as AMS-IX, LINX, and DE-CIX. In terms of rationalizing connectivity within a sub-national region or MSA by providing a single or binary loci of interconnection markets, the scope of the platform infrastructure is an effective boundary for evaluating stability.

Remote node and remote peering models extend the boundaries of a platform. Recall these offer the option to peer seamlessly on remote platforms over dedicated transport. Within the same IX provider, this extends the logical platform beyond the conventional boundary of the MSA. In contrast, there are also some IX providers that have facilitated peering amongst participants across administrative domains. In effect, a single logical platform is jointly provisioned from platforms in two or more administrative domains. On instance of this is the relationship between the NL-ix and AMS-IX. Another instance is between France-IX, TOP-IX, and Lyon-IX. In both cases L2 boundaries remain in place, but these collaborations reduce administrative burden on the participant and provision attendant connectivity services. In the case of France-IX, membership fees provision finite capacity e between the administrative domains. Additional capacity must be provisioned, through sanctioned providers, if participants wish to exchange more than the threshold e. In terms of the options logic developed in Section 6.1, this facilitates exploring the potential value of options in the other administrative domain and investing in more specific assets when demand is more certain. In the case of France-IX, this regional federation of administratively heterogeneous platforms corresponds to the geographic clustering in Section 6.1, illustrated in Figure 6-3.

[608] Figure 5-4 describes the topology of delegation structures (hierarchies) that can be developed in response to operational demand for resources. The resultant parcelizations are the boundaries of the number system. Recall the distinction between structure and parameters in Section 5.6.3. Structure provide the decision *points* in the of range topologies boundaries can assume. Parameters, such as minimal delegation size or allocation window, determine the range of delegations that can occur at those points.

and their claim on ensuring the broader public interest.[609]

In a post-depletion NRS the delegation mode will be dominated by transfers. RFC 7020, discussed in Section 5.7.1, asserts that registry accuracy, not the historical norm of number conservation, is now the RIR's primary objective. Historically the need for additional addresses was the selective incentive to jointly provision (maintain) accurate registry information. Absent that incentive, boundaries may easily deteriorate. Boundaries may deteriorate as a result of shirking, not updating the registry, and, subsequently, those blocks being increasingly susceptible to security externalities such as hijacking.

The transition to a transfers dominated IPv4 stock presents two threats to the integrity of boundaries in the NRS. First, unfettered transfers that are not accurately documented in the registry threaten particular boundaries. Following Ostrom's principle above, this threatens to destabilize property rights. Even absent reputation issues, actors may attempt to invoke squatters rights on blocks that appear to be unused. In some resource systems, this is perfectly legitimate and has been sanctioned.[610]

The second threat is the threat of "bad neighborhoods" cropping up in topologies where boundaries have deteriorated due to poor registry management. Absent well-defined boundaries, the self-help role of the registry is substantively diminished. Further, absent well-defined boundaries, demarcations of responsibility for these resources, is diminished, potentially creating escalation problems with anti-abuse actors that use these boundaries to accurately attribute reputation. As described in Section 7.5.2.2, unfettered transfers coupled with the attribution of reputation may create markets of "damaged goods."[611] Further, as the value of numbers in this market become severely diminished, unfettered transfers will feed more number resources with less diminished value into this market. In terms of boundaries, both unfettered transfers potential effects on registry accuracy and the combination of unfettered transfers and reputation diminish boundaries.

Maintaining the integrity of these boundaries while not infringing on routing practices has been a longstanding ethos of the RIRs as number resource managers. Among other effects, RPKI has the potential to automatically validate assertions regarding boundaries. As an exercise of access rights, RPKI enhances the rights to deny a set of boundaries are valid. In contrast, unrestricted transfers have the potential to erode boundaries defined by the delegation process. Disintermediating the RIR opens the door to eroding the integrity of resource boundaries.

Disintermediating the RIR for an alternative resource manager or a "market" so-

[609]For elaboration on the mechanics, return to the discussion of RPKI in Section 5.7.4, in particular the discussion of predatory rule.

[610]See Clay and Wright (2012) for a discussion of squatters rights in the California gold rush.

[611]Numbers used for abuse will inevitably be attributed with negative reputation. Unfettered transfers make numbers *much more* fungible. Once imbued with negative reputation, these numbers will become, as per the dynamics described in Section 7.5.2.2, highly fungible assets with diminished downstream rights. Such assets will only have value to *a*) either those with unique downstream uses not affected by the reputation attributed or *b*) to abusive actors. The result is a market for damaged numbers that benefit from *deteriorated* boundaries. Absent an accurate registry, remediation becomes much, much more difficult.

lutions has a number of trade-offs that affect boundaries and other design elements of the CPR. The degenerate case of a market solution would facilitate frictionless transfers, with no interference from any institutional body. This would not work because, as developed in Chapter 2, the Internet relies on unique origination rights guaranteed by some documentation of those origination rights. Multiple actors have likened the minimal registry *role* to a titles office. In its most basic form, the core function of the registry is to record who holds basic appropriation rights bundles for a given prefix or set of prefixes. In other words, the registry's primary function is to record boundary demarcations. Previously, the need for additional addresses served as an incentive. The challenge to the stability of the number system is whether the community can identify a feasible combination of transfers, RPKI, and joint registry provisioning incentives of sufficient value to sustain an accurate registry.

8.2.1.3 Boundaries to Route Access

In terms of the stock of routes, like the notion of the global routing table, the complete set of boundaries cannot be observed by any single observer. Rather, only the local boundaries, the local subset of all routes provisioned by neighbors, can be observed from any given vantage point. The local boundaries here refer to the routes that may be legitimately appropriated by a given actor.

Access to boundaries is a critical element of route and traffic exchange. Recasting the IX regime as a social order, the objective was to expand the boundaries of local access to routes beyond that offered and controlled by transit bundles. Transit sets access conditions (i.e., offers lowest-cost routes rather than a selection of routes) to suit the economics of the transit provider. IXes increase a participants' *access* to other actors' boundaries, *facilitating* but not guaranteeing more efficient provisioning of routes. In effect, IXes as a management facility expanded access to local interconnection options. Following the characterization of boundaries above, IXes expanded the boundaries and have maintained efficient platforms for facilitating access, but, as per their constitutional norms, do not interfere in route exchange, the decision regarding who can appropriate which resources within its bounds.

Route availability and attendant capacity is an increasingly important regulatory topic. In particular:

1. Are currently appropriated routes provisioning access to significant consumer markets congested?
2. Do alternate routes *exist*?
3. If so, are those alternate routes available, i.e., are they accessible?
4. If so, are they available at various loci of interconnection?
5. If so, why are they not being appropriated to remediate congestion?

These questions can be condensed to: "What is the contractual closure of interconnection relations affecting the congested service?"[612] This closure draws clear boundaries around *a*) actors jointly provisioning routes across privately managed

[612]In other words, what subset of the value network provisioning this service is contributing to diminished quality of experience and is the problem rooted in infrastructure capacity or bargaining

networks and *b*) the expectation to provision sufficient capacity in the data plane to satisfy demand for services by end consumers. This shifts the focus to precisely who is provisioning routes affecting a particular service. This is distinctly different than broad sweeping rules under an under-specified rubric of "network neutrality." Here the focus is on the *context-specific* conditions of congestion, if there are alternative, what those alternatives are, and why they have not been selected. This analysis is rooted in identifying boundaries, but rather than the boundaries of the system as a whole, it draws on notions of congruence (Principle 2, Section 8.2.2), accountability and graduated sanctions (Principles 4 and 5, Section 8.2.3), and nested enterprises (Principle 8, Section 8.2.5).

In the context of this work, options to invest in interconnection diversity modifies the congestion question. Rather than simply asking whether a route is congested, the question becomes, "Is the routed congested and, if so, is there an alternate route that is not congested?" In other words, is congestion a function of available routes or the choice to pursue various routes? A consequence of a healthy interconnection market is that actors have a diverse set of interconnection options, facilitating coordination that avoids diminished consumer utility. *Observable boundaries* refocuses the regulatory discussion from the red herring of specific interconnection contracts to whether actors are leveraging resources efficiently and efficaciously. Narrowed to boundaries, the question becomes one of whether actors experiencing congestion have diverse mutual boundaries, which routes are provisioned on those boundaries, and whether congestion is a function of access and capacity, or, in a healthy interconnection market, the more likely culprit, bargaining failures. Thus, the root cause can be identified by a combination of infrastructure indicators (congestion) and consumer indicators (quality of experience) but is not necessarily an infrastructure issue. Rather, it is an instance of industry actors' bargaining failures creating strategic externalities in the infrastructure, and often clouded by limited purview into that infrastructure, but manifest as diminished consumer utility.

As a completely artificial (man-made) system, boundaries are an artifact of resource structure and are quite mutable. A combination of fit to industry conditions and collective management processes continuously (re)shape these boundaries. In the context of boundaries, constitutional rules establish, among other things, the possible range of boundaries and the basis on which they may be established. Operational rules establish particular boundaries. Congruence, the topic of the next section, is especially important for ensuring boundaries and the processes that set and change them, are aligned with industry structure and incentives.

8.2.2 Adaptation Through Congruence and Collective Choice

The "fit to resource mechanics," or in Ostrom's terms, congruence, of the function-specific CRIs, is one of the major strengths of the NRS. Ostrom defines the "congru-

failures within that subnetwork? Focusing on the conjunction of the value network and diminished quality of experience focuses on the context and the economics of the downstream service in question, intrinsically highlighting the incentives and contention at play.

ence between appropriation and provision rules and local conditions" (E. Ostrom, 1990, p. 91) as

> [a]ppropriation rules restricting time, place, technology, and/or quantity of resource are related to local conditions and to provision rules requiring labor, materials, and/or money. (E. Ostrom, 1990, p. 92)

Ostrom goes on to indicate that congruence means that "rules reflect the specific attributes of the particular resource," (E. Ostrom, 1990, p. 92). Following the parallels with the diversity of water rights drawn out in Chapter 3, Ostrom's definition of congruence concludes with:

> No single set of rules defined for all irrigation systems in the region could deal with the particular problems in managing *each of these broadly similar, but distinctly different, systems*. (E. Ostrom, 1990, p. 92, emphasis added here)

Each of the studies in Part II established how its respective institutions, in particular its rules structures, contribute to resource management. In the RIR system limiting inefficient use has been achieved via needs-based assessment and delegation; it is currently being challenged by "depletion-era" rule-making and how to develop transfer markets that are sufficiently fluid to meet demand but with registry participation incentives sufficient to preserve integrity. In the IXes, congruence can be summarized as *a*) how well the IX meets participants' demand for port and exchange capacity overall and *b*) how well its node deployment strategies lower barriers to participation. Congruence in the anti-abuse community means that reputation indicators *a*) when applied minimize abuse externalities, *b*) minimize false positives when attributing reputation indicators, and *c*) promulgate BCPs for effectively navigating the legitimate message sending space shaped by these indicators.

The second half of Ostrom's definitions contextualize congruence to "local conditions." Recall that each of the resource systems in this study are man-made and that both the CRI and the NRS managed are quite mutable. As an industry, the Internet infrastructure industry is not as high a clockspeed an industry as consumer electronics but the industry's clockspeed is substantively higher than conventional infrastructure such as rail, roadways, water systems, or power infrastructure.[613] Continuously changing industry structure, demand, and dynamics means CRIs must be sufficiently agile to adapt their rules not only to ensure system integrity but also to preserve their own utility. Ostrom's notion of congruence parallels the way Lake uses the notion of social order. A high level of congruence is a necessary, but not

[613]In all fairness, extending capacity for many actors dependent on Internet infrastructure often means appropriating colocation space and transport services. These actors do not have direct capital investments in Layer 1 infrastructure and can simply appropriate "infrastructure" services on the market. At Layer 1, while capital intensive, laying fiber in underground conduits or laying submarine cable systems does not have the growth limitations of, say, widening a roadway, digging a new canal, or building new long-distance power transmission lines. Moreover, short of cable cuts, fiber does not wear out, and subsequently require maintenance and repair, in the same way that roadways, rail, water systems, sewage systems, and others do.

sufficient, indicator of a functioning, contingent social order. If a *proposed* social order is not congruent with industry needs, actors will now longer have an incentive to subordinate themselves, or, in the language of common resource management, contribute resources to the joint provision of facilities maintaining this social order. In contrast, formal-legalistic models are not as susceptible to congruence. Given authority is delegated to a role, social order is can be *imposed*, regardless of congruence rather than a product of a direct social contract.

The issues sections of each of the CRI studies can be framed as evidence of the community adapting to ensure congruence. Congruence for each CRI means the criteria above hold at most points in the CRI's operation. Congruence is clearly not a story of static efficiency. Given the dynamic character of Internet infrastructure, congruence is one factor in the adaptive capability of the NRS. Collective choice (Section 8.2.2), monitoring, and graduated sanction (Section 8.2.3 complete a feedback loop that Chapter 9 will frame as a valuable form of anticipatory adaptation complementary to state authority. Congruence in the NRS means the rules are not only *capable of adapting* as local conditions change, but do adapt under the aegis of participant demand.[614]

Moreover, the congruence component of adaptive capacity means the object of adaptation is fit to the resource. Recall the object of RIRs' collective choice rules are operational rules, the object of IXes' collective choice rules are constitutional rules, and anti-abuse are the guidelines for navigating the legitimate sending space. In each case, these are congruent with the facility managed.

Adaptation can and does happen ad hoc—the Pakistan-YouTube story is a canonical instance, shored up by more systematic information sharing processes discussed in Chapter 4 on knowledge sharing arenas and Chapter 7's discussion of reputation provision and attribution. A more systematic model of adaptation implies that congruence is a *product* of consensus-based decision making. Consider Ostrom's description of collective choice arrangements:

> Most individuals affected by the operational rules can participate in modifying the operational rules . . . CPRs that use this principle are better able to tailor their rules to local circumstances, because the individuals who directly interact with one another and with the physical world can modify the rules over time so as to better fit them to the specific characteristics of their setting. (E. Ostrom, 1990, p. 93)

In a contingent social order, collective choice rules need to be congruent with their respective resource system, here their respective CRI. For instance, RIR collective choice rules adapt operational rules, IX collective choice rules adapt constitutional rules, and anti-abuse collective choice rules update non-binding operational guidelines that track adaptations in the MVN.[615]

[614]Recall the discussion of Hart's (1994) primary and secondary rules in Section 3.4.

[615]The distinction between "individuals affected by the operational rules" and "individuals who directly interact" with the resource system was made in Section 3.1's discussion of resource rights ($\mathcal{B}$) from downstream rights ($\mathcal{D}$). Participants in the routing system, those appropriating basic rights

Every adaptation does not require an act of collective choice, though. Ostrom illustrates congruence in terms of how rules affect information about the performance of the system and, subsequently, the value that can garnered under existing appropriation rules. Each of the studies in Part II concludes with a discussion of how organizations have dealt with changes in their environment. In effect, these sections describe particular issues where congruence was re-evaluated or re-articulated as a result of consensus-based decision processes.

For instance, debates over IPv4 delegation rules regarding transfers and needs-based justification have been perennial issues as the NRS approached complete free-pool depletion.[616] These debates and the resultant resource policy are an exercise in adapting rules to ensure fair access to the remaining free pool. Regardless of contracts with RIRs, adherence to a social order remains quasi-voluntary. Many in the NOG and RIR communities argue that transfers will happen with or without the sanction of the RIR. The question for the integrity of the number system is whether these will occur in a way that preserves the existing social order, i.e. the well-defined boundaries described in the previous section, or will short-term incentives lead to the "anarchy" of unfettered transfers and markets for damaged goods? The transfers debates over needs-based justification (Section 5.7.2) and the scope and limits of alienation rights (Section 5.7.3) are debates over which selective incentives are congruent with a post-depletion world. Again, the challenge for the RIR is provisioning facilities supporting a sufficiently low-cost social order that preserves registry accuracy while supporting a sufficiently fluid transfers market. If this social order fails and IPv6 deployment remains low, the likely alternative is diminished RIR authority, with it diminished support from number system participants, and likely intervention by states as the risk manager of last resort that has the capability to impose, rather than propose, order.

Amongst associational membership IXes, the evolution of platform architectures and reach, in particular reseller programs and remote nodes, highlight congruence as the product of constructive coordination supporting more diverse and accessible loci of interconnection. The product was not only a more sophisticated topology that lowered barriers to entry, but also a more nuanced appreciation of neutrality.[617] Anti-abuse is perhaps the most obvious case of adaptation, the legitimate sending space is far from static. Congruence as adaptation means continuously balancing *a*) keeping participants appraised of normative (anti-abuse) interpretations of messaging indicators while *b*) ensuring these do not teach abusive actors how to game and/or subvert reputation indicators or the processes that produced these

---

bundles $\mathcal{B}$ and/or affecting the value of those rights are direct participants. Those garnering value from downstream uses ($\mathcal{D}$) but not affecting actual infrastructure operations are indirect users—at their most proximate, they are consumers of services provisioned by NRS appropriators. Indirect users are typically not a part of operational communities managing the NRS.

[616]Recall that the IANA free-pool has been depleted, but each RIR has been delegated its last /8. All but AFRINIC are delegating from their last /8, under their respective /8 run-out policy, as of this writing.

[617]See the discussion of associational membership IXes' family of neutrality norms in Section 6.4.1.2, in particular how these norms tracked the growing geographic reach of the IX platform.

indicators.

Like the other CPR principles, congruence is necessary, but not sufficient to ensure stability. In the larger NRS, CRI-specific congruence may also be a weakness, especially considering NRS stability writ broadly. While CRI-specific congruence is characteristically adaptive, organizational professionalization narratives in each of the studies highlight resource management firms' role in ensuring rules are durable with respect to both endogenous *and* exogenous deviations from CRI rules. Endogenous durability can create stability but it can also lead to locking in particular norms and rules. For instance, one argument against needs-based assessment in the RIR system is that it is a vestige of the conservation norm that moderated the delegation of a finite, scare resource (IPv4 addresses) and that, in a post-depletion world where all delegations are transfers, the market mechanisms is a more effective allocation function. On the surface, this is true, but digging into precisely how needs-based and conservation works, there are tacit dependencies in the RIR system's audit and registry accuracy mechanisms. One should, to use the old euphemism, "be careful not to throw the baby out with the bathwater." Following the transfers discussion above, needs-based was a long standing selective incentive tightly bound to audit as means to ensure broader registry accuracy. Here again the question is what selective incentive is sufficiently congruent and low-cost to achieve a similar audit function in a post-depletion world? More generally, the confluence of congruence and collective choice is about a constituency identifying, evaluating, and credibly committing themselves to a set of selective incentives imposed by a CRI administrator. Anti-abuse is also facing issues of endogenous durability challenging congruence. The use and interpretation of messaging and reputation indicators has been strongly advocated for because of their role incenting network actors to endogenize externalities. Instrumental application of messaging and reputation indicators may be a sign of these indicators having become *too* durable. Both the anti-abuse operational rules on the dynamics of indicator development and navigational BCPs consistently encourage the use of indicators as *signals* that sending practices should be re-evaluated from first principles. The perception that M^{3}AAWG is becoming a deliverability conference rather than a conference on anti-abuse portends displacement of more general norms, perhaps even the image of consent-based messaging, with durable indicators of deliverability that are less costly to adhere to in the short term. In effect, the character of the social order may be changing and with it, those that hold substantive-purposive authority in that new order.

Without sufficient coordination, *local* congruence can, and has, lead to externalities and *operational* contention in the larger system. Revocation of number use rights based on abusive is perhaps the primary instance in this work.[618]

[618]Under the RIR system, when numbers were delegated from the free pool, the set of downstream uses is unrestricted. Recall that, in comparison to the market of "damaged goods" discussed earlier, the RIR free pool comprise numbers that had never been delegated. These numbers are "clean" in the sense they have never been used in the routing system, thus have no reputation, either positive or negative. Appropriation of and use of reputation information provisioned by reputation indicators limits what the anti-abuse community has deemed abusive behavior, potentially creating collateral damage in the course of graduated escalation. Amongst network operators in the RIR community,

RPKI is, on the surface, an instance of endogenous durability. The internal trade-offs related to RPKI were foreshadowed in Footnote 192 and discussed in at length in Section 5.7.4. RPKI is certainly a means to secure origination rights and a stepping stone to BGP path security. That latter serves the routing security community, but both limits the discretion of network operators and creates a durable loci of control outside the existing mechanisms. That latter is a form of exogenous control. The precautionary argument discussed in Section 5.7.4, spearheaded by in the RIPE region by Malcolm Hutty, warn against appropriation of security tools by government actors. In the language developed here, the risk is that endogenous durability will facilitate predatory rule; early LEA requests to "shut off" parts of the Internet are evidence of the demand for these capabilities from government actors. Given such a durable control point, endogenous capabilities may be appropriated by external administrators, potentially turned to broader political ends. Under these circumstances, to preserve the RIR system's reputation as a non-discriminatory institution, the RIR system is now faced with leveraging political capital to defy predatory states. The RIR system, in particular the RIPE NCC, has the social capital amongst its constituents but it is unclear that it *currently* has the broader political capital to deny states what they may well argue is a right to manage network connectivity within their national jurisdictions.

Appropriation by states would diminish the RIR system's social order, shifting rule making, at first in this case, back into the hands of the conventional "rulers."

As an evaluation of broader NRS stability, congruence is a testament to the power of function-specific regimes *and* their potential to let parochial interests threaten the integrity and stability of the larger system. In a number of interviews, actors in leadership roles, in particular the RIRs and anti-abuse, have accepted this tension as the status quo. Accepting the status quo has worked thus far, under the hood and away from exogenous threats to existing (formal-legalistic) authority structures, but it is an increasingly dangerous short-term strategy. First, Internet security issues, in particular infrastructure security and resilience, are only becoming more salient to state actors. IXes' scope allow them to make strong claims to resilience and security, shielding them in part from regulatory intervention. RIR participants and the anti-abuse community, though, each manage different elements of the NRS. Moreover, RIRs' vociferous rejection of enforcement roles outside registry accuracy (recall interviewees in leadership roles are adamant that they do not wish to be the "Internet police"[619]) will serve as arguments by would-be-principles that gaps in inter-CRI comity can, and will, lead to instability.

---

escalation is further damage. In contrast, amongst credible senders in the MVN, *gradual* escalation is a signal triggering remediation mechanisms and dialog with the reputation aggregators attributing negative reputation.

[619]When considering fundamental notions of authority, rejecting enforcement is a troubling indicator of weak authority. As will be developed in Chapter 9, adapting the notion of relational authority rooted in social order offered by Lake (2006), authority means that the "ruler" has right to use coercion, i.e. enforcement mechanisms, and the "ruled" considers that use legitimate. The threshold criteria for this abdication of "freedoms" is that the social order provisioned by the ruler, here effective number resource management, is sufficiently valuable to warrant particular degrees of subservience.

A key element of adaptation, as the combination of congruence and collective choice in a contingent social order, is the credible commitment to rules. Ostrom discusses this combination in terms of compliance:

> Agreeing to follow rules ex ante is an easy commitment to make. Actually following rules ex post, when strong temptations arise, is the significant accomplishment.
>
> The problem of gaining compliance to the rules—no matter what their origin—often is assumed away by analysts positing all-knowing and all-powerful *external* authorities who enforce agreements. In the cases described here, no external authority has had sufficient presence to play any role in the day-to-day enforcement of the rules in use. Thus, external enforcement cannot be used to explain these high levels of compliance. (E. Ostrom, 1990, p. 93, emphasis in original)

Compliance is fostered through the joint provision of rules by those beholden to those rules. Compliance is a result of a credible commitment to a social order, not an external coercive force or, in Ostrom's terms above, an "all-knowing and all-powerful *external* authority." As per the discussion of social order, credible commitment is contingent on both the value conferred by the social order and ensuring rulers do not create too fragile a social order.[620]

A frequently used strategy by RIR PDP moderators is to ask a dissenting participant to elaborate on what would constitute an acceptable compromise and the scope of the compromise in the text of the draft under discussion. There are many cases where the dissenting party does not fundamentally disagree with the spirit of the proposal, but with particular localized nuance that is perceived to cause problems. In the IXes, one particular instance of navigating the compromise space was the decision by the LINX membership to sanction developing a platform in the US. The initial proposal had sufficient support for the firm to move forward, but, in the face of distinct dissent, the leadership decided to address the dissent and re-present the proposal at the next member meeting. Some, but not all, of the dissenting parties contention with the plans put forward by the LINX were assuaged and the proposal garnered even more support.

Ostrom's rationale for credible commitment, that participants livelihoods depend on the integrity of the resource is supplemented in her collective choice discussion with the role of monitoring and sanction.

Consensus strategies serve two purposes: one immediate, one longer term. The immediate strategy is simply to resolve conflicts by identifying the precise point

[620]Reconsider the notion of a compromise space in the context of contingent social order. Collective choice rules are the means to navigate the compromise space amongst the preference sets of nominally competing actors. Not all points in the compromise space will yield sufficient compliance across CRI sub-constituencies to sustain a CRI's social order. Those points that do yield sufficient compliance, feasible compromise spaces, are not necessarily adjacent. Consensus-based adaptation is a dynamic process of moving from one feasible compromise to another. The various collective-choice processes described in Part II are CRI-specific mechanics that foster this kind of adaptation.

of disagreement and focusing the operational community on that point. The second purpose is more subtle, deriving from the high threshold of support necessary to "call" consensus and the willingness to address minority concerns. Relational authority requires a greater affirmation of commitment before dedicating limited resources to a particular endeavor. This higher threshold is intended to ensure the social order is of sufficient value to as broad a set of participants, incenting compliance, even when, in Ostrom's terms "strong temptations arise."

An artifact of these two characteristics is that, in contrast to majoritarian voting, where support may constitute anything from unanimity to "51% for and 49% against," consensus aspires to upwards of 75%, often claiming closer to 90%. Active and passive consensus explicitly attempt to reconcile differences amongst dissenters. In effect, it attempts to adapt the proposed solutions in such a way that ex ante commitment is not satisficing to move on, but addresses dissenters reasoning, ensuring the "new" social order resulting from the change is not so substantively diminished in value to the dissenters that they will defect. While the contingent character of the relational authority at play cannot, and should not, be eliminated, the rules produced can be made more durable. As a result, the social order these actors depend on is durable, but still accountable to participants. Accountability is further reinforced by efforts at monitoring and enforcement. In the next section, monitoring and graduated sanctions, along with Lake's notion of coercion under relational authority, will be used to explain endogenous monitoring and enforcement that *does* exist and the weaknesses signaled by its absence.

8.2.3 Monitoring and Graduated Sanction

Monitoring and graduated sanction are offered as two distinct analytic principles, but they are tightly interleaved in practice.

> Monitors, who actively audit CPR conditions and appropriator behavior, are accountable to the appropriators or are the appropriators.
>
> Appropriators who violate operational rules are likely to be assessed graduated sanctions (depending on the seriousness and context of the offense) by other appropriators, by officials accountable to these appropriators, or by both. (E. Ostrom, 1990, p. 94)

In the network operator and anti-abuse community, monitoring is considered a necessary criteria for the professional, reliable provision of downstream services. Ostrom indicates that "[m]any of the ways that work teams are organized in the Swiss and Japanese mountain commons also have the result that monitoring is a natural by-product of using the commons," (1990, p. 95). As discussed in Sections 2.1 and 2.2 on the mechanics of the NRS, a certain degree of awareness of the appropriation practices of one's neighbors is an inevitable consequence of observing AS-paths.

In a number of cases amongst network operators managing routing elements and connectivity, monitoring has been made durable via protocols and common communication mechanisms. Recall the discussion of route flap in Section 2.2.2.1:

early on it could be monitored and attributed to particular actors, later a protocol for dampening route flap was identified, and through information sharing amongst the operational epistemic community, empirically informed parameterizations were identified. Route flap in particular is an interesting case of automating graduated sanction. Other instances are the use of network operator e-mail lists and backchannels. The outages lists comprise a particular instance of reporting, tracing, and validating observed network failures. Network failure reporting is largely monitoring and remediation—*most* failure modes adversely affect the operator experiencing the failure and no further incentive for remediation is necessary.[621]

Monitoring and information sharing accrues public and private benefits.[622] Public benefits in the routing system include immediate pressure to resolve a failure. Public monitoring information also incents prompt response to potentially solve the problem before it is reported or at least gain the reputation as a credible, prompt remediator. Ostrom argues that the private benefits accrue "[i]n repeated settings in which appropriators face incomplete information, appropriators who undertake monitoring activities obtain valuable information for themselves that can improve the quality of the strategic decision[s] they make," (1990, p. 96).

Amongst the CRIs, operational rules in the anti-abuse regime are almost exclusively about monitoring *and* graduated sanction. Returning to the characterization of anti-abuse BCPs, in this case the two characterization can be framed here as *a*) explicating how to perform effective messaging and reputation indicator development and monitoring and *b*) explicating how to navigate the graduated sanctions imposed when actors approach (or cross) indicator thresholds. As per the discussion of the intentionally hazy boundaries of the legitimate sending space, different (reputation) monitors and aggregators may (and do) apply differentiated graduated sanction disciplines. In the anti-abuse regime, these monitors and aggregators are held accountable by the market for messaging and reputation indicators. Actors with high false positives, aggressive sanctions, or that are themselves extractive are not utilized and thus do not garner the network effects sufficient to achieve their goals or warrant their use.

Ostrom's parallel to the diversity monitoring and sanctions in the anti-abuse community is that

> [b]ecause the appropriators tend to continue monitoring the guards, as well as each other, some redundancy is built into the monitoring and sanctioning system. Failure to deter rule-breaking by one mechanism does not trigger a cascading process of rule infractions, because other mechanisms are in place. (1990, p. 96)

[621]The dispute between Comcast and Netflix is an interesting exception. Congestion was attributed to Comcast, but, as it turns out, Cogent, contracted for transit by Netflix, was prioritizing retail customers to the detriment of Netflix traffic. For some time there was not evidence of Cogent as the cause of the congestion, resulting in substantive contention between Netflix and Comcast, drawing regulatory attention.

[622]Benefits are public in the sense that once information on infractions or failures are shared, any actors, including free-riders, may consume and benefit from the efforts of monitors.

Recall the notion of reputation images discussed in Section 7.2.4.1. The degenerate case of the local reputation image is parameterized only by local data; external inputs both supplement local indicators and provide additional information. The diverse set of monitors and aggregators in the anti-abuse community is an extreme form of Ostrom's redundant monitors. In the MVN, ESPs and IPBLs serve are a market for "guards" that both limits "cascading rule infraction" and selects the most effective, and often complementary, monitors and aggregators.[623] In effect, a subnetwork of the MVN is a robust market for monitoring and reputation tools.

Graduated sanction is also evident in the IX regime. For instance, when a new participant joins AMS-IX, they start in quarantine to ensure they are not violating network hygiene rules.[624] In the RIRs, simple rule infractions do not necessarily result in immediate deregistration. A series of warnings and service limitations are imposed, diminishing service but not completely and permanently revoking, access and appropriation rights as a means to get the attentions of the deviant actor.[625] In both cases, sanction is graduated, not simply a wholesale revocation.

The power of graduated sanction is the discretion to consider extenuating circumstances when imposing sanctions. For instance, Sections 7.3 and 7.4 argue that swift listing signals the efficacy of monitoring and the capability to sanction. Once the reputation aggregator has the sanctioned actor's attention, the aggregator can use their knowledge of the context and discretion to determine where in the legitimate sending spectrum the externality lies: is it an instance of naïve operational externalities or is it composite extractive abuse? If the former, sanction may be replaced with education or a pointer to a reputable ESP that can help the sender better implement good sending practices. In the latter, a stronger sanction may be imposed along with information sent to concerned law enforcement. In terms of relational authority, discretion in graduated sanction is an instance of facilitating self-help via leadership.

Earlier, discussions of Ostrom's framing of monitoring argued for accrual of costly information that confers strategic advantage. This is the case for both pro-

[623]As evidence of the complementary character, recall the role of ESPs as monitors that appropriate reputation indicators from a variety of aggregators and monitors. Further recall that some of these actors primary role is that of a receiver but provide monitoring mechanisms such as well-developed industry standardized feedback loops.

[624]Once the firm confirms the new participant is "clean" it moves the participant to the main peering LAN. Technically it is all the same hardware, the quarantine is implemented via a VLAN. It is not guaranteed that, once configured, a network will continue to adhere to network hygiene norms. This is especially the case for actors with little more experience than setting up a transit relation. These are cases of what was labeled naïve operational externalities in Section 2.2: they are not intentional, often due to inexperience, but still create costs for external actors. When actors violate network hygiene, depending on the extent and severity of the externality, they may be placed back in quarantine rather than wholesale disconnected. These participants then work with the IX to debug and solve the problem. In small-to-medium IXes with less formally structured processes, informal organization and management accommodates discretion in graduated sanctions.

[625]For instance, if an ARIN member does not pay its fees for six months, number rights will be revoked. The numbers revoked are held for the next six months in the event that the actor remediates and pays their fees. At the end of twelve months if the actor has not yet paid, numbers are returned to the free pool.

fessional monitors, such as ESPs, as well as reputation aggregators, such as IPBLs. Given a diverse set of clients across the legitimate sending spectrum, monitors accrue data on an equally diverse store of negative externalities as they help clients remediate. This diverse store contributes to the monitors capabilities to more efficiently identify externalities, assess how ingrained in the sender's practices the offending processes, and offer either generic or tailored solutions. Under this mechanism, monitoring enhances the actors knowledge of the larger ecosystem but also makes the monitor a better monitor. Again, the discretion facilitated by the experience from monitoring and graduated sanction facilitates self-help via leadership.

Ostrom also highlights the feedback between accruing knowledge and the design of rules:

> When appropriators design at least some of their own rules (design principle 3), they can learn from experience to craft enforceable rather than unenforceable rules. (1990, p. 96)

Within the anti-abuse community, learnings from these diverse experiences are transposed into BCPs often championed by individuals from credible firms in the MVN. Across the CRIs, information from monitoring feeds in into continuous updates to operational rulesIn the IXes, the health of the platform is almost exclusively about monitoring utilization and enforcement of network hygiene. In all of these cases, the operational community serves to continuously improve, to continuously *adapt* operational rules and supporting facilities. Moreover, it is arguable that these are the actors with information best suited to engage in this continuous adaptation.

In the hands of an operational epistemic community serving as a common resource management collective, monitoring and graduated sanction is about much more than enforcing rules. Monitoring is a means to collect information about infractions as a means to both monitor rights, but also to monitor the efficacy and efficiency of the rules protecting those rights. The actors that are breaking those rules are also members of the operational epistemic community and may well have very valid technical, operational, and economic reasons for violating rules. Taking these rationales into account, graduated sanction provides the discretion necessary for adaptation that emerges, perhaps counter-intuitively, from patterns of infractions rather than patterns of compliance.

8.2.4 Conflict Resolution Mechanisms

At a very fundamental level, common resource management is an ongoing process of navigating the compromise-space amongst participants, the overall goal an attempt to improve aggregate participant welfare while disenfranchising as few as possible. Collective choice processes are framed as constructive conflict.Ostrom's characterization of conflict explicitly addresses "low-cost arenas to resolve conflicts among appropriators or between appropriators and officials," (E. Ostrom, 1990, p. 100) followed by discussion of interpretation and ambiguity in operational rules. Conflicts include *a*) bilateral conflicts between CRI participants over *interpretations* of rules, *b*) conflicts between participants and CRI firms over firm decisions

and interpretations of rules, *c*) conflicts between participants over business relationships that are either a direct consequence of CRI participation or for which the CRI resources are a key input, *d*) general contention amongst actors in various CRI arenas, *e*) pre-emptive identification of conflicts or inconsistencies in the rules. Conflict resolution is typically scoped to a particularistic transaction amongst two or more parties (a–c). Point (*d*) addresses the contention in *the course of* a collective action process, but does not frame collective action processes themselves as conflict resolution. Instances are norms related to PDP engagement, in particular norms related to engaging in discussion fora and e-mail lists, and facilitation in the anti-abuse community. The focus is to mitigate contention, turning the differences surfaced to the productive process of developing consensus. Each of the RIRs has a dispute mechanism, either referring to a local, established arbitration mechanism or maintaining its own arbitration mechanism. These mechanisms are intended to resolve conflicts between participants and the RIR firm.[626]

As a pre-emptive strategy, ARIN also has a PDP performance review. After policies are implemented in the ARIN region, a review of these rules' performance is presented by ARIN. This process highlights inconsistencies that can potentially lead to conflicts in the interpretation of the rule. This follows the spirit of conflict resolution offered by Ostrom's discussion, that ambiguity will inevitably crop up and conflict resolution mechanisms are necessary. The PDP performance review identifies ambiguity from the perspective of the firm (resource manager) implementing rules.

Both sides have incentives to engage in proactive rule evaluation. First, it avoids a formal appeal process by an individual member.[627] For the firm, each time this rule must be invoked, the firm must cope with ambiguity.[628] The firm can adopt an interpretation and apply it consistently. The consistency solution is an exercise in discretion that usurps the management rights of the collective, a technical violation of the social contract between the firm and the constituency. Thus the incentive for the firm is multifaceted: *a*) the firm does not usurp authority of the constituency to exercise management rights, thus not suffering the attendant audience costs; *b*) the firm can provide the community (constituency) with an opportunity to correct the ambiguity or grant the firm the discretion to handle the ambiguity on its own; *c*) the potential differences across interpretations can be eliminated, streamlining

[626]ARIN's RSA (ARIN, 2014b, Section 14(k)) indicates the America n Arbitration Association (AAA) is the preferred arbiter for disputes, including appeals to number resource requests (ARIN, 2009a). The RIPE maintains is own board of arbiters, comprised of community members (RIPE, 2014a). In AFRINIC, appeals are adjudicated by the AFRINIC board (AFRINIC, 2014a, Section 13). APNIC's dispute resolution process is described in its by-laws (APNIC, 2014a, Articles 73–81); in particular, it lays out a timeline for pursuing a dispute, the option for disputants to select an arbiter, failing a selection by the disputants a selection will be made by the Institute of Arbitrators of Australia. In LACNIC, disputes are referred to the Center of Conciliation and Arbitration of the Uruguayan Chamber of Commerce (LACNIC, 2014g, Article 11).

[627]In the aggregate the transaction cost for an individual actor is low. Relative to that actor, though, that transaction cost may be high.

[628]Host masters are the actors in the ARIN firm that evaluate resource requests; in the face of ambiguity, different host masters may interpret the rule in a different way.

firm resource delegation processes.

Associational membership IXes have a number of different positions on conflict resolution, ranging from none, to explicitly referencing court jurisdiction, to providing a dispute resolution process. The two simple instances are France-IX and the AMS-IX. France-IX explicitly indicates conflicts will be resolved by the courts of the applicable jurisdiction (France IX, 2011, Article 30). The AMS-IX does not have a documented dispute resolution process outside of its SLA dispute reporting mechanisms (AMS-IX, 2014). These positions are one interpretation of IX neutrality: the IX will not interfere in conflicts between participants.

The LINX takes a different stance, leveraging its role as a neutral third party to serve as an alternate mediator that can offer more efficient dispute resolution that addresses root cause rather than legal positions. The LINX has an explicitly defined dispute resolution process (LINX, 2014a). The LINX dispute resolution mechanism is cognizant of its role relative to statutory mechanisms:

> A satisfactory conclusion resulting from this dispute resolution procedure will avoid the referral of complaints to statutory regulators and demonstrate a regime of responsible industry self regulation. (LINX, 2014a)

The LINX dispute resolution mechanism is completely voluntary and scoped to conflicts arising in the core activities of the IX.[629] That said, the premise of the dispute resolution process is that this mediated procedure is less costly than involving the legal system and lawyer's time preparing a formal case. Further, consider the bullets

[629]LINX (2014a) goes on to list a number of advantages of the LINX dispute resolution mechanism relative to legal mechanisms:

> Potential advantages are:
>
> - Promptness and speed of resolution.
> - Substantial savings in legal fees and other litigation expenses and in the time and energy of executives.
> - Creative, business driven "win-win" solutions not available in a court of law.
> - Solutions based on parties' real interests, not just legal positions.
> - Preservation of the business relationship.
> - Privacy and confidentiality.
> - Parties of different nationalities often are reluctant to litigate or arbitrate in each other's country. They should be less reluctant to operate this procedure.

The mediation process may be initiated by one or all of the parties involved but, to proceed, it must be agreed to by all involved.

Once initiated, the LINX dispute resolution procedure has the following characteristics. A LINX staff member is appointed mediator if the parties in dispute do not select a mediator themselves. Each step of the process is documented in written agreements between the parties. The LINX mediator has final say over the procedure. Confidentiality in these proceedings is paramount, all information shared with the parties in dispute and the mediator is to remain confidential. The mediation process is intended to be as swift as possible: it is expected that the mediator will set an expeditious meeting schedule amongst the parties in dispute and parties are expected to make good faith efforts to follow this schedule. In the LINX procedures, there is a burden to provide materials pertinent to the dispute and to bear the costs of introducing external experts.

above, in particular "[s]olutions based on parties' *real* interests, not just legal positions" and "[p]reservation of the business relationship," (LINX, 2014a, emphasis added). Such proposition is more in line with the values of an operational epistemic community: identifying the "real," which is read here to be the technical, operational, or economic root causes rather than legal-formalistic rationales rooted in black-letter contract law or simple liability arguments that address to symptom, but not the root cause. Here the LINX dispute resolution mechanism is an instance of informal mechanisms within the operational epistemic community being professionalized to match the reliance of actors on NRS facilities.

There is no official conflict resolution mechanism amongst the CRIs. Aside from engagement within convening fora, no formal conflict resolution exists *between* the CRIs. The overlap between the NOGs, RIRs, and IXes as well as the longstanding separation between number policy (RIRs) and routing policy (private policies amongst NOG, RIR, and IX participants) limits conflicts amongst the communities. Moreover, many of the IX federating agents emerged out of the NOG and RIR communities and thus have both a tacit understanding of and respect for the number and routing policy separation. While the status quo has worked in the past, current changes warrant additional comity amongst CRIs. In particular, the combination of RPKI, transfers, and reputation attribution have been considered as largely independent issues, but, especially with the depletion of the free pool, will become increasingly integrated.

8.2.5 Rights to Organize and Nested Institutions

The rights for CRO's to form organizational entities and the nested, or multilevel character of the NRS and adjacent resource management regimes are tightly interleaved and, as such, will be discussed together. Ostrom describes the first of these principles, principle 7, as:

> The rights of appropriators to *devise their own institutions* are not challenged by external governmental authorities. (E. Ostrom, 1990, p. 101)

This principle encompasses both the mechanisms to form an organization in a particular jurisdiction, but more importantly, the language "devise their own institutions" speaks to *authoritative management* of a common resource system. As per the discussion in Section 2.1, the origins of the NRS can be traced to the original Network Working Group and the IETF RFC series as the activities that produced the resource units discussed here: *a*) the stock of IPv4 addresses, *b*) the stock of ASNs, *c*) the specification of BGP as the canonical mechanism for provisioning and appropriating routes.

The modern NRS has been framed as an institutional complex comprising a number of institutions contributing to a network operations knowledge commons and producing function-specific resource management facilities used by NRS participants to manage, enhance, and diminish number rights. Principle 8 addresses the distribution of CPR functions within a particular CPR system:

> Appropriation, provision, monitoring, enforcement, conflict resolution, and governance activities are organized in multiple layers of nested enterprises... the Spanish *huertas*, for example, irrigators are organized on the basis of three or four nested levels, all of which are also nested in local, regional, and national government jurisdictions. (E. Ostrom, 1990, p. 101, emphasis in original)

The NRS as a whole comprises the functions above, organized in a non-hierarchical, relational topology of loosely federated CRIs, not a strict hierarchy. Hierarchy is in the NRS is limited to the production of number resources and the hierarchical delegation hierarchy described in Section 5.2.

In terms of principle 7, the right to form organizations has historically been the case in the NRS. Operating "underneath the hood" these organizations did not attract attention but did ensure the integrity of common resources relied upon by more well-known downstream services, such as the world wide web (WWW), e-mail, and over-the-top (OTT) video rely on. Most of the organizations discussed in these studies are either a form of non-profit organization or a trade association. As organizational entities, most have had clear rights to operate within their respective jurisdictions.[630]

"[D]evis[ing] their own institutions" speaks to the substantive-purposive authority to manage resources. Substantive-purposive authority complements arguments that participants are more credibly committed than external actors. Credible commitment is not only rooted in a dependence on the resource for its livelihood, but also rooted in experience to maintain and adapt valuable social order. Many of the conventional, natural resource CPRs discussed by E. Ostrom (1990) are part of the natural environment, within a particular state jurisdiction. The operational rules structuring appropriation may or may not conflict with those of state jurisdictions. Such a scenario can create problems for the CPR. If a government "presume[s] that only they have the authority to set the rules" then it will be difficult for CPR participants to sustain a CPR of their design (E. Ostrom, 1990, p. 101). Alternately, if a sub-constituency of the CPR is unhappy with the rules, they may appeal to the government as a "higher" authority, an authority that albeit not recognized by the larger CPR constituency, may have state-based (formal-legalistic) authority to act as the next higher jurisdiction in the hierarchy. Again, the threat is jurisdiction shopping, undermining endogenous authority by appealing to formal-legalistic.

IXes and the anti-abuse community are distinctly non-hierarchical. IXes share a common set of constitutional norms, and these norms are promulgated by both individual IX providers and coordinating agents such as Euro-IX and Open-IX, but associational membership IXes are typically beholden exclusively to their memberships. IXes may borrow interpretations of norms from one another, such as the development of the France-IX constitution based on lessons from the AMS-IX and LINX constitutions, but there is not a global hierarchy of IXes. Some IXes have a

[630]Exceptions are jurisdictions with telecommunication service licensing regulation that includes Internet communication, limiting the development of IXes. See Kende and Hurpy (2012) for an instance in Kenya. Often this is framed as regulatory capture.

federated structure, such as CABASE in Argentina, but CABASE is an exceptional case. As noted earlier, the IX firm is incorporated in a state jurisdiction, but, short of regulatory capture of licensing mechanisms by conventional telecommunications actors whose revenue is threatened by IXes' rationalization of interconnection and connectivity markets, IXes are effectively private firms with the "rights... to devise their own institutions," (E. Ostrom, 1990, p. 101).

The anti-abuse community and its institutions are similarly diverse. Anti-abuse as a regime comprises the constituencies discussed in Section 7.2: senders, receivers, and vendors. Actors from the former two, credibly committed to anti-abuse norms and best practices, certainly contribute to monitoring and enforcement, but enforcing anti-abuse norms is not the *primary* value proposition of either group. In contrast, for vendors such as reputation monitors and, in particular, reputation aggregators, monitoring and enforcing anti-abuse norms *is* their primary value proposition. Reputation aggregators such as Spamhaus have been brought to court in lawsuits, the largest being e360Insight.[631] Although not explicit, Spamhaus does have the sanction of LEAs; it has contributed to a number of investigations.[632] Spamhaus has been presented the Cyber Crime Fighter award by the National Cyber-Forensics and Training Alliance (NCFTA) and lists the FBI as a main working partner (Spamhaus, 2015a). Moreover, Spamhaus' objectives and outcomes, in particular tracing abusive externalities to their root cause and sanctioning those actors, are quite well-aligned with LEAs.

8.2.5.1 Hierarchical Number Rights Delegation and Authority

Returning again to the fundamental production narrative of the NRS in Section 2.1, IPv4 addresses and ASNs were created under a US contract, managed initially by Jon Postel in his role as the IANA. According to the Security and Stability Advisory Committee (2014), the two primary numbers related roles of the current IANA are

2. Internet Numbers Registry Management and
3. Protocol Parameters Registry Management, including management of the "Address and Routing Parameter Area" (.ARPA) TLD. (2014, p. 5)

The IANA role, in particular number delegation functions, have been fulfilled under a number of organizational forms: *a*) by Postel on an ad hoc basis under a series of Department of Defense funded projects, *b*) under contract with a number organizations (for instance the Defense Data Network Network Information Center, DDN-NIC), *c*) later by Network Solutions, and *d*) most recently ICANN. Initially, the IANA's role was both coordinative *and* operational. Over the course of the IANA role's evolution, here in particular focusing on the number management and delegation functions, the IANA has subsequently delegated *most* of its operational number delegation function to the RIR system. In terms of Figure 5-4, the IANA has delegated all by $L1$ delegations to the RIR system.

[631] See Spamhaus (2011) and Masnick (2011) for the report of the case by Spamhaus and TechDirt.
[632] Base on interviews with Spamhaus contributors.

Recall that Postel as the IANA managed all number delegation; one simply asked for a block, provided a reasonable rationale, and was delegated a block of numbers with no contractual limitations; as discussed in Section 5.2 and 5.6.3, this is the origin of legacy delegations. Until the early 1990's, number delegation was handled by the IANA. In 1992 and 1993, respectively, RIPE NCC and APNIC were created. Each was created as a regional NIC that, among other responsibilities, were to delegate number resources in their respective regions. At the same time, the notion of an IR was explicitly developed in the IETF's RFCs.[633] The IANA delegated number blocks to these two RIRs to subsequently delegate within their regions. In 1997, ARIN was created to handle number delegation in regions not currently serviced by the RIPE NCC and APNIC. Shortly after this what remained of the IANA role as a whole was delegated to the newly created ICANN under a contract from the NTIA.

The number resource delegation role of the IANA under the NTIA contract with ICANN is quite different from the earlier roles of delegating numbers directly to actors using those numbers. The creation of the RIPE NCC and APNIC was a centralization of management rights into independent regional IRs. This centralization created a coordination mechanism that facilitated authoritative delegation of number rights. Acknowledgment of the RIPE NCC and APNIC IRs by the IANA and the IETF, along with regular $L1$ delegations to these RIRs (and later subsequent RIRs) together constituted an effectively permanent alienation of management rights. Authority in the IRs is rooted in contingent specific-purposive authority, the value of that social order to appropriators of number resources. Authority is not a result of the conferral of rights by the IANA, it would not exist but for the conferral of those rights and the rights to subsequently delegate number rights to members. Conventional formulations of principal-agent relations specify that authority delegated by the principle can be rescinded by the principal.[634] If present at all, the IANA's authority to rescind the authority to operate as an RIR is implicit.

Further contrast the pragmatics of conventional principal-agent relations, such as the delegation of authority from legislative bodies such as the US Congress to specialized agencies such as the FCC. These are governed in principle by the delegation of authority, but pragmatically by the principal's control of agency funding. The RIRs, in their most fundamental registry maintenance and delegation functions, would not exist but for $L1$ delegations of resources by the IANA—there's no point in a resource registry if one has no resources to delegate and register. That said, *operational decision-making authority* has since rested in the firm, adapted by collective choice processes described in Section 5.6.2. Further, the source of RIR funding is membership fees, not the IANA. In further contrast to a supposed principal-agent relationship between the IANA and the RIR system, for a brief period of time in late 2000, the IANA role was not under contract by any entity—the existing RIRs, the RIPE NCC and the APNIC, directly funded USC/ISI to maintain IANA function.[635]

[633]As indicated in Chapter 2, RFC 2050 (Hubbard et al., 1996) has been the historically most frequently referenced definition of an IR.

[634]See the discussion of principal agent in the international arena, discussed at length in the first chapter of (Hawkins et al., 2006).

[635]According to Security and Stability Advisory Committee (2014):

With the creation of ARIN in 1997 and the subsequent formalization of the NTIA IANA Functions in the contract with the newly formed ICANN in 1998, rights to perform all but $L1$ delegations rested with the RIR system. From that point until the depletion of the free pool in 2011, the only coercive mechanism available to the IANA was to *a*) withhold subsequent $L1$ delegations from an RIR or the RIR system as a whole and/or *b*) withhold updates to the .ARPA TLD. The result is an interestingly interdependent, non-cumulative distribution of resource rights in the number resource system. Post-depletion, the IANA function can be said to perform, with respect to IPv4 address[636] and ASN management, an almost exclusively coordinative role in managing the .ARPA TLD. Of the NRS CRIs, the RIR system is the only CRI whose relational authority could be said to have elements of a hierarchical organization that could be considered to follow the pattern of "nested enterprises" referenced by Ostrom's principle 8. As discussed in Section 5.2, *resource rights delegations* flow hierarchically, but RIRs derive authority largely from their constituency.

Based on this discussion and the discussion of constituencies and the dynamic character of the legitimate sending space in Chapter 7, the anti-abuse regime is clearly not hierarchical. Rather than framing the coordination and cooperation structures as "nested," this work re-operationalizes these structures as relational. Each of the CRI's is based on a contingent social order. Relational coordination and cooperation serves to sustain the broader social order in response to demand for that order.

Mattli and Woods (2009b) frames non-state "private" authorities as responding to private industry's demand for regulation. Endogenously, the relational authorities discussed here also provide order. Mattli and Woods (2009b, loc. 212) defines regulation as "the organization and control of economic, political, and social activities by means of making, implementing, monitoring, and enforcing rules." Often these private authorities find themselves framed as transnational entities *in competition with* state authorities. In contrast, CRIs' strategy is to develop the diplomatic capabilities necessary to retain their operational authority, provide complementary support and policy advice to states' regulatory efforts, while not becoming subordinate to those state authorities.

> Following the termination of DARPA funding for the Tera-Node project (which included funding for the IANA Functions) and prior to the institution of the NTIA IANA Functions Contract there was a short period of time during which the IANA Functions had no explicit funding. During this period, the RIRs that existed at the time (RIPE-NCC and APNIC) provided funding directly to USC/ISI to fund IANA Functions operations. (2014, p. 37)

This is one instance of a community contributing their own resources to support a social order, here a coordination mechanism, that, while limiting that community's range of behaviors, is of value to that community. Albeit short lived, this is an instance of temporary joint provisioning.

[636] As noted before, this work does not spend a substantive amount of time addressing IPv6 issues. Succinctly, the lines of authority regarding IPv6 production and delegation are much more clear than the path-dependent character of IPv4 production, delegation, and the devolution of the IANA function (back) to the community.

8.3 Out From Under the Hood

The NRS can no longer remain under the hood. To function in the global political arena, it must learn to engage with external authorities. External authorities not only includes states, but IGOs as their proxies and NGOs. Further, the NRS would benefit from learning to coordinate with its new peers, NGOs and other private and relational authorities. The following develops some of the engagement strategies developed thus far.

8.3.1 Engagement Strategies

CRIs' engagement within the NRS has been framed as coordination loosely structured by the two common images of in integrity in routing system management. External engagement has historically been largely reactive. Ostrom's principle 7 indicates that "appropriators[’] ... own institutions are not challenged by external government authorities," (E. Ostrom, 1990, p. 101). Drawing from traditional international relations, authority that is not rooted in the state is seen as competing. Here, principle 7 implies a compromise, where multiple authorities co-exist, although as per the previous discussion, when nested, the state can *impose* authority when a resource system's social order does not align with that of the state.

Principle 8 highlights this nesting and the threats of exogenous "imposed" authority, ultimately diminishing the authority, and subsequently the adaptive character of, common resource rule-making. In "small" CPRs such as those discussed by E. Ostrom (1990),[637] the resource system typically exists within a state jurisdiction and must either *a*) hew to existing state-based rules, developing common resource rules within the private sphere or *b*) "challenge" the state's jurisdiction. The latter often requires understanding how social capital amongst participants managing the common resource can be used to create political capital in the external political arena, in the canonical cases, a state jurisdiction An instance of this latter is contention over Thai forestry management between historical local (common resource) institutions managing forests and government-based national forest agency (Birner

[637]Cases in (E. Ostrom, 1990) largely deal with small and/or rather old common resource institutions. Consider the critique Dolšak and Ostrom (2003a) indicate is frequently leveraged against the study of common resource systems:

> "Why are you wasting your time with the study of small, unimportant, local resources and outdated institutions? Don't you know that local resources are boring? Common-property institutions are a thing of the past. Common-property institutions will wither away within the next few decades." ...The basic message of the challenger is that common-pool resources are insignificant in modern times and that common-property institutions are not worth studying because they will not survive. (2003a, loc. 125–128)

The NRS is clearly a counter-example, along with the offered in the studies collected in Dolšak and Ostrom (2003b). The challenge for these resource systems, developed by the works in Dolšak and Ostrom (2003b) and in this study of the NRS, is how these distinctly different management authorities can engage with existing state actors.

& Wittmer, 2003). Local interests want to preserve the resource; these local participants are a canonical instance of those that depend on the resource for their livelihood and thus being more credibly committed to long-term integrity. These local common resource management regimes pushed back against national administrators that wanted to introduce external appropriators, for instance the "provision of commercial plantations" by outside developers. This was perceived as potential exploitation by relatively *external* appropriators that, unlike *local* management, do not depend on the resource for their livelihood.

The Thai forestry narrative focuses on how social capital within the common resource management regime can be used to create political capital in the broader political arena. In early fieldwork and interviews, actors stressed the social component of network operations management. The operational epistemic community coheres around common images of system operation and social capital amongst actors perceived to have substantive-purposive authority in that epistemic domain. In the Thai case, the issue-space in question is national forestry management.[638] In the NRS, the common resource issue-space in question is global management of the Internet's routing system.

The NRS is clearly a transnational institutional complex that, while not as beholden to a single state as smaller, terrestrial common resources are, still faces the threat of exogenous, would-be principles. RIRs have historically eschewed government engagement. A number of external government actors have related their experience engaging the RIR arenas. LEAs in particular were met with skepticism. Early in ARIN's history, the FBI requested ARIN "shut off" particular network blocks—the FBI was under the impression that ARIN had this capability. ARIN staff spent some amount of time explaining to the FBI that this was not possible. Ultimately, this initial encounter led to a productive cooperative relationship. One outcome was educational outreach by ARIN to train LEAs how to access registry information in support of investigations. As discussed in Section 5.5.2 on Government Arenas, ARIN's APWG, the RIPE RoundTables, and APNIC's LEA outreach are all efforts to keep government officials appraised of the role of RIRs in the Internet infrastructure, RIR activities, and how to access information in the registry. Initially these activities were reactive, but now they are proactive informational activities.

The anti-abuse community's monitoring and enforcement activities are also proactive. More than simply proactively providing *general* information on CRI activities, a number of actors offer explicit support to LEA prosecution efforts against actors engaging in extractive, often composite extractive, abuse externalities. A noted earlier, Spamhaus has a history of coordinating with LEAs and providing evidence for prosecutions. Recall the discussion of listing practices in Section 7.3.2.1. Building a case for attributing reputation is essentially an investigation comprising *a*) reports of or indicators of deviance (abuse), *b*) collection of evidence from as many sources as available, and *c*) weighing the available evidence to determine if attribution of reputation is warranted. This parallels an LEA investigation, highlighting the poten-

[638] Studies in (Dolšak & Ostrom, 2003b) also address global commons, the most well-known being environmental issues and function-specific management issues within that commons.

tial for alignment of reputation aggregators and LEAs. Here the common interest of the anti-abuse community is aligned with LE as a state-sanctioned social order clearly operating in the public interest.

RIR activities and reputation aggregator activities represent two general modes of proactive engagement along a spectrum of historical engagement modes. The degenerate cyber-libertarian mode has been to eschew engagement with external actors in fear of predatory rule. In the most innocuous case external actors are framed as meddlesome politicians whose lack of technical and operational knowledge limit the value of participation in collective choice processes.. In the worst case, external actors are framed "as the enemy," with whom engagement should be avoided and denounced. Variants of the cyber-libertarian model still hold in some network operator communities, but a more recent generation of leadership, what this work refers to as third generation leadership,[639] treat operations and common resource management as a business proposition that, while still avoiding undue regulation, must learn to both acknowledge and engage with government actors.

Reactive engagement is the case where CRI external relations actors' primary role is to engage when external actors request information. The initial framing of the request, be it for information, services, or actions perceived to be in the purview of the CRI, was historically constructed based on regulators' understanding of the infrastructure and the external actors' political motivations. While the content of information provided may affect external actors' agenda, the CRI may be faced with multiple problems: *a*) correcting fundamental operational misperceptions held by external actors; *b*) reframing the original request given corrections to misperceptions; *c*) servicing the reframed request. This process is difficult given the external actor's utility may have rested on the political implications of the original framing.

Engagement in this mode is further compounded by the need to avoid making regulatory actors, many of whom are non-technical, "look stupid" especially amongst their peers. Within the constructively critical operational epistemic community, critique, sometimes acerbic in nature, is par for the course. As per earlier discussion of community norms, this is type of critique is one way the community polices the quality of contributions. That said, this mechanism does not work amongst regulators—regulators are not participants in the community and need community knowledge transposed into actionable items salient to that regulator's issue space. In terms of engagement, this is a skill that even experienced policy entrepreneurs with substantive social capital within the community will need to learn before engaging externally. The root cause of this distinction is not merely different community cultures; Chapter 9 will highlight how differences the types of author-

[639]First generation leadership comprises the actors that designed the Internet and its early operational management institutions such as Vint Cerf and Jon Postel. Second generation comprises the network operators that managed the transition from NSFNet to the early commercial Internet. Third generation comprises actor that have developed their careers largely in the context of the modern CRIs. These actors learned from second generation, but, amongst the sample of leadership interviewed for this work, third generation has demonstrated an ethos that synthesizes substantive-purposive authority and business objectives less than the cyber-libertarian ideals that have been attributed to earlier Internet governance leadership.

ity give rise to these differences in engagement and problem solving. Under the reactive engagement scenario, the CRI is faced with the challenge of indicating to external actors they are wrong, correcting them, then providing information that, relative to their original request, may seem to have little to no utility.

A more proactive approach has been applied by all three of the CRIs to mitigate this. As per the comparison above, there are two coarse-grain modes observed, presented here as ideal forms referred to as proactive informational and proactive supportive. The latter is a superset of the former, but has distinct differences in regulatory agenda setting. Current CRI engagement lies on a spectrum with elements of both ideal forms. *Proactive informational engagement* communicates current CRI activities, changes in operational policies, and services. As stated here, proactive informational is a general information strategy—the scope of the content is broad and it is not necessarily tailored to the needs of the audience. In a number of cases, the opportunity to engage proactively has described "just getting to be at the table" as an accomplishment.

Proactive supportive engagement provides specific information, scoped to the needs of the intended consumer. Supportive engagement *can be* the result of reactive engagement or proactive engagement. As noted in the discussion above, reactive supportive engagement is burdened with reframing the problem presented by the would-be consumer. Proactive supportive engagement tailors anticipatory engagement strategies to preemptively correct, or even completely avoid, fundamental misperceptions in the context of regulatory agenda items that either address, or are adjacent to, Internet infrastructure issues. LEA requests are an canonical instance. Early requests were fundamentally misinformed regarding the RIRs' functions and capabilities. After meeting with RIR staff, these misperceptions were corrected. In this case, reactive engagement ultimately led to proactive supportive engagement, one scoped to a particular need. Later, as LEAs needed registry information, the RIRs developed a more proactive supporting program for training LEAs on how to use the registry.[640]

8.3.2 Transitioning to External Engagement

As discussed in Section 3.1, the NRS is valuable as a common resource supporting the Internet as a an infrastructure because it is an input to a growing set of public, private, and social goods. External engagement is one means to signal effective management of the NRS to external actors charged with managing public and social goods built atop, and dependent upon, this Internet infrastructure. This chapter has argued that, individually, the CRIs are facing challenges, but are largely stable. That said, there are substantive differences between how the CRIs operate and the character of their decision-making processes. In some cases, a combination of operational and superficial normative differences have created tensions between CRIs, here in particular the RIR system and the anti-abuse system. While analytically

[640]For instance, for a time, ARIN provided FBI agents with some training on how to use the registry for investigations. RIPE has engaged with LEAs via its RoundTable functions (RIPE NCC, 2014f). APNIC has also engaged in LEA training, see (APNIC, 2014e).

consonant, these tensions are a potential threat to ongoing efforts at developing explicit assurances.

The key capabilities of the NRS valuable to state regulators is its operational capabilities, credible knowledge assessment processes, and adaptability. The next chapter frames these as political capital. To understand how this political capital can be leveraged, the challenges to the sources of this political capital, knowledge assessment problems, are presented. With these challenges in mind, the remain sections highlight evidence of existing ad hoc cooperation and potential paths to developing explicit assurances of the alignment between the common interests of the NRS and states' commitment to the public good.

Chapter 9

Authoritative Knowledge Assessment

POLITICAL CAPITAL is defined as the "resources used by an actor to influence policy formation processes and realize outcomes that serve the actor's perceived interests," (Birner & Wittmer, 2003, loc. 3556–3557). In the NRS, political capital is rooted in common goods produced by its constituent epistemic communities: *a*) epistemic domain knowledge, *b*) credible knowledge assessment capabilities, and *c*) the resultant adaptive capabilities. As developed in the previous chapter, these goods have been produced by the NRS as a distinct social order in its pursuit of maintaining the integrity of the routing system. The last dilemma to be considered by this work is how this resource management regime will adapt to necessary engagement in the global political arena. As posed in Chapter 1's concluding section, are the incentives and resources of the NRS commensurate with the potential losses resulting from partial, or worse yet, systemic, failure? This includes evaluating the strengths and the limits of the NRS's adaptive capability as well as leveraging that capability as political capital. This chapter will explore this question by exploring how the NRS is leveraging its existing political capital to support external authorities while sustaining the processes and NRS social capital that fundamentally engender those capabilities.

Discussions of *a*) the norms and ethos manifest in Chapter 4 NOGs as the social substrate of operational epistemic communities; *b*) NOG-like communities amongst the RIRs, IXes, and SISCs in anti-abuse for discussing common problems; and *c*) collective choice problems in the RIRs, IXes, and anti-abuse. Each contributes to how information is shared to reduce uncertainty and remediate externalities. These are also manifestations of social capital. The notion of "capital" is used

> to be able to analyze the social world as an accumulated history that cannot be reduced to a sequence of mechanical equilibria [(Bourdieu, 1992, p. 49)]. [Bourdieu] defines social capital as the totality of all actual and potential resources associated with the possession of a lasting network of more or less institutionalized relations of knowing and respecting each other. [(Bourdieu, 1992, p. 63)][641]

[641]This quote is from (Birner & Wittmer, 2003, loc. 3505–3507). The original text of Bourdieu (1992) is in German.

The common goods produced by the NRS are the product of effectively mobilizing the social capital of NRS participants to create and maintain the resources supporting the integrity of the routing system.

Thus far, the NRS has operated quietly under the hood. Social capital has been leveraged to serve common interests. Moreover, CRIs have been explicitly cognizant to avoid impinging on public policy. In contrast to conventional international regimes prone to expansion of their scope, the NRS has self-limited to the scope of its substantive-purposive authority, namely, the routing system and messaging integrity. As such, the NRS's common interests have not run counter to the public interest. This is a weak form of harmonious alignment of interests. A path-dependent history of harmonious alignment between a commons and the public does not carry the explicit assurances of alignment resulting from explicit, credible commitment to coordination and cooperation efforts.

Following the comparison of authority types in Chapter 8, would-be principals in the global political arena lack both the social capital to manage the Internet's routing structure, much less the whole of the infrastructure. Synthesizing the theoretical threads developed thus far, social capital in the NRS is the accumulated capability to manage a quasi-voluntary complex of privately managed networks that cohere under loosely federated, function-specific common resource management regimes. Analytically, the substantive-purposive authority engendered by the NRS is complementary to state-base formal-legalistic authority. This chapter will argue that it is not only analytically complementary, but that it *can* function as a complement to existing external authorities rather than further raising the specter of "Internet governance" as a competitor in the transnational regulatory space. Independently, neither class of authority, the NRS nor state-based authorities, comprise a sufficient mix of authority *and* capabilities to manage the Internet as it exists today:

The alternatives are suboptimal. Purely technical management abdicates a responsibility to the public interest, in particular the implications of infrastructure decisions on downstream public, private, and social goods. As alluded to above, absent explicit assurances, continuing to rely on "weak" alignment creates political uncertainty for external authorities that are legitimately accountable to their publics' interests. Ceding resource management to states will arguably lead to fragmentation, at minimum between the economically liberal states that benefit from a global communications platform and those that wish to control it and the transnational public policy issue coalitions that it facilitates. Adding NRS stewardship to a state's portfolio of domestic regulatory interests will expose management processes to powerful short-term interests that will inevitably weaken, if not eliminate, extant credible knowledge assessment and adaptive capabilities. In effect, such aggressive predatory rule would likely eliminate precisely the characteristics that make the NRS a valuable steward of a high clockspeed infrastructure. Rather, to maintain a variant of the global Internet, while not unduly aggravating notions of transnational competition, the prescriptions outlined here focus on *a*) the means by which the NRS can proactively contribute its operational capabilities and knowledge base to state interests *without b*) overstepping the NRS's mode of authority, but while *c*) avoiding subordination that would compromise the knowledge assessment and

adaptive capabilities that make it valuable to regulatory and state authorities in the first place.

9.1 Knowledge Problems and Social Capital

Knowledge problems are endemic in the management of complex systems. In Section 3.2.1 knowledge problems were discussed in terms of coping with uncertainty in common resource systems in. Knowledge problems are present when policy making requires expert knowledge to understand potential outcomes but decision-makers face uncertainty when identifying *sources* of this knowledge. A conventional misconception is that objectivity is synonymous with isolation from political influence. In particular, the notion of objectivity has been appropriated from idealized recipes for progress: *a*) scientific progress leads to *b*) technological development that *c*) yields immediate social utility. This section outlines the problems faced by knowledge as a product of the social capital of NRS communities: *a*) misperceptions of how knowledge is applied, *b*) credibility of knowledge as a product of operational epistemic (social) capital to external authorities, *c*) commonalities with threats to other epistemic communities, and *d*) the need for state sanction to empower acting on boundary issues such as security and IPv6 deployment, where progress is in the interest of both NRS participants and state authorities.

Real systems combine complex feedback between technological development and operational knowledge. Feedback loops in the Internet infrastructure play out between Internet protocols and operations, between operations and CRI firms, and between CRIs, amongst CRI participants, and between NRS participants and industry suppliers. For resource management entrepreneurs within the NRS, effectively mobilizing this complex of feedback loops is an application of social capital, constrained by the contingent character of relational authority developed in the last chapter. Unfortunately, ideal images of objectivity perpetuate unrealistic and untenable assumptions about the generation of pragmatic knowledge in industrial and information societies. Rather, as presented in Section 3.2.4, knowledge assessment "considers the scientific and *other factual ... bas[es]* for policy choice," (McCray, 2003, Footnote 1, p.1, emphasis added). Further, McCray (2003) goes on to indicate "it reflects the collective judgment of a set of experts in the pertinent domains of knowledge, and not just the views of a single individual." Leadership in these fora have developed the credibility, the social capital, within their community to convene competing experts necessary for credible knowledge assessment.

In contrast, knowledge problems for capital "S" science do not have the same credibility problems, but do have credible assessment problems when placed in political arena. Recall from Section 3.2 that the "canonical" epistemic community comprises academics leveraging their "cognitive authority" as a source of power in the political arena. These epistemic communities also create endogenous social capital rooted in substantive-purposive order. In the political arena, the product of that social capital is domain-specific knowledge, much the same as the product of operational epistemic communities. Ultimately these two communities face similar

challenges. In the case of the NRS, operational epistemic communities must first demonstrate more fundamental credibility and a willingness to cooperate. Given that, they must then cope with the political interests faced by conventional epistemic communities.

Within the community, collective choice processes and less formally structured problem solving processes are credible because they are well-understood and have demonstrated utility. In the terms of the previous chapter, they contribute to the value of the social order. The problems faced by the NRS is the credibility of those processes to external actors. Thus, while these actors do have valuable credible knowledge assessment processes, they are not necessarily empowered to leverage them in the political process.

Part of the knowledge problem is the perception that scientific knowledge intrinsically begets technological advancement. As stated, the first part of the problem is rooted in skepticism of knowledge that is not conferred the mantle of "big S" science. The second part is the perceived deterministic character, that technological deployment and success is just a function of having knowledge to do a thing.

Social capital in the operational epistemic communities is rooted in substantive-purposive authority. Within such a social order, authority is not conferred, but, as per the definition, inheres in the actors that have capabilities of value to the "ruled." Formal-legalistic authority seeks justification by posing as substantive-purposive, backed by the objectivity of capital "S" science. Appropriation of knowledge becomes a political process that is leveraged to enhances an actor's position and interests. This type of politicization creates intrinsic tension with credibility. Complicating the problem further is uncertainty intrinsic in scientific knowledge. The result is a potentially vicious cycle pitting one group's ideologically colored expert against another. Not only does this cloud assessments of credibility, but in the worst cases may even erode institutions with the potential for credible assessment.

Recent public science and technology policy issues include *a*) food safety issues such as genetically modified foods, evaluation of mad cow, margarine; *b*) drug and medical safety; *c*) factors in climate and environmental change; *d*) disease externalities; *e*) particulate matter; all have, as above, a basis in conventional processes of generating scientific knowledge. Conventional processes often include academia and the perception of objective, third party knowledge assessment as a legitimate source. A wide variety of public policy issues are not only rooted in scientific processes, but also *operational* processes. Operational knowledge, created by operational epistemic communities, is often unique to communities' experience. Further, firms have valuable knowledge but may not be incented to either share that knowledge or have the tools and experience necessary to extract tacit knowledge and present it in a way useful to policy development processes.

Counter to the idea of ideal objectivity, a counterintuitive solution is to facilitate constructive conflict amongst contending "authorities." The contending actors operate *as* authorities, these actors hold substantive-purposive authority. Substantive-purposive authority characterizes epistemic communities, in particular operational epistemic. Discussions of consensus in the studies described and explained collective choice as an endogenous knowledge assessment process, but did not elaborate

the challenges of *maintaining* consensus as a knowledge assessment. The next section discusses the threats to knowledge assessment processes.

Knowledge assessment, both in the epistemic consensus described by P. M. Haas (1990) and collective choice, relies on a balance between contention and common, authoritative, image of order. Contention serves a number of functions in the assessment process:

1. paints a broader picture of the potential knowledge space,
2. highlights common base knowledge,
3. highlights gaps in common knowledge base and points of contention,
4. provides insights into which elements of the contentious domain is in the hands of which actors and their incentives.

Constructive conflict helps identify the sources of uncertainty, but as described, is, in its degenerate form, unstructured and opportunistic at best. The role of institutions facilitating credible assessors is to make constructive contention a repeatable, dependable process for informing and adapting policy. Such institutions build on what Ostrom referred to as an authoritative image, what Haas referred to as "the collective beliefs of transnationally organized networks of knowledge based communities" (P. M. Haas, 1990, p. 347).

Key challenges for contributing to public policy is to *a*) making knowledge accessible to policy makers, *b*) demonstrating the legitimacy of the process, *c*) sustaining an institutions authoritative position as the source of knowledge, and *d*) establishing comity between two distinctly different modes of authority. The former two are a matter of diplomacy and engagement, discussed in terms of modes of engagement in Section 8.2.5's. Providing information is not sufficient, though, it does not in and of itself to create subsequent commitments for either the current stewards of the NRS, nor the stewards of the public goods that increasingly rely on the NRS.

9.2 Barriers to Credible Assessment

What is referred to here as a knowledge problem has factors that have been framed under multiple disciplines. One component of the knowledge problem is the politicization of claims. Adaptive regulatory processes discussed by McCray and Oye (2006) speak to the capacity for anticipation and adaptation in domestic and international rule making. Uncertainty in "turbulent" technical environments mean that "norms, rules and procedures embodied in an international regime or domestic regulatory system will be inevitably wrong," (McCray & Oye, 2006, p. 3). Rather than framing policy as strictly authoritative, McCray and Oye (2006) frame rule making and experience in the exercise those rules as "policy experiments" and that "[p]roperly designed regulatory systems harvest information generated by policy experience, and use that information to revise and update policies," (2006, p. 3). McCray and Oye (2006) offers studies that evaluate how effectively rule makers

use policy experience—where does adaptation occur in conventional governance regimes and under what conditions?

The opposite end of this spectrum is knowledge creation in ad hoc communities. The notion of an epistemic community offered by P. M. Haas (1992) is often framed in the context of professional or scientific communities' power contributing to policy. Despite this common application, Haas also offers a notion of epistemic community that is rooted in intersubjective understandings, rooted in "a shared way of knowing" and "shared patterns of reasoning," (P. M. Haas, 1992, Footnote 5 on p. 3). In terms of knowledge problems, epistemic community members may be the sources of knowledge contested in conventional regulatory agencies addressed by McCray and Oye. These actors may also leverage their access and control over knowledge as a form of power derived from their role an authoritative source of knowledge. The range between legitimate source and authoritative knowledge broker is in part a function of the "turbulence" of the knowledge domain and the degree of contestation within that domain.

While epistemic communities may share common mental models, and even a common paradigm,[642] their prescriptions may differ. Amongst actors with a shared mental model of that knowledge generation process, uncertainties intrinsic in those prescriptions may be acceptable within a given domain. Borrowing from the language of common resource management that will be addressed shortly, having a common image of what constitutes management goals can offset conflicting interests. Such differences may create contention with in the community, but do not necessarily limit the scope of influence.

Consider the narrative of the Med Plan offered by P. Haas (1989). Increasing pollution of the Mediterranean was a collective action problem—more precisely, P. Haas (1989) characterizes it as a set of localized collective goods problems.[643] Lacking knowledge of marine ecology and other salient fields, UNEP (United Nations Environment Programme) enlisted the academics from a variety of fields related to marine pollution: "geology, oceanography, ecology, meteorology, chemistry,"[644] among others. P. Haas (1989) describes the function of ecology in the sustenance of an epistemic regime:

> [E]cology is fundamentally a framework in which other disciplines may be assembled... it facilitated the formation of coalitions among scientists, because most contending views about what are important research questions and the appropriate levels and methods of analysis may be integrated within such a broad framework. By promoting the adoption of a very broad definition of "pollution" which emanated from an

[642]The notion of a paradigm is in the sense of Kuhn (1993). Actors may share a common mental model about effective means to create knowledge (methods, approaches, and ethical considerations) while differing on the paradigms that emerge from the application of these mental models.

[643]This is salient in general because much of the analysis of the number resource system draws on ideas rooted in the management of collective goods. It also avoids the common pitfall of claiming, like some climate and environmental issues, that this is a public goods problem rather than a collective goods problem.

[644]At P. Haas (1989, p. 385), originally from (World Resources Institute, 1987, pp. 163).

> ecological perspective, UNEP and members of the ecological epistemic community were able to encompass more parochial interests under its umbrella. (P. Haas, 1989, pp. 385–386)

The "holistic" point of view offered by ecology provided sufficient common objectives, a common image of order that gave primacy to contributing to those objectives rather than quibbling over methodological differences.

Jasanoff offers a different narrative of science in domestic and international arenas. Placed in a different context, uncertainties amongst differences may be exploited by political actors expert in binding authority to rules that further legitimize one prescription over another. Jasanoff discusses this process in terms of the conventionally considered epistemic community, scientists:

> [E]xposés of uncertainty and disunity in science undermine public confidence and raise troublesome questions about whether scientists really deserve the symbolic and material rewards they have claimed from society in this century. Scientists, especially those whose work impinges on policy, thus have much to lose unless they can safeguard the classic normative view of science against charges of excess indeterminacy. To shore up their claims to cognitive authority, scientists have to impose their own boundaries between science and policy, thereby coming into potential conflict with policy-makers pursuing opposing interests. (Jasanoff, 1987, pp. 198–199)

This is a contrast to the success of the Med Plan presented by Haas.

Haas highlights the prestige in the domestic arena that comes from having participated in the Med Plan and UNEP proceedings:

> The external support from UNEP enhanced the scientists' domestic prestige and strengthened their domestic political base. Although their work was only loosely coordinated by UNEP, the knowledge gained through collaborative efforts established or reinforced their authority in the issue-area of marine pollution control. (P. Haas, 1989, p. 387)

Prestige and funding opportunities are selective incentives. These further enhance fidelity to (and of) the common objects framed under the umbrella of ecology in this particular scenario. Not only does this process contribute to coherence, but it also confers cognitive and policy authority onto knowledge claimants cohering around the common edicts of ecology both scientists and policy-makers identify with.

Haas argues one of the most compelling elements of the Med Plan was coordinated domestic application. In contrast, Jasanoff warns that political actors will identify "uncertainty and disunity." Through processes of deconstruction of knowledge, rulemaking processes can reduce the legitimacy of particular epistemic community members or groups of these actors that "com[e] into potential conflict" with existing interests. In conventional political arenas, epistemic community members are exposed to political actors that will leverage differences in interpretation, intrinsic uncertainty in scientific attestations, and methodological differences to both

serve their pre-existing interests, to make way for the (re)construction of a scientific narrative that fits their goals, or to simply undermine competitors. Depending on the domain, Haas's common "umbrella reigning in parochial interests" may make the collective more cohesive, but may not mitigate exploitation of uncertainty. This not only has ramifications for immediate outcomes, but, as suggested by Jasanoff above, it may damage the scientific (epistemic) community itself.

Haas and Jasanoff provide two narrative of epistemic actors engaging in policy making activities. Haas offers a story of power wielded by a coherent epistemic community unified under the cognitive authority of ecology. Jasanoff offers a narrative of scientists defending the cognitive authority of science from actors that wish to appropriate the cognitive authority and legitimacy of "objective science" to reinforce the "rational justification" of regulatory decisions. Haas focuses on epistemic communities as regimes, the unit of analysis is the epistemic community and its members. Outcomes are framed in terms of the success or failure of the epistemic community influencing policy (P. Haas, 1989, pp.) For Jasanoff, the unit of analysis is science and the process of deconstruction and reconstruction at play in rulemaking processes. Jasanoff "examines three contested boundaries that have acquired special visibility," namely *a*) 'science' in contrast to 'trans-science' or 'science policy,' *b*) distinctions between 'risk assessment' and 'risk management,' and *c*) notions of 'peer review.' Each of these speak to how the process of doing science, here generalized to creating knowledge, affect distinctions between the former two and the role of the latter as decision criteria (Jasanoff, 1987, p. 200).

These two elements are not the only factors in solving knowledge problems. As a discussion of knowledge problems in complex systems, the arena[645] is yet another component. Here, the arena is explored empirically through institutionalized consensus processes in operational epistemic communities managing Internet number resources (Section 5.6.2). Ostrom's notion of an action arena comprises participants and a situation. Haas speaks to the epistemic character of participants and the Med Plan as a particular situation. Jasanoff speaks to the effects contestation within the situation (a rulemaking process) has on a particular scientific endeavor.

The arena itself is a more durable institutional setting in which participants engage and the attendant rules shape how knowledge problems are treated. The narratives offered by Haas and Jasanoff are particular analyses of how different arenas, both across epistemic communities and political arenas, intersect. The primary arena of interest here is the set of institutions that are developed by epistemic communities (be they scientists or other knowledge claimants) that sustain that set of beliefs and expertise within that community. Most notably here are the NRS institutions, but Jasanoff's reference to peer review highlights the role of academic institutions. These arenas comprise actors and groups that serve as knowledge sources. The canonical source is academic actors that engage in the study of a domain. In the case of the NRS and a number of historical and contemporary CPRs, epistemic communities comprise operational actors that distill experience with a (resource) system into a shared body of knowledge.

[645]Here the notion of an arena refers to an action arena in the sense of (E. Ostrom, 2005).

In contrast to epistemic communities, McCray (2003) offers a characterization of arenas purposely constructed to explicitly assess the credibility of knowledge that may serve as the basis of policy making decisions and rationales. As alluded to earlier, the notion of the "objective" oracle insulated from political influence is an unfortunately appealing fiction. McCray's analysis identifies "channels for supplying knowledge for policy" (McCray, 2003, p. 2) that are relevant to analyzing the NRS.

1. The first channel is agency staff members, which may either be active members of an epistemic community or ex-patriots of epistemic communities that apply specialized domain knowledge as credibility assessors. McCray highlights two limitations, external skepticism of impartiality and keeping staff up-to-date on the latest research. In terms of the epistemic community framing, one immediate question is how closely tied to the epistemic community in question is the staff member? Is the staff member full time or, in the case of many academic research centers, are they an "expert affiliate" that is still a very active member in the epistemic community? Either case can be attacked on impartiality grounds. Whether a contributor is up-to-date may be a factor of proximity to, in terms of participation and engagement with, the community as the source of knowledge. This is especially the case in operational communities where information may be shared through semi-formal presentation, informal back-channels, or may simply be a function of densely interconnected collectives of actors continuously engaged in problem solving activities. This notion of proximity and participation is a key element of the analysis of NRS institutions as both sources of knowledge and credible assessors.
2. Notice and comment provisions are the second channel highlighted by McCray. Notice and comment periods "during which policy options are laid out for comment by interested and affected parties," (McCray, 2003, p. 2). In the NRS, in particular in the RIRs and the IXes, various forms of notice and comments periods are used to elicit reaction to and evaluation of potential policy and strategy proposals. In the RIRs, these are explicitly codified in the formal structure of policy development processes. In the IXes, proposals are often presented in membership meetings, and commentary is collected via the variety community communication channels, including informal community interaction and back-channels. In many cases, consensus is an informal mechanisms for gauging community interest in and support for a formal change requiring a vote. McCray (2003) indicates "the real and perceived effectiveness of this channel depends on the agency's in re-examining such claims fairly." A number of factors affect that alacrity, among them how closely the agency's responses are followed and the capability and willingness of actors to sanction the agency for poor performance in responding to feedback.
3. The third channel is an impact statement that describes "the anticipated effects of policy choices," (McCray, 2003, p. 2). McCray indicates these are "on the wane" amongst agencies in his studies. In the NRS institutions, policy and strategy development processes include explicit impact statements as well

as history for evaluating consensus decisions in terms of outcomes' effects on NRS institutions and the (number) resource system as a whole. Two modes of impact are evident in the NRS: effects on the system itself and effects on facilities that manage and enhance utilization of the system by participants. NRS policies are often require a problem statement, the attendant policy for which proposes a solution and its impact on system participants. As such, the former impact is often part of the comment periods in consensus and deliberation processes. The latter impact analysis is performed by actors that maintain the management resource, such as the registry or the IX fabric; further, these are often in terms of operational burden and costs that will ultimately be born by the epistemic community, eliciting an evaluation of whether the problem and solution pair warrant the cost.

4. McCray's final channel is for-profit and non-profit contractors. This essentially outsources the analysis, distancing it from the agency using the analysis as the rationale for a particular policy and its expected outcomes. In the epistemic community framing, a key question is why one group is chosen to perform the analysis over staff or some other group. In the NRS, it is not uncommon to see a task forces or working group established to perform a particular analysis. This is not outsourcing to an external actor, but it follows the essential delegation of authority to an agent that is considered better suited to the task, here a credible assessment of a particular issue. Within the NRS, RIRs and anti-abuse communities like M^3AAWG have both standing and ephemeral, situation-specific working groups that could be considered to perform a similar role as a contractor.

Above, McCray's channels provide an entry point for comparing the elements of knowledge assessment with the operational epistemic communities in the NRS. In all of the narratives from the literature thus far, most expertise is references communities of scientists. McCray's definition admits "other factual information" . In other communities, expertise, more generally knowledge, is derived from systematic investigation and analysis that does not conform to conventional scientific norms, to informal processes, and experience with the system. The notion of science as an impartial endeavor apart from partisan ideologies gives it some of the privilege and symbolic rewards Jasanoff suggests in general, but also specifically in the quote above. Moreover, the privileged place of science often comes with the assumption that it is knowledge created *in service of* the good of society. This argument is also manifest in the prestige acquired by participants in Haas's narrative of the Med Plan. One result is that science is offered up as de facto credible and legitimate and should be shared for the betterment of mankind, one means of doing so being informing policy.

An unfortunate result is that other sources of knowledge are considered second class, if not actively undermined by actors in the scientific community. Perhaps the most well-known articulation of the problem of one class of intellectuals subordinated to another is Snow's Two Cultures (Snow, 1964). Although this is a much larger topic in general, the lesson for this discussion is that sources of knowledge

are pluralistic. Other sources of knowledge are considered valuable to private use, thus, even though this knowledge may be valuable for the development of public policy, it remains largely inaccessible. Yet other information and modes of creating knowledge are tacit in the operational experience of actors engaging with the system.

The contrast with other modes of credible assessment provides a framework for reasoning about how modes of specialized knowledge creation outside conventional scientific arenas can be used in service of *public* policy. The common pool resource literature has identified the role of community knowledge. The canonical scenario is longstanding resource management regimes such as the management of pastureland, fisheries, and water resources. Evaluated in the context of resource management (operational) rules these informal institutions have been characterized by specialized knowledge necessary to manage those systems. Rather than evaluating knowledge based on the process by which it was created, knowledge is evaluated in terms of outcomes. This does not mean that science is not evaluated on outcomes, but rather the process is often the focus in terms of credibility. In common resource systems, institutional arrangements contribute to ensuring outcomes are observable to those charged with monitoring and enforcing operational rules (resource policy) as well as resource users beholden to that policy.

Knowledge created by users of CPRs are often embedded in the operational rules. In some systems, knowledge accumulates with use and experience. For instance, Ostrom indicates that:

> Rules devised by resource users are based on years, decades, and sometimes centuries of experience in using a common-pool resource. Such information is gleaned while engaging in everyday harvesting activities. Fishers learn which spots in a fishing ground are most productive and which areas of the grounds are most compatible with various types of gear, by fishing day after day. Consequently, the rules that resource users devise are well matched to the physical environment in which they will be used. (1996, loc. 2758–2761)

Experience is a valuable tool referenced again and again in the CPR literature, but it does not necessarily have to be ad hoc.

In CPRs, and the NRS in particular, credibility may be considered a function of effective outcomes that demonstrably preserve the integrity of the system. In conventional CPRs, this means the resource is not exploited and/or depleted. Indicators of exploitation and depletion are often embedded in the operational rules. As a set of pragmatic (teleological) rules, the objective is to ensure outcomes the community recognizes from experience do not damage the resource. The continued observation of effective function is often the criteria for adjudicating these rules.

Another property of CPRs, in particular the NRS, is the role of eliciting knowledge as a means to updating rules. Earlier a distinction was made between information that may be locked away for private use versus information and knowledge that can be gained that can improve public policy. Collective choice rules in CPRs necessarily elicit knowledge of the system to create effective (pragmatic) rules that

govern utilization of that system and the types of modifications that may be made to enhance utilization.[646] These rules often require eliciting knowledge that has been passed down from generation to generation and the application of this knowledge to change at hand as a means to evaluate foreseeable outcomes. Part of this process requires actors either distill or share private knowledge related to resource utilization in contributing to modifying operational rules. The degree to which actors share information varies based on the character of the system and whether CPR managers of institutionalized data collection and evaluation processes are incorporated into operational rules.

The result of teleological knowledge assessment in CPRs is a form of continuous monitoring not only of the resource, but also of the efficacy of the operational rules. This is arguably an implicit form of adaptation. For instance, weather is a source of uncertainty in ancient and modern irrigation systems. CPR-based systems have one set of rules for the "common" case, but rules that have been developed as corner cases were also identified by communities. For instance, an irrigation system reported by (E. Ostrom, 1990, loc.) reports that in times of drought a set of allocation rules rooted in minimizing crop loss may temporarily displace rules based on a periodic or semi-random appropriation schedule. The former recognizes some crops are more susceptible to drought than others or may be at a more susceptible point in their growth. The latter applies a principle equitable distribution based on impact and need rather than simply applying a rules driven by a strict schedule and pre-defined volume allotments. These more nuanced notions of equity reflect experience with losses from draconian application of "fair" allocation practices ill-fit to particularistic conditions. A shared knowledge of which crops are sensitive to drought have all contributed to this adaptive set of operational rules.

McCray and Oye (2006) discuss knowledge problems in terms of anticipation and adaptation by regulatory agencies. While the there are substantive differences between modern agencies and common pool managers of rural irrigation systems, there are lessons to be shared. Operational rules can be roughly partitioned into two groups: *a*) normal operating conditions and rules and *b*) extraordinary. In the case of irrigation systems, extraordinary includes, among other things, drought conditions. Irrigation management rules are, in this case, adaptive. The learning process and application to extend operational rules from normal to include extraordinary, is evidence of anticipation, in particular learning from the experience of previous operational experiments.

The question for anticipation in this particular scenario, described by McCray and Oye (2006) and McCray et al. (2010), is whether anticipation is ad hoc or whether is it actively planned. Ad hoc approaches may be reactive. In the irrigation instance above, experience with a drought stimulated resource users to reconsider the efficacy of rules intended for normal operating conditions. Planned adaptation

[646]Enhancing utilization typically falls under the umbrella of management rights, those rights to change the structures that facilitate use of a resource or altering the structure of the resource itself, such as altering the coastline that provides access to a fishery or developing wells and water storage systems. Management rights are discussed extensively by); for this work it is sufficient to note these rely on maintaining teleological knowledge of the system.

involves anticipation of and provisioning for reviews of existing rules. One mode of anticipation is routinized reviews. For instance, McCray and Oye (2006) and McCray et al. (2010) describe the Clean Air Act (2010, pp. 4–5) having explicit, periodic reviews of rules. "The Act, as amended, specifies that each of the[] standards [regulated under the Act] be subjected to *fresh scientific reviews* every five years," (2010, p. 4, emphasis added). As a mode of anticipation, resources must be dedicated to "fresh scientific reviews."

Anticipation does not necessarily mean routinized. McCray et al. (2010) goes on to highlight a different mode of planned adaptation in air transportation safety:

> The "planning" in this form of Planned Adaptation is not a general knowledge assessment tied to a scheduled review cycle. Instead, it is the provision, in advance, of ample investigatory capacity (seen in the NTSB "go-teams" that are quickly mobilized to examine every civil aviation accident) to enable diagnosis of problems when they arise.

"Go-teams" comprise experts whose job is to not only evaluate the proximate causes of an aviation accident, but also to distill information collected into knowledge that contributes to improving safety regulations.

Anticipatory provisioning highlights that ad hoc evaluation and planned adaptation is not a strict dichotomy. In both the case of irrigation systems and aviation, hazards, droughts and aviation accidents, are uncertain events. Anticipatory provisioning has the effect of proactively preparing to maximize the information and knowledge generated when hazards manifest. The result is that the reaction not only copes with the hazard itself, but, as an observation in the policy experiment, can contribute to adapting rules. In case of aviation hazards, unforeseen hazards from faulty equipment or poor operational practices can be identified, evaluated, documented, and reflected in safety rules, potentially avoiding future hazards altogether.[647] In the case of drought, mitigating the occurrence of the hazard itself is out of the scope of operational rules, but mitigating the effect of the hazard *when it does* manifest, such as recognized alternative allocation mechanisms, is well within the scope of operational rules. In this latter case, anticipation provisions for learning from uncertain and unavoidable hazards in order to better mitigate subsequent hazards.

The NRS offers yet another mode of anticipatory provisioning rooted in management by operational epistemic communities. Each of the NRS institutions is largely comprised of members of a particular epistemic community in the larger Internet infrastructure operations space. Consensus processes are sets of collective choice rules that continually re-evaluate operational rules that set out resource utilization rights.

In discussing the substantive variance in credible assessment organizations, McCray (2003) highlights that credible assessment may be one of a number of functions performed by the organization. In the NRS, credible assessment in consensus processes

[647] In the dissertation, avoidable hazards rooted in faulty equipment and operational practices are classified as operational externalities. Here, depending on the incentives at play, may be more appropriately characterized as safety externalities.

is a joint endeavor by the community and the firm managing an particular resource. Credibility is susceptible to a number of attacks. These organizations do not necessarily claim to be engaging in scientific analyses[648] and may even eschew that model as ineffective in contrast to more pragmatic approaches that are known to yield effective outcomes. In the NRS and similar contexts, credibility must demonstrate accountability and that assessments are not susceptible to capture by particularistic interests. In the NRS, threats to credibility mean capture by industry sub-sectors (such as access, CDNs, colocation, etc.), governments, or other NGOs in the larger global political arena.

9.3 Scope of Authority

The NRS clearly has adaptive capability in the technical realm. A number of contemporary issues have pushed on the scope of NRS authority, testing the boundaries of these adaptive capabilities. These boundaries do not diminish the capabilities, but do highlight where the engagement with external authorities is perceived as a potential threat to stability or where engagement has been confounded by the fundamentally different perspectives of NRS and external authorities. RPKI is an instance that has, as per Section 5.7.4 confounded the community: fears of predatory rule have been discussed extensively. The RIPE NCC's surprising actions during Operation Ghost Click[649] is a particularistic, but telling choice between cooperation and common interests.

Some cases, such as LEA engagement with the RIRs, started with contention, but has developed into effective relationships. Other boundaries illustrate quite complementary capabilities. IPv6 illustrates selective instances of leveraging governments convening power to promulgate infrastructure development, a positive. LEAs and anti-abuse also have a rather constructive relationship. These issues are revisited here as illustrations of the scope of NRS authority, where state engagement has empowered NRS institutions, and where there is still work to be done.

9.3.1 Training and Development

Each of the RIRs does some degree of training. The baseline training is largely geared toward educating the community what the registry does and how to use the registry. In effect, these efforts train participants to update the registry as part of their responsibilities as a steward of their number resource delegations. Training includes engaging with registry facilities (services), IPv6 protocol training, and

[648] In particular, one analytics group is quite adamant about distinguishing its analysis from science, arguing the focus is pragmatic information provided to the community in service of consensus processes, but necessarily requires an experienced actors prior knowledge for effective interpretation. This is developed more in the dissertation, but should be one entry point in reasoning about the difference between teleological knowledge creation and the privilege enjoyed by modes that are classified as scientific and, subsequently, have lower bars for credibility in some dimensions and amongst those wishing to leverage those analyses as rationales for partisan policy making.

[649] This is discussed with supporting evidence in Sections 9.3.3.1 and 9.3.3.2.

training on how to use RIR provisioned RPKI services. In regions with a number developing states, namely AFRINIC, LACNIC, and APNIC, training also includes operations training that parallels and/or supplements the education initiatives in the NOGs. In these regions the RIRs and NOGs are filling gaps in these economies' technical capabilities.

In the LACNIC and APNIC regions, these RIRs have developed relationships with support organizations and offer resources supporting development in the region. For instance LACNIC has a number of development projects it supports (LACNIC, 2015). Project FRIDA is the Regional Fund for Digital Innovation in Latin America and the Caribbean. In its capacity promoting IPv6, LACNIC has recently met with a number of state incumbents and state governments to educate and elicit support for IPv6 deployment;[650] the mechanics of IPv6 development are discussed in the next section. In the AP region, APNIC has worked with the ITU-D, in particular the "ITU Regional Office and its Centre of Excellence to organize a series of workshops on IPv6 migration strategies," (APNIC, 2015a). APNIC has also worked with ITU-D on regional development. In addition to the ITU, APNIC has relationships with a number of other AP region and international organizations, including: Asia Pacific Telecommunity (APT), the Asia-Pacific Economic Cooperation (APEC) Telecommunications and Information Working Group (TEL) as an observer, Association of Southeast Asian Nations (ASEAN) Telecommunications and IT Ministers Meeting (TELMIN) "where guidelines APNIC helped develop with APEC TEL in 2012" were adopted, as a founding member of the Internet Technical Advisory Committee (ITAC), and the Secretariat for Pacific Community (SPC) (APNIC, 2015c).

In terms of engagement between NRS authority and state authority, development efforts such as these are quite complementary. Development efforts such as promoting IPv6 are a canonical role of an epistemic community in general. Many of APNIC's engagements above are in APNIC's role as an invited IPv6 expert. As will be discussed in the next section, much of the focus of this work is to inform government actors that IPv4 is now effectively depleted and that they should encourage IPv6 deployments in their states. In terms of the modes of engagement, this is proactive, supportive, but largely general, engagement. In general, this is mutually beneficial: APNIC provides valuable insight into IPv6 deployment, using this to encourage deployment. In effect, APNIC is encouraging state actors to leverage their authority to promote IPv6.

9.3.2 Convening IPv6 Promotion

Howard and Sowell (2014) is study of IPv6 deployment case studies. Howard and Sowell (2014) identified multiple configurations of state-based and private actors (actors in the IPv6 operational epistemic community) that have contributed to IPv6 deployment: *a*) government on its own, *b*) NRS participants acting as catalysts, and *c*) government as a legitimating convener, (2014, pp. 17–24) "[W]ith the ex-

[650]LACNIC has recently visited Trinidad and Tobago and Peru (LACNIC, 2014d), Ecuador (LACNIC, 2014b), Colombia (LACNIC, 2014a), Panama (LACNIC, 2014c), Venezuela (LACNIC, 2014e),

ception of Singapore's funding of IPv6 development," (2014, p. 19) the incentives in the United States, the Czech Republic, the European Union, and Brazil are largely driven by purchasing and contracting influence. The efforts of LACNIC, the Brazilian NIR NIC.br, and APNIC, such as those described above, are instances of NRS participants acting as catalysts.

The most interesting is the third, government as a *legitimating* convener. Consider the go6 Institute, founded by Jan Žorž in Slovenia. Žorž and go6 are regulars at RIPE and other NOGs, both as a source of and contributor to IPv6 deployment as its epistemic domain. On coordination with government officials, the assumption was that go6 was seeking funding. Rather, go6 indicated what would be more valuable would be the government's role as a legitimizing convening agent, to encourage participation by C-level actors engineers and NOGs do not necessarily have access to.[651]

As per Howard and Sowell:

> A critical element of this mode of engagement is that Žorž convinced the agency to participate *without* driving an agenda. The government sent invitations to major content companies to participate in an IPv6 Summit. Since the invitations came from government, they were routed to C-level executives, who had to ask their engineers for information and advice about IPv6. C-level executives wished to participate and look informed. Go6 held semi-annual or annual meetings, at which each participant would describe their IPv6 status. Firms with nothing to report lost prestige. To garner prestige, participants would attempt to have something positive to report. (2014, p. 19, emphasis in original)

A number of complementary elements are at play here. Government lent its legitimacy to facilitate an operational epistemic community's common interest and its own public interests. Coordination made engagement and shared interests more explicit. C-level actors became aware of the issue; moreover, engagement with engineers may have facilitated longer lasting effects within the firm. Finally, recall Haas's discussion of prestige in the MedPlan. Here, regular meetings to report IPv6 status was an opportunity to either garner prestige in the IPv6 operational epistemic community or lose reputation.

The case of Belgium is another instance of members of the IPv6 operational epistemic community engaging with regulators to improve the operational environment. Éric Vyncke and Gunter van der Velde from Cisco engaged with "the communications regulator, the economic ministry, and the federal policy to argue that [carrier grade network address translation (CGN)] was an inferior solution to IPv6," (Howard & Sowell, 2014, p. 20). Vyncke and Velde appealed to the public interests of those agents, arguing that *a*) CGN, having higher operational costs than public IPv4 or IPv6 addresses, was unfair to new entrants (economic argument) and *b*) CGN impinges on LEA investigations, making it more difficult to attribute log information to individual end users. The solution was to limit CGN to 16 users per

[651] This is not a unique situation, it will be revisited shortly.

live address, incenting ISPs and mobile carriers to deploy IPv6 instead. In the small Belgian IPv6 community, Vyncke and Velde would share the success of ISPs with others, fostering competition for prestige (Howard & Sowell, 2014, p. 20) .

Two additional cases reported by Howard and Sowell depict combination of RIR and government efforts in Bolivia and Costa Rica. In the case of Bolivia, the state initially engaged the RIR to participate in IX development. At the time, Bolivia's infrastructure market was fragmented amongst local and regional providers. Leveraging the potential for engagement, LACNIC informed the regulator of the dearth of IPv6 deployment in the region. The result was meeting the regulator made mandatory for a number of the larger ISPs. LACNIC first presented IPv6, then IX development strategies to the regulator as a neutral convener alongside high ranking engineers and finance staff from ISPs. Afterwards, a number of ISPs began experimenting with IPv6 deployment. In Costa Rica, the story is similar: the convening agent was the local regulator and a longstanding ISOC participant, convening the meeting at the local university as a neutral location.

Part of this story highlights a deficit in the community that both APNIC and LACNIC are trying to fill: both institutions have substantive access to the operational epistemic community, the engineers, but the IPv6 advocacy message does not seem to be reaching C-levels through these channels. Borrowing government legitimacy is one strategy. As noted in the previous section LACNIC has made an effort to contact governments and C-levels in a number of countries; similarly, APNIC staff have done a "tour" of AP region states to encourage C-levels of large backbone providers and regulators to encourage the deployment of IPv6. In these cases, prestige and economic competition play out at the state level. For instance, APNIC staff held up a Bangladesh backbone provider as a case where they could easily "turn on" IPv6, but had not yet had the incentive. Within a few weeks of the visit from APNIC, this provider had activated IPv6. This kind of state based competition has been described in the Middle East as well, where RIR staff share the IPv6 success stories of countries with other regulators in the region.

The issue of empowerment by governments, especially with their access to corporate decision makers is becoming increasingly salient. One RIR board member lamented the what he considered a waste of resources on "feel-good" propaganda such as IPv6 stickers and similar promotional materials at the RIRs. Although a seemingly low lying issue, it highlights prestige and a problem with the limits of prestige to the operational epistemic community. The stickers, typically placed on the lids of laptops, are visible to all the other engineers supporters engage with, in the fora that are already advocating. Such advocacy is not reaching salient external authorities. This is a simple, but easy to understand illustration that social capital does not immediately translate to political capital. In contrast, RIR efforts to proactively engage and support government efforts to deploy IPv6 is an early instance of developing and leveraging political capital.

9.3.3 Engagement with LEAs

RIRs, IXes, and anti-abuse have had experiences engaging with LEAs, each different in tenor and intent. RIRs initially had a rocky start, with LEAs demanding capabilities the RIR simply did not have. IXes have been the focus of data collection, but helped LEAs understand tapping at the IX would be overkill and there are points closer to the subject better suited for the purpose. Anti-abuse has the good relationship with LEAs, in no small part because their common interest is quite well-aligned with the public interest served by LEAs. The following sections review the relationships with each, highlighting points of friction and complementary relationships.

9.3.3.1 RIRs

A number of the RIRs have developed effective relationships with the LEA community. ARIN, RIPE, and APNIC all have well documented engagement with the LEAs that will be discussed here. The relationships were not all that rosy at the outset, though. Early on, the relationship between ARIN and the FBI started out a bit rocky, but has since developed into a complementary balance. In other regions, an FBI agent that frequents various RIR meetings and ICANN related that the community was at first suspicious of why he was engaged, but has since accepted him as a trusted contact with the LEA community.

Consider an early engagement between ARIN and the FBI. The FBI had a case for which it wanted to shut down a particular network. A couple of agents went to ARIN and demanded that it shut down that network. ARIN responded that first, it did not have that power and, if it did, it, as the RIR firm, did not have the authority to shut down that network. After some tense discussions, ARIN staff were able to convey to the FBI that ARIN's primary technical function was the delegation of identifiers. ARIN stressed that it did not control route dissemination or traffic flows.

ARIN did use this situation as a teaching experience. ARIN explained to the agents what it did do, in particular the use of the numbers registry as a "contact list" for networks that could disable particular number ranges. In time, the relationship with the FBI and other LEAs improved. The AGWG was developed as a forum to keep agents of the state appraised of ARIN activities and policy changes related to their (public) interests. Further, for a time, ARIN provided the FBI with training on how to use the number registry in its investigations. This has been the topic of at least two discussions in ARIN membership meetings. In 2005, Bobby Flaim, an FBI agent, explained at ARIN XVI how the FBI used registry information and the training on registry information different FBI squads have (ARIN, 2005). In that discussion Flaim indicated that:

> the way that [the number registry is] set up now we can actually get a lot of good information from what's on there now. And I think one of my main points was that that's good. Obviously in a perfect world part of our job is to get the most information we can as expeditiously as we can. And what we have now is we're able to get a lot of public

> information in a quick and efficient manner and obviously we want to keep that. (ARIN, 2005)

A more recent instance of engagement is at ARIN XXX[652] in the discussion of a proposal to make certain registry data private, and thus less available to LEAs. This proposal was presented by its ARIN Advisory Council shepherd, Scott Leibrand. As with many proposals, it was contrasted with its active counterpart in the APNIC region. That said, the proposal did not receive much support; AC member Stacy Hughes immediately asked if Flaim was in the room and could comment from the perspective of LE. Flaim indicates that:

> As international enforcement, the reason why we review the Whois and an open and accurate Whois as kind of an integral triage tool for our investigations is that that's kind of where it can start off.
>
> So once you put IP addresses or domain names on the domain name side under a proxy system, which is kind of what this is, or private system, you're adding another layer to our cases. (ARIN, 2012)

A number of experienced operators in the community voice support, indicating that the proposal should be shelved. In particular, one actor, Kevin Blumberg went as far to indicate that

> But my most important point to all of this is we are a self-regulating industry. And my biggest concern is if we use a big hammer and swing privacy one way on this, we can find ourselves going from self-regulated to government-regulated WCIT or whatever the case may be, so that really scares me.
>
> So I'd rather be fair than regulated. (ARIN, 2012)

The net result was that the policy did not pass, but the contribution of both Flaim as a representative of LEAs is an illustrative instance of state agents participating in the RIR process. Further, many of the other comments balanced the common interest, such as Blumberg's, to avoid regulation, but also recognized the LEA's operational need. It should be noted that one of the elements of Flaim's comment was, if this were to pass, it would like result in more subpoenas, raising the specter of additional costs for network operators.

Shifting gears to a particular instance of engagement, consider the interaction between the RIPE NCC and the Dutch Police. On 8 November 2011 the Dutch police issued an order for the RIPE NCC to freeze the registration records of 4 IP blocks documented in the RIPE number registry from 8 November 2011 to 22 March 2012 (RIPE NCC, 2011a). The freeze was part of an international effort, Operation Ghostclick, a coordinated effort to replace rogue DNS servers supporting the DNSChanger malware. To perform this replacement, the IP addresses identifying those servers were to be re-delegated to "replacement" servers hosted and

[652]ARIN 30, ARIN uses Roman numerals in its conference enumeration.

managed by the ISC. The rationale for replacement is that millions of users were infected with DNSChanger and simply shutting down those servers would result in name resolution failure for those users, creating chaos for those users and their upstream ISPs receiving support calls.

The RIPE NCC complied with the freeze, but quickly contested the legality of the freeze on grounds that the mechanism used, the Police Act, was not the appropriate mechanism. According to the RIPE NCC:

> The RIPE NCC is receiving independent legal advice and is in discussion with the appropriate authorities. It is the intention of the RIPE NCC to pursue this matter further in Dutch court to establish a precedent so that it is certain of its rights regarding such orders.
>
> In the interest of transparency, the RIPE NCC is working on full disclosure of the background documents. The RIPE NCC will update the community if and when any other publishable materials become available.
>
> The RIPE NCC is committed to acting in the best interest of its membership and will continue to inform the Internet community on this matter as it progresses.
>
> The RIPE NCC has not withdrawn, removed or reclaimed the address blocks in question; it has temporarily locked a registration of address blocks. (RIPE NCC, 2011b)

The positive interpretation of this effort is that, as above, the NCC is clear regarding its rights in future situations. The mechanism used was the Police Act. If precedent were set that this was a valid action, there is some fear that the police may arbitrarily freeze blocks more frequently in the course of investigations. In the extreme, abuse of this power could result in substantive degradation of the number registry as a tool for remediating externalities.

The negative interpretation is a conflict of authority. Regardless of the mechanism at play, the immediate purpose was in service of remediating DNSChanger. The RIPE NCC received external legal advice that that "the police order had no sufficient legal grounds to force the RIPE NCC to execute the order," (RIPE NCC, 2012a) and the blocks were unlocked on 10 January 2012. The contract with the member holding two of the blocks, `93.188.160.0/21` and `85.255.112/20` expired and those blocks were placed in quarantine for six weeks, then subsequently reallocated according to RIPE policy. The justification was a combination of legal standing, extant RIPE region policies, and impending IPv4 exhaustion. The former, legal standing, was an appeal to external formal-legalistic authority. The latter two are signals to the community that their social order, prioritizing the integrity of regional policies and ensuring numbers are available as the community approaches depletion, was the RIPE NCC's priority. On 13 February 2013 "the Dutch court had deemed the RIPE NCC's case 'inadmissible' and that it had been dismissed," (RIPE NCC, 2013c).

The contention between the RIPE NCC and the Dutch police is confounding, especially in the context of Operation Ghostclick. The RIPE NCC has engaged in a

variety of efforts at engaging with LEAs before and after the event, through specialized task forces and meetings and the more general Roundtable events described below. As an illustration of the difference between modes of engagement, one hypothesis is that reactive and proactive general engagement and informational sessions are "easy" commitments to comity—ex ante, neither party really has to alter their existing activities in the moment.

Recall Ostrom's discussion of credible commitment from Section 8.2.2:

> Agreeing to follow rules ex ante is an easy commitment to make. Actually following rules ex post, when strong temptations arise, is the significant accomplishment. (E. Ostrom, 1990, p. 93)

Most of the engagement between the RIPE and LEAs is educational, proactive engagement that has been tailored to some degree, but is largely an adaptation of more general "how to use the registry" processes. This is extraordinarily valuable, but does not in and of itself engender a credible commitment, an explicit assurance, of mutual cooperation in the future. The is not a pejorative critique of either side, simply an assessment of the state of cooperation. Here, despite engagement between the RIPE NCC, the Dutch Police, and other LEAs, when the Dutch Police needed to move on an action, they chose to use their formal-legalistic instruments of coercion. In contrast, ARIN has actively worked with LEAs to ensure court orders are structured such that it satisfies the LEA's need for a court order, garners access to information or services necessary, while complying with ARIN's larger responsibilities. In this case, ARIN has found common ground with LEAs that satisfy both images of order, allowing for a more credible commitment. That said, ARIN does have it easier than the others (and this is articulated by ARIN staff as well as others) having to deal largely with only US and Canadian LEAs, not the diverse set of LEAs found in other regions.

9.3.3.2 Anti-Abuse

Actors on the receiving end of negative reputation attribution have portrayed the anti-abuse as vigilantes, actors operating outside the law. While there have been rogue reputation aggregators in the past, the anti-abuse market has, as discussed in Chapter 7, weeded those out. In terms of relational authority, these rogue actors did not contribute to the social order. The actors remaining serve as what Chapter 7 refers to as credible reputation aggregators. More than being credible within the community, a number of those that have proven credible in the anti-abuse community have also proven to be credible complements to law enforcement.

The strength of the anti-abuse community is the purview it has into the sources of abusive activities. These actors have demonstrated investigative capabilities. Moreover, the anti-abuse social order has actively honed investigative capabilities to cope with a market demanding accountable reputation indicators with low false positives. The anti-abuse relational authority, as argued in Chapter 8, follows the leadership mode of creating social order. The resulting common interest has, as argued earlier, created externalities of similar magnitude to those of public interests.

When asking the question of whether the resources available to the anti-abuse community are commensurate with potential social loss, the answer is mixed. While the anti-abuse community does embrace self-regulation, it has also recognized two elements of the larger security problem that cannot be dealt with through reputation alone: *a*) the security escalation problem and *b*) the limitation of reputation. The former, escalation, is the consistent adaptation of abusive actors to reputation and enforcement strategies of anti-abuse actors. As described in Chapter 7, abusive actors ranging from satisficers to extractive abusers are engaged in continuously, instrumentally exploring the legitimate sending space, identifying and exploiting vulnerabilities. What's more, many of these actors are repeat offenders.

Anti-abuse efforts do have an immediate remediative effort. As per the arguments for graduated sanction as a signaling mechanism, some naïve operational and satisficing abusers may adopt anti-abuse norms and principles to become credible participants in the community. Others may not adopt anti-abuse norms, either continuing as satisficers or becoming extractive abusers. Reputation attribution forces satisficers bordering on extractive as well as extractive abusers to be continuously "on the move." The result is what the criminology literature refers to as displacement: when the police begin patrolling a neighborhood with recent burglaries, the criminals recognize increased enforcement and move to a different neighborhood. Similarly, as evidenced by the many cases on Spamhaus's ROKSO (Spamhaus, 2014e), especially egregious composite extractive abusers constantly move from address space to address space.[653]

While reputation is not sufficient, the information and capabilities necessary to create reputation is aligned with the interests of LE agencies. Further, the anti-abuse community's common image and that of LE is aligned: both have a strong commitment to protecting end users. Anti-abuse actors roots this common image in consent based messaging. LE roots its common image in formal-legalistic authority serving the common interest. While the root of these images differs, the pragmatics of the alignment are beneficial.

Of the studies in this work, the coordination between LE and anti-abuse is perhaps the most naturally complementary. A particular instance of cooperation and a general instance of cooperation between Spamhaus and LEAs are presented to illustrate proactive, supportive engagement that has not resulted in subordination. The particular instance is Operation Ghostclick, the collaborative effort between actors in the anti-abuse communities and LE to remediate the DNSChanger malware infection. The general mechanism was described earlier in the context of the RIPE NCC's engagement with the Dutch Police. Here, a brief account of and explanation of the origins of that collaboration is presented.

Microsoft's Digital Crime Consortium (DCC) is amongst the NOG-like anti-abuse arenas where the community convenes to share information. After the meeting, sub-group of the anti-abuse community, along with a couple of members of the FBI that have developed rapport with the anti-abuse community, met to discuss some of the contemporary issues considered most pressing to the community. This meeting

[653]Recall the discussion of making reputation durable in Section 7.2.2.

has since been referred to as the Bar Meeting given that, like many informal engagements amongst operators, the meeting took place away from the formal venue at a local bar in Redmond. The issue of DNSChanger came up and the FBI agents attending took notes. In the LEA engagement narrative referencing Bobby Flaim, Flaim indicated in interviews that he spent substantive time learning the technology from the communities in order to *a*) better understand the issues facing the community and why they were problematic and *b*) relate this issues back to the FBI as a means to incent support for LE engagement. In the case of this latter, the ideal support from LE is to develop a case or provide the authority necessary for more effective remediative efforts. The FBI agents supporting Operation Ghostclick are Flaim's counterparts in the anti-abuse community. These actors have longstanding engagement with the community and in M^{3}AAWG. They have developed an understanding of the technologies and implications sufficient to understand how anti-abuse mechanisms described in Chapter 7 work. More importantly, these actors, can make the case to their superiors that a particular issue warrants an investigation.

In the case of Operation Ghost Click, the salient support relative to the NRS was the development of the court order necessary to freeze address blocks used by DNSChanger servers. Cooperation between the Department of Justice and Estonian law enforcement led to the arrest of Estonian nationals behind DNSChanger and the seizure of the elicit DNS servers that facilitate the fraud. As noted earlier, given millions of users affected by DNSChanger would have been adversely affected by simply shutting those DNS servers down, the number resources were temporarily delegated to ISC to run replacement DNS servers that would mitigate these costs. The freeze was critical because it ensured that those numbers would not be re-delegated to other actors, potentially allowing another abusive actor to pick up where DNSChanger left off. This was certainly possible: even though the DNS servers were removed, the infections remained. If a malicious actor were able to reclaim those number resources they could recreate the problem, thus recreating the problem again.

Operation Ghost Click was considered a substantive achievement in cooperation between international law enforcement and the anti-abuse community. In both interviews and fieldwork it has been held up as *the* instance of collaboration that should be learned from and repeated. The cooperation with the anti-abuse community and LE is a fantastic instance of proactive, supportive engagement (FBI, 2011). As noted earlier, largely following its policy authority, the RIPE NCC unlocked two of the blocks locked on 10 January 2012. A leader in the DCWG, the DNS Changer Working Group, reported on the RIPE NCC's choice to unlock these blocks. Under Operation Ghost Click, the DCWG hosted clean DNS servers (via ISC) to help remediate DNS Changer from 9 November 2011 to 9 July 2012.

On 10 August 2012, DCWG participants detected that two blocks the RIPE NCC had unlocked had now been reallocated and were being actively routed. According to Greene:

> It was assumed that these blocks would remain in limbo until all the court proceedings were completed. The "assumption" was not correct.

> Alarms that many service providers and security professionals put in place to watch for miscreant "hijacking" of these IP address block were triggered early on Friday. What was a surprise was that these blocks were not "hijacked," but re-allocated. . . . DNS Changer was not the only malicious activities inside these Netblocks. Who ever controls these netblocks can hijack computers that are still infected with DNS Changer and other malware. The way these netblocks are reallocated and surprise at the speed of the reallocation exacerbates the concern. This move by RIPE surprised many in the security industry, law enforcement, and participants of the DCWG. Given the surprise, network operators should be very cautious with any traffic to/from these netblocks. (Greene, 2012a)

A few days later on 15 August the RIPE NCC posted a response, clarifying the legal advice and policy choices discussed in the context of the Dutch Police order. Greene (2012b) goes on to critique the RIPE NCC's action, in particular the clarification of action (RIPE NCC, 2012b).

As with the analysis of the Dutch Police issue here, Greene recognizes the RIPE NCC's reaction and the need to maintain what is here referred to the as RIPE community and the RIPE NCC's authority. Greene responds to the RIPE NCC's clarification (RIPE NCC, 2012a):

> [T]he general behind the scenes reaction is surprise. This statement by RIPE has not helped the situation. In fact, it might have made it worse given what RIPE has communicated in this statement... Finally, where is RIPEs accountability?... RIPEs leadership and staff knows many of the people in DCWG. Many of the DCWG participants are RIPE members. It would have been simple for RIPE to communicate officially or unofficially that they were moving beyond unlocking to using RIPE 541 to re-allocated... Given the number of organizations within the RIPE community who participated in the DNS Changer clean up, it can only be assumed that RIPE intentionally decided to NOT communicate. There was on-going dialog between members of the DCWG and ARIN/RIPE leadership to find ways to responsibly manage all IPv4 blocked involved with the Rove Digital/DNS Changer activity. At no time was it communicated from RIPE to their membership who were part of DCWG that contractual relationship for `93.188.160.0/21` and `85.255.112/20` ended and that these would then be put back into the allocation pool. RIPE could have clued in the community when the contract ended. RIPE could have clued in the community when the addresses went back into the allocation pools quarantine period. RIPE could have communicated when the addresses were live in the allocation pool. (Greene, 2012b)

This problem does not help the existing tension between the anti-abuse community and the RIR system. Rather, it is a very specific instance of operational tension between *a*) the concerted efforts of anti-abuse and LEAs and *b*) the common interest of a particular RIR, in this case the RIPE NCC. While the catalyst may well have been

a coordination failure on the part of Dutch Police and the RIPE NCC, and Greene's recognition of the NCC's responsibility to preserve its authority, the issue of communicating within the community remains. In interviews, the issue was presented exclusively in terms of the issue with the Dutch Police. Given unlocking the block was well after the initial order and that the block was previously held by known abusive actors and the overlap in the communities cited by Greene, the reallocation is surprising.

To the knowledge of the author, the reallocation did not cause direct harm. That said, it is a coordination failure. Such a failure exacerbated the existing tension between the anti-abuse community and the RIPE community in particular. Further, it unfortunately sends the signal that community coordination is not reliable.

In contrast, reputation aggregators, here in particular Spamhaus, has developed a consistent, proactive, supportive relationship with a number of international LEAs. As noted in the anti-abuse chapter (7), Spamhaus participates in a number of closed security groups and SIEs. Not only that, Spamhaus has substantive data collection resources and partners providing abuse data feeds. Not every case identified by Spamhaus and other closed communities with relationships with LEAs can be acted on. For some cases, there is insufficient information or it is difficult to find a combination of LEA and prosecutors willing to act on a case. In some cases, there was not sufficient information. In others, offices in the appropriate jurisdictions could not be convinced of the validity of the case. This latter scenario, in particular with the FBI, this a reflection of the limited resources dedicated by LEA to understand and discriminate between feasible and infeasible cases.

That said, such as in the case of Operation Ghost Clock, there are success cases. With the investigative capabilities, data resources available via SIRs and SIEs, and the cooperation of LEAs, select cases can and have been pursued and successfully supported the development of LEA investigations and prosecutions in multiple jurisdictions. Anti-abuse actors have engaged in what is referred to as *parallel construction*. While much of the data collected by Spamhaus is based on its network of collaborators and resources, much of it can be reproduced when one knows where to look. Parallel construction is the process by which an actor such as Spamhaus provides LEAs with the means to collect similar information on its own, thus satisfying evidentiary requirements and keeping Spamhaus at arm's length from the formal investigation, in particular the compulsion to testify. One actor indicated that LEA, in order to make an effective case for an official investigation and prosecution, wanted a "gift box" of evidence: *a*) samples of malicious traffic; *b*) documentation of particular elements of that sample that are illegal; *c*) a cause of action and the law that is being broken; *d*) traffic logs; *e*) names of the players, who they are, what jurisdiction they are in. In effect, LEAs want the "gift box" to contain any information that Spamhaus can provide that would further an investigation. Included in the gift box is how to reproduce its contents, ISPs to contact, etc. For instance, CenturyLink maintains a very detailed FAQ on LE engagement; (CenturyLink, 2014) if CenturyLink were a source of information, Spamhaus would indicate CenturyLink as a provider and suggest the LEA follow the FAQ to acquire the data in a manner that satisfies necessary evidentiary procedure.

Parallel construction is an instance of proactive, supportive engagement that, via the informal character of the arrangements and the reproducible "gift box," avoids subordination to the agency. In Paul Vixie's testimony to the Subcommittee on Crime and Terrorism of the Senate Judiciary Committee, he also indicated that collaboration resolving Conficker and Operation Ghost Click were "hand-shake deals":

> Each of these examples shows an ad-hoc public/private partnership in which trust was established and sensitive information including strategic planning was shared without any contractual framework. These takedowns were so-called "handshake deals" where personal credibility, not corporate or government heft, was the glue that held it together and made it work. And in each case the trust relationships we had formed as members of M[3]AAWG were key enablers for rapid and coherent reaction. (Vixie, 2014)

While this illustrates the informal character of the relationship, the argument made here based on this testimony and interviews with Spamhaus, is that informality is a means to shelter the substantive-purposive authority developed in the M[3]AAWG community, as above, from the political travails of a formal-legalistic authority. The personal credibility above is reputation and trust in the anti-abuse communities, especially amongst members of closed security and trust groups. As per Chapter 1's introduction to the Pakistan-YouTube case, the seemingly ad hoc character is animated by well-defined, albeit less-document, rights and obligations amongst stewards of the anti-abuse community.

9.3.4 RPKI and Predatory Rule

As per discussion of RPKI in Sections 5.2 and 5.7.4, there is substantive tension, especially in the RIPE region, regarding how RPKI may be used by external authorities. Recall from the discussion that the objective of RPKI, as expressed in the RFCs, is to serve as authoritative assertions of origination rights. These assertions are made durable in cryptographically signed assertions in the RPKI hierarchy. The RPKI hierarch*ies* are currently regional: each RIR maintains an authoritative RPKI root. As per the RPKI discussions, RPKI does not change the substance of the rights conferred by the RIRs but do make enforcement more durable and more *immediately* enforceable.

Security mechanisms such as RPKI are perceived to have the potential to eliminate routing externalities such as prefix hijacks. Such an effect would require two things: *a*) broad deployment and implementation of RPKI by RIR participants and *b*) exclusive trust in those assertions. Under these circumstances, RPKI would ensure that if a network originates a prefix, any actor appropriating that route would be able to immediately and automatically either *a*) confirm legitimate origination rights of the AS in question or *b*) identify invalid advertisements.

These capabilities are possible without RPKI. For instance actors can acquire bulk access to registry data, develop their own query system, and integrate that system

with their routing policy tools. This is a custom mechanism to confirm origination rights. Such custom mechanisms endogenize the costs of externalities such as prefix hijacking externalities. As an alternative, RPKI is an automated mechanism that is built into routers, making the endogenization of costs supposedly lower (less customization) and easier (automated in the router actors will be buying anyway).

The fear expressed in the RIPE community is that such a tool is *a)* not as lightweight as advertised by its designers and proponents and *b)* is at risk of appropriation by external authorities. The former is born out by experience and is nominally a problem that is clearly in the wheelhouse of the operator community. If the technology proves valuable, if it contributes to the social order that sustains the integrity of the routing system, the community has demonstrated a willingness and an ability to invest effort and information. This was the case with the route damping protocol and community identification of parameters that worked well in the wild. While current RPKI implementations have been reported by community members as buggy, needing work by vendors, the community also has a track record of providing that feedback to improve these systems. In effect, the implementation by improving the technology, through policy experiments and experience is an iterative, and importantly, like the route damping protocol experiments, a *reversible* experiment to run.

In contrast, the latter issue, appropriation by external authorities is neither iterative nor is it reversible. In a panel discussion with RPKI architect Steve Kent (a proponent of RPKI deployment), Malcolm Hutty made a pre-cautionary argument. Hutty argued that the community is in a position, as a respected body of experts, to either sanction RPKI or reject it as an accepted, legitimate tool. The danger, he argued, is of the community putting forth RPKI as the "right" technical tool for solving routing security problems. Hutty argued that a single root RPKI has the potential to be appropriated by governments that want greater regulatory or enforcement control over network interconnectivity.

Although Hutty did not use Levi's term predatory rule, he was essentially arguing that precise point. Recall ARIN's early rocky relationship with the FBI and RIPE's recent issues with Dutch Police. The former is an instance of predatory rule: an agent of a "ruler" appropriating any tool or resource that it perceives can reinforce its authority. In the case of ARIN, the FBI was mistaken, ARIN did not have the power or the authority to "shut off" networks in the registry. In the case of the Dutch Police order, the police were willing to use the Police Act to achieve its goals.

Consider a hypothetical. If RPKI, as envisioned by the RFCs, was in place when ARIN was initially visited by the FBI, the outcome may have been quite different. If RPKI had been in place, ARIN could not have argued it did not have the power to disconnect the target network. Under the strong application of RPKI, revocation of origin rights would result in all rights assertions in routes to be rejected. Absent the provision of routes, the revoked rights result in disconnection. ARIN, or any RIR administering the strong form of RPKI, could not deny it had the power to disconnect networks.

Strong RPKI is an attractive tool for regulatory authorities. Regulators can attempt to appropriate strong RPKI through (formal) legal authority. Such an effort

would force external authorities and RIR authority into a potentially contentious, even conflicting, engagement. Hutty's precautionary stance is an attempt to avoid both this conflict and the potential for appropriation.

As implied at the outset of this discussion, a critical problem with RPKI and predatory appropriation is that unlike most NRS policy, the experiments are largely reversible. These are reversible in the sense that they are largely endogenous technical decisions such as utilization criteria in the RIRs and deployment strategies in the IXes. There is evidence both have been quite mutable, some being reversed, others simply modified. If a policy does not work as expected, it can be changed, it can be updated to work more effectively. Reversibility is a tacit requirement for adaptation. A character of sustaining contingent social order is cultivating dependence on assets specific to that order. Until recently, that asset was subsequent delegations of number resources. In the absence of ongoing delegation as a selective incentive to maintain registry accuracy, RPKI is the mechanism for assuring registry accuracy, especially in the face of a delegation regime dominated by transfers. The result is that the RIR system has a dilemma tied up in RPKI as a mechanism for sustaining accuracy in a delegation regime dominated by transfers.

At present, RPKI is being encouraged by the RIRs, but it is not widely deployed. Proponents of RPKI make largely technical counterarguments. Strong proponents argue that RPKI information is one of a number of parameters used in the decision to select a route for traffic forwarding or not. In effect, they argue that RPKI is an apolitical technology. Here apolitical is used in the sense of the question posed by Winner (1980), "do artifacts have politics?" In a political vacuum, RPKI is a fantastic tool for durably enforcing number rights. Unfortunately, if RPKI gains critical mass, regulators may likely see RPKI as a durable tool, as well.

9.4 Developing Explicit Assurances

The general notion of credible assessment in the political context described in Sections 9.1 and 9.2 provide the analytic basis for challenges to operational domain knowledge. McCray et al. (2010) also provide illustrative instances of knowledge assessment and adaptation from non-Internet cases. The cases from the NRS in the previous section illustrate how that cooperation between NRS and state authority is possible, but many of these are, like Operation Ghost Click, ad hoc. This section further draws on both the non-Internet and Internet cases to illustrate where in the regulation development processes NRS participants can engage. Using the ANIME framework presented by Abbott and Snidal (2009) this section offers prescriptions for developing more durable, proactive, supportive engagement backed by explicit assurances. In effect, these prescriptions are a means to create the accumulative effect described in the discussion of social capital: create mutual political capital between the NRS and state authorities backed by explicit assurances.

9.4.1 Deploying Political Capital

Abbott and Snidal (2009) provide the ANIME framework in the context of their evaluation of private authorities as, to appropriate their language, a way to evaluate "the new governance game in town." Within their typology of private authorities, they identify a number of configurations of public and private authorities. For instance, the configurations identified in by Howard and Sowell (2014) are represented in this typology as well as the development relationships cultivated by LACNIC and APNIC. Each of these configurations has strengths and weaknesses in terms of their capabilities. To highlight these, Abbott and Snidal (2009) decomposes the regulatory process into agenda setting (A), negotiation (N), implementation (I), monitoring (M), and enforcement (E).

> To be sure, the five stages are neither as distinct nor as neatly ordered as [ANIME] suggests: for example, some activities support multiple stages, and each stage is typically carried out with an eye to previous and subsequent steps. (Abbott & Snidal, 2009, loc. 1567–1569)

The NRS's potential to contribute to agenda setting is considered the most immediately viable, although these contributions will address other phases where appropriate. The strengths and limitations of various CRIs contributions will be offered, tempered with a discussion of the scope of authority.

This latter, the scope of authority, is particularly important to ensure CRIs are, as discussed above, supportive of, but not impinging on, public policy making processes. This criteria requires some adaptation of the phases to reflect a supportive role. For instance, the assumption of private authorities is that they are performing the function of a regulator, thus need public attention and support. In contrast, the NRS institutions' constituency comprises the operational epistemic community. CRIs' social capital is the means to marshal the essential competencies of independence, representativeness, expertise, and operational capacity (Abbott & Snidal, 2009, loc. 1293) at play in support of the phases below.

In terms of developing mutual political capital, agenda setting is considered a key entry point for NRS institutions. As the CRIs learn to navigate the global political arena, in particular the politics of more direct engagement, contributing to agenda setting is currently the most similar to their current engagements while also developing more explicit assurances. Abbott and Snidal indicate that "[a]genda-setting requires an ability to capture public attention, frame issues in politically powerful ways, gather and disseminate information, and formulate appropriate ways to proceed," (Abbott & Snidal, 2009, loc. 1580–1581). Public attention is the purview of the state and regulator; the CRI is not a public policy making institution, nor is it in the business of eliciting public sentiment. The NRS performs is concerned with reaching its constituents and would-be constituents (members of the operational epistemic community, but not necessarily the broader public. The latter three elements of agenda-setting are salient to the NRS, especially when engaging with their government counterparts.

"Framing issues in politically powerful ways" is a double-edged sword. Such a framing could easily slip the scope of the NRS, impinging on the authority of state or other external authorities better suited to handle a particular situation. For instance, the promotion of IXes is an instance were making too broad a claim to IX benefits can diminish credibility. In one of the earliest IX interviews Henk Steenman, the CTO of the AMS-IX, warned against promoting the IX as a panacea. In the context of IXes, arguing the market effects and lower-barriers to entry is arguably both more effective as a long-term strategy and appeals to regulators interests in healthy, functioning markets. Open-IX made it clear that its intention is to undermine what it considered an interconnection monopoly through standard-setting, improving the market. This is clearly a common interest in line with DOC's interest in well-functioning markets.

IPv6 and development efforts are, like anti-abuse and LE to be discussed shortly, an instance of largely harmonious alignment. When RIR staff were traveling to promote IPv6, the "political framing" was of an existing issues, IPv4 depletion and IPv6 deployment, that will have effects on the public interest. The problem was presented in that context, but not as an assertion of authority in the public policy-making space. Here the RIRs' role is to provide authoritative information regarding the state of the resource stock that many of that state's infrastructure providers, both domestic and international, will be coping with in the near future. In effect, the RIR is providing forewarning of a resource shortage the government may have been (and often are) otherwise unaware of. As noted by Abbott and Snidal, "[e]ffective advocates must possess information and expertise about the business context in which the issue arises (business expertise)," (Abbott & Snidal, 2009, loc. 1581–1583). Presentations to officials shared with the author provide *a*) basic statistics on the stock of IPv4; *b*) why IPv4 is running out and at what rate; *c*) comparison with the size of IPv6 space available; and *d*) high level plans outlining the critical path to IPv6 deployment for various types of infrastructure (backbone versus access).

Anti-abuse, already largely aligned with LE objectives, can and does easily frame its efforts as supportive of LE investigations and case development efforts. To this point, recall Spamhaus has received a number of awards from LEAs. Further note that a large portion of Vixie's testimony cited in the previous section is a description of how to ensure these types of collaboration can continue to be successful. In an agenda setting mode, M^{3}AAWG produces a number of informational reports that are often used as supporting materials in comments offered to government agencies, such as comments on implementation of CSRIC III Cybersecurity (M^{3}AAWG, 2014b) and comments to the ITU's request for consultation on combating spam are both illustrations of M^{3}AAWG's Public Policy Committee's (M^{3}AAWG, 2015) efforts at framing technical issues in ways that can be consumed and acted on by those setting regulatory agendas in domestic and international arenas. In the course of interviews with M^{3}AAWG leadership, the common strategy of providing supporting information, but not normative advice impinging on broader public policy was repeated and stressed.

RPKI is the issue that has the most clear and obvious "politically powerful" framing. Hutty's precautionary framing is less about the RIRs engaging in agenda setting, but rather, about the implications of putting the weight of the RIR's legitimacy as operations experts behind RPKI as the community sanctioned solution. In effect, Hutty is warning of how such a sanction by the community would be appropriated by governments in regulatory efforts. Although not explicit in Hutty's comments, the precautionary framing runs counter to the apolitical framing that RPKI is just a tool, it does not intrinsically have politics of its own.

The threat landscape for RPKI is currently rather nuanced. It has not seen nearly the deployment sufficient to have the network effects that would make it pragmatically attractive. Further, RPKI has liability issues that have been pointed out by Huston et al. (2015) and is currently in discussion amongst on ARIN's PPML (Gallo, 2014). The two issues are that a single error in an RPKI configuration can have cascading effects, obviating its intended value. Another issue is that ARIN has shifted liability for RPKI failures from ARIN as the firm, leaving it to rest with those members it is trying to entice to use RPKI. Further, a number of network operators have commented that RPKI implementations remain buggy. Taken together, the RIRs have time to figure out how to handle framing RPKI, but, given governments' history of predatory appropriation, they should be careful how they advertise RPKI's success if it does begin to garner network effects.

As contributors to the agenda-setting process, perhaps one of the NRS's greatest strengths is its access to information that can better inform the policy making process. Moreover, the diverse sources contributing to and vetting this information contribute to validity and credibility in much the same way the consensus process filters for capture. IXes have a unique perspective into the interconnection market in a region and can provide insightful measures that respect the privacy of individual members flows, giving local regulators a useful view into the health of a metropolitan, regional, or even national market. For instance, PTT Metro in Brazil has a government sponsored measurement infrastructure "built in" to its national Internet exchange to monitor the quality of the Internet in Brazil. Organizations such as Euro-IX, IX-F, and Open-IX have an opportunity to further develop this information capability as a means to enhance their relationship with government agencies.

In the RIR system, number resource information and statistics is an obvious information asset. Another substantive asset is the RIPE Atlas infrastructure, the largest distributed Internet measurement platform in the world, with nodes on six continents and actively distributed to ensure diversity at the metropolitan, intrastate region, by state, and by RIR region. The RIPE NCC encourages participation in the Atlas program, use of the probes, use of the data, and sharing of interesting analyses. As a source of information, selection amongst the probes to collect information at various regional granularities is a power source of information to make available to policy makers. That said, these tools are deployed by *volunteers*, participants in the RIPE community. Many of these deployments are in private networks. In some cases, networks have multiple probes, trading the contribution to the common interest of collecting global data for the ability to use the on-net probes for private measurements. When considering what elements of this infrastructure to

make available to government analysts, these private uses should be considered.

Within the anti-abuse community, reports generated by organizations like those provided by M^{3}AAWG offer valuable neutral insight into the MVN. For instance, Andersen, O'Reirdan, Upton, and Knecht (2014), the M^{3}AAWG Bot Metrics Report, is presented as "the first industry report with data provided directly by service operators and ISPs detailing the number of subscribers identified as having a system infected by malware, also known a 'bot,' and the percentage of those subscribers notified of the problem." Another instance is M^{3}AAWG's E-mail Metrics Reports, covering a diverse set of at least 100 million mailbox's, the last report covering more than 400 million (Andersen et al., 2014). The e-mail metrics program has been running since December of 2005. Participation in this program is voluntary. As noted earlier, the data collection programs fostered by communities like M^{3}AAWG support both their efforts at commenting on public policy issues and have the potential to be leveraged more directly in public policy evaluations through mutual cooperation between organizations like M^{3}AAWG and states.

Coordination between Spamhaus and LE is nuanced instance of gathering and disseminating information. In this case, the agenda setting component is to encourage greater investment in LE agents that can better coordinate with actors such as Spamhaus, making better use of the vast data sets and potential investigations on offer. Recall from earlier that while there are a number of very successful investigations and prosecutions, a number of cases go uninvestigated for a variety of reasons. This is a characteristic instance of resource deficiency on both sides: Spamhaus is expending resources building gift boxes and there is a need for more investigators that know what to do with these gift boxes. As an instance of empowerment, this phenomena may be used to encourage additional investment in LEA agents trained in anti-abuse techniques, effectively increasing the collateral benefits of coordinated investigations.

Obviously the NRS has a surfeit of information that it can offer, the challenge is, as with framing issues in politically powerful ways, how to make use of this without impinging or without damaging the common interest of those provisioning these data sources. As per the argument framing this chapter, the danger is that state interests will aggravate the common interests necessary to provision this data, thus diminishing the resource it is trying to appropriate. Abbott and Snidal (2009, loc. 1580–1581) argue a final component of agenda setting is the ability to "formulate appropriate ways to proceed." As a joint process, this becomes a question of how NRS institutions can approach external authorities, providing salient data and analyses without appearing to have a public policy agenda. NRS institutions certainly have a political interest in preserving the management regimes.

As alluded to earlier, in especially in the discussion of Haas's epistemic consensus, the creation of a common, authoritative image in relation to a particular issue is critical to establishing comity. This common image also serves as the foundation for explicit assurances between the NRS and external authorities. Here the agenda setting phase is the vehicle to identify existing and potential issues in which such a collaboration may be beneficial. The issues in the last section are certainly starting points. For instance, Greene's surprise and disappoint at the RIPE NCC's

choice to reallocate without making an explicit effort to alert the DCWG is a case of failed "weak alignment." Rather, while DCWG is a great instance of cooperation with state authorities, on the NRS side it does appear that it relied too much on the "hand-shake" agreements referenced in Vixie's testimony.

Recall that a sub-theme of the RIRs, IXes, and anti-abuse story has been one of maturation and professionalization. This theme was made most explicit in the IX narrative. As participants became more dependent on the IX, they demanded more professionalized, consistent guarantees of service. Explicit assurances of alignment are the manifestation of this process in the global political arena. The new customer is external actors dependent on the infrastructure provisioned. The new challenge for the NRS is to develop the capability to transition existing, ad hoc relationships that demonstrate there is value and potential in cooperative relationships to more explicitly defined relationships. One implementation may be a memorandum of understanding more specific than "we pledge to work towards Internet development," but less so than a finely articulated consulting contract.

Consider each of the specific issues in the previous section and the following hypothetical instances of more explicit assurances. In the case of development, there have been one-off instances of cooperation, typically training events. Under the remit of the RIRs in developing regions, part of their mission is to support Internet development in the region. One step forward may be to establish joint development team with a particular economy, say an LDC, that would benefit from a longer term relationship with distinct development milestones and reports to both the RIR constituency and the broader development community. Such an effort is still within the remit of the RIR, but creates a more credible commitment to outcomes than proactive general engagement or one-off proactive supportive engagement. To mix expertise in this arena and avoid impinging on international development actors, actors such as APNIC may engage with state governments and existing development organizations work in the region.

A similar strategy could be used to foster IPv6 development. For instance, many regions experiencing Internet growth are precisely those that received the fewest IPv4 allocations late in the path-dependent history of IPv4 distribution. This is an opportunity to *a*) actively support demand for Internet connectivity and *b*) help developing regions leap-frog developed regions that are locked into large IPv4 deployments. Again, the strategy is to coordinate not just with state governments, but also with development agencies that have expertise with both the region at hand and the local institutions and infrastructure (writ large, meaning power, water, etc.).

Returning to the relationship between Spamhaus and LE, an immediately intuitive idea for developing explicit assurances would be to make parallel construction and the process of developing gift boxes a more well-defined process. That said, recall that informal relationships often shelter the CRI from the trappings of my formal relationships that would require commitments incompatible with CRI interests. For instance, while Spamhaus investigators have testified as expert security witnesses, they typically do so in other roles, such as employees of their primary employer or as

members of more public SISCs, not as Spamhaus affiliates.[654] Rather, given the existing attacks on Spamhaus, both the recent DDOS and lawsuits, public testimony is considered a liability. While Spamhaus's techniques are legal data collection mechanisms in collaboration with sources, those relationships are also based on trust and anonymity, much like any another closed SISC. If these partners saw Spamhaus engaging in collaborations that increase the risk these trust relationships would be exposed, creating liabilities and threats for the group as a whole, that would be a violation of those relationships and diminish Spamhaus's credibility. Again, this is not because those relationships are illicit, but, like the DDOS on Spamhaus, there are plenty of actors that are willing to launch DDOS or other attacks against anyone known to affiliate with Spamhaus.

9.4.2 Preserving the NRS's Social Order

Part of explicit assurances between the NRS and external authorities is to ensure the ongoing function of the routing system. For the near term, this means preserving the NRS's social order. As noted earlier, part of the state's obligation is to promote and sustain resources contributing to the public good. While not yet explicitly aligned with the public interest, NRS authority does not impinge on it. Moreover, it is currently the basis for an ever-increasing set of public, private, and social goods.

Analytically, the social order developed by the NRS thrives on endogenous, constructive contention amongst its participants but is not fundamentally in contention with the state. Rather, those issues that are in contention with state authority are those that play out atop the Internet infrastructure. The NRS is a transnational social order, but like telephone communication and commercial air travel, it facilitates transnational coalition building, but, as a non-discriminatory infrastructure, it does not directly contribute to issue-specific politics. That does not mean the NRS does not have politics. Rather, it is clearly interested in preserving a social order that has demonstrably contributed to the infrastructure industries' common interests. Further, from the perspective of NRS participants, the self-limitation of scope, while creating a weak form of alignment, has differentiated the NRS from typically expansionist regimes. This is a distinct difference from "broad Internet governance" fora such as ICANN and that IGF, two institutions that embrace and encourage the confluence of transnational public policy issues with notions of "Internet governance."

While the limited scope has served the NRS well, there are places it could improve, in particular to demonstrate further maturation of the NRS as a whole. The early instances of cooperation above provide evidence that cooperation is possible, but the NRS still need to work towards resolving operational tensions between RIR participants and the anti-abuse community. While the NRS does have substantive operational adaptability, demonstrated by its ability to keep pace with and foster Internet growth, reconciling the RIR/anti-abuse tensions would demonstrate diplomatic capabilities that would serve it well in the global political arena. Moreover, the confluence of RPKI, transfers, and reputation is a very dangerous situation for

[654] Exceptions are actors in public facing leadership positions, such as Alan Murphy.

the RIRs in particular. If the "nightmare scenario" referenced in Section 7.5.2.2 emerges, where negative reputation promulgates amongst fungible IPv4 address blocks, this could substantively damage the reputation of the RIRs as stewards of the number system. Although the ITU seems to have become less aggressive in the recent Plenipotentiary than during the WCIT, such predatory appropriators would quickly politicize the nightmare scenario as a rationale for stronger state intervention.

The NRS had demonstrated many of the laudable characteristics of common resource systems. It has created an adaptive social order that has given rise to a global communications infrastructure. The NRS has developed not just one credible knowledge assessment process, but a family of three consensus-based processes that have contributed to the growth and success of the modern Internet. This modern Internet now faces two challenges that are critical to its ongoing success. The topic of this chapter, developing explicit assurances with complementary external authorities is challenging, but there is evidence of early successes and the potential to build on its success.

The second challenge is, perhaps ironically, the more difficult of the two. Despite a common image of consensus based decision making and the attendant constructive conflict, potentially destructive conflict remains between the RIR community and the anti-abuse community. Further, this conflict is rooted in operational practices—analytically, the overarching common images are consonant. That said, there is no common forum in which to remediate this tension. As illustrated in Chapter 8's analysis of the NRS as a social order and in the confluence of transfers and reputation, although the CRIs are largely independent, they are bound together in a common number resource system. While there is certainly evidence of the potential for cooperation with the state, if the strategies above are successful, explicit assurances with state authority will inevitably demand a more coherent NRS whose institutions are not only analytically consonant, but cohere around their own explicit assurances of cooperation.

Appendix A

Fieldwork and Research Subjects

Data collected for this work comprised a combination of archival data analysis, fieldwork, and interviews. Archival data analysis is documented via references in the body of the dissertation, as used. Fieldwork comprised passive observation and rapport building within the network operator and anti-abuse communities. A list of fieldwork locations is provided in the next section, in Table A.1. A list of research subjects that did not wish to remain anonymous is provided in Section A.2.

A.1 Fieldwork

Table A.1 lists the conferences, organizations, and locations at which fieldwork took place. If only a start date is listed, that event only lasted a single day.

Table A.1: Fieldwork

Start	End	Conference	Organization	Location	Study
10/9/2011	10/11/2011	NANOG53	NANOG	Philadelphia, PA	NOG
10/12/2011	10/14/2011	ARIN28	ARIN	Philadelphia, PA	RIR
10/31/2011	11/4/2011	RIPE63	RIPE NCC	Vienna, Austria	RIR
2/5/2012	2/7/2012	NANOG54	NANOG	San Diego, CA	NOG
2/21/2012	2/23/2012	M^{3}AAWG24	M^{3}AAWG	San Francisco, CA	Anti-Abuse
2/27/2012	3/2/2012	APNIC33	APNIC	New Delhi, India	RIR
3/19/2012	3/21/2012	GPF	GPF	New Orleans, LA	NOG
4/16/2012	4/20/2012	RIPE64	RIPE NCC	Lubljana, Slovenia	RIR

Continued on next page

Start	End	Conference	Organization	Location	Study
4/24/2012	4/27/2012	Counter-eCrime Operations Summit (CeCOS VI)	APWG	Prague, Czechoslovakia	Anti-Abuse
5/21/2012	5/22/2012	LINX77	LINX	London, UK	IX
5/23/2012	5/24/2012	More IP	AMS-IX	Amsterdam	IX
6/3/2012	6/6/2012	NANOG55	NANOG	Vancouver, British Columbia	NOG
6/14/2012	6/15/2012	Euro DIG	Euro DIG	Stockholm, Sweden	Government Engagement
6/24/2012	6/29/2012	ICANN44	ICANN	Prague, Czechoslovakia	All
6/30/2012	7/15/2012	Interview visit	RIPE NCC	Amsterdam, Netherlands	RIR
6/30/2012	7/15/2012	Interview visit	AMS-IX	Amsterdam, Netherlands	IX
8/21/2012	8/31/2012	APNIC34	APNIC	Phnom Penh, Cambodia	RIR
9/3/2012	9/7/2012	Interview Visit	APNIC	Brisbane, Australia	RIR
9/17/2012	9/21/2012	Interview Visit	ARIN	Washington, DC	RIR
10/3/2012		IX Manchester Meeting 2	LINX	Manchester, UK	IX
10/9/2012		UKNOF23	UKNOF	London, UK	NOG
10/22/2012	10/25/2012	M^{3}AAWG26	M^{3}AAWG	Baltimore, Maryland	Anti-Abuse
10/28/2012	11/1/2012	LACNIC	LACNIC	Montevideo, Uruguay	RIR
11/6/2012	11/9/2012	IGF	IGF	Baku, Azerbaijan	Government Engagement
1/17/2013		UKNOF24	UKNOF	Newark, UK	NOG
2/19/2013	2/21/2013	M^{3}AAWG27	M^{3}AAWG	San Francisco, CA	Anti-Abuse
3/4/2013		Post-WCIT discussion	Oxford	Oxford, UK	
5/5/2013	5/10/2013	LACNIC	LACNIC	Medellin, Colombia	RIR
8/14/2013		Open-IX	Open-IX	New York, NY	IX

Continued on next page

Start	End	Conference	Organization	Location	Study
10/27/2013	10/29/2013	Euro-IX	Euro-IX	Helsinki, Finland	IX
2/10/2014	2/12/2014	NANOG60	NANOG	Atlanta, GA	NOG
2/17/2014	2/20/2014	M^{3}AAWG30	M^{3}AAWG	San Francisco, CA	Anti-Abuse
9/5/2014		Open-IX	Open-IX	New York, NY	IX
10/20/2014	10/23/2014	M^{3}AAWG32	M^{3}AAWG	Boston, MA	Anti-Abuse
2/16/2015	2/19/2015	M^{3}AAWG33	M^{3}AAWG	San Francisco, CA	Anti-Abuse

A.2 Research Subjects

Table A.2 comprises the set of research subjects that were willing to be listed as interviewees in this work. A number of research subjects were willing to be interviewed, but did not want to be listed. Table A.2 does not imply attribution of any facts or responsibility for the content of the work found in the body, i.e. Chapters 1–9, of this dissertation to any particular individual. Unless the body of this work directly attributes a fact to a given subject at that point in the dissertation body, it should not be assumed that any particular subject contributed that fact. In many cases, interviews served as validation of fieldwork observations and pointers to archival data that can be more effectively referenced.

Table A.2: The following individuals participated in one or more interviews. The accuracy of content in the body of the dissertation is the sole responsibility of the author.

Name	Roles
Tobias Knecht	abusix CEO, M^{3}AAWG participant, RIPE Anti-Abuse WG
Job Witteman	AMS-IX CEO
Henk Steenman	AMS-IX CTO
Cara Mascini	AMS-IX Chief Marketing Officer
Christian Kaufmann	AMS-IX Executive Board Chairman, RIPE NCC Executive Board, RIPE Measurement, Analysis and Tools Working Group Co-Chair, Founding Member of France-IX
Bastiaan Goslings	AMS-IX Governance and Policy Officer
Steven Bakker	AMS-IX Senior Network Engineer
Richard Brown	APNIC Business Director
Geoff Huston	APNIC Chief Scientist
Sanjaya	APNIC Director of Operations and Services
Akinori Maemura	APNIC Executive Council Chairman

Continued on next page

Name	Roles
German Valdez	APNIC External Relations Program Director
Philip Smith	APNIC Learning and Development Director
Andy Linton	APNIC Policy SIG Chair, ICANN ASO AC
Adam Gosling	APNIC Senior Policy Specialist
George Michaelson	APNIC Senior Research and Development Scientist
Byron Ellacott	APNIC Technical Director
John Sweeting	ARIN Advisory Council Chair
Stacy Hughes	ARIN Advisory Council
Robert Seastrom	ARIN Advisory Council
Scott Bradner	ARIN Board of Trustees
Paul Vixie	ARIN Board of Trustees, Anti-Abuse Community
John Curran	ARIN CEO
Nate Davis	ARIN Chief Operations Officer
Susan Hamlin	ARIN Director of Communications and Member Services
Leslie Nobile	ARIN Director of Registration Services
Cathy Handley	ARIN Executive Director Government Affairs and Public Policy
Einar Bohlin	ARIN Senior Policy Analyst
Hernan Seoane	CABASE Gerente General
Arnold Nipper	DE-CIX CTO
Bijal Sanghani	Euro-IX Secretary General
Mike Hughes	Former LINX Staff
Gaurab Raj Upadhaya	Founder Nepal IX, Founder SANOG, APNIC Executive Council
Arturo Servin	LACNIC CTO
John Souter	LINX CEO
Patrick Gilmore	LINX Council of Management, SIX Board, form NANOG Board
Derek Cobb	LINX CTO
Keith Mitchell	LINX Executive Chairman and Founder (1994-2000)
Malcolm Hutty	LINX Head of Public Affairs
Chris Roosenraad	M^3AAWG Chairman
Michael Goldman	M^3AAWG Facilitation Trainer
Michael O'Reirdan	M^3AAWG Malware Co-Chair
Melinda Plemel	M^3AAWG Participant
Christine Borgia	M^3AAWG Participant
Dennis Dayman	M^3AAWG Participant, Founding Participant
Aaron Hughes	NANOG community member
Paul Ebersman	NANOG Development Committee
Greg Dendy	NANOG Program Committee
Tony Tauber	NANOG Program Committee
Kurt Erik Lindqvist	Netnod CEO

Continued on next page

Name	Roles
Patrick Fälström	Netnod Head of Research and Development
Brian Nisbet	RIPE Anti-Abuse Co-Chair
Richard Barnes	RIPE Co-Chair Measurement, Analysis, and Tools Working Group
Richard Barnes	RIPE Measurement, Analysis and Tools Working Group Co-Chair
Serge Radovcic	RIPE NCC Chief Communications Officer
Jochem de Ruig	RIPE NCC Chief Financial Officer
Andrew de la Haye	RIPE NCC Chief Operations Officer
Paul Rendek	RIPE NCC Director of External Relations
Remco van Mook	RIPE NCC Executive Board
Nigel Titley	RIPE NCC Executive Board Chairman, Former LINX Staff
Axel Pawlik	RIPE NCC Managing Director
Carlos Watson Carazo	SJO-IX Founder
Alan Murphy	Spamhaus Project

Bibliography

Abbott, K. W., & Snidal, D. (2009). The Governance Triangle: Regulatory Standards Institutions and the Shadow of the State. In W. Mattli & N. Woods (Eds.), *The Politics of Global Regulation* (pp. 44–88). Princeton, N.J: Princeton University Press.

Aben, E. (2014, 31 March). *A RIPE Atlas View of Internet Meddling in Turkey.* RIPE NCC. Retrieved from https://labs.ripe.net/Members/emileaben/a-ripe-atlas-view-of-internet-meddling-in-turkey

Acosta, A., Antoniello, N., Moonesamy, S., Onyango, D., Ramirez, M., Yamanishi, M., & Smith, P. (2011, 20 February). *prop-097: Global Policy for post exhaustion IPv4 allocation mechanisms by the IANA.* Retrieved from http://www.apnic.net/policy/proposals/prop-097

Advanced Network Technology Center, University of Oregon. (2015). *University of Oregon Route Views Project.* Retrieved from http://www.routeviews.org/

AFRINIC. (2014a). *AFRINIC Service Agreement 2014.*

AFRINIC. (2014b). *Current Policies.* Retrieved from http://www.afrinic.net/en/library/policies/current

AFRINIC. (2015). *AFRINIC Government Working Group.* Retrieved from http://meeting.afrinic.net/afgwg/

Ager, B., Chatzis, N., Feldmann, A., Sarrar, N., Uhlig, S., & Willinger, W. (2012, 13-17 August). Anatomy of a Large European IXP. In *SIGCOMM'12.* Helsinki, Finland: ACM. Retrieved from http://www.eecs.qmul.ac.uk/~steve/papers/ixp-sgcm.pdf

Alaettinoglu, C., Villamizar, C., Gerich, E., Kessens, D., Meyer, D., Bates, T., ... Terpstra, M. (1999, June). *Routing Policy Specification Language (RPSL)* (No. 2622). RFC 2622 (Proposed Standard). IETF. Retrieved from http://www.ietf.org/rfc/rfc2622.txt (Updated by RFC 4012)

AMS-IX. (2013a). *AMS-IX General Terms and Conditions.* Retrieved from http://goo.gl/3IEZBO

AMS-IX. (2013b). *Board Member Code of Conduct.* Retrieved from http://goo.gl/GgD3tA

AMS-IX. (2013c). *Wargaming Interactive.* Retrieved from http://goo.gl/bASGSF

AMS-IX. (2014). *Service Level Agreement.* Retrieved from https://ams-ix.net/services-pricing/service-level-agreement

AMS-IX. (2015). *Partner Program.* Retrieved from https://ams-ix.net/connect-to-ams-ix/partner-program

Andersen, K., O'Reirdan, M., Upton, J., & Knecht, T. (2014, 20 October). *M³AAWG Bot Metrics Report.* Retrieved from https://www.maawg.org/dm3z/category/botnets

Andree Toonk. (2014, 29 March). Turkey Hijacking IP Addresses for Popular Global DNS Providers. *BGPMON Blog.* Retrieved from https://www.bgpmon.net/turkey-hijacking-ip-addresses-for-popular-global-dns-providers/

APNIC. (2008, August). *Operational Policies for National Internet Registries in the APNIC Region.* Retrieved from http://www.apnic.net/policy/operational-policies-nirs

APNIC. (2011, 9 May). *Policies for IPv4 address space management in the Asia Pacific region.* Retrieved from http://www.apnic.net/policy/add-manage-policy

APNIC. (2013, 26 March). *APNIC PPAC Inaugural Session Transcript.* Retrieved from http://conference.apnic.net/35/program/public-policy-advisory-committee-ppac/transcript

APNIC. (2014a). *By-laws of APNIC.* Retrieved from http://www.apnic.net/publications/media-library/documents/corporate/by-laws

APNIC. (2014b). *Current Policies.* Retrieved from http://www.apnic.net/community/policy/current

APNIC. (2014c). *Mailing list archive: wg-government.* Retrieved from http://mailman.apnic.net/mailing-lists/wg-government/

APNIC. (2014d, May). *Policies for IPv4 address space management in the Asia Pacific region.* Retrieved from http://www.apnic.net/policy/add-manage-policy

APNIC. (2014e). *Recent Presentations Given by APNIC: LEAs.* Author. Retrieved from http://www.apnic.net/events/apnic-speakers/presentations/lea (Retrieved from http://www.apnic.net/events/apnic-speakers/presentations/lea.)

APNIC. (2014f). *Recording Network Assignments.* Retrieved from http://www.apnic.net/apnic-info/whois_search/using-whois/updating-whois/network-assignments

APNIC. (2014g). *SIG Guidelines.* APNIC. Retrieved from http://www.apnic.net/community/participate/join-discussions/sigs/sig-guidelines

APNIC. (2014h). *Working Groups.* Retrieved from http://www.apnic.net/community/participate/join-discussions/sigs/wgs

APNIC. (2015a). *APNIC and ITU collaboration.* Retrieved from http://www.apnic.net/community/ecosystem/intergovernmental-stakeholders/working-with-the-itu

APNIC. (2015b). *Assignment Window - FAQ.* Retrieved from http://www.apnic.net/services/services-apnic-provides/helpdesk/faqs/assignment-window

APNIC. (2015c). *Intergovernmental Organizations.* Retrieved from http://www.apnic.net/community/ecosystem/intergovernmental-stakeholders

ARIN. (2005, 27 October). *ARIN XVI Public Policy Meeting Minutes Day 2, 27 October 2005.* Retrieved from https://www.arin.net/participate/meetings/reports/ARIN_XVI/ppm_minutes_day2.html

ARIN. (2009a, 10 December). *ARIN Appeal Process.* Retrieved from https://www.arin.net/resources/resource_requests/appeal_process.html

ARIN. (2009b). *ARIN Policy Development Process.* Retrieved from https://www.arin.net/policy/archive/pdp_archive_20090107.html

ARIN. (2009c). *Draft Policy ARIN-2009-3 (Global Proposal): Allocation of IPv4 Blocks to Regional Internet Registries.* Retrieved from https://www.arin.net/policy/proposals/2009_3.html

ARIN. (2009d). *Internet Assigned Numbers Authority (IANA) Policy for Allocation of ASN Blocks (ASNs) to Regional Internet Registries.* Retrieved from https://www.arin.net/policy/proposals/2009_6.html

ARIN. (2011). *Draft Policy ARIN-2011-6: Returned IPv4 Addresses.* Retrieved from https://www.arin.net/policy/proposals/2011_6.html

ARIN. (2012, 25 October). *ARIN XXX Public Policy Meeting Day 2 Draft Transcript - 25 October 2012.* Retrieved from https://www.arin.net/participate/meetings/reports/ARIN_XXX/ppm2_transcript.html

ARIN. (2012, 30 January). *Legacy Registration Services Agreement.* Retrieved from https://www.arin.net/resources/agreements/legacy_rsa.pdf

ARIN. (2014, April). *Advisory Council.* Retrieved from https://www.arin.net/about_us/ac.html

ARIN. (2014a, 21 January). ARIN Number Resource Policy Manual [Computer software manual]. Chantilly, VA. Retrieved from https://www.arin.net/policy/nrpm.html

ARIN. (2014b). *Registration Services Agreement.* Retrieved from https://www.arin.net/resources/agreements/rsa.pdf

ARIN. (2015). *Requirements and Responsibilities.* Retrieved from https://www.arin.net/about_us/bot_requirements.html

ARIN Member Services. (2011, 21 February). *Draft Policy 2011-6: Returned IPv4 Addresses, staff assessment.* Retrieved from http://lists.arin.net/pipermail/arin-ppml/2011-February/020085.html

ARIN-PPML. (2014). *Internet Fairness Blah Blah.* Retrieved from http://lists.arin.net/pipermail/arin-ppml/2014-December/029506.html (Thread discussion re-initiated by Meuller on "fairness.")

Augustin, B., Krishnamurthy, B., & Willinger, W. (2009). IXPs: Mapped? In *Proceedings of the 9th ACM SIGCOMM conference on internet measurement conference* (pp. 336–349). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/1644893.1644934 doi: 10.1145/1644893.1644934

Avižienis, A., Laprie, J.-C., Randell, B., & Landwehr, C. (2004, January). Basic concepts and taxonomy of dependable and secure computing. *IEEE Transactions on Dependable and Secure Computing, 1*(1), 11–33. doi: 10.1109/TDSC.2004.2

Bargisen, N., Field, B., Gilmore, P., Hickson, N., Rogan, B., Tal, G., & Temkin, D. (2012, 23 October). CDN. In *NANOG 55.* NANOG. Retrieved from https://www.nanog.org/meetings/nanog55/agenda

Barnes, B., & Edge, D. (Eds.). (1982). *Science in Context.* MIT Press.

Becker, L. (1977). *Property Rights: Philosophic Foundations*. Routledge and Kegan Paul.

Behjat, A. (2010, December). *Join The Global Passive DNS (pDNS) Network Today & Gain Effective Tools To Fight Against Cyber Crime.* Retrieved from https://www.isc.org/blogs/join-the-global-passive-dns-pdns-network-today-gain-effective-tools-to-fight-against-cyber-crime/ (The pDNS network is managed by the Internet Systems Consortium.)

Bellovin, S. M., & Gansner, E. R. (2003). Using Link Cuts to Attack Internet Routing. Retrieved from http://academiccommons.columbia.edu/catalog/ac:126877

Bianchi, N. M. (2013, 2 December). The Return of Open Relays. *Spamhaus News*. Retrieved from http://www.spamhaus.org/news/article/706/the-return-of-the-open-relays

Birner, R., & Wittmer, H. (2003). Using Social Capital to Create Political Capital: How Do Local Communities Gain Political Influence? A Theoretical Approach and Empirical Evidence from Thailand. In N. Dolšak & E. Ostrom (Eds.), *The Commons in the New Millenium: Challenges and Adaptation* (chap. 10). Cambridge, MA: The MIT Press. (Kindle Edition.)

Blackstone, W. (1979 [1766]). *Commentary on the Laws of England* (Vol. 2). Chicago, IL: University of Chicago Press.

Blau, P. M. (1964). *Exchange and Power in Social Life*. John Wiley.

Blomquist, W. (2012). A Political Analysis of Property Rights. In D. H. Cole & E. Ostrom (Eds.), *Property in Land and Other Resources* (chap. 12). Cambridge, Mass.: Lincoln Institute of Land Policy.

Blunk, L., Damas, J., Parent, F., & Robachevsky, A. (2005, March). *Routing Policy Specification Language next generation (RPSLng)* (No. 4012). RFC 4012 (Proposed Standard). IETF. Retrieved from http://www.ietf.org/rfc/rfc4012.txt

Bono, V. J. (1997, 26 April). *7007 Explanation and Apology*. Merit. Retrieved from http://www.merit.edu/mail.archives/nanog/1997-04/msg00444.html

Bourdieu, P. (1992). Die verborgenen Mechanismen der Macht (The Hidden Mechanisms of Power). In (pp. 49–79). Hamburg: VSA-Verlag.

Bradner, S. (1997, March). *Key words for use in RFCs to Indicate Requirement Levels* (No. 2119). RFC 2119 (Best Current Practice). IETF. Retrieved from http://www.ietf.org/rfc/rfc2119.txt

Brown, M. (2008, 24 February). Pakistan Hijacks YouTube. *Renesys Blog*. Retrieved from http://www.renesys.com/2008/02/pakistan-hijacks-youtube-1/

Bull, H. (1977). *The Anarchical Society: A Study of Order in World Politics*. New York: Columbia University Press.

Bush, R. (2012, August). *A Final IP Address Policy Proposal.* (See http://conference.apnic.net/34/policy for an initial description, see http://conference.apnic.net/34/policy/transcript for the transcript.)

Bush, R., Pelsser, C., Kuhne, M., Maennel, O., Mohapatra, P., Patel, K., & Evans, R. (2013, January). *RIPE Routing Working Group Recommendations on Route Flap Damping*. Retrieved from http://www.ripe.net/ripe/docs/ripe-580

Büthe, T., & Mattli, W. (2011). *The New Global Rulers: The Privatization of Regulation in the World Economy*. Princeton, NJ: Princeton University Press.

Butler, K., Farley, T., McDaniel, P., & Rexford, J. (2010). A Survey of BGP Security Issues and Solutions. *Proceedings of the IEEE, 98*(1), 100–122. doi: 10.1109/JPROC.2009.2034031

CABASE. (2015). *NAPs en Funcionamiento*. Retrieved from http://www.cabase.org.ar/wordpress/naps-en-funcionamiento/

Carstensen, S. (2014, 29 March). Google's Public DNS Intercepted in Turkey. *Google Online Security Blog*. Retrieved from http://googleonlinesecurity.blogspot.co.il/2014/03/googles-public-dns-intercepted-in-turkey.html

Castro, I., Cardona, J. C., Gorinsky, S., & Francois, P. (2014, 2–5 December). Remote Peering: More Peering Without Internet Flattening. In *CoNEXT'14*. Sydney, Australia.

CenturyLink. (2014, 18 September). *FAQs For Law Enforcement Agencies*. Retrieved from http://www.centurylink.com/static/Pages/AboutUs/Legal/LawEnforcement/agencyFAQ.html

Cerf, V. (1990, August). *IAB recommended policy on distributing internet identifier assignment and IAB recommended policy change to internet "connected" status* (No. 1174). RFC 1174 (Informational). IETF. Retrieved from http://www.ietf.org/rfc/rfc1174.txt

Cerny, P. G. (2010). *Rethinking World Politics: A Theory of Transnational Neopluralism*. New York, NY, USA: Oxford University Press.

Chartrand, G., & Oellermann, O. R. (1993). *Applied and Algorithmic Graph Theory*. New York, NY: McGraw-Hill.

Chharia, R. (2010, August). *Proposal for Government Advisory Committee (GAC) in APNIC*. Retrieved from http://meetings.apnic.net/__data/assets/text_file/0017/23327/gac-proposal.txt

Ciriacy-Wantrup, S. V., & Bishop, R. C. (1975). Common Property as a Concept in Natural Resources Policy. *Natural Resources Journal, 15*, 713. Retrieved from http://heinonline.org/HOL/Page?handle=hein.journals/narj15&id=731&div=&collection=journals

Claffy, K., & Clark, D. (2013). Platform Models for Sustainable Internet Regulation. In *Proceedings of the 41st Research Conference on Communication, Information and Internet Policy*. George Mason Uinversity, Arlington, VA.

Clark, C. W. (1980, January). Restricted Access to Common-Property Fishery Resources: A Game-Theoretic Analysis. In P.-T. Liu (Ed.), *Dynamic Optimization and Mathematical Economics* (pp. 117–132). Springer US. Retrieved 2014-12-31, from http://link.springer.com/chapter/10.1007/978-1-4684-3572-6_7

Clark, D., Lehr, W., & Bauer, S. (2011, 23–25 September). Interconnection in the Internet: The Policy Challenge. In *Proceedings of the 39th Research Conference on Communication, Information and Internet Policy*. George Mason Uinversity, Arlington, VA.

Clark, D. D., Partridge, C., Ramming, J. C., & Wroclawski, J. T. (2003). A Knowledge

Plane for the Internet. In *Proceedings of the 2003 conference on applications, technologies, architectures, and protocols for computer communications* (pp. 3–10). New York, NY, USA: ACM. Retrieved 2014-02-03, from http://doi.acm.org/10.1145/863955.863957 doi: 10.1145/863955.863957

Clay, K., & Wright, G. (2012). Gold Rush Legacy: American Minerals and the Knowledge Economy. In D. H. Cole & E. Ostrom (Eds.), *Property in Land and Other Resources* (chap. 3). Cambridge, Mass.: Lincoln Institute of Land Policy.

Coase, R. (1988). *The Firm, the Market, and the Law*. Chicago, IL: University of Chicago Press.

Coase, R. H. (1960, October). The Problem of Social Cost. *Journal of Law and Economics, 3*, 1–44. Retrieved from http://www.jstor.org/stable/724810

Cole, D. H., & Ostrom, E. (2012a). *Property in Land and Other Resources*. Cambridge, Mass.: Lincoln Institute of Land Policy.

Cole, D. H., & Ostrom, E. (2012b). The Variety of Property Systems and Rights in Natural Resources. In *Property in Land and Other Resources* (chap. 2). Cambridge, Mass.: Lincoln Institute of Land Policy.

Constant Contact. (2015). *Send Beautiful Emails. Get Real Results.* Retrieved from http://www.constantcontact.com/

Cowie, J. (2008, 11 November). Brazil Leak: If a Tree Falls in the Rainforest.... *Renesys Blog*. Retrieved from http://www.renesys.com/2008/11/brazil-leak-if-a-tree-falls-in/

Cowie, J. (2013, 19 November). The New Threat: Targeted Internet Traffic Misdirection. *Renesys Blog*. Retrieved from http://www.renesys.com/2013/11/mitm-internet-hijacking/ (Retrieved from http://www.renesys.com/2013/11/mitm-internet-hijacking/)

Crouch, A., Khosravi, H., Doria, A., Wang, X., & Ogawa, K. (2010, October). *Forwarding and Control Element Separation (ForCES) Applicability Statement* (No. 6041). RFC 6041 (Informational). IETF. Retrieved from http://www.ietf.org/rfc/rfc6041.txt

Cutler, A. C. (2003). *Private Power and Global Authority: Transnational Merchant Law in the Global Political Economy*. Cambridge, UK: Cambridge University Press.

Cutler, A. C., Haufler, V., & Porter, T. (Eds.). (1999). *Private Authority and International Affairs*. Albany, NY: State University of New York Press.

Daigle, L. (2004, September). *WHOIS Protocol Specification* (No. 3912). RFC 3912 (Draft Standard). IETF. Retrieved from http://www.ietf.org/rfc/rfc3912.txt

Davie, B., & Farrel, A. (2008). *MPLS: Next Steps* (Vol. 1). Morgan Kaufmann.

DE-CIX. (2015). *GlobePARTNER*. Retrieved from https://www.de-cix.net/products-services/de-cix-frankfurt/globepartner/

De Neufville, R. (1990). *Applied Systems Analysis: Engineering Planning and Technology Management*. New York: McGraw-Hill Companies.

Demsetz, H. (1967, May). Toward a Theory of Property Rights. *The American Economic Review, 57*(2), 347–359. Retrieved 2014-04-18, from http://www.jstor.org/stable/1821637

Doering, G. (2014, 6 November). *[anti-abuse-wg] Hijack Factory: AS201640 / AS200002.* Retrieved from http://www.ripe.net/ripe/mail/archives/anti-abuse-wg/2014-November/002712.html (E-mail response on the RIPE community's Anti-Abuse Working Group e-mail list.)

Dolšak, N., & Ostrom, E. (2003a). The Challenges of the Commons. In N. Dolšak & E. Ostrom (Eds.), *The Commons in the New Millenium: Challenges and Adaptation* (chap. 1). Cambridge, MA: The MIT Press. (Kindle Edition.)

Dolšak, N., & Ostrom, E. (Eds.). (2003b). *The Commons in the New Millenium: Challenges and Adaptation.* Cambridge, MA: The MIT Press. (Kindle Edition.)

Doria, A., Salim, J. H., Haas, R., Khosravi, H., Wang, W., Dong, L., ... Halpern, J. (2010, March). *Forwarding and Control Element Separation (ForCES) Protocol Specification* (No. 5810). RFC 5810 (Proposed Standard). IETF. Retrieved from http://www.ietf.org/rfc/rfc5810.txt

Dul, A. (2006, 17 February). *IPv6 HD-Ratio.* ARIN. Retrieved from https://www.arin.net/policy/proposals/2005_5.html

ECIX. (2015). *ecix european commercial internet exchange.* Retrieved from https://www.ecix.net/

Eggertsson, T. (1992). Analyzing Institutional Successes and Failures: A Millennium of Common Mountain Pastures in Iceland. *International Review of Law and Economics, 12*(4), 423—437.

Eggertsson, T. (2012). Opportunities and Limits for the Evolution of Property Rights Institutions. In *Property in Land and Other Resources* (chap. 1). Cambridge, Mass.: Lincoln Institute of Land Policy.

Ellickson, R. C. (1991). *Order Without Law: How Neighbors Settle Disputes.* Cambridge, MA: Harvard University Press.

Elmer-DeWitt, P., & Jackson, D. S. (1993, December). First Nation in Cyberspace. *Time, 142*(24), 62. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=9311307512&site=ehost-live

ENOG. (2015). *Eurasian Network Operator's Group: About.* Retrieved from http://www.enog.org/about/

Epstein, R. A. (2012). Playing by Different Rules? Property Rights in Land and Water. In D. H. Cole & E. Ostrom (Eds.), *Property in Land and Other Resources* (chap. 11). Cambridge, Mass.: Lincoln Institute of Land Policy.

Euro-IX. (2012). *What is an IXP?* Retrieved from https://www.euro-ix.net/what-is-an-ixp

Euro-IX. (2013a). *Euro-IX Peering Matrix.* Retrieved from https://www.euro-ix.net/tools/peering_matrix

Euro-IX. (2013b). *Euro-IX Peering Matrix.* Retrieved from https://www.euro-ix.net/tools/peering_matrix

Euro-IX. (2014). *European Internet Exchange Association.* Retrieved from https://www.euro-ix.net/

Euro-IX. (2015). *Euro-IX ASN Filter.* Retrieved from https://www.euro-ix.net/tools/asn_filter

Falk, J., & Kucherawy, M. (2012, June). *Creation and Use of Email Feedback Reports: An Applicability Statement for the Abuse Reporting Format (ARF)* (No. 6650).

RFC 6650 (Proposed Standard). IETF. Retrieved from http://www.ietf.org/rfc/rfc6650.txt

Faratin, P., Clark, D., Gilmore, P., Bauer, S., Berger, A., & Lehr, W. (2007). Complexity of Internet Interconnections: Technology, Incentives and Implications for Policy. In *Proceedings of the 35th Research Conference on Communication, Information and Internet Policy.*

FBI. (2011, November). *Operation Ghost Click International Cyber Ring That Infected Millions of Computers Dismantled.* Retrieved from http://www.fbi.gov/news/stories/2011/november/malware_110911/malware_110911

Feldman, M., & Chuang, J. (2005). The Evolution of Cooperation Under Cheap Pseudonyms. In *Proceedings of the Seventh IEEE Conference on E-Commerce Technology.*

Feldman, M., Papdimitriou, C., Chuang, J., & Stoica, I. (2006, May). Free-Riding and Whitewashing in Peer-to-Peer Systems. *IEEE Jorunal on Selected Areas in Communications, 24*(5), pp. 1010-1019.

Fight Club. (1999). Retrieved from http://www.imdb.com/title/tt0137523/

Flaim, B. (2009, 28 April). *ARIN Government Working Group Update.* ARIN. Retrieved from https://www.arin.net/participate/meetings/reports/ARIN_XXIII/ppm2_transcript.html#anchor_8

Flathman, R. E. (1980). *The Practice of Political Authority: Authority and the Authoritative.* Chicago: Univ of Chicago Pr.

France IX. (2011). *France IX Services.* Retrieved from https://www.franceix.net/media/cms_page_media/817/France_IX_Services_statutes.pdf

France-IX. (2015a). *Frequently Asked Questions.* Retrieved from https://www.franceix.net/en/solutions/faq/

France-IX. (2015b). *Missions: The End of the French Exception?* Retrieved from https://www.franceix.net/about-france-ix/missions/

Friedman, E. J., & Resnick, P. (2001, Summer). The Social Cost of Cheap Pseudonyms. *Journal of Economics & Management Strategy, 10*(2), pp. 173–199.

Frischmann, B. M. (2012). *Infrastructure: The Social Value of Shared Resources.* New York, NY, USA: Oxford University Press.

Gallo, A. (2014, 4 December). *[ARIN-PPML] RPKI Relyin Agreement.* Retrieved from http://lists.arin.net/pipermail/arin-ppml/2014-December/029377.html (Gallo's message is the entry point into a thread discussion RPKI and liability in the ARIN region.)

Gereffi, G., Humphrey, J., & Sturgeon, T. (2005, February). The governance of global value chains. *Review of International Political Economy, 12,* 78–104. Retrieved from http://www.tandfonline.com.libproxy.mit.edu/doi/abs/10.1080/09692290500049805 doi: 10.1080/09692290500049805

Gerich, E. (1992, October). *Guidelines for Management of IP Address Space* (No. 1366). RFC 1366 (Informational). IETF. Retrieved from http://www.ietf.org/rfc/rfc1366.txt (Obsoleted by RFC 1466)

Gill, P., Arlitt, M., Li, Z., & Mahanti, A. (2008). The Flattening Internet Topology: Natural Evolution, Unsightly Barnacles or Contrived Collapse? In M. Claypool

& S. Uhlig (Eds.), *Passive and Active Network Measurement* (Vol. 4979, pp. 1–10). Springer Berlin / Heidelberg. Retrieved from http://www.springerlink.com/content/1255p8g3k6766242/abstract/

Gilmore, P. (2014, 15 January). *best practice for advertising peering fabric routes.* Retrieved from http://mailman.nanog.org/pipermail/nanog/2014-January/063506.html (Response by Gilmore to a NANOG post regarding advertisement of IX prefixes.)

Greene, B. (2012a, 10 August). *Beware! DNS Changer's IP Blocks Are Reallocated and Advertised!* Retrieved from http://web.archive.org/web/20120813224320/http://www.senki.org/archives/930

Greene, B. (2012b, 15 August). *RIPE NCC Responds to the Rove Digital/DNS Changer Re-allocations.* Retrieved from http://web.archive.org/web/20130801035523/http://www.senki.org/archives/948

Grundemann, C. (2013, June). IPv6 Security Myths. In *NANOG58.* New Orleans, LA. Retrieved from https://www.nanog.org/sites/default/files/best_current_operations_practices.mp4 (Introductory comments by Aaron Hughes.)

Haas, P. (1989, Summer). Do Regimes Matter? Epistemic Communities and Mediterranean Pollution Control. *International Organization, 43*(3), 377–403.

Haas, P. M. (1990, December). Obtaining International Environmental Protection through Epistemic Consensus. *Millennium: Journal of International Studies, 19,* 347–363.

Haas, P. M. (1992). Introduction: Epistemic Communities and International Policy Coordination. *International Organization, 46*(1), 1.

Haas, R. (2010, March). *Forwarding and Control Element Separation (ForCES) MIB* (No. 5813). RFC 5813 (Proposed Standard). IETF. Retrieved from http://www.ietf.org/rfc/rfc5813.txt

Haleplidis, E., Koufopavlou, O., & Denazis, S. (2011, September). *Forwarding and Control Element Separation (ForCES) Implementation Experience* (No. 6369). RFC 6369 (Informational). IETF. Retrieved from http://www.ietf.org/rfc/rfc6369.txt

Haleplidis, E., Ogawa, K., Wang, W., & Salim, J. H. (2010, November). *Implementation Report for Forwarding and Control Element Separation (ForCES)* (No. 6053). RFC 6053 (Informational). IETF. Retrieved from http://www.ietf.org/rfc/rfc6053.txt

Halpern, J., & Salim, J. H. (2010, March). *Forwarding and Control Element Separation (ForCES) Forwarding Element Model* (No. 5812). RFC 5812 (Proposed Standard). IETF. Retrieved from http://www.ietf.org/rfc/rfc5812.txt

Hanna, S., Folke, C., & Mäler, K.-G. (Eds.). (1996). *Rights to Nature: Ecological, Cultural, and Political Principles of Institutions for the Environment.* Washington, DC: Island Press. (Kindle Edition)

Hardin, G. (1968, December). The Tragedy of the Commons. *Science, 162*(3859), 1243–1248. Retrieved from http://www.jstor.org/stable/1724745

Hart, H. (1994). *The Concept of Law* (2nd ed.). Oxford University Press.

Hawkins, D. G., Lake, D. A., Nielson, D. L., & Tierney, M. J. (Eds.). (2006). *Delegation and Agency in International Organizations*. Cambridge University Press.

Hess, C., & Ostrom, E. (n.d.). Introduction: An Overview of the Knowledge Commons. In *Understanding Knowledge as a Commons: From Theory to Practice* (chap. 1).

Hess, C., & Ostrom, E. (2003, January). Ideas, Artifacts, and Facilities: Information as a Common-Pool Resource. *Law and Contemporary Problems, 66*(1/2), 111–145. Retrieved from http://www.jstor.org/stable/20059174

Hoffman, P. (Ed.). (2012). *The Tao of the IETF: A Novice's Guide to the Internet Engineering Task Force*. IETF. Retrieved from http://www.ietf.org/tao.html

Hohfeld, W. N. (1917). Fundamental Legal Conceptions as Applied in Judicial Reasoning. *Yale Law Journal*(26), 710–770.

Holton, G. (1999, Fall). A Vision of Jeffersonian Science. *Issues in Science and Technology Online*. Retrieved from http://www.issues.org/16.1/holton.htm

Housley, R., Curran, J., Huston, G., & Conrad, D. (2013, August). *The Internet Numbers Registry System* (No. 7020). RFC 7020 (Informational). IETF. Retrieved from http://www.ietf.org/rfc/rfc7020.txt

Howard, L., & Sowell, J. H. (2014, 12–14 September). A Comparison of Public Policy Approaches to the IPv4-IPv6 Transition. In *Proceedings of the 42nd Research Conference on Communication, Information and Internet Policy*. Arlington, VA: Telecommunications Policy Research Consortium.

Hubbard, K., Kosters, M., Conrad, D., Karrenberg, D., & Postel, J. (1996, November). *Internet Registry IP Allocation Guidelines* (No. 2050). RFC 2050 (Best Current Practice). IETF. Retrieved from http://www.ietf.org/rfc/rfc2050.txt

Hurricane Electric. (2014). *Hurricane Electric Acceptable Use Policy (AUP)*. Retrieved from https://www.he.net/aup.html

Huston, G., Michaelson, G., Martinez, C., Bruijnzeels, T., Newton, A., & Aina, A. (2015, 25 January). *RPKI Validation Reconsidered*. Retrieved from https://tools.ietf.org/html/draft-ietf-sidr-rpki-validation-reconsidered-01

Huston, G., & Smith, P. (2010, 10 February). *IPv4 Address Transfers*. APNIC. Retrieved from https://www.apnic.net/policy/proposals/prop-050

Hutty, M. (2011, 3 May). *[address-policy-wg] 2008-08 (Initial Certification Policy in the RIPE NCC Service Region) going to Last Call*. RIPE NCC. Retrieved from http://www.ripe.net/ripe/mail/archives/address-policy-wg/2011-May/005737.html

ICANN. (2001, 4 June). *ICP-2: Criteria for Establishment of New Regional Internet Registries*. ICANN. Retrieved from http://www.icann.org/en/resources/policy/global-addressing/new-rirs-criteria

ICANN. (2009, 6 March). *Global Policy for the Allocation of the Remaining IPv4 Address Space*. Retrieved from https://www.icann.org/resources/pages/remaining-ipv4-2012-02-25-en

ICANN. (2012a, 6 May). *Global Policy for Post Exhaustion IPv4 Allocation Mech-*

anisms by the IANA. Retrieved from https://www.icann.org/resources/pages/allocation-ipv4-post-exhaustion-2012-05-08-en

ICANN. (2012b, 25 February). *Internet Assigned Numbers Authority (IANA) Policy For Allocation of IPv4 Blocks to Regional Internet Registries.* IANA. Retrieved from https://www.icann.org/resources/pages/allocation-ipv4-rirs-2012-02-25-en

ICANN. (2014, 20 May). *Remaining IPv4 Addresses to be Redistributed to Regional Internet Registries — Address Redistribution Signals that IPv4 is Nearing Total Exhaustion.* Retrieved from https://www.icann.org/news/announcement-2-2014-05-20-en

Jasanoff, S. (1987). Contested boundaries in policy-relevant science. *Social Studies of Science, 17*(2), 195–230.

Jasinska, E., Lucente, P., van Dussen, B., Vijn, A., & Hughes, A. (2012, 23 October). Traffic Accounting. In *NANOG 56.* NANOG. (Retrieved from https://www.nanog.org/meetings/abstract?id=2016.)

Karrenberg, D., Pawlik, A., Band, A., Hutty, M., Freedman, D., Murphy, S., & Kent, S. (2011). *Plenary and RPKI Discussion Panel Transcript.* RIPE NCC. Retrieved from https://ripe63.ripe.net/archives/steno/5/

Karrenberg, D., Ross, G., Wilson, P., & Nobile, L. (2001, December). Development of the Regional Internet Registry System. *The Internet Protocol Journal, 4*(4), 17–29. Retrieved from http://www.cisco.com/web/about/ac123/ac147/archived_issues/ipj_4-4/regional_internet_registries.html

Kelling, G. L., & Wilson, J. Q. (1982, March). Broken Windows. *The Atlantic.* Retrieved from http://www.theatlantic.com/magazine/archive/1982/03/broken-windows/304465/

Kende, M., & Hurpy, C. (2012, April). *Assessment of the Impact of Internet Exchange Points—Empirical Study of Kenya and Nigeria* (Report for the Internet Society No. 20945-144). Internet Society. Retrieved from http://internetsociety.org/ixpimpact

Keohane, R., & Victor, D. (2010, January). *The Regime Complex for Climate Change* (Discussion Paper No. 10-33). Harvard University. Retrieved from http://belfercenter.hks.harvard.edu/files/KeohaneVictor_Final.pdf

Keohane, R. O. (1971, April). The Big Influence of Small Allies. *Foreign Policy*(2), 161–182. Retrieved 2014-12-24, from http://www.jstor.org/stable/1147864 doi: 10.2307/1147864

Keohane, R. O. (2005). *After Hegemony.* Princeton University Press.

Klensin, J. (2008, October). *Simple Mail Transfer Protocol* (No. 5321). RFC 5321 (Draft Standard). IETF. Retrieved from http://www.ietf.org/rfc/rfc5321.txt

Kohl, U. (2007). *Jurisdiction and the Internet.* Cambridge University Press.

Kuerbis, B., Asghari, H., & Mueller, M. (2013, 21–23 September). In the Eye of the Beholder: The Role of Needs-Based Assessment in IP Address Market Transfers. In *Proceedings of the 41st Research Conference on Communication, Information and Internet Policy.* Arlington, VA: Telecommunications Policy Research Consortium.

Kuhn, T. S. (1993). *The Structure of Scientific Revolutions* (3rd ed.). University of Chicago Press.

Labovitz, C., Ahuja, A., Bose, A., & Jahanian, F. (2000). Delayed Internet Routing Convergence. In *Proceedings of the conference on applications, technologies, architectures, and protocols for computer communication* (pp. 175–187). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/347059.347428 doi: 10.1145/347059.347428

Labovitz, C., Ahuja, A., & Jahanian, F. (1999, June). Experimental Study of Internet Stability and Wide-area Network Failures. In *Proc. International Symposium on Fault Tolerant Computing.*

Labovitz, C., Iekel-Johnson, S., McPherson, D., Oberheide, J., & Jahanian, F. (2010, 30 August–3 September). Internet Inter-Domain Traffic. In *SIGCOMM'10.*

LACNIC. (2014a, February). *LACNIC Meets with Colombian Authorities and ISPs to Discuss the New Internet Protocol.* Retrieved from http://www.lacnic.net/en/web/anuncios/2014-visitas-informativas-colombia

LACNIC. (2014b, February). *LACNIC Meets with Ecuatorian Authorities and ISPs to Discuss the New Internet Protocol.* Retrieved from http://www.lacnic.net/en/web/anuncios/2014-visitas-informativas-ecuador

LACNIC. (2014c, February). *LACNIC Meets with Panamanian Authorities and ISPs to Discuss the New Internet Protocol.* Retrieved from http://www.lacnic.net/en/web/anuncios/2014-visitas-informativas-panama

LACNIC. (2014d, February). *LACNIC Meets with T&T and Peruvian Authorities and ISPs to Discuss the New Internet Protocol.* Retrieved from http://www.lacnic.net/en/web/anuncios/2014-visitas-informativas-peru-tyt

LACNIC. (2014e, February). *LACNIC Meets with Venezuelan Authorities and ISPs to Discuss the New Internet Protocol.* Retrieved from http://www.lacnic.net/en/web/anuncios/2014-visitas-informativas

LACNIC. (2014f, 25 March). *LACNIC Policy Manual* (v2.1 ed.). Retrieved from http://www.lacnic.net/documents/10834/21254/manual-politicas-en-2.1.pdf

LACNIC. (2014g). *Registration Services Agreement.* Retrieved from http://lacnic.net/docs/rsa-en.pdf

LACNIC. (2015). *LACNIC Initiatives / Projects.* Retrieved from http://www.lacnic.net/web/lacnic/iniciativas

Lake, D. A. (2006, September). *Relational Authority in the Modern World: Towards a Positive Theory of Legitimacy* (SSRN Scholarly Paper No. ID 1004424). Rochester, NY: Social Science Research Network. Retrieved 2014-01-28, from http://papers.ssrn.com/abstract=1004424

Lake, D. A. (2009, November). Relational Authority and Legitimacy in International Relations. *American Behavioral Scientist, 53*(3), 331–353. Retrieved from http://abs.sagepub.com.libproxy.mit.edu/content/53/3/331 doi: 10.1177/0002764209338796

Lake, D. A. (2010). Rightful Rules: Authority, Order, and the Foundations of Global Governance. *International Studies Quarterly, 54*, 587–613.

Layton, E. T. (1979, January). Scientific Technology, 1845-1900: The Hydraulic

Turbine and the Origins of American Industrial Research. *Technology and Culture, 20*(1), 64–89. Retrieved from http://www.jstor.org/stable/3103112

Lepinski, M., & Kent, S. (2012, February). *An Infrastructure to Support Secure Internet Routing* (No. 6480). RFC 6480 (Informational). IETF. Retrieved from http://www.ietf.org/rfc/rfc6480.txt

Lessig, L. (1999). *Code and Other Laws of Cyberspace*. New York, NY: Basic Books.

Levi, M. (1989). *Of Rule and Revenue* (Kindle Edition ed.). University of California Press.

Libecap, G. D. (2012). Water Rights and Markets in the U.S. Semiarid West. In D. H. Cole & E. Ostrom (Eds.), *Property in Land and Other Resources* (chap. 13). Cambridge, Mass.: Lincoln Institute of Land Policy.

LINX. (2013). *History of LINX*. Retrieved from https://www.linx.net/about/history-of-linx.html

LINX. (2014a). *Dispute Resolution Procedure*. Retrieved from https://www.linx.net/govern/index.html

LINX. (2014b). *LINX Fees Schedule*. Retrieved from https://www.linx.net/govern/servicesfees.html

LINX. (2015a). *The Consultation Process at LINX*. Retrieved from https://www.linx.net/about/consult.html

LINX. (2015b). *LINX from Anywhere*. Retrieved from https://www.linx.net/join/linxanywhere.html

LINX. (2015c). *Memorandum of Association of London Internet Exchange Limited and Articles of Association of London Internet Exchange Limited*. Retrieved from https://www.linx.net/govern/manda.html

LONAP. (2014a). *Our Network Infrastructure*. Retrieved from http://www.lonap.net/network.shtml

LONAP. (2014b). *Route Servers*. Retrieved from http://www.lonap.net/routeservers.shtml

M^{3}AAWG. (2008, June). *Trust in Email Begins with Authentication* [Technical Report]. Retrieved from https://www.m3aawg.org/sites/maawg/files/news/MAAWG_Email_Authentication_Paper_2008-07.pdf

M^{3}AAWG. (2011a). *MAAWG Sender Best Communications Practices*. Messaging, Malware and Mobile Anti-Abuse Working Group (M^{3}AAWG). Retrieved from http://www.maawg.org/sites/maawg/files/news/MAAWG_Senders_BCP_Ver2a-updated.pdf

M^{3}AAWG. (2011b, November). *Messaging Anti-Abuse Working Group (MAAWG) Vetting Best Common Practices (BCP)*. Messaging, Malware and Mobile Anti-Abuse Working Group (M^{3}AAWG). Retrieved from http://www.maawg.org/sites/maawg/files/news/MAAWG_Vetting_BCP_2011-11.pdf

M^{3}AAWG. (2013, October). *M^{3}AAWG Best Current Practices for Building and Operating a Spamtrap* (Tech. Rep. No. M3AAWG075). Author. Retrieved from http://www.maawg.org/sites/maawg/files/news/M3AAWG_Spamtrap_Operations_BCP-2013-10.pdf

M^{3}AAWG. (2014a, June). *Help—I'm on a Blocklist*. Messaging, Malware and Mobile

Anti-Abuse Working Group (M³AAWG). Retrieved from https://www.m3aawg.org/sites/maawg/files/news/M3AAWG_Blocklist_Help_BP_2014-06.pdf

M³AAWG. (2014b, 26 September). *M³AAWG Comments on Implementation of CSRIC III Cybersecurity Best Practices.* Retrieved from https://www.maawg.org/sites/maawg/files/news/M3AAWG_FCC_CRSIC_III_Cybersecurity_2014-09.pdf

M³AAWG. (2014c, June). *M³AAWG Document Development Guidelines.* (Internal BCP development documentation for M³AAWG members.)

M³AAWG. (2014d, November). *M³AAWG Email Metrics Report* (Report No. 16). Retrieved from https://www.maawg.org/sites/maawg/files/news/M3AAWG_2012-2014Q2_Spam_Metrics_Report16.pdf

M³AAWG. (2014e, February). *M³AAWG Feedback Reporting Recommendation.* Messaging, Malware and Mobile Anti-Abuse Working Group (M³AAWG). Retrieved from https://www.maawg.org/sites/maawg/files/news/M3AAWG_Feedback_Reporting_Recommendation_BP-2014-02.pdf

M³AAWG. (2015). *Public Policy Committee.* Retrieved from https://www.maawg.org/activities/public-policy-committee

Mao, Z. M., Govindan, R., Varghese, G., & Katz, R. H. (2002). Route Flap Damping Exacerbates Internet Routing Convergence. In *Proceedings of the 2002 conference on applications, technologies, architectures, and protocols for computer communications* (pp. 221–233). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/633025.633047 doi: 10.1145/633025.633047

MAPS. (2004, June). *Introduction to the Realtime Blackhole List (RBL) Servers.* Retrieved from http://web.archive.org/web/20040701065615/http://www.mail-abuse.com/support/wp_introrbl.html

Markoff, J., & Perlroth, N. (2013, March). Online Dispute Becomes Internet-Snarling Attack. *The New York Times.* Retrieved 2014-01-30, from http://www.nytimes.com/2013/03/27/technology/internet/online-dispute-becomes-internet-snarling-attack.html

Masnick, M. (2011, 6 September). *e360's $11 Million Win Against Spamhaus... Now Reduced To Just $3 (Not $3 Million, But Just $3).* Retrieved from https://www.techdirt.com/articles/20110903/00560215803/e360s-11-million-win-against-spamhaus-now-reduced-to-just-3-not-3-million-just-3.shtml

Mattli, W., & Woods, N. (2009a). In Whose Benefit? Explaining Regulatory Change in Global Politics. In W. Mattli & N. Woods (Eds.), *The Politics of Global Regulation* (pp. 1–43). Princeton, N.J: Princeton University Press.

Mattli, W., & Woods, N. (Eds.). (2009b). *The Politics of Global Regulation.* Princeton, N.J: Princeton University Press.

Mayer-Schonberger, V. (2002). Shape of Governance: Analyzing the World of Internet Regulation, The. *Virginia Journal of International Law, 43,* 605. Retrieved from http://heinonline.org/HOL/Page?handle=hein.journals/vajint43&id=615&div=&collection=journals

McCray, L. (2003, August). *Doing Believable Knowledge Assessment for Policymaking:*

How Six Prominent Organizations Go About It.

McCray, L., & Oye, K. A. (2006, October). Adaptation and Anticipation: Learning from Policy Experience. In *NSF-EPA Trans-Atlantic Uncertainty Colloquium.*

McCray, L. E., Oye, K. A., & Petersen, A. C. (2010, July). Planned adaptation in risk regulation: An initial survey of US environmental, health, and safety regulation. *Technological Forecasting and Social Change, 77*(6), 951–959. Retrieved from http://www.sciencedirect.com/science/article/pii/S0040162509001942 doi: 10.1016/j.techfore.2009.12.001

McKean, M. A. (1996). Common-Property Regimes as a Solution to Problems of Scale and Linkage. In S. Hanna, C. Folke, & K.-G. Mäler (Eds.), *Rights to Nature: Ecological, Cultural, and Political Principles of Institutions for the Environment.* Washington, DC: Island Press.

McPherson, D., & Patel, K. (2006, January). *Experience with the BGP-4 Protocol* (No. 4277). RFC 4277 (Informational). IETF. Retrieved from http://www.ietf.org/rfc/rfc4277.txt

MENOG. (2015). *The Middle East Network Operators Group.* Retrieved from http://www.menog.org/

Meyer, D., & Patel, K. (2006, January). *BGP-4 Protocol Analysis* (No. 4274). RFC 4274 (Informational). IETF. Retrieved from http://www.ietf.org/rfc/rfc4274.txt

Meyer, D., Schmitz, J., Orange, C., Prior, M., & Alaettinoglu, C. (1999, August). *Using RPSL in Practice* (No. 2650). RFC 2650 (Informational). IETF. Retrieved from http://www.ietf.org/rfc/rfc2650.txt

Mitchell, K. (2005, 25 May). The UK Network Operator's Forum. In *UKNOF 1 Meeting.* London, UK. Retrieved from http://www.uknof.org.uk/uknof0/UKNOF-intro.pdf

Moss, D. A. (2004). *When All Else Fails: Government as the Ultimate Risk Manager.* Harvard University Press.

Mueller, M. (2002). *Ruling the Root.* Cambridge, MA: The MIT Press.

Mueller, M., Kuerbis, B., & Asghari, H. (2013, September). Dimensioning the Elephant: An Empirical Analysis of the IPv4 Number Market. *info, 15*(6), 6–18. Retrieved from http://www.emeraldinsight.com/doi/abs/10.1108/info-07-2013-0039 doi: 10.1108/info-07-2013-0039

NaMeX. (2015). *Nautilus Mediterranean Exchange Point.* Retrieved from https://www.namex.it/

NANOG. (2012, August). *About NANOG.* Retrieved from http://www.nanog.org/about/ (Retrieved from http://www.nanog.org/about/.)

NANOG. (2015a). *Guidelines for Presenting at a NANOG Meeting.* Retrieved from https://www.nanog.org/meetings/presentation/guidelines

NANOG. (2015b). *The History of NANOG.* Retrieved from https://www.nanog.org/history

NANOG. (2015c). *How to Give a Presentation at NANOG.* Retrieved from https://www.nanog.org/meetings/presentation

NANOG. (2015d). *NANOG and Internet Governance.* Retrieved from http://nanog.org/governance/home

NANOG. (2015e). *Tutorials.* Retrieved from https://www.nanog.org/resources/tutorials

Netnod. (2014, June). *Netnode IX Standard Contract.* Retrieved from http://www.netnod.se/sites/default/files/Netnod-IX-Contract-20140625.pdf

Netnod. (2015). *Netnod Reach.* Retrieved from http://www.netnod.se/ix/reach

NL-ix. (2015). *About Us.* Retrieved from https://www.nl-ix.net/about/company/

North, D. (1990). *Institutions, Institutional Change, and Economic Performance.* Cambridge University Press.

NRO. (2003a, 24 October). *NRO Memorandum of Understanding.* Retrieved from http://www.nro.net/documents/nro-memorandum-of-understanding

NRO. (2003b, 24 October). *NRO Memorandum of Understanding FAQ.* Retrieved from http://www.nro.net/documents/nro-memorandum-of-understanding-faq

NRO. (2003c, 6 November). *Open Letter to ICANN from the Regional Internet Registries.* Retrieved from http://www.nro.net/news/open-letter-to-icann-from-the-regional-internet-registries

NRO. (2011). *Free Pool of IPv4 Address Space Depleted.* Retrieved from https://www.nro.net/news/ipv4-free-pool-depleted

NRO. (2013, 30 April). *NRO Executive Council Appoints Permanent NRO Executive Secretary.* Retrieved from http://www.nro.net/news/nro-executive-council-appoints-permanent-nro-executive-secretary

NRO. (2014a). *Delegated-extended Statistics FIle.* Retrieved from http://www.nro.net/pub/stats/nro/delegated-extended

NRO. (2014b). *NRO: Number Resource Organization.* Retrieved from http://www.nro.net/

NRO. (2014c). *Resource Certification (RPKI).* Retrieved from http://www.nro.net/about-the-nro/technical-coordination/certification

NRO. (2014d, 17 April). *RIR Comparative Policy Overview.* Retrieved from http://www.nro.net/rir-comparative-policy-overview/rir-comparative-policy-overview-2014-01

NRO. (2014e). *Statistics Format.* Retrieved from http://www.nro.net/wp-content/uploads/nro-extended-stats-readme5.txt

NRO. (2015). *The NRO Number Council.* Retrieved from https://www.nro.net/about-the-nro/the-nro-number-council

Nye, Jr., J. S., & Keohane, R. O. (1971). Transnational Relations and World Politics: An Introduction. *International Organization, 25*(3), 329–349.

Olson, M. (1965). *The Logic of Collective Action.* Harvard University Press.

Open-IX. (2014a). *Data Center Technical Standards - OIX2.* Retrieved from http://www.open-ix.org/standards/data-center-technical-requirements/

Open-IX. (2014b). *IXP Technical Standards OIX-1.* Retrieved from http://www.open-ix.org/certification/ixp-technical-requirements/

Ostrom, E. (1986, January). An Agenda for the Study of Institutions. *Public Choice, 48*(1), 3–25. Retrieved from http://link.springer.com/article/10.1007/BF00239556 doi: 10.1007/BF00239556

Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge; New York: Cambridge University Press. ((Kindle Edition))

Ostrom, E. (2005). *Understanding Institutional Diversity*. Princeton University Press.

Ostrom, E., & Schlager, E. (1996). The Formation of Property Rights. In S. Hanna, C. Folke, & K.-G. Mäler (Eds.), *Rights to Nature: Ecological, Cultural, and Political Principles of Institutions for the Environment* (pp. 127–156). Washington, DC: Island Press.

Ostrom, V., Tiebout, C. M., & Warren, R. (1961). The Organization of Government in Metropolitan Areas: A Theoretical Inquiry. *American Political Science Review, 55*, 831–842.

PacNOG. (2005). *First PacNOG Meeting and Educational Workshop*. Retrieved from `https://www.pacnog.org/pacnog1/`

PacNOG. (2015). *The Pacific Network Operators Group - Organisational Structure*. Retrieved from `https://www.pacnog.org/org.html`

Pelsser, C., Maennel, O., Mohapatra, P., Bush, R., & Patel, K. (2011, January). Route Flap Damping Made Usable. In N. Spring & G. F. Riley (Eds.), *Passive and active measurement* (pp. 143–152). Springer Berlin Heidelberg. Retrieved from `http://link.springer.com/chapter/10.1007/978-3-642-19260-9_15`

Peterson, L., & Davie, B. (2011). *Computer Networks: A Systems Approach*. Morgan Kaufmann. (Kindle Edition.)

Pickett, J. P. (Ed.). (2002). *The American Heritage College Dictionary* (Fourth ed.). Boston, MA: Houghton Mifflin Company.

PLNOG. (2015). *About*. Retrieved from `https://warszawa.plnog.pl/about/` (Retrieved from `https://warszawa.plnog.pl/about/`.)

Pollard, J. (2013, 23 April). How to Delist Your IPs at Spamhaus. *Return Path Blog*. Retrieved from `http://blog.returnpath.com/blog/john-pollard/how-to-delist-your-ips-at-spamhaus`

Porter, T. (1999). Hegemony and the Private Governance of International Industries. In A. C. Cutler, V. Haufler, & T. Porter (Eds.), *Private Authority and International Affairs* (chap. 10). Albany, NY: State University of New York Press.

Postel, J. (1981, September). *Internet Protocol* (No. 791). RFC 791 (Standard). IETF. Retrieved from `http://www.ietf.org/rfc/rfc791.txt` (Updated by RFCs 1349, 2474)

Rekhter, Y., Li, T., & Hares, S. (2006, January). *A Border Gateway Protocol 4 (BGP-4)* (No. 4271). RFC 4271 (Draft Standard). IETF. Retrieved from `http://www.ietf.org/rfc/rfc4271.txt` (Updated by RFCs 6286, 6608)

Rekhter, Y., Moskowitz, B., Karrenberg, D., de Groot, G. J., & Lear, E. (1996, February). *Address Allocation for Private Internets* (No. 1918). RFC 1918 (Best Current Practice). IETF. Retrieved from `http://www.ietf.org/rfc/rfc1918.txt`

Resnick, P. (2014, April). *On Consensus and Humming in the IETF* (Internet Draft). Internet Engineering Task Force. Retrieved from `https://datatracker`

.ietf.org/doc/draft-resnick-on-consensus/ (Retrieved from https://datatracker.ietf.org/doc/draft-resnick-on-consensus/.)

Return Path. (2014a). *About Sender Score.* Retrieved from https://www.senderscore.org/senderscore/

Return Path. (2014b). *Reputation Network Blacklist.* Retrieved from http://www.returnpath.com/wp-content/uploads/resource/reputation-network-blacklist/Return-Path-Reputation-Network-Blacklist.pdf

Return Path. (2015). *See More. Know More. Matter More.* Retrieved from http://www.returnpath.com/

RIPE. (2014a). *Cooperation Working Group.* RIPE NCC. Retrieved from http://www.ripe.net/ripe/groups/wg/coop

RIPE. (2014b). *The History of RIPE.* Retrieved from http://www.ripe.net/ripe/about/the-history-of-ripe (Retrieved from http://www.ripe.net/ripe/about/the-history-of-ripe.)

RIPE. (2014, February). *IPv4 Address Allocation and Assignment Policies for the RIPE NCC Service Region* (No. RIPE-606). Retrieved from http://www.ripe.net/ripe/docs/ripe-606

RIPE. (2014a). *RIPE NCC Conflict Arbitration Procedure.* Retrieved from http://www.ripe.net/ripe/docs/ripe-613

RIPE. (2014b). *RIPE Policies.* RIPE NCC. Retrieved from http://www.ripe.net/ripe/docs/current-ripe-documents/ripe-policies

RIPE. (2015). *Presentation Guidelines.* Retrieved from https://ripe69.ripe.net/submit-topic/guidelines/

RIPE NCC. (2011a, 9 November). *RIPE NCC Blocks Registration in RIPE Registry Following Order from Dutch Police.* Retrieved from http://www.ripe.net/internet-coordination/news/about-ripe-ncc-and-ripe/ripe-ncc-blocks-registration-in-ripe-registry-following-order-from-dutch-police

RIPE NCC. (2011b, 16 November). *RIPE NCC Intends to Seek Clarification from Dutch Court on Police Order to Temporarily Lock Registration.* Retrieved from http://www.ripe.net/internet-coordination/news/about-ripe-ncc-and-ripe/ripe-ncc-to-seek-clarification-from-dutch-court-on-police-order-to-temporarily-lock-registration

RIPE NCC. (2012a, 15 August). *Clarification on Reallocated IPv4 Address Space Related to Dutch Police Order.* Retrieved from http://www.ripe.net/internet-coordination/news/announcements/clarification-on-reallocated-ipv4-address-space-related-to-dutch-police-order

RIPE NCC. (2012b, 10 January). *RIPE NCC Unlocks Registration in RIPE Registry.* Retrieved from http://www.ripe.net/internet-coordination/news/about-ripe-ncc-and-ripe/ripe-ncc-unlocks-registration-in-ripe-registry

RIPE NCC. (2013a). *EIX Working Group.* Retrieved from http://www.ripe.net/ripe/groups/inactive-working-groups/eix

RIPE NCC. (2013b, 26 April). *IPv6 Transition Mechanisms.* Re-

trieved from http://www.ripe.net/lir-services/training/e-learning/ipv6/transition-mechanisms

RIPE NCC. (2013c, 14 February). *The RIPE NCC's Case Against the State of the Netherlands Dismissed.* Retrieved from http://www.ripe.net/internet-coordination/news/about-ripe-ncc-and-ripe/ripe-nccs-case-against-the-state-of-the-netherlands-dismissed

RIPE NCC. (2014a). *Atlas Results.* Retrieved from https://atlas.ripe.net/results/

RIPE NCC. (2014b). *Frequently Asked Questions.* Retrieved from https://atlas.ripe.net/about/faq/

RIPE NCC. (2014c, 14 January). *Improving the Readability of RIPE Documents.* Retrieved from http://www.ripe.net/ripe/readability

RIPE NCC. (2014d). *List of Arbiters.* Retrieved from http://www.ripe.net/lir-services/ncc/legal/arbitration/list-of-arbiters

RIPE NCC. (2014e). *RIPE Atlas - Map Visualisations.* Retrieved from https://atlas.ripe.net/results/maps/

RIPE NCC. (2014f, 30 December). *RoundTable Meetings.* Retrieved from http://www.ripe.net/ripe/meetings/roundtable

RIPE NCC. (2014g). *What is RIPE Atlas?* Retrieved from https://atlas.ripe.net/about/

RIPE NCC. (2015a). *Certification Statistics.* Retrieved from http://certification-stats.ripe.net/

RIPE NCC. (2015b). *Measurement, Analysis and Tools (MAT) Working Group.* Author. Retrieved from http://www.ripe.net/ripe/groups/wg/mat

RIPE NCC. (2015c). *RIPE NCC Executive Board - Functions and Expectations.* Retrieved from http://www.ripe.net/lir-services/ncc/executive-board/ripe-ncc-executive-board-functions-and-expectations

Russell, I. (2013, 6 August). Using the Broken Windows Theory to Tackle Antisocial Behaviour. *The Guardian.* Retrieved from http://www.theguardian.com/housing-network/2013/aug/06/antisocial-behaviour-broken-window-theory

Ryerse, S. (2014, 4 April). *Subject: [ARIN-PPML] ARIN-PPML DIGEST, VOL 106, ISSUE 8.* ARIN. Retrieved from http://lists.arin.net/pipermail/arin-ppml/2014-April/028174.html (ARIN PPML message dated 4 April 2014.)

Sanghani, B. (2013). *Euro-IX Twin Program of IXPs in Need.* Retrieved from https://www.euro-ix.net/documents/673-twin-ixp-program-pdf?download=yes

SANOG. (2015a). *SANOG Partnerships.* Retrieved from http://www.sanog.org/partnerships.htm

SANOG. (2015b). *South Asia Network Operators Group Committees.* Retrieved from http://www.sanog.org/committee.htm

Seastrom, R. (2006, 17 February). *Policy 2005-7: Rationalize Multi-Homing Definition and Requirement.* ARIN. Retrieved from https://www.arin.net/policy/proposals/2005_7.html

Security and Stability Advisory Committee. (2014, 15 August). *Overview and His-*

tory of the IANA Functions (Tech. Rep. No. SAC067). ICANN. Retrieved from https://www.icann.org/en/system/files/files/sac-067-en.pdf

SendGrid. (2014). *Warming Up an IP Address.* Retrieved from https://sendgrid.com/docs/User_Guide/warming_up.html

SendGrid. (2015). *Email Delivery. Simplified.* Retrieved from https://sendgrid.com/

Shafranovich, Y., Levine, J., & Kucherawy, M. (2010, August). *An Extensible Format for Email Feedback Reports* (No. 5965). RFC 5965 (Proposed Standard). IETF. Retrieved from http://www.ietf.org/rfc/rfc5965.txt (Updated by RFC 6650)

Sharma, V., So, N., Ali, Z., Bolt, G., Claise, B., HIggins, S., & Akhter, A. (2012, 23 October). Network-Centric Performance Management: "So Near and Yet So Far?". In *NANOG 56.* NANOG. Retrieved from https://www.nanog.org/meetings/abstract?id=2010

Simon, H. A. (1996). *The Sciences of the Artificial.* Cambridge, Mass.: MIT Press.

Smith, P., & Meynell, K. (2014). *Agenda: BGP Routing (IPv4 & IPv6).* Retrieved from https://nsrc.org/workshops/2014/pacnog16-ws/wiki/Track1Agenda

Snow, C. (1964). *The Two Cultures and A Second Look.* Cambridge University Press.

Solum, L. B. (2008). *Models of Internet Governance* (Tech. Rep. No. Law & Economics Research Paper No. LE08-027). University of Illinois Law. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1136825

Sowell, J. H. (2013, 21–23 September). Framing the Value of Internet Exchange Participation. In *Proceedings of the 41st Research Conference on Communication, Information and Internet Policy.* Arlington, VA: Telecommunications Policy Research Consortium.

Spamhaus. (2011). *Case Answer: e360Insight vs. The Spamhaus Project.* Retrieved from http://www.spamhaus.org/organization/statement/003/case-answer-e360insight-vs.-the-spamhaus-project

Spamhaus. (2014a). *BGPf FAQ.* Author. Retrieved from http://www.spamhaus.org/faq/section/BGPf%20FAQ

Spamhaus. (2014b). *The Domain Block List.* Retrieved from http://www.spamhaus.org/dbl/

Spamhaus. (2014c). *Exploits Block List.* Retrieved from http://www.spamhaus.org/xbl/

Spamhaus. (2014d). *The Policy Block List.* Retrieved from http://www.spamhaus.org/pbl/

Spamhaus. (2014e). *Register of Known Spam Operators.* Retrieved from http://www.spamhaus.org/rokso/ (Retrieved from http://www.spamhaus.org/rokso/.)

Spamhaus. (2014f). *SBL Delisting Procedure.* Retrieved from http://www.spamhaus.org/sbl/delistingprocedure/

Spamhaus. (2014g). *The Spamhaus Block List.* Retrieved from http://www.spamhaus.org/sbl/

Spamhaus. (2014h). *Spamhaus Botnet Controller List.* Retrieved from http://www.spamhaus.org/bcl/

Spamhaus. (2014i). *The Spamhaus Don't Route Or Peer Lists.* Retrieved from http://www.spamhaus.org/drop/

Spamhaus. (2014j). *zen.spamhaus.org.* Retrieved from http://www.spamhaus.org/zen/

Spamhaus. (2015a). *About Spamhaus.* Retrieved from http://www.spamhaus.org/organization/ (See discussion of LEA engagement in second half of page.)

Spamhaus. (2015b). *Glossary.* Retrieved from http://www.spamhaus.org/faq/section/Glossary

Spamhaus. (2015c). *What do the 127.*.*.* Return Codes mean?* Retrieved from http://www.spamhaus.org/faq/section/Spamhaus%20DBL#291

Stanojevic, R., Castro, I., & Gorinsky, S. (2011). CIPT: Using Tuangou to Reduce IP Transit Costs. In *Proceedings of the seventh Conference on emerging networking EXperiments and technologies* (pp. 17:1—17:12). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/2079296.2079313 doi: 10.1145/2079296.2079313

Steffan, S. (2012, 18 April). *What is Consensus?* Retrieved from https://ripe64.ripe.net/programme/meeting-plan/address-policy-wg/

Stigler, G. J. (1971). The Theory of Economic Regulation. *The Bell Journal of Economics and Management Science, 2*(1), 3–21. Retrieved from http://www.jstor.org/stable/3003160

stopbadware. (2015). *stopbadware.* Retrieved from https://www.stopbadware.org/

Stophaus. (2015). *Stophaus Project.* Retrieved from http://stophaus.com/

Tandon, N., Nair, S., & Lee, Y. (2010). *Proposal for Public Policy Advisory Committee (PPAC) at APNIC: Draft Version 2.* Retrieved from http://www.apnic.net/__data/assets/text_file/0018/32139/PPAC-at-APNIC.txt

UKNOF. (2014). *Respect at UKNOF.* Retrieved from https://wiki.uknof.org.uk/Respect_at_UKNOF

Underwood, T. (2005, 24 December). Internet-Wide Catastrophe—Last Year. *Renesys Blog.* Retrieved from http://www.renesys.com/2005/12/internetwide-nearcatastrophela/

Underwood, T. (2006, 23 January). Con-Ed Steals the 'Net. *Renesys Blog.* Retrieved from http://www.renesys.com/2006/01/coned-steals-the-net/

Villamizar, C., Chandra, R., & Govindan, R. (1998, November). *BGP Route Flap Damping* (No. 2439). RFC 2439 (Proposed Standard). IETF. Retrieved from http://www.ietf.org/rfc/rfc2439.txt

Vixie, P. (2014, 15 July). *Testimony of Paul Vixie, Chairman & CEO Farsight Security, Inc. before the Subcommite on Crime and Terorism United States Senate Commite on the Judiciary Hearing on Taking Down Botnets: Public and Private Eforts to Disrupt and Dismantle Cybercriminal Networks.* (Retrieved from http://www.judiciary.senate.gov/imo/media/doc/07-15-14VixieTestimony.pdf.)

Wang, F., Mao, Z. M., Wang, J., Gao, L., & Bush, R. (2006). A Measurement Study on the Impact of Routing Events on End-to-end Internet Path Performance. In *Proceedings of the 2006 conference on applications, technologies, architectures, and protocols for computer communications* (pp. 375–386). New York,

NY, USA: ACM. Retrieved 2014-02-04, from http://doi.acm.org/10.1145/1159913.1159956 doi: 10.1145/1159913.1159956

Weller, D., & Woodcock, B. (2013, 31 January). *Internet Traffic Exchange: Market Developments and Policy Challenges* (Tech. Rep.). Organisation for Economic Co-operation and Development, Directorate for Science, Technology and Industry, Committee for Information, Computer and Communications Policy.

Wilson, P., Plzak, R. A., Echeberria, R., & Pawlik, A. (2002, 10 October). *RIR Blueprint for Evolution and Reform of Internet Address Management.* NRO. Retrieved from http://www.nro.net/wp-content/uploads/nrr-blueprint-20021010.pdf

Wilson, P., Plzak, R. A., & Pawlik, A. (2002a, 20 June). *Regional Internet Registries' Submission to the Committee on ICANN Evolution and Reform.* NRO. Retrieved from http://www.nro.net/wp-content/uploads/rir-statement-20020621.pdf

Wilson, P., Plzak, R. A., & Pawlik, A. (2002b, 20 May). *Regional Internet Registry Joint Statement on ICANN Evolution and Reform.* NRO. Retrieved from http://www.nro.net/wp-content/uploads/rir-statement-20020508.pdf

Winner, L. (1980). Do Artifacts Have Politics? *Daedalus, 109*(1), 121–136.

Woodcock, B., & Adhikari, V. (2011, May). *Survey of Characteristics of Internet Carrier Interconnection Agreements* (Tech. Rep.). Packet Clearing House. (Retrieved from https://www.pch.net/resources/papers/peering-survey/PCH-Peering-Survey-2011.pdf.)

World Resources Institute. (1987). *World Resources.* New York: Basic Books.

Yang, L., Dantu, R., Anderson, T., & Gopal, R. (2004, April). *Forwarding and Control Element Separation (ForCES) Framework* (No. 3746). RFC 3746 (Informational). IETF. Retrieved from http://www.ietf.org/rfc/rfc3746.txt

Young, O. R. (1996). Rights, Rules, and Resources in International Scoiety. In S. Hanna, C. Folke, & K.-G. Mäler (Eds.), *Rights to Nature: Ecological, Cultural, and Political Principles of Institutions for the Environment.* Washington, DC: Island Press. (Locations refer to Kindle edition of book.)

Zittrain, J. (2005). *Jurisdiction.* Foundation Press.