15.0 Custom Integrated Circuits

Academic and Research Staff

Prof. J. Allen, Prof. L. Glasser, Prof. B.R. Musicus, Prof. P. Penfield, Jr., Prof. J.L. Wyatt, Jr.

Graduate Students

R.C. Armstrong, D.G. Baltus, C. Bamji, L.M. Brocco, C. Hauck, H. Jeong, S.L. Lin, S.P. McCormick, K. Nabors, A. Malamy, P. O'Brien, C. Selridge, D. Standley, B. Thompson, H. Wright

15.1 Custom Integrated Circuits

U.S. Air Force - Office of Scientific Research (Grant AFOSR-86-0164)

Robert C. Armstrong, Donald G. Baltus, Cyrus Bamji, Lynne M. Brocco, Charles Hauck, Shing Lih Lin, Steven P. McCormick

The research goal of this project is to devise CAD techniques for the performance-directed synthesis of digital VLSI circuits, focused on digital signal processing applications. The main goal is to develop expert system techniques that are designed to yield optimal performance in these signal processing systems through a fundamental approach to the design task. An overall design involves the specification of multiple constraint domains corresponding to various levels of representation. These facets include the function, architecture, logic, circuit, layout, and device levels. They must be coherently related so that each is a consistent and correct projection of the complete design onto a single facet or type of representation. This viewpoint has led to three major concerns:

1. Appropriate representations and constraint mechanisms for each design level must be discovered and built. This task is particularly difficult for those levels that are far from the physical reality of the circuit itself.

2. Expert systems must be constructed to achieve optimal design. This approach depends on and affects the representational issues mentioned above.

3. Performance optimization must apply both within and across representational domains. Hence, it must provide both local and global optimization.

The intent in performance-directed synthesis is to avoid long synthesis phases followed by comprehensive analysis. Instead, this research aims to drastically shorten the synthesis-analysis loop, providing for both rapid generation of designs at the various levels, as well as their fast and accurate assessment by new, optimized analysis tools. The design techniques developed under this grant can be viewed as components of an overall expert system for high-performance design. Effective algorithms are combined within heuristic frameworks to capture, and even exceed, the capability of experienced and talented designers.

A major goal of high-performance design is to squeeze the corresponding layout into the smallest possible area. Under this grant, Lin and Allen[1] developed a constraint graph-based compactor that minimizes total area on each of the layout levels, subject to a designer-specified weighting factor. In addition, Reichelt[2] developed the necessary data representations for hierarchical compaction. A parameter specified in this representation allows the designer to compact designs hierarchically with direct control over the amount of overlap between adjacent cells. As a result, hierarchical designs can be compacted with only abbutting cells, or with overlapping cells, to the extent permitted by the design rules. This new representation is now being exploited to build a high-performance hierarchical compactor that utilizes both symbolic and direct geometric input of layouts. In this way, the compacter can be used for technology tracking and direct realization of high-performance layouts from a symbolic design representation.

The next level above symbolic layout is the circuit level of representation. Baltus[3] has built a new expert system for the conversion of a circuit's net representation to an optimized layout. This strategy allows the user to deal selectively with varying aspect ratios, constraints due to input/output pins, constraints imposed by different circuit styles, and the difficulties introduced when widely varying device sizes (due to speed optimizations) are utilized. High-quality layouts have already been obtained in both NMOS and CMOS for many different circuit macros, and effective packing is obtained by introducing groups of transistors according to similar sizes rather than their structural affinity in the original design. This new level of representation has proven very useful, and serves as an appropriate interface to the symbolic compactor mentioned above to yield final layout.

The creation of an effective floor plan, or global placement of modules in a minimal area that allows optimal interconnect in terms of both length and time delay, has been achieved by S. Weiner.[4] This program dynamically chooses and switches between several floor planning strategies using problem-specific facts as a guide. This has been an exceedingly good area for expert systems, since no single strategy is uniformly optimal, and since it is possible to characterize the floor plan problem at a representational level that can be effectively exploited to choose the best strategy.

In order to complement the Regular Structure Generator[5] previously developed under this grant, L. Brocco[6] has developed a new program for the estimation of delay through circuit blocks in a VLSI design. This program achieves accuracy within 5% of SPICE for generalized inverter-type structures in CMOS. Using abstract, functionally based macro models that permit the introduction of nonlinearities (which are essential for high-accuracy modeling), transmission gate circuits have been analyzed effectively with errors well below 10% of the corresponding SPICE results. This is the first time that transmission gate circuits (which are characterized by two distinct time constants) have been adequately modelled for delay purposes in a unified manner with inverter-like circuits. Not only are the delays obtained accurately, but the computational time is approximately 0.001 of the magnitude of SPICE runs for circuits as complex as array multipliers. This program permits the rapid assessment and exploration of design alternatives, and, hence, makes performance-directed synthesis a viable design option for a wide class of high-performance applications.

In a major new development, formal grammars have been introduced to permit the formal unification of several different levels of design representation. For example, context-free grammars are used for circuit-level representations, and regular grammars

are used for layout. A context-free grammar has been demonstrated that readily generates and accepts all classical CMOS circuits. C. Bamji has developed a new technique to elaborate all legitimate designs, subject to a user-specified set of input constraints, which can be examined in a design exploration phase. The individual grammars for each level of design representation are unified by a coupling grammar which pulls together the separate grammars. It is important to note that these grammars can generate all legitimate representational couplings, and can be used with a parser to analyze conjoined representational descriptions of designs for correctness.

Formal grammars can also be utilized as the fundamental data objects for design editing, in order to maintain "alignment" between the various levels of representation. When a designer edits one level of representation (e. g., layout), changes must be automatically propagated to the corresponding effects at other levels of representation (e. g., circuit or logic). Since the formal grammars for these levels of representation can be coupled as previously mentioned, R. Amstrong has developed a new mechanism that provides a strong linguistic base for all design editing, regardless of the "design facet" being changed. The two projects cited that utilize formal grammars provide a new basis for design representation and manipulation which can be used as an extendable kernel, around which the various performance-directed algorithms can be attached.

All of these CAD programs are being developed in the context of high-performance 68020-based workstations with high-resolution color displays, a minimum of 8 megabytes of memory, local area network connections, a minimum of a 130 megabytes of disk storage, utilization of the UNIX operating system, and the provision of the x-windows display system. This computational environment, coupled with the performance-directed synthesis tools described here, leads to higher levels of capability in the correct, quick, and economical generation of high-performance VLSI designs.

References

[1] S. Lin and J. Allen, "MINPLEX -- A Compactor that Minimizes the Bounding Rectangle and the Individual Rectangles in a Layout," Proceedings of the 23rd Design Automation Conference, July 1986, p. 123.

[2] M. Reichelt and W. Wolf, "An Improved Cell Model for Hierarchical Constraint-Graph Compaction." Proceedings of the International Conference on Computer-Aided Design, November 1986, p. 482.

[3] D.G. Baltus, "Generating Efficient Layouts from Optimized MOS Circuit Schematics." M.S. Thesis, Dept. of Electr. Eng. and Comp. Sci., M.I.T., Cambridge, Mass., 1987.

[4] S.S. Weiner, "APRICOT: A Priority-Conscious Tool for VLSI Floorplanning," M.S. thesis, Dept. of Electr. Eng. and Comp. Sci., M.I.T., Cambridge, Mass., 1987.

[5] C.S. Bamji, C.E. Hauck, and J. Allen, "A Design-by-Example Regular Structure Generator," Proceedings of the 22nd Design Automation Conference, June 1985, p. 16.

[6] L.M. Brocco, "Macromodelling CMOS Circuits for Timing Simulation," M.S. Thesis, Dept. of Electr. Eng. and Comp. Sci., M.I.T., Cambridge, Mass., 1987, R.L.E. Tech. Report No. 529.

15.2 High Performance Circuit Design

Lance A. Glasser, A. Malamy, C. Selridge, B. Thompson

This year significant progress was made in the three complementary areas of theory, experiment, and computer-aided design (CAD).

Theoretical progress was made in the classical area of linear circuit analysis. Tight bounds on the highest natural frequency of oscillation of circuits composed of linear devices have been derived. A catalog analogy best describes the work. Given a linear model of each type of element in a catalog, we can predict exactly the highest frequency of oscillation of any network built of any number of these elements, taken in any multiplicity (one can order as many parts from the catalog as one wants), impedance scaled by any positive real number (if the catalog has a 10 Ohm resistor one can order 100 and 2 Ohm resistors), and connected together with ideal wire and ideal transformers. This work has application to device and circuit design.

In the experimental area, we have successfully demonstrated the transmission of power to integrated circuits without pins. We have used magnetic coupling and on-chip inductors and rectifiers to couple several hundred microwatts of power at several volts--enough power to drive low power CMOS circuits. This work has potential application to smart credit cards and biomedical implants.

In the CAD area, progress has been made on two programs. We have a working program to discover the maximum frequency of oscillation of complex linear circuits, as discussed above. This program is written in LISP. We have also continued our work on RELIC, a unique reliability simulator for integrated circuits. This program enables the simultaneous simulation of several different reliability mechanism for VLSI circuits.

15.3 Extracting Masks from Optical Images of VLSI Chips

Bruce R. Musicus, Hong Jeong, Rosalind Wright

One of the chief dificulties in studying image modeling and image understanding is that it is difficult to find useful models to aid in interpreting unconstrained images. In order to better understand the role of image modelling, we are therefore concentrating on the particular problem of "reverse-engineering" a VLSI chip given a micro-photograph of the chip. An enormous amount of information is available concerning the design of VLSI chips; they are deposited in layers of known composition and optical appearance, the images are formed from strips of material delineated by clear, though ragged, boundaries, the strips must form electrical circuits with known characteristics. Given all this *a priori* information, including knowledge of minimum line and feature widths as well as rules about the composition of layers making up the chip, our goal is to build an efficient analysis system for reconstructing the masks that were used to manufacture the chip.

Our initial work in this area focused on low level image processing issues, such as compensating for improper focusing, imbalanced lighting, and texture, while trying to accurately segment the image into line strips. It was found that "local" analysis meth-

ods, analyzing small windows to decide if they contained an edge or not, worked reasonably well and were relatively insensitive to lighting and texture fluctuations.

Our latest work is to address the back end of the analysis system. Given a clean line drawing representing a section of the VLSI chip, how do we piece together the various strips into the masks that formed the chip? At present we are developing various algorithms which deduce all possible legal interpretations of a given line drawing using only information about the edges. Using a database of rules of VLSI design, the programs start with edges and vertices, piece together "paths" marking the edges of strip in some mask, assign the paths to layers, and label the layers. The most difficult part is to correctly infer where mask strips cross over each other, and to properly interpret "accidental edges," where several strip edges coincide in the image.

Publications

Jeong, H., Ph.D. Diss., M.I.T., Cambridge, Mass., expected 1987.

Jeong, H. and B. Musicus, "Mask Extraction from Optical Images of VLSI Circuits," I.E.E.E. International Conference on Acoustics, Speech, and Signal Processing, Dallas, Texas, April 1987.

Wright, R., "Highly Distributed Image Understanding Algorithms Applied to Segmentation of VLSI Images," S.M. Thesis, M.I.T., Cambridge, Mass., 1986.

15.4 Algorithmic Fault Tolerance on Digital Signal Processing

Bruce R. Musicus, William Song

Information theory suggests a specific strategy for communicating information reliably across a noisy channel. Using a model of the noise process in the channel, a coder spreads the information in the signal across the channel bandwidth in such a way that a noise spike may destroy part of many bits, but not an entire bit. The decoder at the receiving end uses redundant information transmitted down the channel to reconstruct the entire message without error. Careful coding allows systems to achieve nearly 100% correct information transfer with only modest overhead for the coding.

Unfortunately, coding ideas do not always translate well into fault-tolerant computer architectures. Some computer modules, such as memory, buses, or networks, are easily protected against transient or permanent part failures by error-coding techniques. These applications, however, are restricted to modules which do not modify the data. Computational modules are not easily protected by error-coding techniques. Instead, conventional approaches for building fault-tolerant processors rely on duplicating, triplicating, or quadruplicating all computational resources, and voting on information transferred across module boundaries in order to discard bad results.

In this project, we are developing a new approach to fault tolerance, in which we can protect certain types of linear computation against processor failure. The basic strategy can be applied when the same computation is applied to many different data items, and when an error coding computation can be found which commutes with the

processing computation. The "coder" distributes data to various independent processors, and also computes various functions of the input data to form inputs for some redundant processors. The "decoder" verifies that the output of the redundant processors is consistent with the output of the other processors; if not, it then isolates which processor (s) have failed, and corrects them (if possible).

One application of this idea uses a bank of analog-to-digital converters operating in round-robin fashion to achieve an overall sampling rate somewhat above the Nyquist rate for the signal. A dither system and digital low-pass filter combine to reduce quantization errors in the front end. This same low-pass, however, can be used to detect and interpolate over temporary or permanent errors in any of the converters, without substantially increasing the total amount of computation. The system is able to trade off additional hardware for greater accuracy and higher levels of fault protection. As converters fail, all that happens is that the effective quantization error increases.

Another application is to the FFT processor system used in range and velocity doppler sonar processing. Here we use a set of processors to process multiple scans of sonar data from a phased-array antenna. Each processor does the same linear FFT processing, but on different sets of range cells. Adding extra processors working on linear combinations of the inputs to the other processors allows simple fault detection and correction. Regardless of the number of processors in the system, detecting K simultaneous failures requires only extra processors; detecting and correcting K simultaneous failures requires only 2K extra processors. When conventional truncation or rounding arithmetic is used, however, then the error checking can only be approximate. In this case, adding more processors improves the accuracy of the fault checking and correction. We are presently working with Draper Labs on the design of a sonar system incorporating these concepts.

15.5 Cellular Array for Image Processing

Bruce R. Musicus, G.N. Srinivassa Prasanna, Hong Jeong, Andrew Fraley, Mitchell Oslick, Edward Schembor, John Deroo, Kevin O'Connor

Low level image processing operations, such as contrast stretching, compensation for lighting variation, noise suppression, or edge enhancement, often rely on highly repetitive processing of the pixels in the image. In conventional image processing architectures, this characteristic is exploited by pipelining the image data through a computational pipeline which repeatedly executes the same instruction on all the data flowing through it. An alternative approach, which we are exploring, is to build a large number of small processors, and use these processors in parallel to execute the same instructions on different parts of the image. The Connection Machine is the best known commercial implementation of this architectural idea. Our goal is to explore much simpler and cheaper implementations, which may be carefully matched to the algorithmic domain in order to achieve high performance at low cost.

To explore hardware, software, and algorithmic issues involved in this approach, we are building a small 16 by 16 array of 256 single-bit processors, packaged on 2 VME boards with data memory, a horizontally microcoded sequencer, and a host interface. Combined with a frame grabber and a 68000 controller card, we will have a very high performance machine capable of extremely high speed computation for a particular

class of signal processing problems. The array is built from four AAP chips from OKI Semiconductor, and operates at a 6.5 MHz rate, performing 256 bit-operations on every clock tick. Both bit-serial and bit-parallel arithmetic are supported. Data memory is specially designed to supply overlapping frames of bit-serial or bit parallel data to the processor array. The machine is programmed with a micro-assembler with high-level control constructs and expression evaluation. Various algorithms for low level image processing and matrix arithmetic are under development.

15.6 Waveform Bounding for Fast Timing Analysis of Digital VLSI Circuits

U.S. Navy - Office of Naval Research (Contract N00014-80-C-0622)
National Science Foundation (Grant ECS-83-10941)

John L. Wyatt, Jr., Paul Penfield, Jr., Lance A. Glasser, Keith Nabors, David Standley, Peter O'Brien, Desmond Kirkpatrick

Our work over the last year has been concentrated on reworking the waveform bounding results of Penfield, et al., originally created for timing analysis of MOS chips, into a form useful for delay estimation in ECL circuits. A smaller effort has also been devoted to the circuit design and layout of a parallel analog MOS VLSI chip for early vision applications.

Peter O'Brien has continued to work on ECL delay estimation for his S.M. thesis, in cooperation with Digital Equipment Corporation. The final product we envision is a timing analyzer, written in C, for use on ECL standard cell designs. The goal remains: accurate estimation of signal propagation delay without resorting to computationally intensive numerical solution of the network equations. In apparent contrast to the earlier MOS-related work, the largest technical problem we've encountered is not estimating signal delay in interconnect, but rather accurately modeling the electrical behavior of the gates themselves. The ECL gates we've studied act like voltage sources with internal resistance when pulling a line high, but like almost ideal current sources when pulling it low. The latter case does not fit easily into the framework of the Penfield approach, which is tailored to voltage-source drives.

Our modeling progress to date consists of a reasonably simple macromodel for the gates themselves, and an approximation to the driving-point impedance of loaded, branched metal interconnect lines that lets us translate a current- source drive into an approximate resulting voltage waveform at the output of the driving gate. This combined procedure gives delay results that agree with detailed SPICE simulations to around 4%, with a savings of about three orders of magnitude in computer time. The only fundamental hurdle remaining is to find a simple but accurate way of approximating the voltage waveform at the input of any driven gate by a simplified waveform, such as a saturated ramp, that permits closed-form calculation of the gate macromodel's response.

Our other project, design of the analog depth-interpolation chip, has progressed to the point that we expect to send the first test chip to MOSIS in November. This design contains a 4x4 pixel array that is directly accessible from the pins, and test structures

with each of the subcircuits needed for the final system, e.g., A/D and D/A converters and differential amplifiers. We'll tell you how it worked in the next progress report.

Publications

Kirkpatrick, D., "Numerical Solution of the RC Mesh," S.B. Thesis, Dept. of Electr. Eng. and Comp. Sci., M.I.T., Cambridge, Mass., 1986.

O'Brien, P.R. and J.L. Wyatt, Jr., "Signal Delay in ECL Interconnect," in Proceedings of the 1986 International Symposium on Circuits and Systems, Santa Clara, Calif., May 1986, p. 755; (M.I.T. VLSI Memo No. 86-296).

Standley, D. and J.L. Wyatt, Jr., "Improved Signal Delay Bounds for RC Tree Networks," M.I.T. VLSI Memo No. 86-317, May 1986.

Wyatt, J.L., Jr., C.A. Zukowski and P. Penfield, Jr., "Step Response Bounds for Large-Scale Linear Systems Described by M-Matrices, with VLSI Application," SIAM Conference on Linear Algebra in Signals, Systems and Control, Boston, Mass., 1986; abstract appeared on p. A35 of Conference program.

Zippel, R.E., P. Penfield, Jr., L.A. Glasser, C.A. Leiserson, J.L. Wyatt, Jr., F. T. Leighton, and J. Allen, "Recent Results in VLSI CAD at M.I.T.," to appear in Proceedings of the Fall Joint Computer Conference, Dallas, Texas, November 1986; M.I.T. VLSI Memo No. 86-341.