

25. Custom Integrated Circuits

Academic and Research Staff

Prof. J. Allen, Prof. L.A. Glasser, Prof. P. Penfield, Jr., Prof. R.L. Rivest, Prof. G.J. Sussman, Prof. J.L. Wyatt, Jr., Dr. H. Shrobe, Jr.

Graduate Students

R. Armstrong, C. Bamji, C. Hauck, S. McCormick, L. Seiler

25.1 Conversion of Algorithms to Custom Integrated Circuits

*U.S. Air Force - Office of Scientific Research (Contracts F49620-81-C-0054 and F49620-84-C-0004)
Jonathan Allen, Paul Penfield, Jr., Ronald Rivest, Gerald Sussman, Howard Shrobe*

We have viewed the design of custom integrated circuits as an overall process of transformations starting at a high-level functional description and ending at a detailed masked specification for a particular process technology. At the highest level, a functional specification for the circuit must be generated, which should be specified in a form that permits the depiction of all of the possible parallelism present in the task. A great many high-level hardware design languages have been specified, but so far none have been utilized for integrated circuit design that provide an appropriate semantic foundation while permitting compilation to a fully parallel graph. In recent years, however, a great deal of research has taken place in the data flow computation area leading to the specification of languages that permit such a compilation to a graphical structure where all of the latent parallelism is revealed. It is important to realize that these results can be exploited for many purposes without the need to make any commitment to utilization of data flow architectures at the implementation level. We are currently building on this experience in order to derive a well founded formalism for functional specifications that is amenable to such optimal compilation techniques. Such a compilation will not only reveal the parallelism between diverse unrelated tasks, but can also unwind iterative loops in a way that reveals parallelism if in fact there is no sequential dependency between these iterations. Once the fully parallel graph structure is revealed, our techniques then focus on a phase of architectural exploration in order to allow the designer to select and utilize the degree of parallelism appropriate to the performance levels required in the application environment. At this stage, a large number of different techniques can be utilized. Techniques for transforming signal flow graphs into a variety of highly parallel architectures (including systolic forms) have been derived, as well as new techniques for retiming that provide the appropriate foundation for the establishment of a timing discipline in the overall system. It is clear that there is a substantial collection of these architectural exploration techniques that have been devised by a variety of researchers interested in various aspects of architectural exploration. In this study, our purpose is to unify these treatments into an overall theory of the interrelatedness of architectural forms and transformations between these

representations. We are also complementing this formalism with new results on area time trade-offs that go beyond results that have been achieved using asymptotic complexity theory. Instead, numerical estimates for constants in these asymptotic bounds are being derived, so that reliable engineering estimates of these performance variables can be made at the level of architectural exploration. We see this work as fundamental, and as comprising the core of an interactive program that permits the designer to systematically explore design alternatives. It is interesting that previously, different architectures for a particular task were often seen as unrelated. These new results, however, indicate that all of these architectural forms are systematically related through transformations, the form of which is just now being derived. For this reason, there is a new opportunity for the derivation of a basic theory of computer architecture having a formal and systematic basis, rather than the current heuristic basis for much of computer architecture.

Once a particular architecture is established, it remains to generate the detailed layout for the constituent cells. During the previous year, we had considerable success in devising a new PLA generator based on the specification of a minimal exemplar of a PLA circuit style, combined with a detailed logical specification for that PLA. We are now going on to generalize this approach by designing a new system that can generate a wide variety of regular structures. In each case, the user need only specify the individual constituent cells and their local connectivity, together with a high-level "concept file" that characterizes the nature of the overall module to be produced. It is interesting to note that if the only constraints provided by the designer at this point are those due to local connectivity, then a substantial (certainly not unique) variety of modules can result. It is the role of the concept file to remove this ambiguity and to place a form on the overall structure that suits the intended task while maintaining the local connectivity constraints. This approach is exceedingly general, but can be used for a wide variety of structures including register arrays, arithmetic logic units, program logic arrays, multiplexors, memory, and even specialized units such as array multipliers. Another facet of this work is the utilization of topological abstractions for the constituent cells, so that a wide variety of design rules, and hence processes, can be utilized. Once the overall module is created by use of the local connectivity constraints and the concept file, then a compaction phase is entered which introduces the technology constraints associated with a particular set of design rules. The result is a single mechanism that provides for a wide class of optimal circuit generators. Associated with each of these generators is a class of performance estimators that characterize the area consumed by the module, the time delay associated with it, and its power dissipation. In addition to these specialized generators, a general purpose processor architecture will also be provided, since we believe that most realistic systems cannot be realized in their entirety by utilization of these specialized function generators. It is, of course, the task of the architectural exploration phase to specify the overall design in terms of the specialized generators as well as the general purpose architecture in a form that satisfies the performance constraints dictated by the user. Two additional activities associated with the function generator phase are also worthy of note. We are designing a new CMOS multiplier, where the target speed for a 16x16 multiply is 50 nanoseconds. This is a regular design being carried out in conjunction with Professor Hoefflinger of Purdue

University. We are also completing the design of a special purpose function generator for a floating point unit, since many of our applications are in the area of digital signal processing which increasingly demands floating point capability on chip.

During this year, we finished work on a high accuracy circuit extraction program. Together with a comprehensive circuit simulator, this tool can be used to carefully evaluate the designs created by the function generators, and hence to derive procedures for the calculation of area, time, and power estimates previously mentioned. For digital signal processing applications, where not only high parallelism but aggressive circuit performance is required, this ability to accurately characterize automatically generated designs is essential. If, as a result of the performance estimates applied to the individual modules of the overall system, the user decides that the overall design will be acceptable, then we may go ahead to a placement and routing phase which uses the PI program which has been under development for several years within our project. This program creates a layout from the constituent modules, where each module is assumed to be rectangular in shape and orthogonally related to the other modules, with interconnect on all four sides. Routing strategies are already highly developed, and several placement packing strategies are under investigation.

From the above discussion, it is clear that the elements of an overall design system are being assembled in a fundamental way. We believe that the utilization of functional languages that characterize the deep semantics of the task and which are suitable for compilation to fully parallel representations provides a fundamental and new initial phase to integrated circuit design systems. We also recognize that a great many efforts involving architectural transformations have taken place within limited spheres of concern, but that these may now be unified into a more general framework, and this is the direction of much of our research. The use of local constraints and the "concept file" idea has permitted us to generate a wide variety of specialized modules that are highly efficient, and hence supplies us with one framework that is easily specialized to the many requisite modules of an overall design. Finally, the placement and routing strategies used employ well-tuned heuristics that can lead to a densely packed chip with little open area.

Lastly, we comment that construction of a hardware design rule checker has been completed, and that we have been able to demonstrate an improvement of two orders of magnitude in speed over conventional software design rule checkers by utilization of a single circuit board of specialized hardware, which is easily added to a standard general purpose computer. In this way, a significant amount of the computational burden in IC design is reduced to a level that is compatible with highly efficient interactive design.

25.2 A Circuit Theory for Digital VLSI Systems

National Science Foundation (Grant ECS81-18160)

Lance A. Glasser, Paul Penfield, Jr., John L. Wyatt, Jr., Isaac Bain, Charles Zukowski

VLSI circuit theory is an extension of classical non-linear circuit theory to networks that are so large as to make complete analysis infeasible. The goal is to allow precise partial circuit analysis, while exploiting many of the typical properties of VLSI circuits to achieve speed. These properties include the use of regular structures, simple design methodologies, and restoring digital signals.

The first area of VLSI circuit theory being pursued is design methodology verification. To deal with complexity, large circuits are often constructed with simple rules that ensure proper operation in some respect. For example, the proper connection of ratioed logic will always eventually produce valid logic signals at the outputs given valid and stable inputs. Therefore, verification that the rules of construction are not violated constitutes a useful partial analysis of a circuit.

Previous work in the area of methodology verification has involved producing programs that check for a particular set of rules. We are currently testing a language that can express very general methodologies. This is being used as an input to a syntax checking program. Circuits are specified for the syntax checker using an experimental circuit design language for expressing a designer's intent.

The second area of VLSI circuit theory being pursued is waveform bounding. Instead of finding the exact response of a network as in classical circuit analysis, the goal of waveform bounding is to merely narrow down the set of possible solutions to the extent necessary. The analysis is still precise, but the amount of information obtained about a circuit is sacrificed for greater speed. Bounding is feasible for digital circuits because signal uncertainty is restored.

For many purposes in the design and analysis of VLSI circuits, exact circuit simulation is impractical. Solving for the exact response of an accurate circuit model is very time consuming for large circuits, and an exact solution is not always essential, e.g., for a first pass design. To combat this problem, new simulators have been developed that calculate the response of much simpler circuit models. Such an approach decreases computation at the expense of accuracy.

In the waveform bounding approach, instead of approximating a more complex model by a simpler one, a simpler model is derived that "bounds" the behavior of the more complex one. Given upper and lower bounds on the response of the complex model, one can produce an approximate response by averaging the bounds, as well as a maximum error. In addition to measuring the errors made in simplifications, bounds can naturally incorporate uncertainties in a simulation, such as device parameter variations and imperfectly characterized input waveforms.

VLSI circuits are often designed to meet various external constraints such as a maximum delay. In this case approximating simulators are never reassuring design verifiers while bounding simulators are adequate in many cases. Given a delay specification, a bounding simulator can produce one of three conclusions: the circuit definitely meets the specification, it definitely does not meet the specification, or further calculation is needed to determine if the specification is met. Assuming that a

bounding simulator can be adjusted to trade speed for accuracy, the fact that maximum errors are calculated allows the computation speed to be reduced only when more accuracy becomes necessary.

A large amount of work has been done on the generation of rigorous bounds on common VLSI subcircuits: RC interconnect models, restoring logic gates, and pass transistor circuits. Circuit theorems have been proved about monotonic relationships between the basic circuit elements and the responses of these sub-circuits. The error produced by various bounds based on these theorems is currently being investigated. The theory of combining bounds on subcircuits into bounds on entire networks is also being pursued. Ultimately, this work could contribute to a new generation of VLSI circuit simulation programs.

Publications

Wyatt, J.L., Jr., C. Zukowski, L.A. Glasser, P. Bassett, and P. Penfield, Jr., "The Waveform Bounding Approach to Timing Analysis of Digital MOS IC's." Proc. IEEE International Conf. on Computer Design/VLSI in Computers, 1983, pp. 392-395.

Zukowski, C., "Circuit Simulation with Iterating Bounds," VLSI Memo, pp. 84-159.

25.3 Very Large Scale Integrated Circuit Research

U.S. Air Force - Office of Scientific Research (Contracts F49620-81-C-0054 and F49620-84-C-0004)

Lance A. Glasser, Mark Matson

Our work in this area focuses on two issues: modelling and optimization algorithms. We are developing precise models to describe the behavior of MOS logic gates. These models deal not with voltages and currents, but rather with delays and loading conditions. For example, our models predict a gate's output waveform shape as a function of input waveform shape and output capacitance. Details such as the logic gate's voltages and currents are implicitly accounted for by the model. By working at this higher level of abstraction, we have considerably improved computational speed while maintaining accuracy.

We have carefully analyzed the characteristics of our design problem, and found it well suited to a divide and conquer strategy where we break the problem into several subproblems which are then solved almost independently. We are applying optimization algorithms, in conjunction with sophisticated data structures, which use this strategy, resulting in improved computational efficiency.

At present we have implemented the modelling and optimization software, and have tested them on simple circuits. The tests were successful; our program demonstrated that it could compute transistor sizes that met a speed specification while minimizing power. Just as importantly, computer execution speed was fast (about 2 cpu seconds per design parameter).

We are now refining these implementations for more accuracy and higher computational speed, and

plan to test them on more complex circuits, such as microprocessors' arithmetic-logic units.

Publications

Glasser, L.A. and J. Hoyte, "Delay & Power Optimization in VLSI Circuits," to be presented at the 21st Design Automation Conference, June 25-27, 1984.

25.4 Waveform Bounding for Fast Timing Analysis of MOS VLSI Circuits

National Science Foundation (Grant ECS83-10941)

John L. Wyatt, Jr., Paul Penfield, Jr., Lance A. Glasser, Charles Zukowski, Han-Ngee Tan, Pearl Yew, Peter O'Brien

This project continues, enlarges, and applies one of the research avenues first uncovered in the work described in Section 25.2 of this report. That section outlined the basic rationale for waveform bounding: here we discuss some of the methodologies in more detail.

To obtain rigorous upper and lower bounds on the voltage response at all the nodes of a dynamic circuit, one has to study the differential equations describing it. Digital MOS circuits are highly nonlinear, but they have special features that make them analytically somewhat tractable. These include monotonicity of key element constitutive relations, the absence of inductors in most circuit models, a limited number of d.c. paths to supply voltage and ground, the restoring character of logic circuits, and the presence of a capacitance from every node to ground in most circuit models.

This research began with Rubenstein, Penfield and Horowitz' derivation of closed-form response bounds for linear network models of MOS interconnect.¹ Subsequent work by Wyatt extended the method to unbranched line models incorporating realistic monotone nonlinearities.² Qingjian Yu and Omar Wing of Columbia University then discovered a way to include branched nonlinear lines.³ Dr. Yu will be in residence with our group during the summer of 1984 in a temporary postdoctoral position.

A current research goal is to tighten the bounds on delay through linear interconnect models given in Ref. 1. One approach is based on a spatial monotonicity property of the time derivative of voltage first noted in Ref. 4 and proved in Ref. 5. Another is based on second-order bounds, i.e., bounds that can be computed in a length of time which grows quadratically with the number of elements in the circuit model. To exploit these second order bounds we have recast the problem in the form of a linear minimum- and maximum-time optimal control problem with linear state constraints. Han Tan hopes to completely solve the optimal control problem this summer (1984) with some assistance from Professor Violet Haas, a visiting professor from Purdue University who specializes in optimal control.

Interconnect is at worst only mildly nonlinear: logic gates are highly nonlinear and require a

different approach. We have found a way to use the first variational equation to study the sensitivity of delay through MOS logic circuits to the shape of the input waveform.^{6,7} The main result is that for inverters the time of maximum sensitivity is the moment when the pulldown makes the transition between the ohmic and saturation regions.

The most fruitful approach to date for bounding the response of MOS logic circuits is to exploit the monotone character of the MOSFET v - i curves. In this way one can model a complex multiple-input gate by a much simpler network and be confident of the sign of the error produced by the model simplification.⁸

Perhaps the most exciting development in this project is an effort to marry waveform bounding techniques to methods of circuit simulation based on waveform relaxation.⁹ At its simplest, the idea is to use waveform bounding to provide a good global initial guess for an exact simulator like RELAX. A more advanced version of the idea is to perform iterations on the bounds themselves.¹⁰ In the interest of speed the iterations need not be carried to convergence: computation can be terminated when the bounds are tight enough. Charles Zukowski is writing his doctoral thesis on this approach to simulation, and John Wyatt will spend August of 1984 at IBM Yorktown Heights collaborating with the developers of RELAX on its implementation.

References

1. J. Rubenstein, P. Penfield, Jr., and M.A. Horowitz, "Signal Delay in RC Tree Networks," *IEEE Trans. Comput. Aided Design*, CAD-2, 3, 202-211 (1983).
2. J.L. Wyatt, Jr., "Monotone Sensitivity of Nonlinear, Nonuniform RC Transmission Lines, with Application to Timing Analysis of Digital MOS Integrated Circuits," *IEEE Trans. Circuits, Syst.*, accepted for publication.
3. Q. Yu and O. Wing, "Waveform Bounds of Nonlinear RC Trees," Proceedings of the IEEE International Symposium on Circuits and Systems, Montreal, Canada, May 1984, pp. 356-359.
4. M.A. Horowitz, "Timing Models for MOS Circuits," Ph.D. Thesis, Department of Electrical Engineering, Stanford University, December 1983.
5. P. O'Brien, "Analytical Results for Waveform Bounding," S.B. Thesis, Department of Electrical Engineering and Computer Science, M.I.T., May 1984.
6. J.L. Wyatt, Jr., "The Sensitivity of Inverter Delay to Details of the Input Waveform: A Variational Approach," VLSI Memo 82-126, November 1982.
7. P. Yew and J.L. Wyatt, Jr., "A Variational Approach to Delay in MOS Logic Circuits," Proceedings of the IEEE International Symposium on Circuits and Systems, Montreal, Canada, May 1984, pp. 852-855.
8. J.L. Wyatt, Jr., C. Zukowski, L.A. Glasser, P. Bassett, and P. Penfield, Jr., "The Waveform Bounding Approach to Timing Analysis of Digital MOS IC's," Proceedings of the IEEE International Conference on Computer Design: VLSI in Computers, Port Chester, New York, November 1983, pp. 392-395.
9. E. Lelarsmee, A.E. Ruehli, and A.L. Sangiovanni-Vincentelli, "The Waveform Relaxation Method for Time-Domain Analysis of Large Scale Integrated Circuits," *IEEE Trans. Comput. Aided Design*, CAD-1, 3, 131-145 (1982).
10. C.A. Zukowski, "Circuit Simulation with Iterating Bounds," VLSI Memo 84-159, January 1984.

