

VLSI RESEARCH

SEMI-ANNUAL
Technical R&D STATUS REPORT
APRIL - OCTOBER 1983

PRINCIPAL INVESTIGATOR

R.W. BRODERSEN
(415)642-1779

FACULTY RESEARCHERS

R.W. Brodersen
N. CHEUNG
A. DESPAIN
D.A. HODGES
C.HU
D. MESSERSCHMITT
R.S. MULLER
A. NEUREUTHER
A.P. NEWTON
W.G. OLDHAM
J. OUSTERHOUT
D.A. PATTERSON
A. SANGIOVANNI-VINCEN TEDLI
C. SEOUIN

SPONSORED BY
DEFENSE ADVANCED RESEARCH PROJECTS AGENCY
ARPA ORDER NO. 4031

MONITORED BY NAVAL ELECTRONIC SYSTEMS COMMAND
UNDER CONTRACT NO. N00039-81-K-0251

The views and conclusions contained in this document are those of the authors and should not/be interpreted as representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the U.S. Government.

ELECTRONICS RESEARCH LABORATORY

COLLEGE OF ENGINEERING
UNIVERSITY OF CALIFORNIA, BERKELEY, CA 94720

AD-A134628

FILE COPY

VLSI RESEARCH

SEMI-ANNUAL
TECHNICAL R&D STATUS REPORT
APRIL - OCTOBER 1983

PRINCIPAL INVESTIGATOR

R.W. BRODERSEN
(415)642-1779

FACULTY RESEARCHERS

R.W. BRODERSEN
N. CHEUNG
A. DESPAIN
D.A. HODGES
C. HU
D. MESSERSCHMITT
R.S. MULLER
A. NEUREUTHER
A.P. NEWTON
W.G. OLDHAM
J. OUSTERHOUT
D.A. PATTERSON
A. SANGIOVANNI-VINCENTELLI
C. SEQUIN

TABLE OF CONTENTS

1. SECOND GENERATION
2. ARCHITECTURE SOFTWARE PROTOTYPING
3. MULTIPROCESSOR CIRCUIT SIMULATION
4. NOVEL, HIGH-PERFORMANCE ARCHITECTURES
5. SPEECH RECOGNITION
6. MACROCELLS FOR SIGNAL PROCESSING
7. RESEARCH IN CIRCUIT SIMULATION AND LAYOUT
8. NEW VLSI TOOLS DISTRIBUTION
9. BERKELEY ADVANCED CMOS
10. TECHNOLOGY

1: a

1. SECOND GENERATION RISCs (Patterson, Sequin)

We have submitted the RISC II processor and instruction cache for fabrication.

The 43,000 transistor RISC II chip uses a compact dual bus register cell, providing 60% more registers in a 25% smaller chip using the same design rules. However, the sharing of the bit lines for reading and writing required a extra pipe stage plus operand forwarding circuits. Since the chip was thoroughly checked with the Crystal timing analyzer, we expect this CPU to come close to our original performance goals [2] [4].

The 46,000 transistor RISC II instruction cache uses several new architectural ideas. For example, the addition of one bit per word enables us to turn off bad words thus doubling the yield of this chip, and an instruction address predictor effectively doubles the speed of the cache memory [5]. Our long term goal is to combine the cache on the same chip with the processor.

Through talks, panels, debates, and papers we have continued to explain the RISC concept [1] [3] [8], and the methodologies used for its design and implementation [6] [9].

- [1] D.A. Patterson: "Microprogramming," *Scientific American*, Vol. 248, No. 3, March, 1983, pp. 36-43.
- [2] J.K. Foderaro, K.S. Van Dyke, and D.A. Patterson: "Running RISCs," *VLSI Design*, Vol. III, No. 5, pp. 27-32, September/October, 1982.
- [3] D.A. Patterson and R.S. Piepho: "Assessing RISCs in High-Level Language Support," *IEEE Micro*, November, 1982.
- [4] D.A. Patterson and C.H. Séquin: "A VLSI RISC," *Computer* Vol. 15, No. 9, pp. 8-21, September 1982.
- [5] D.A. Patterson, P. Garrison, M. Hill, D. Lioupis, C. Nyberg, T. Sippel, and K. Van Dyke: "Architecture of a VLSI Instruction Cache for a RISC," Tenth Annual Symposium in Computer Architecture, June 14-16, 1983.
- [6] C.H. Sequin and D.A. Patterson: "Design and Implementation of RISC I", *Computer Science Report No. UCB/CSD 82/106*, Oct. 1982, 23 pages.
- [7] R.M. Fujimoto and C.H. Sequin: "The Impact of VLSI on Communications in Multiprocessor Networks", *Proc. COMPSAC 82*, pp. 231-238 Chicago, Nov. 10-12, 1982.
- [8] Y. Tamir and C.H. Sequin: "Strategies for Managing the Register File in RISC", accepted for publication in *IEEE Trans. on Computers*, (~ Oct. 1983), 35 pages.
- [9] C.H. Sequin: "Managing VLSI Complexity: an Outlook", *IEEE Proceedings* Vol 71, No 1, pp 149 - 166, Jan. 1983.

2. ARCHITECTURE FOR SOFTWARE PROTOTYPING (Patterson)

Our first step towards building a machine to run exploratory programming environments is to make a version of Smalltalk-80 that runs on the VAX under Berkeley UNIX. We have completed that first step, called Berkeley Smalltalk. [1] This system will serve as the 'guinea pig' for our architecture experiments and will also be the software kernel for our new machine. In addition, although there are three other implementations of Smalltalk-80 using various operating systems and languages on the VAX, Berkeley Smalltalk is the fastest. Our initial architecture investigations have found that several of the ideas from RISC work well with Smalltalk. For example, the register window studies for RISC I found that the best buffer size was 8 windows. We have completed a similar study for Smalltalk and found that the knee in the curve is also at 8 windows. Perhaps the nicest discovery is that the reduced instruction set seems compatible with Smalltalk. Even though an operator such as '+' can mean a wide variety of operations (e.g., floating point add, logical or, set union in addition to integer addition) that cannot be determined until runtime, we found that on 93 % of the time both operands are simple integers. Similarly, the full blown semantics of Smalltalk requires a table lookup for each procedure call to see which procedure is to be activated, but 95 % of the cases go to the same procedure as last time. Thus we plan to short circuit the table lookup to accelerate execution. Our current architecture viewpoint is to use the same RISC philosophy to design and build a machine to that runs fast for the common cases yet is easy to implement. It will be interesting to see if this approach will once again yield a machine with surprisingly high performance.

[1] D.M. Ungar and D.A. Patterson: "Berkeley Smalltalk: Who Knows Where the Time Goes," *Smalltalk-80: Bits of History, Words of Advice*. Addison-Wesley, Summer '83.

3. MULTIPROCESSOR CIRCUIT SIMULATION (Messerschmitt)

Code generators for LU decomposition in a SPICE circuit simulator running on a multiprocessor architecture have been completed. These have been run on a multiprocessor simulator, and their performance evaluated. Improved scheduling algorithms which take into account interprocessor communications latency have been developed and evaluated, and yield a 30 to 50 percent speedup. Further improvements are being pursued.

Processor interconnection topology and routing algorithms are being pursued. Automatic generation of topology from traffic statistics has been implemented using a clustering algorithm. Measures of interconnection hardware complexity and speed performance are being developed in order to evaluate alternate interconnection topologies.

4. NOVEL, HIGH-PERFORMANCE ARCHITECTURES (Despain, Patt and Baden)

We have determined that functional programming languages provide a natural framework for programming highly parallel computers [1]. Our goal was to adopt a programming style that allows the programmer to express parallelism without being overburdened with low-level hardware dependencies.

We developed a compiler for Backus' functional Language, FP [2] called Berkeley FP; it has been released on 4.2 BSD UNIX*. A recent study of the Berkeley FP code generator has shown that FP's underlying algebra can be used to make dramatic improvements in the quality of the code generated.

- [1] S.B. Baden, D.R. Patel, "Berkeley FP - Experiences with a Functional Programming Language", *COMPCON '83*, February 28 - March 3 1983, San Francisco, California.
- [2] S.B. Baden, "Berkeley FP User's Manual, Rev. 4.1", December 15, 1982, *available on 4.2 BSD*.

5. SPEECH RECOGNITION (Brodersen)

The goal of this project is to design an MOS-LSI speech recognition system that is capable of accurately recognizing a large vocabulary of isolated words in real time. This system is based on two custom-designed integrated circuits, memory, and a control microprocessor [1].

The algorithm used is similar to that used in other dynamic-time-warped, template based automatic speech recognition (ASR) systems. It maintains a dictionary of model words or templates to which all input words are compared. The template that is most similar to the word just spoken is recognized and some associated action is performed. If no words are similar enough to the input, no action is performed. This dictionary can either be filled by the user in a training phase prior to usage (speaker-dependent mode) or it can be used in a speaker-independent mode by having multiple templates of each word which span the various types of speakers.

The first of the custom chips circuit analyzes the input speech by the use of a filter bank, decides when a word was spoken, and passes on an internal representation of the word to a word comparator. The word comparator is the second custom chip and can process up to 1000 vocabulary words (500 seconds of speech) in real time. The microprocessor collects the outputs of the word comparators and uses them along with other syntactic or semantic information to make its recognition decision. The microprocessor then passes on the recognized word to the host system or performs some further application system task.

The system can also be configured to recognize connected speech. In this mode it uses the isolated-word recognizer to spot words inside the phrase and then it ties these words together using an algorithm performed by the control microprocessor.

A TTL version of the recognition system has been working for some time which is based on 3 circuit boards. Recently the chip to perform the word comparison has been completed which has basically replaced one of these boards [2]. A second version of this chip is now in design that will include some of the glue logic that is now being done in TTL.

The chip to perform the front end functions of spectral analysis and end point detection (for isolated speech) is now in layout using a macrocell approach to the design.

- [1] H. Murviet, M. Lowy and R.W. Brodersen, "A 1000 Word Recognition System Using MOS-LSI" *Proc. of VLSI in Signal Processing* Nov. 1982, pp. 90-95.
- [2] M. Lowy, H. Murviet, D. Mintz and R.W. Brodersen, "An Architecture for a Speech Recognition System", *ISSCC Technical Digest*, New York, Feb. 1983,

6. MACROCELLS FOR SIGNAL PROCESSING (Brodersen)

Design of a 23,000 transistor NMOS integrated circuit implementing a LPC vocoder function has been completed. Testing and verification of the device is currently under way. The device implements a lattice analyzer, pitch/voicing analyzer, and lattice synthesizer. The tenth-order adaptive lattice analysis filter uses an algorithm similar to that described by Fellman [1]. The pitch/voicing analyzer uses Gold's algorithm [2]. The synthesizer consists of an excitation generator based on a voiced/unvoiced speech model, and a lattice synthesis filter. Alternatively, the excitation may be taken from an external source, allowing use in baseband encoding schemes.

The device interfaces directly to A/D, D/A and host microcomputer devices, allowing a low-parts-count system implementing a full-duplex 2400 bits/sec. speech transmission system conforming to DARPA and DOD speech communications protocols.

A macrocell library for signal processing circuits has been developed to facilitate the design of a number of signal processing LSI circuits including the LPC vocoder described above [3]. This library supports the rapid design of semi-custom circuits with applications in speech processing, image processing and data communications. Processors, control sequencers and other large functional units may be configured from this library into multiprocessor circuits with high functional throughput.

Three circuit designs are in progress in addition to the LPC vocoder: a 16-channel filter-bank circuit; a word endpoint detector circuit; and a variable-bit-rate formant speech synthesizer.

- [1] Fellman, R.D., "An MOS-LSI Adaptive Linear Prediction Filter for Speech Processing", *UCB/ERL Memo. M82/82*, University of California, Berkeley Ca., Nov. 1982.
- [2] B. Gold, L. Rabiner, "Parallel Processing Techniques for estimating Pitch Periods of Speech in the Time Domain", *J. Acoust. Soc. of Amer.*, V. 46, No. 2 (part 2), 1969, pp. 442-448.
- [3] S.P. Pope, R.W. Brodersen, "Macrocell design for Concurrent Signal Processing", *Proc. 3rd Annual Conf. on VLSI, California Inst. of Tech.*, Pasadena Ca., March 1983.

7. RESEARCH IN CIRCUIT SIMULATION AND LAYOUT (R. Newton and A. Sangiovanni-Vincentelli)

We focused in this funding period on three topics: accurate electrical simulation with relaxation-based algorithms, PLA optimal design and channel routing. Channel routing is a new area for us and we had some interesting preliminary results reported in [6].

7.1. Circuit simulation

We concentrated on extensions and improvements of algorithms in the class known as Waveform Relaxation. In these methods, large circuits are decomposed into many loosely coupled small circuits (subcircuits), and then the transient response waveform for each subcircuit is calculated by "guessing" the behavior of the surrounding subcircuits. The calculated responses for each subcircuit are used to improve the "guesses", and the transient response waveforms are recalculated. The procedure is iterated until the waveforms converge. We implemented the waveform relaxation algorithm for the special case of MOS circuits. A test program, RELAX, calculated accurate transient responses for large subcircuits up to 60 times faster than SPICE.

RELAX2[1], a new program for the accurate simulation of MOS circuits, has been written in C. RELAX2 handles a broader class of MOS digital circuits than the RELAX program, and includes several new techniques to reduce computation time. In RELAX the user had to describe his circuit with predefined MOS subcircuits, such as NAND gates, NOR gates, or Flip-Flops, but RELAX2 allows the user to define and use his own subcircuits. Also, the RELAX program converged slowly when used to simulate circuits with logical feedback (e.g. finite-state machines). The RELAX2 program solves this problem by allowing the user to break the simulation up into several "windows" of time. This technique effectively breaks the logical feedback loop, and increases the speed of convergence of the waveform relaxation method. Another speed-up technique involves changing the accuracy of the calculation of the subcircuit waveforms with the iteration, so that for the first few iterations the waveforms are calculated quickly and approximately, but by the final iteration they are calculated accurately. We have been able to prove rigorously that the convergence properties of the waveform relaxation methods did not change if these speed up techniques are used [2]. One way to approximate the calculation of the subcircuit waveforms is to use a simple resistor and switch model for the MOS devices, and then change to the more complex Shichman-Hodges model as the waveforms approach convergence. Experimental results from this technique have not been as good as expected. A more successful approach was to change the amount of error allowed by the numerical integration algorithm used to solve for the subcircuit waveforms. Here, unlike changing the device models, it is possible to increase the accuracy of the calculation of the waveforms at each iteration by screwing down this error. The results from this approach were better, providing a factor-of-two speedup in many cases. In addition, the RELAX2 program provides information about the topology of the circuit, which will aid the investigation of the available parallelism in the waveform relaxation method.

7.2. Optimal topological design of PLAs

We continued our work on efficient algorithms for optimal topological design of PLAs by improving the algorithms for constrained multiple folding. In particular, we defined a graph theoretic interpretation of the multiple PLA folding problem[3,5] and then we defined a constrained optimization problem to achieve minimal silicon area occupation with constrained positions of electrical inputs and outputs[4,5]. We implemented the algorithms in a new version of

PLEASURE, a program for constrained/unconstrained simple/multiple folding[4,5].

- [1] J. White and A. L. Sangiovanni-Vincentelli, "RELAX2: A Modified Waveform Relaxation Approach to the Simulation of MOS Digital Circuits" *Proc. 1983 International Symposium on Circuits and Systems*, Newport Beach, California, May 1983
- [2] E. Lelarsmee and A. Sangiovanni-Vincentelli, "Some New Results on Waveform Relaxation Algorithms for the Simulation of Integrated Circuits," *Proc. of 1982 IEEE Int. Large Scale System Symposium*, Virginia Beach, Oct. 1982, pp. 371-376.
- [3] G. De Micheli and A. Sangiovanni-Vincentelli, "Multiple Folding of Programmable Logic Arrays" *Proc. 1983 Int. Symp. on Circ. and Syst.*, Newport Beach, May 1983.
- [4] G. De Micheli and A. Sangiovanni-Vincentelli, "PLEASURE: A Computer Program for Simple/Multiple Constrained/Unconstrained Folding of Programmable Logic Arrays", *Proc. 1983 Design Automation Conference*, Miami Beach, June 1983.
- [5] G. De Micheli and A. Sangiovanni-Vincentelli, "PLEASURE: A Computer Program for Simple/Multiple Constrained/Unconstrained Folding of Programmable Logic Arrays", ERL Memo, ERL 82/56, December 1982.

8. NEW VLSI TOOLS DISTRIBUTION

We have assembled a new release of several of our VLSI Tools. The new release includes about 25 programs. It contains updated versions of Caesar and Mextra and other old programs, as well as several previously-unreleased programs, such as Lyra, Crystal, Peg, and Tpack. The 1983 release was sent to eight beta test sites in January, and began general distribution on April 1.

8.1. Tpack: A System for Combining Graphics and Procedures (Ousterhout, Newton)

We have developed a style of module generation called *tile packing*. In this style, a module generator consists of two parts: a collection of graphical tiles that define building blocks, and a program that arranges the building blocks into modules such as PLAs, ROMs, and decoders. The tile packing approach has the advantage of separating the technology and design rule information (kept in tiles) from the arrangement information (kept in the program and in truth tables). This makes it easy to construct simple module generators, and ensures that more complex module generators are not obsoleted by changes in design rules. We have built a new PLA generator, Tpla, using the tile packing approach, and have retargeted it for several different NMOS design rules. Work is underway to retarget the same program to CMOS, and to build splitting and folding PLA generators.

8.2. Caddy - A New IC layout System

We have undertaken the development of a new IC layout system called Caddy. The system has three overall goals, based on problems experienced with our earlier systems. The first goal is to integrate design rule and circuit information into the layout editor in order to provide incremental design rule checking and circuit extraction. This additional expertise will permit interactive compaction and stretching of layouts. The second goal is to move away from fabrication details by eliminating the need for designers to specify implants and wells and contact details explicitly. These layers will be generated automatically by the system (the result is much like "sticks" except that it is fleshed out). The third goal is to provide interactive semi-automatic routing aids. In this respect, our goal is not to invent new algorithms and paradigms, but to find powerful ways of embedding existing techniques into an interactive design environment. Initial discussions were held in the Spring and Fall of 1982, during which the underlying data structures and algorithms ("corner stitching") were developed. Between January and April of 1983 the detailed design of the system was completed and implementation was begun. We expect to use a bare-bones system in the development of the SOAR chip in the Spring of 1983.

- [1] J.K. Ousterhout: "Crystal: A Timing Analyzer for nMOS VLSI Circuits," *Third Caltech Conference on VLSI*, Computer Science Press, 1983, pp. 57-70.
- [2] R.N. Mayo and J.K. Ousterhout: "Pictures with Parentheses: Combining Graphics and Procedures in a VLSI Layout Tool," *20th Design Automation Conference*, to appear.

9. BERKELEY ADVANCED CMOS (Neureuther)

Fabrication of the first round of devices using the Berkeley Advanced CMOS Process is virtually complete. A combined mask with typical active devices and Latch-Up structures is being used. Delays were encountered due to the remodeling of the IC Laboratory and problems in pattern generation of the masks. Only the contact etching and metalization remain before electrical characterization can be started.

9.1. Simulation Aids for Viewing Wafer Topography from Layout

A top down design of the topography simulator SEMPLE has been defined and dry lab tested using the Berkeley advanced CMOS. The dry lab testing consisted of walking through 49 processing steps following the evolution of the two dimensional cross section of the device and determining the processing information needed at each step for defining the profile. The dry lab device profile has been displayed on KIC and photographed as an example of the kind of display which can be generated. The approach for implementing SEMPLE consists of working from CIF layout files to produce profile description files. This choice gives maximum portability and allows SEMPLE and the local layout system to be implemented as subsets of each other. The programming in "C" is being initiated and will proceed from Manhattan, to arcs, and finally SAMPLE and SUPREM linked features.

10. TECHNOLOGY (Oldham, Brodersen, Hu and Cheung)

10.1. Soft Error Studies

Computer analysis of the effects of surface condition on the is surrounded by reflecting surface, charge collected may be assumed an absorbing surface model which assumed an absorbing surface. A fuller presentation of this work has been submitted for publication. We have also made preliminary considerations of the soft error problem in GaAs IC's. There are theoretical reasons to expect significant soft error problems in them.

10.2. Hot Carrier Currents in Scaled MOSFETs

We have published the finding that photons are generated in the high field region in the channel and these photons subsequently create minority carriers in the substrate [2]. We have carried out further theoretical and experimental studies and now believe that bremsstrahlung is the mechanism of photon generation. We have also photographed the light emission from operating MOSFETs and used this technique to study the uniformity of the electric field along the width of MOSFETs. This work will be reported in June [3]. We have published the study of the breakdown of MOSFET [4] and the characteristics of MOSFET near and in the breakdown regime [5]. A full paper on the later subject will appear soon [6]. So will a paper on the punchthrough of MOSFETs [7]. We have demonstrated a universal correlation between the gate and the substrate and minority leakage currents. These results were reported at the 1983 ISSCC [9]. Finally these correlations have been found to hold in 0.15 μm n-channel MOSFETs.

10.3. Thin Oxide Studies

Special system, software, and measurement techniques have been developed to characterize the generation and filling of traps in SiO_2 and at the Si/ SiO_2 interface. Degradation of 10 nm-gate MOSFETs due to Fowler-Nordheim tunneling was published [11]. This work has been extended to degradation due to substrate hot electrons and due to channel hot electron. Several papers are being prepared to report the findings. Modeling of the I-V characteristics of very thin oxides will be published in April in the Netherlands [12] and in May in San Francisco [13]. Recently, we have completed a literature review on the subject of time-dependent breakdown of oxides. We have postulated and are examining a physical model for this phenomenon.

10.4. Latchup Studies

Test structures have been developed to investigate the latchup problem. We have measured the distribution of potential at or near the surface around a point source of current. The purpose is to quantify the dependence of the triggering current and latchup susceptibility on layout, scaling, and the guard-ring design. The results are supplemented with simulations and simple theoretical analysis. A paper is being prepared. Experiments are underway to study the impact of field implant and the use of epitaxial substrates and a novel way of replacing the epitaxial substrate with a blank, high-energy deep implant.

- [1] K.W. Terrill, C. Hu, A.R. Neureuther, "Company Analysis of the Significance of Surface Boundary Conditions on the Collection of α -Induced Charge," *Solid State Electronics*, 26, pp. 15-18, January 1983.
- [2] S. Tam, F.C. Hsu, P. Ko, C. Hu, and R.S. Muller, "Hot-Electron Induced Excess Carriers in MOSFETs," *IEEE Electron Devices Letters*, pp. 376, December

1982.

- [3] S. Tam, P. Ko, and C. Hu, "Light Emission and Field Uniformity in MOSFETs," to be presented at Device Research Conf., June 1983.
- [4] F.C. Hsu, P. Ko, S. Tam, R.S. Muller, and C. Hu, "An Analytic Breakdown Model for Short Channel MOSFETs," IEEE Trans. Elect. Dev., pp. 1735-1740, November 1982.
- [5] F.C. Hsu, R.S. Muller, and C. Hu, "Characteristics of Short-Channel MOSFETs in the Breakdown Regime," Digest of 1982 IEDM, pp. 282-285, December 1982.
- [6] F.C. Hsu, R.S. Muller, and C. Hu, "A Simplified Model of MOSFETs in the Breakdown Regime," to appear in IEEE Trans. Elect. Dev., June 1983.
- [7] F.C. Hsu, R.S. Muller, and C. Hu, "A Model of MOSFET Punchthrough Voltage," accepted by IEEE Trans. Elect. Dev.
- [8] S. Tam, P. Ko, C. Hu, and R.S. Muller, "Correlation Between Substrate and Gate Currents in MOSFETs," IEEE Trans. Electr. Dev., pp. 1740-1744, November 1982.
- [9] C. Hu, S. Tam, F.C. Hsu, R.S. Muller, and P.Ko, "Correlating the Channel, Substrate, Gate, and Minority-Carrier Currents in MOSFETs," 1983 IEEE ISSCC, pp. 88-89, February 1983.
- [10] S. Tam, F.C. Hsu, C. Hu, R.S. Muller, and P. Ko, "Hot Electron Currents in Very Short Channel MOSFETs," submitted to IEEE Elec. Dev. Letters.
- [11] M.S. Liang, M.S. Yeow, Y.T. Chang, C. Hu, and R.W. Brodersen, "MOSFET Degradation Due to Stressing of Thin Gate Oxide," 1982 IEDM, pp. 50-53, December 1983.
- [12] C. Chang, R.W. Brodersen, and C. Hu, "Direct and F-N Tunneling in Thin-Gate Oxide MOSFET Structures," Inter'l Conference on Insulating Films on Semiconductors (INFOS 83), April 1983.
- [13] C. Chang, M.S. Liang, C. Hu, and R.W. Brodersen, "Charge Tunneling and Impact on Thin Oxide Device Reliability," Electro Chemical Society Meeting, San Francisco, May 1983.

10.5. As Implant Anomalies

Anomalous negative threshold shifts are observed following annealing of thin oxides implanted with arsenic. Thermally grown oxide (650 Å) is implanted with arsenic at 25 Kev ($R_p=150$ Å, $AR_p=50$ Å) with a dose of $10e^{13}$ cm⁻². After annealing a negative threshold shift of up to -5 volts is observed for n+ polysilicon gate structures. Other experiments suggest penetration of oxide by arsenic, despite predictions of LSS scattering and arsenic diffusion [1].

10.6. Low Temperature MOS Electronics - CODMOS

We present a new depletion mode n-channel device structure which functions over the temperature range 77° - 400° K. This charged-oxide depletion MOS device (CODMOS) uses cesium ion implantation of silicon dioxide of the MOS structure. Cesium ions have shown great stability in SiO₂ under temperature bias stressing. The presence of these positive ions in the gate oxide turns on a channel under zero gate voltage. Overcoming of the donor freeze out problem observed in conventional depletion devices, and improved subthreshold and substrate sensitivity behaviors are expected using this structure [2].

A Standard local oxidation n-channel process has been used for fabrication of test devices. Low energy (40 keV, $R_p=190 \text{ \AA}$, $\Delta R_p=50 \text{ \AA}$) Cs^+ ions are implanted after growing 850 \AA gate oxide in 1000 C dry O_2 with doses of 3.48×10^{12} and $6.92 \times 10^{12} \text{ cm}^{-2}$. Conventional depletion devices are also fabricated by (120 keV, 1.488×10^{12} and $2.8 \times 10^{12} \text{ cm}^{-2}$) arsenic implantation of SiO_2 .

It is observed that the activation of cesium implanted ions depends on the post implantation annealing cycles. Cs^+ implanted devices have been fabricated with a room temperature threshold voltage of -6.4 volts which shifts to -3.2 volts at 77° K. This is compared with 99 % positive shift for conventional depletion devices. It is suspected that freeze out of the interface states generated by Cs^+ ion implantation is the cause of threshold shifts in the CODMOS devices. Better substrate sensitivity of CODMOS devices is observed over conventional depletion devices.

Successful CODMOS device test results confirm the feasibility of this design technique in fabrication of high performance n-channel depletion mode devices especially when low temperature operation is desired.

- [1] R. Kazerounian and W.G. Oldham, "Threshold Shift from As-Implantation of SiO_2 ", Proceedings of the Electrochemical Society Meeting, May 1983, San Francisco.
- [2] R. Kazerounian and W.G. Oldham, "CODMOS A Depletion Device Using Fixed Oxide Charge", submitted to Device Research Conference, June 1983 Burlington.

10.7. Channeling Effect of Low Energy Boron Implant in (100) Silicon

The unintentional ion channeling in low energy ion implantation of boron into (100) silicon results in much deeper junctions than predicted by LSS theory even for wafers tilted 8 degrees from the normal. This partial channeling effect imposes a limitation for the achievable p-type shallow junction depth. Both B^+ and BF_2^+ implants show the effect. The characteristics of this partial channeling profile are under investigation. An empirical formula was found to describe the enhancement of the junction depth. Experiments also showed that the deep penetration tail could be suppressed by a silicon pre-implant to drive the silicon surface layer amorphous.

- [1] T. M. Liu and W. G. Oldham, "Channeling Effect of Low Energy Boron Implant in (100) Silicon", to be published in Electron Device Letters, March 1983.

10.8. Al/Si Contact Electromigration

An important issue in VLSI technology development is to achieve reliable shallow junctions where Al/Si contact spiking is a major concern. In NMOS technology, this concern can often be relieved by introducing phosphorous plug into n+ contact. However, in CMOS technology, the introduction of plug would increase the complexity of the already complicated process. A study of Al/n+ and Al/p+ contact electromigration would be necessary to find out the optimum CMOS device structure with shallowest possible junction.

A automatic measurement setup based on IBM Personal Computer is configured. The materials of test packages are also chosen to operate at temperature about 250C without failure.