VLSI TECHNOLOGY

V.L. Rideout

Rapporteurs: Mr. K. Heron
Mr. J.G. Givens

Abstracts:

1. Limits to Improvement of Silicon Integrated Circuits

The limitations to improvement of silicon integrated circuits will be surveyed, primarily through an examination of the physical and electrical limitations to reducing the dimensions of devices and structural elements. The number of components on a silicon integrated circuit chip increased exponentially over the fifteen-year period from 1960 to 1975, doubling every year. Over the past five years, however, the rate of improvement has diminished, and presently the component count doubles about every two years at best. The present rate of improvement in lithographic capability (halving every six years) can alone double the component count every three years at most. The factors that will constrain the integration level of the future silicon technology will be assessed. For the purpose of discussions, limits to improvement will be categorised as physical, technological, and complexity limits, with the primary emphasis placed on technological limits.

2. Trends in Silicon Processing

The advent of VLSI will require substantial progress in all aspects of silicon technology - processing, lithography, modelling, design tools, chip architecture and applications. Current trends in silicon integrated circuit fabrication will be surveyed, focussing on new developments and outstanding problems. Progress in both bipolar and MOSFET technologies will be considered. Silicon fabrication techniques will be described in terms of the repetitious application of operations that are additive (oxidation, doping, deposition), selective (lithography) and subtractive (etching). The objective of these operations is a reliable and predictable device structure. Device structures will be described in terms of isolation areas, devices, contacts (interconnection vias), wiring (interconnection lines) and passivation. Immediate problems in isolation size, device performance, contact resistance and wiring topography will be identified. Future needs for improved structures will be indicated. Promising new approaches such as lightly-doped drain FETs and silicide-on-polysilicon (polycide) wiring will be described. Throughout this discussion the importance of process modelling will be emphasised.

LIMITS TO IMPROVEMENT OF SILICON INTEGRATED CIRCUITS

V. Leo Rideout
IBM Corporate Headquarters
Armonk, N.Y. USA 10504
Presented at the Joint International Seminar on the
Teaching of Computing Science - VLAI, September 9, 1981.

ABSTRACT

The technological limitations imposed on very large scale integration (VLSI) will be examined in terms of the laws of physics, technological capability, design complexity, and economic constraints. The primary emphasis will be placed on technological capability (e.g., physical and electrical limitations to reducing the dimensions of devices and structural elements) and on design complexity (e.g., difficulties associated with the design, testing and application of VLSI). Probably the best single measure of integrated circuit progress is the number of components on a commercially manufactured chip. From 1960 to 1975, the number of components per chip increased exponentially, doubling every year (1). Over the past five years, however, this rate of improvement has diminished and now only doubles every two years at best (2,3). This talk will show that over the next ten years, the most optimistic rate of improvement attainable will be a doubling time of four or five years and that will depend primarily on progress in lithography.

Fundamental limits are determined by the laws of physics and include the velocity of light, entropy (irreversibility) (4), the uncertainty principle, and thermal energy (5). Such fundamental phenomena present barriers to switching speed and power dissipation that cannot be surpassed. For dimensional comparison, a 16 Kbit dynamic RAM utilizes an average fabricated feature size of about 6 micrometers (6). Using this dimension as a yardstick, physical limits establish ultimate feature sizes of about 0.01 to 0.02 micrometers, suggesting an ultimate density improvement of 1,000's of times over that attained by present-day IC technology.

Technological limits concern materials constants, fabrication techniques and electrical parameters. The constraints imposed by these considerations can often be circumvented by use of structural changes, new materials, lower operating temperatures (7), better cooling techniques, and other forms of device and circuit cleverness (6). Materials constants include electrical and thermal conductivity (8), mobility, dielectric constants, saturation velocity (9), and dopant solubility. Limits associated with fabrication techniques

involve doping fluctuations (10), processing radiation (11), defects, layer thickness uniformity, and pattern edge roughness (12), bias and tolerance. Reduction of the high temperature processing cycle (i.e., the diffusion coefficient-time product) will also be considered. Constraints relating to electrical parameters include oxide (13) and junction (14) breakdown, tunneling (13), hot electron injection (15), avalanche multiplication (16,17) punch through conduction (18), and small geometry effects (19). The implications of nuclear radiation in the form of alpha particles (20) or cosmic rays (21) will be considered. Technological concerns indicate a practical feature size of about 0.1 to 0.2 micrometers (13,14,22) with a potential density improvement of about 100 times over current technology.

Device scaling, a concept that concerns coordinated changes in dimensions, voltages and doping concentrations, will be explained. Scaling is more easily applied to MOSFET's (23,24) than to bipolar transistors (25). Parameters such as sub-threshold conduction and the resistance of interconnection lines and contact areas defy the scaling approach. Low temperature operation offers a means of overcoming some of these "nonscaling" parameters (26). The scaling of parasitics and processing tolerances will be briefly discussed.

Complexity limits are difficult to quantify, but they relate to our inability to design circuitry involving very large numbers of components. These could also be thought of as conceptual or human limits. Complexity includes product definition, design time, engineering changes, testing, on-chip redundancy, computer-assisted design, and packaging. Complexity problems already apparent in microprocessor design (2) suggest that difficulties will arise in the 1 to 2 micrometer range. This will be within a factor of 10 of the component density of today's most complex IC chips.

As a monitor of progress in photolithography, the NMOS, polysilicon-gate, one-device, dynamic memory cell provides a good vehicle (6). Due to its high volume market, the dynamic RAM represents the technological leading edge of IC component density. Within the chip, the memory cell is an effective measure of a manufacturer's basic groundrule capability because the linewidth resolution, overlay misalignment, and edge bias tolerances are all brought into play. By the term groundrule we mean the minimum exposable linewidth or linespacing of a photoresist image, a single dimensional figure of merit that describes the photolithographic capability.

Figure 1 plots the number of minimum photolithographic squares per memory cell for a number of commercial polysilicon-gate 4 and 16 Kbit dynamic RAM chips relative to the year of

announcement (6). By summing the diffused, polysilicon and metal linewidths and linespacings in the cell and dividing by six, an average fabricated feature size is obtained which can serve as a figure of merit. The average feature size decreased from 8.2 micrometers in 1974 to 5.2 micrometers in 1978. By an empirical rule-of-thumb, the average feature size is about 1.5 times larger than the minimum exposable line width which in turn has decreased from about 5.0 to 3.5 micrometers in the same four-year period. Progress in cell design cleverness (6) is measured in units of lithographic squares ($25\mu m^2$ in 1974). The larger single polysilicon diffused bit line cell requires about 45 squares while the smallest double polysilicon cell requires about 25 squares. The number of squares per cell remains constant with time for a given cell layout.

The size of a dynamic RAM chip is directly proportional to the cell size because the cell utilization, i.e., the fraction of the chip area devoted to cells, is relatively constant. This feature is illustrated in Table 1 which also lists some of the areal parameters of recently announced 64 Kbit dynamic RAMS and two exploratory 256 Kbit chips. A great deal of device and circuit cleverness has gone into the design of memory cells. Since 1977, however, a barrier of 25 lithographic squares per cell has emerged, which has not been surpassed by either the 64 or 256 kbit dynamic RAMS (see Table 1).

Commercial attempts to break through the cell size barrier have failed, including CCDs and VMOS RAMs. Barring the unforeseen success of some new, vertical, memory cell structure which, for example, can place the MOS storage capacitor over the FET switch, dynamic RAMs appear destined to remain at 25 squares per cell or larger. Static RAM cells which are larger and more complex (e.g., six devices per cell) have shown more susceptability to reduction by circuit design cleverness (e.g., polysilicon load resistors located over the FETs) and now require about 100 squares per cell (27). Nevertheless, the leading product in terms of components per chip, the dynamic RAM, now appears to have reached its zenith when measured by the dimensionless density parameter of lithographic squares per cell.

Figure 2 plots the lithographic progress in manufacturing groundrule (deduced from Figure 1) against data from other sources. There is generally good agreement, particularly with regard to the yearly rate of progress. The slope of Figure 2 indicates that photolithographic linewidth has halved every six years. Taken alone, this rate of lithographic improvement could provide at most a doubling of the number of components per chip every three years for a fixed chip size. Due to shortcomings in registration and in pattern transfer, not all of this advantage can be claimed. Hence a doubling time of four to five years due to lithographic progress is more reasonable.

If the present rate of improvement continues, one can expect a minimum linewidth capability of 3 micrometers in 1979, 2 micrometers in 1982, and 1 micrometer in 1988. A 3 micrometer image capability is sufficient to fabricate a 64 Kbit dynamic RAM with 150,000 components in a chip area of 32μm^2 (50,000 $mils^2$).

Major perturbations in the fabrication environment result from an increase in wafer size or a change in lithographic technique because implementation of either activity generally requires an entirely new facility. Figure 3 indicates the progress in lithographic techniques. The development of new optical techniques has been mostly evolutionary, but the adoption of electron-beam direct write or X-ray exposure will be more revolutionary requiring new photoresists and new pattern transfer techniques (28). Such techniques should be commercially available for production by about 1988 and will provide lithographic image capability of one micrometer.

Figure 4 plots the number of nonredundant components (i.e., transistors, capacitors, diodes, or resistors) per chip versus the year of introduction of a subsequently volume-manufactured part (1). The figure describes the time rate of improvement of the number of components per IC chip. The highest points on the curve are associated with dynamic RAMs. The yearly doubling rate during the 1960 to 1975 timeframe may be decomposed into device and circuit cleverness (doubling every two years), chip size increase (doubling every four years), and lithographic linewidth improvement (doubling every four years). As we have discussed, for dynamic RAMs the improvement due to device and circuit cleverness went to zero in about 1977 and is expected to remain there. Obviously some increase in chip (die) size can be expected in the future and the increasing interest in novel packaging improvements is evidence of this. However, to rough approximation we can assume that the chip size improvement will go about to zero in 1980, which leaves lithographic linewidth improvement as the primary factor to drive component per chip improvements throughout the coming decade.

In terms of memory, Figure 4 indicates that even with only lithographic progress to carry the improvement in components per chip, we should achieve a doubling every four to five years and by 1990 we can expect commercial dynamic RAM chips with one million components, e.g., a 512 Kbit dynamic RAM or a 256 Kbit static RAM. In a mature industry, even this will be a remarkable accomplishment. Historically, custom microprocessors have lagged about a factor of four below dynamic RAMs in terms of components per chip (see Figure 4). The application of enhanced wirability through double polysilicon, silicide-on-polysilicon, and double level metal may close this gap somewhat, but this will probably be offset

by design complexity difficulties. Nevertheless, by 1990 we can expect a microprocessor chip in mass production containing 250,000 components (i.e., about 50,000 equivalent gates) and rivaling the circuit complexity of one of the smaller IBM 370 computers.

In reviewing the factors that may limit the rate of progress in integrated circuits, one observes that physical laws of nature (i.e., fundamental limits) provide limits in the range of 0.01 to 0.02 micrometers with component densities thousands of times smaller than any attainable today. Technological challenges (i.e., fundamental problems) suggest practical limits in the range of 0.1 to 0.2 micrometers with densities hundreds of times greater than in present production. Design complexity, however, presents some fundamental facts of life, which suggest that difficulties should start to appear in the one to two micrometer range attainable in the next five years. Already, the design time for microprocessors is seen to be scaling directly with number of components per chip, another constraint which should help to establish the doubling time of four to five years.

Finally, in terms of economic constraints, it is obvious that the price or value of the integrated circuit chip must exceed its fabrication and design cost. In 1970, the price to cost ratio was six to one, while by 1980 it had shrunk to two to one. This reduced return on investment means that an increasingly larger investment must be made to insure profitability. Then an even more careful balance between projected volume, technical capability, and market potential must be made which in turn tends to dictate the types of products designed and manufactured. In this sense, the economic limits on the rate of technological improvement are already being experienced.

ACKNOWLEDGEMENTS

Helpful comments from colleagues R. R. Troutman, F. H. Gaensslen, H. N. Yu, A. Reisman, L. M. Terman, F. H. Dill, R. W. Keyes, R. H. Dennard and J. E. Tomko of IBM are gratefully acknowledged.

REFERENCES

1. G. E. Moore, "Progress in Digital Integrated Electronics," IEEE Internat. Electron Dev. Meet. Tech. Digest, pp. 11-13, Wash. D.C., (Dec. 1975).

2. G. Moore, "VLSI: Some Fundamental Challenges," IEEE Spectrum, pp. 30-37, (April 1979).

3. V. L. Rideout, "Physical and Electrical Limitations to Improvement of Silicon Integrated Circuits,"

Microcircuit Engineering 79 Technical Digest, pp. 144-152, Aachen, W. Germany, (September 1979) and IEEE Compcon 80 Tech. Digest, pp. 2-6, San Francisco, (Feb. 26, 1980).

4. R. W. Landauer, "Irreversibility and Heat Generation in the Computing Process," IBM J. Res. Develop., 5, 183 (1961).

5. R. W. Keyes, "Physical Limits on Computer Devices," IEEE Compcon 78 Tech. Digest, pp. 294-296, San Francisco, (February 28, 1978).

6. V. L. Rideout, "One-Device Cells for Dynamic Random-Access Memories: A Tutorial," IEEE Trans. Electron. Devel., vol. ED-26, pp. 839-852, (June, 1979).

7. R. W. Keyes, E. P. Harris, and K. K. Konnerth, "The Role of Low Temperatures in the Operation of Logic Circuitry," Proc. IEEE, 58, pp. 1914-1932, (December, 1970).

8. R. W. Keyes, "Physical Problems of Small Structures in Electronics," Proc. IEEE, 60 pp. 1055-1062, (September, 1972).

9. J. Ruch, "Electron Dynamics in Short Channel FET's," IEEE Trans. Electron Dev., ED-19 pp. 652-654, (1972).

10. W. Oldham, "Si Materials Advances--The Challenge of High Resolution," Talk Presented at AIME Electronic Materials Symp., Palo Alto, (March 30, 1977).

11. R. A. Gdual, "The Effects of Processing on Radiation Damage in SiO^2," IEEE Internat. Electron Dev. Meet. Tech. Digest, pp. 148-150, Wash. D.C., (December, 1977).

12. J. T. Wallmark and S. M. Marcus, "Minimum Size and Maximum Packing Density of Nonredundant Semiconductor Devices," Proc. IRE, 50, pp. 286-298 (March, 1962).

13. B. Hoeneisen and C. A. Mead, "Fundamental Limitations in Microelectronics - I. MOS Technology," Solid State Electron., 15, pp. 819-829 (1972).

14. B. Hoeneisen and C. A. Mead, "Fundamental Limitations in Microelectronics - II. Bipolar Technology," Solid State Electron., 15, pp. 891-897 (1972).

15. T. H. Ning, "Hot-Electron Emission Currents in N-Channel IGFET's," IEEE Internat. Electron Dev. Meet. Tech. Digest, pp. 144147, Wash. D.C. (December, 1977), and also see papers in Solid State Electron., (January, 1978).

16. D. P. Kennedy and A. Phillips, "Source-Drain Breakdown in an Insulated Gate Field-Effect Transistor," IEEE Internat. Electron. Devel. Meeting Tech. Digest, pp. 160-163, Wash. D.C., (December, 1973).

17. R. R. Troutman, "Low-Level Avalanche Multiplication in IGFET's," IEEE Trans. Electron. Dev., ED-23, pp. 419-425, (April, 1976).

18. R. R. Troutman, "Subthreshold Design Considerations for IGFET's," J. Solid State Cir., SC-9, pp. 55-60, (April, 1974).

19. F. H. Gaensslen, "Geometry Effects of Small MOSFET Devices," IEEE Internat. Electron Dev. Meet. Tech. Digest, pp. 512-515, Wash. D.C., (December, 1977).

20. T. C. May and M. H. Woods, "Alpha-Particle-Induced Soft Errors in Dynamic Memories," IEEE Trans. Electron Dev., ED-26, pp. 2-9, (January, 1979).

21. J. Ziegler and W. A. Lanford, "The Effect of Cosmic Rays on Computer Memories," Science, 206, pp. 776-788, (November 16, 1979).

22. R. H. Dennard, "Miniaturized Fet's - Status and Ultimate Limits," IEEE Compcon 78 Tech. Digest, pp. 183-185, San Francisco, (February 29, 1978).

23. R. H. Dennard, F. H. Gaensslen, L. Kuhn, and H. N. Yu, "Design of Micron MOS Switching Devices," IEEE Internat. Electron Dev. Meeting, Talk 24.3, Wash. D.C., (December, 1972).

24. R. H. Dennard, F. H. Gaensslen, H. N. Yu, V. L. Rideout, E. Bassous, and A. LeBlanc, "Design of Ion Implanted MOSFET's with Very Small Physical Dimensions," IEEE J. Solid State Cir., SC-9, pp. 256-268, (October, 1974).

25. P. M. Solomon and D. D. Tang, "Bipolar Circuit Scaling," IEEE Internat. Solid-State Cir. Conf. Tech. Digest, pp. 86-87, Philadelphia, (February, 1979).

26. F. H. Gaensslen, V. L. Rideout, E. J. Walker, and J. J. Walker, "Very Small MOSFET's for Low Temperature Operation," IEEE Trans. Electron. Dev., ED-24, pp. 218-229, (March, 1977).

27. V. L. Rideout, "Future Directions in MOSFET Static RAMs," IEEE ICCC 80 Tech. Digest, pp. 4-11, Rye, N. Y., (October 1, 1980).

28. See papers presented at Session 18: "Device Technology-Electron Beam Technology and Applications," IEEE Internat. Electron Dev. Meeting Tech. Digest, pp. 431-445, Wash. D.C., (December, 1976) and IEEE Trans. Electron Dev., Vol. ED-26, "1μm MOSFET VLSI Technology, Parts I-V," pp. 318-378, (April, 1979).

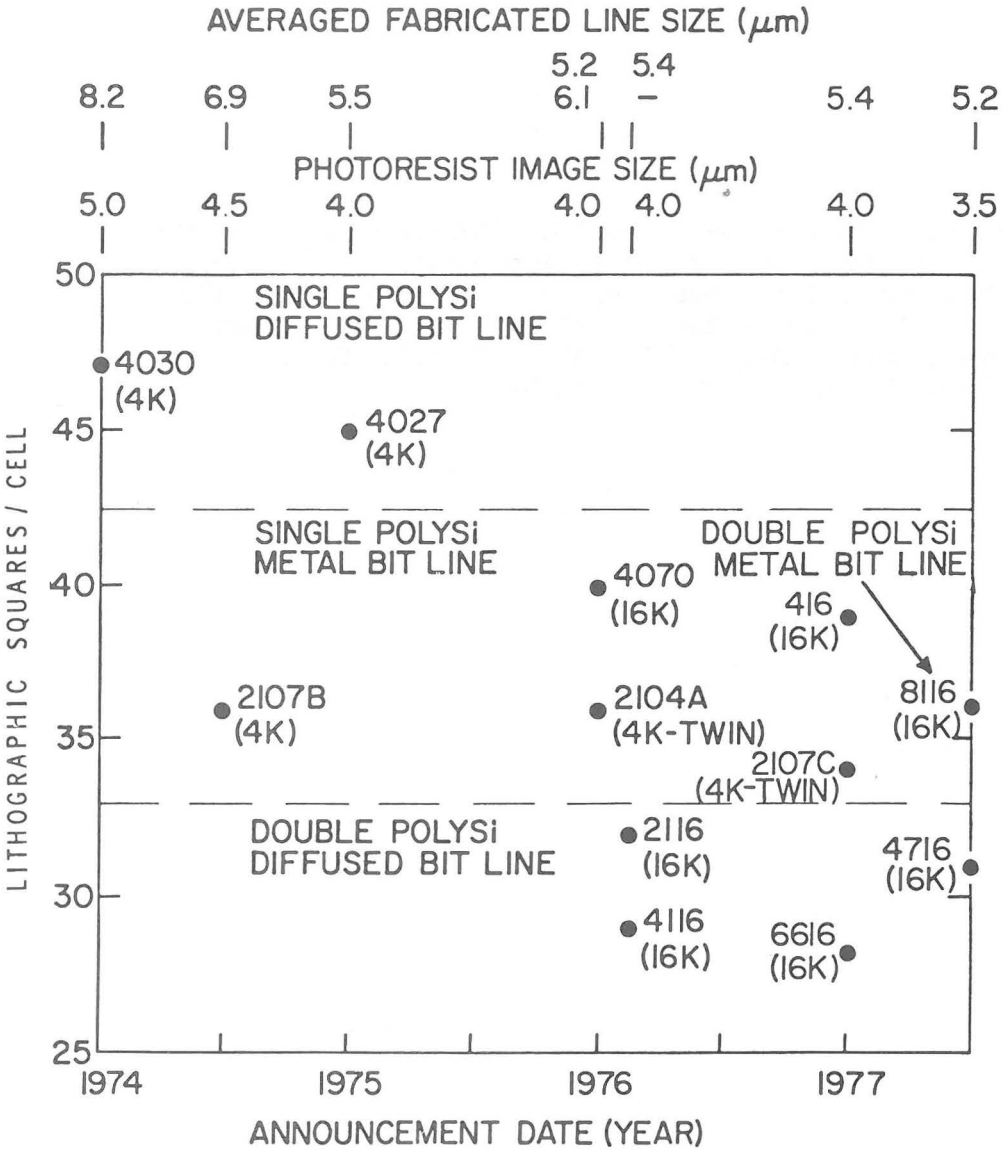

FIGURE 1

LITHOGRAPHIC SQUARES PER CELL VERSUS YEAR OF ANNOUNCEMENT DATE FOR ONE-DEVICE CELL DYNAMIC RAMs.

TABLE 1: DENSITY PARAMETERS OF DYNAMIC RAMs

FIRM	BITS PER CHIP	BIT LINE MATERIAL	CELL AREA (μm^2)	CHIP AREA (mm^2)	CHIP UTIL. (%)	AVERAGE* FEATURE (μm)	LITHO.* LINEWITH (μm)	LITHO. SQUARES PER CELL
A	64K	METAL	144	21.8	44	3.2	2.1	32
B	64K	DIFFUSION	192	27.4	46	3.1	2.1	44
C	64K	METAL	183	29.3	41	3.8	2.6	28
D	64K	POLY-Si	202	30.9	43	3.2	2.1	45
E	64K	DIFFUSION	180	23.0	51	4.0	2.7	25
F	64K	METAL	195	23.6	54	3.5	2.4	34
G	256K	POLY-Si	71	41.6	45	-	1.5	32
H	256K	METAL	69	34.4	53	-	1.5	31

*Lithographic Linewidth = Average Fabricated Feature/1.5

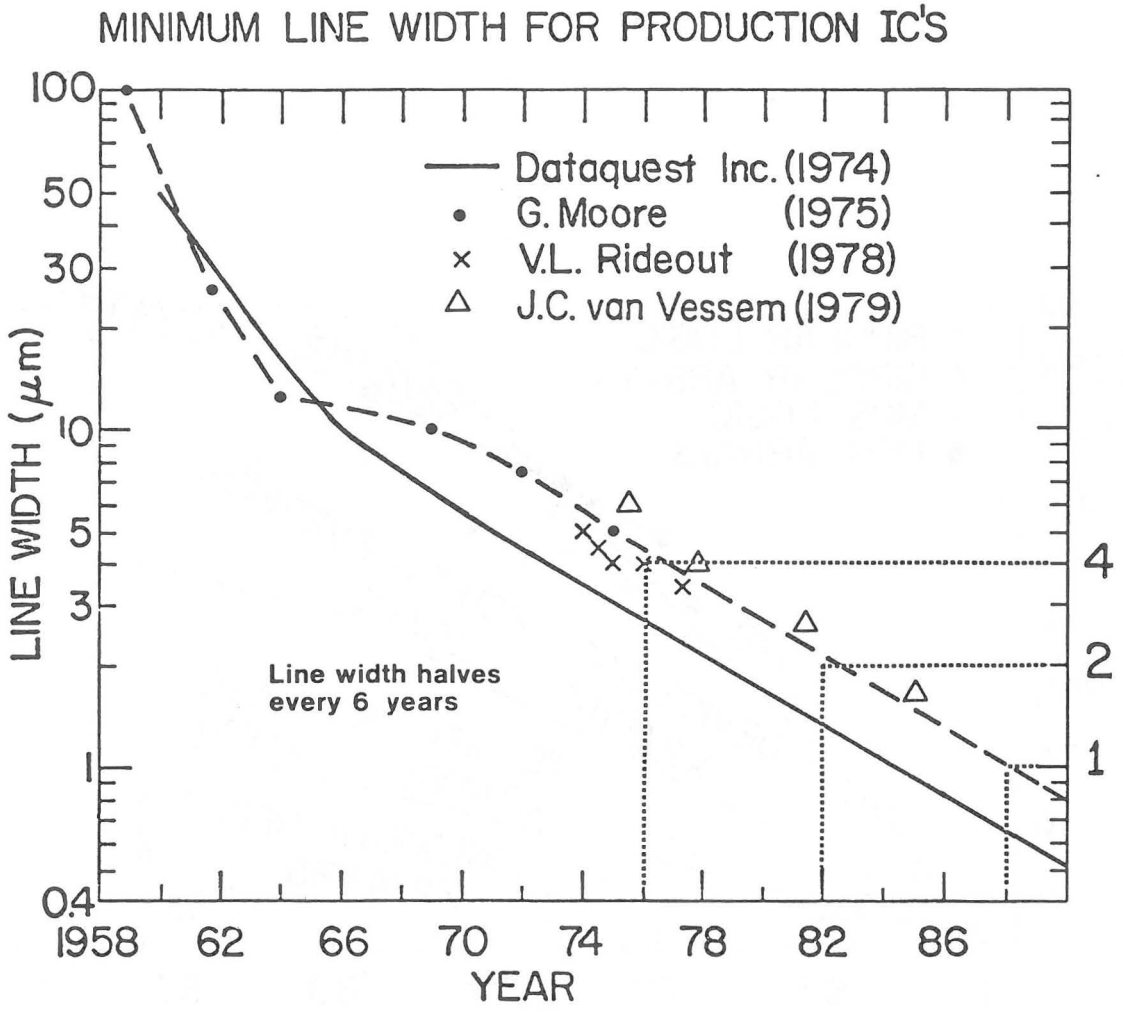

FIGURE 2: YEARLY PROGRESS IN MINIMUM PHOTOLITHOGRAPHIC IMAGE CAPABILITY

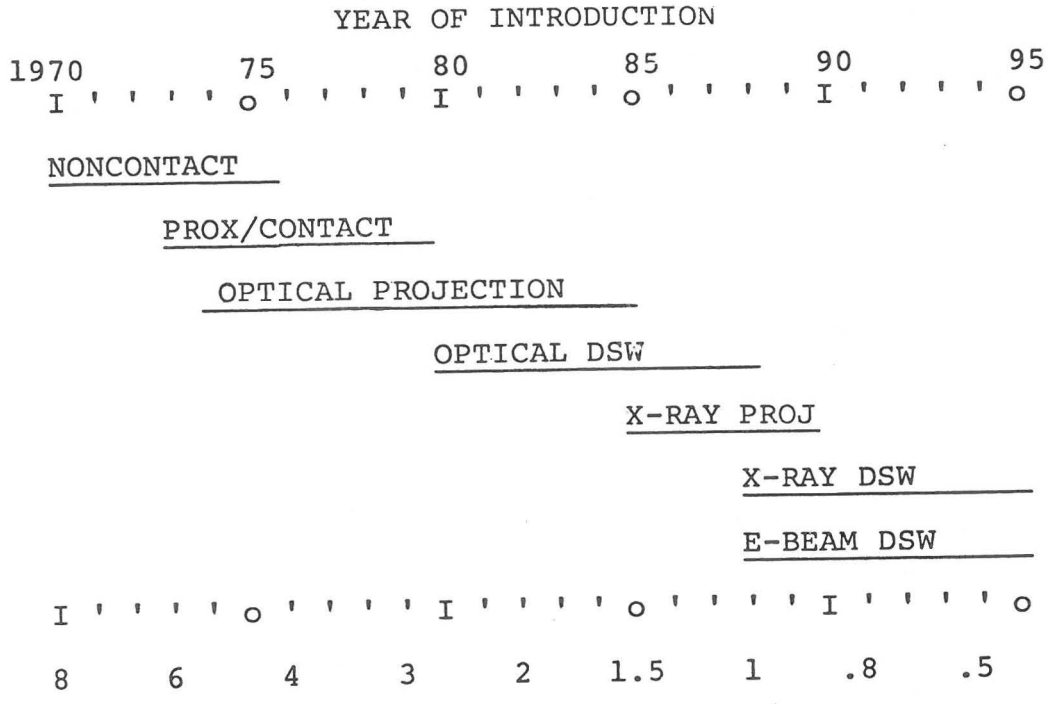

FIGURE 3: PROGRESS IN LITHOGRAPHIC LINEWIDTH

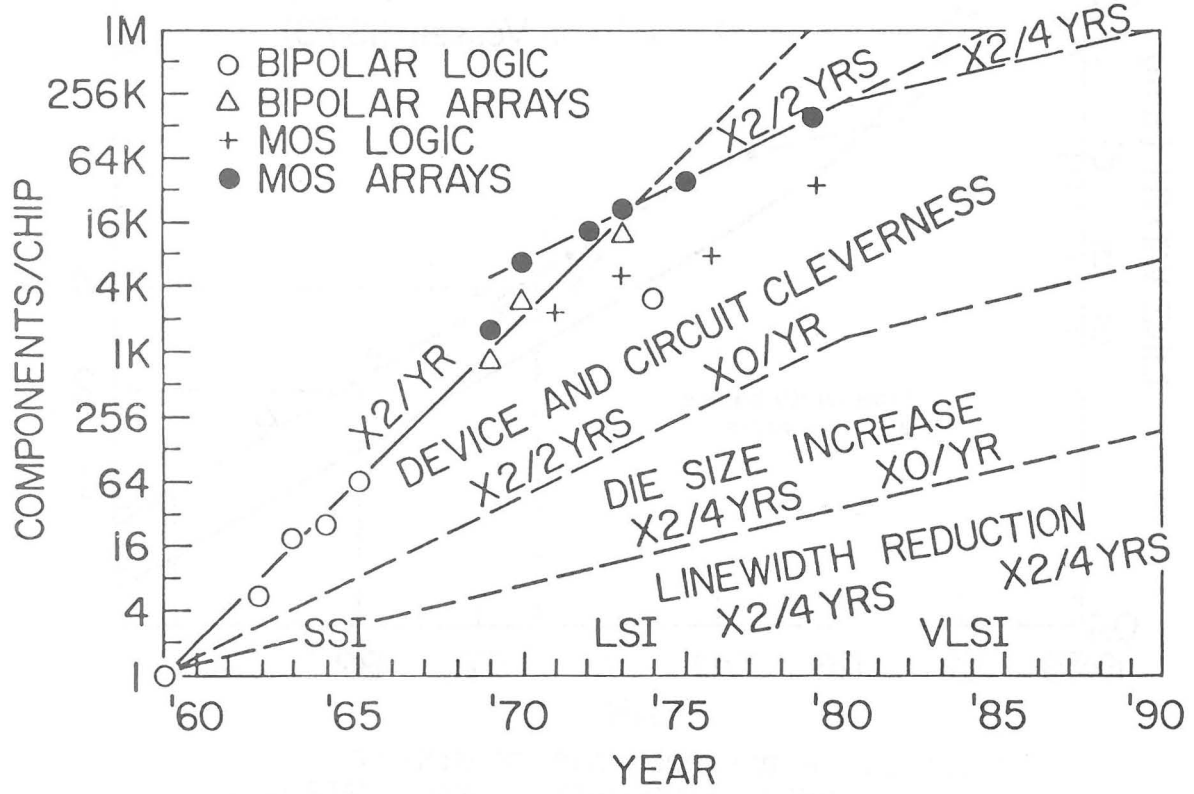

FIGURE 4: YEARLY PROGRESS IN NUMBER OF COMPONENTS PER CHIP

TRENDS IN SILICON PROCESSING

V. Leo Rideout
IBM Corporate Headquarters
Armonk, New York, USA, 10504
Presented at the Joint International Seminar on the
Teaching of Computing Science - VLSI, September 9, 1981

Abstract

The advent of very large scale integration will require substantial progress in all aspects of silicon technology: processing, lithography, modeling, design tools, chip architecture, and applications. This paper will survey current trends in silicon integrated circuit fabrication, focusing on new developments and outstanding problems. Progress in both bipolar and MOSFET technologies will be considered. Silicon fabrication techniques will be described in terms of the repetitious application of operations that are additive (oxidation, doping, deposition), selective (lithography), and subtractive (etching). The objective of these operations is a reliable and predictable device structure. Device structures will be described in terms of isolation areas, devices, contacts (intraconnection vias), wiring (interconnection lines), and passivation. Immediate problems in isolation size, device performance, contact resistance, and wiring topography will be identified. Future needs for improved structures will be indicated. Promising new approaches such as lightly-doped drain FETs and silicide-on-polysilicon (polycide) wiring will be described. Throughout this discussion the importance of process modeling will be emphasized.

1.0 INTRODUCTION

The advent of very large scale integrated circuits (ICs) will require substantial progress in all aspects of silicon technology: processing, lithography, modeling, design tools, chip architecture and applications. This paper will survey current trends in silicon integrated circuit fabrication including processing techniques, lithographic tools, and process modeling. As illustrated in Figure 1, the level of integration in mass manufactured IC chips has reached 160,000 components on a 64 Kbit dynamic random-access memory (RAM) chip, 80,000 on a 16 Kbit static RAM, and 40,000 on a 16 bit custom microprocessor (1). By the end of this decade we can expect 512 Kbit dynamic RAM's and 128 Kbit static RAM's with one million components per chip, and 64 bit microprocessors with 250,000 components. By the term component we mean an integrated transistor, resistor, capacitor, or diode.

In 1979, the worldwide sale of semiconductors was $11.1 billion, most of it in silicon IC's. In 1980, semiconductor manufacturers are expected to invest well over $2.0 billion in new fabrication plant and equipment. Including captive silicon suppliers, the total investment will be in excess of $3.0 billion. Obviously silicon processing is a big business. Presently, the most advanced bipolar and MOSFET circuits are mass produced with photolithographic groundrules of 2.5 to 3.0 microns. This is expected to improve to 2.0 microns in 1982, and to 1.0 microns by about 1988 (see Figure 2). New processing tools and techniques needed to maintain this progress and novel device structures evolving at smaller dimensions will be discussed in this paper.

2.0 ADDITIVE OPERATIONS

Silicon IC fabrication may be thought of as the repetitious application of materials processing techniques that are additive, selective, and subtractive in nature. Typically, only one lithographic masking operation or step is associated with each loop of this repetitious manufacturing procedure. It will be assumed that the reader is at least basically informed about the fabrication process for silicon integrated circuits (2).

2.1 OXIDATION

The ability to thermally grow a layer of silicon dioxide either locally or globally on a silicon wafer is one of that semiconductor's most important properties. The oxides are usually grown in dry oxygen (thin layers/slow growth) or in the presence of water vapor (thick layers/fast growth). Chlorine ions, usually in the form of HCl vapor, can be added to reduce oxide charge and improve capacitor characteristics, particularly in MOSFETs. Oxidations are typically performed in the 900-1100°C range for times of 10 to 100 minutes.

A major direction in oxidation techniques is toward lower temperatures (less than 1,000°C) and higher pressures (more than one atmosphere). The desirability of lower fabrication temperature is quite general because the drive for continually shrinking lateral dimensions (i.e., device scaling) has led to thinner vertical layers and reduced processing times. The processing time reduction can become so severe (only a few minutes) that uniformity control is impaired. In order to maintain or even lengthen processing times for better control, lower processing temperatures are desired. Other advantages of lower processing temperature include reduction of out-diffusion, grain growth, defect generation, and wafer warpage.

During oxidation, both the linear and parabolic rate constants increase with the partial pressure of the gaseous oxidant (3) (see Figure 3). This advantageously leads to a much faster

oxidation rate at a given temperature or the same oxidation rate at a much lower temperature. For every increase of one atmosphere, the oxidation rate doubles. Alternately, if a fixed rate is desired, every increase of one atmosphere allows a reduction of 30°C in oxidation temperature. Potential advantages include reduced thermal-induced damage (e.g., wafer warpage and oxidation-induced stacking faults), lower surface-state density, and reduced boron depletion. Presently, pressurized furnaces are commercially available for dry and wet oxidations at pressures up to 25 atmospheres. This technique can be expected to be incorporated into production in the near future.

2.2 CHEMICAL VAPOR DEPOSITION

Thin layers can also be chemically vapor deposited (CVD) from gaseous sources in an RF induction-heated furnace to provide epitaxial silicon (900-1300°C), polysilicon or SiO_2 or Si_3N_4 (600-1000°C), and passivation (600°C) layers (4). Dopants can be incorporated into the chemical deposition process. A major trend is to use low pressure CVD (0.1 to 40 Torr) which affords improved thickness control (especially for polysilicon layers), reduced auto-doping, and higher throughput. The deposition temperatures listed above could potentially be reduced by 100 to 200°C. Probably the most difficult deposition process to improve on is single crystal silicon-on-silicon epitaxy which is essential to bipolar processing. Two interesting research techniques that address this problem are molecular beam epitaxy or MBE (5) and solid phase epitaxy or SPE (6) which utilize substrate temperatures of 400 to 600°C.

The chemical vapor deposition rate can be enhanced by presence of an RF plasma. The gaseous reactants (e.g., nitrogen, ammonia, and silane) interact to form a solid film product and other gaseous by-products (7). The reaction is sustained by the RF plasma rather than by external hot-wall heating. Low deposition temperatures (200 to 400°C) and highly conformal films are the result. Plasma deposition of silicon nitride is now widely used for final passivation as a replacement for phosphorus-doped silicon dioxide. A related future activity is photo-excited CVD of silicon nitride, silicon dioxide and, possibly, epitaxial silicon.

Molecular beam epitaxy (5) utilizes a vaporized beam in ultra high vacuum. The MBE technique has exhibited excellent thin film quality (one micrometer thickness) but is hindered by throughput limitations and equipment cost. With solid phase epitaxy, a doped amorphous silicon layer is deposited onto the substrate, and an epitaxial film is produced by heating the composite either locally or in a furnace. N-channel MOSFETs have been fabricated in SPE-grown epitaxial layers with

channel mobilities of 360 to 480 cm^2/V-sec (6). In addition to low temperature growth for improved layer thickness control and reduced dopant redistribution, SPE offers selected-area epitaxy which can be attractive for defining isolation regions. This can be achieved by depositing onto a masked substrate, or by local heating with a laser or electron beam.

2.3 BEAM HEATING

In several processing operations it is necessary to subsequently heat or anneal the wafer. One such example is the annealing of regions doped by ion implantation to remove local stress and to activate the dopant. Traditionally a RF heated furnace with an inert gas ambient is used for annealing in the 800-1000°C range for about 30 minutes. Significant dopant redistribution can occur during the annealing step, but laser or electron beam heating offers the potential for fast local heating that avoids this. Obviously, the beams can be scanned to anneal the entire wafer. Potential advantages are reduced processing time and lower cost. Thus far beam heating techniques are not reliable enough to anneal active semiconductor regions for production devices. The future applications of beam annealing in order of acceptance probably are:

- inducing backside damage gettering,
- forming silicide layers,
- activating doping in polysilicon layers,
- annealing contact regions,
- growing epitaxial layers from deposited amorphous films,
- annealing implanted device areas,
- relieving stress in silicon-on-sapphire or in local isolation regions.

2.4 DOPING

Ion implantation is steadily replacing solid and gaseous diffusion as the primary means of doping silicon because it offers much better areal doping uniformity (better than 1%) as well as profile tailoring. The uses of ion implantation for backside gettering, channel threshold adjustment, source/drain or emitter/collector doping, and resistor fabrication, are well known (2,8). Machine capability is steadily being improved, particularly for higher currents up to 10mA, and higher ion energies up to 400 KeV. These higher throughput machines give rise to concerns with heat dissipation in the wafer.

One trend in ion implantation is toward very low energy (less than 10 KeV) implants for shallow distributions which will be needed for threshold adjustment of micrometer-sized MOSFETs.

The uniformity of low dose/low energy implants is still a significant problem. Another concern associated with such low energy implants is anomalous channeling which degrades the distribution. Yet another difficulty is grain boundary channeling, particularly of As or P through polysilicon gate electrodes (9). Another new implant application uses extremely high energies (2-3 MeV) to implant deep buried layers for alpha particle collection grids. Other novel applications of ion implantation include enhanced etching of oxide and silicon regions, contact via hole doping, silicide formation, and double-diffused (DMOS) FET's.

Focused ion beams (10) offer one means for combining additive, selective, and subtractive processing operations. Potentially, this technique could selectively dope the substrate, expose resist patterns, or sputter etch thin films. The goal is to reduce processing steps and eliminate masking operations. Narrow linewidths and precise registation will be required, however. To date, in the research laboratory, Ga ion beams have been used for doping and machining selected regions. Focused boron and arsenic beams have also been demonstrated. The most likely initial application of focused ion beams is in special applications requiring customized fabrication.

2.5 METALLIZATION

By far, the most popular material for metallic low resistance interconnection lines in silicon integrated circuits is aluminum. Aluminum is abundant, inexpensive, easy to evaporate and to pattern, self-passivating, and adheres well to both silicon and silicon dioxide. The most common metal deposition techniques are:

- evaporation from a RF heated source,
- evaporation by electron-beam heating, and
- sputtering.

Of these, the RF heating approach offers the least radiation or surface damage, particularly for FET fabrication. DC magnetron sputtering also has low associated radiation. The deposition of other metals for rectifying contacts (11) (e.g., Pt, W, Ta, Nb) often requires electron-beam heating with higher risks of radiation damage.

One of the most important trends in metallization is the development of two, three, or even four layers of metal wiring paths. Most of the difficulty centers around the insulating layers between metal levels (sputtered quartz, nitride, oxide-nitride or polymide) and the contact holes through these layers. The insulating layers must be deposited at low temperatures so as not to degrade the first metal (aluminum) interconnections. As more layers are added, the topography of

the structure becomes less and less planar, and correspondingly the linewidth control degrades. For example, in a triple metal system, the third level metal lines may have to be double the width of the first level lines. A goal is to develop planarizing techniques and via refill steps to improve planarity and line control. FETs with two metal levels and bipolars with three are now commonplace.

Aluminum, gold, and silver do not form intermetallic compounds with silicon, but many other metals such as Pt, Pd, Ti, Ta, Mo, and W do. These intermetallic compounds, called silicides, provide intimate metal-semiconductor contacts which have a number of useful properties including: high barrier heights, high eutectic formation temperatures, and low resistance. Applications of silicides include both rectifying and ohmic contacts, and, potentially, interconnection lines on top of polysilicon or diffused regions. A new area is the use of ion or laser beams to form silicides.

Low formation temperature silicides like PtSi and Pd_2Si have long been used as rectifying contacts and more recently studied as a means for reducing contact via resistance between Al lines and Si regions. Such techniques are applied late in the process after all high temperature steps have been completed. The development of silicide layers that can withstand high processing temperatures is now one of the most active research and development areas for IC metallurgy.

It has been proposed that a high formation temperature silicide such as tungsten, tantalum or molydenum silicide could be used to reduce the sheet resistance of polysilicon or diffused silicon regions in FETs (12). For process groundrules of over three micrometers, diffusion depths and polysilicon thicknesses are large enough that sheet resistances of 15 to 30 ohms/square can be obtained. As dimensions are reduced, however, sheet resistances rise degrading performance. A major objective is to cover the thin polysilicon layer with a low resistance silicide layer yielding a composite "polycide" layer with the gate electrode properties of polysilicon but with a sheet resistance of one to five ohms/square (13, 14). The layered gate electrode/interconnection line material must withstand high temperature oxidation and annealing steps and be easy to pattern and etch selectively (13). The etching step is particularly troublesome as polysilicon and metallic silicides have quite different chemical behavior. To date a simple preferred polycide technique has not been disclosed and a compromise between sheet resistance, etching behavior, and durability must be achieved. Polycides are also important for advanced bipolar structures that use thin polysilicon layers for wiring and as the source of shallow emitter doping (15).

2.6 PASSIVATION

A passivation layer is required over the silicon chip to inhibit damage from mobile ions (Na, Cu, etc.) and water vapor. Historically, sputtered quartz, phosphorus-doped glass, and oxide-nitride coatings have been used. More recently a move to plasma enhanced CVD nitride and to polyimide has occurred (7). The primary attractions are lower cost and more conformal coatings.

Polymide is particularly attractive in cases where multilevel metal wiring is employed. An interesting research development is a photosensitive polymide which could eliminate resist masking and etching of the passivation layers. Organic coatings for absorption of package-generated alpha particles are now popular for dynamic memory chips.

3.0 SELECTIVE OPERATIONS

Selective operations involve the exposure and development of lithographic patterns into a photosensitive layer (e.g., a resist). Over the past 20 years, the lithographic improvement traced in Figure 2 has progressed from proximity to contact to projection printing, all with full wafer exposure. As indicated in Figure 4, other developments in lithographic technique are expected in the future. Presently, the manufacturing capability of full-field projection printing is about 2.5 ± 1.0 micrometers. This lithographic groundrule refers to the resist patterns, not the final fabricated feature sizes on the silicon wafer which are approximate 1.5 times larger (16) (e.g., about 3.5 to 4.0 micrometers). A natural extension to shorter wavelengths (250 nm or "deep" UV) is occurring with the necessary transition from glass to quartz mask plates. When compared to standard UV (320 nm), deep UV enables the proximity gap between mask and wafer to be widened to reduce mask damage while maintaining system resolution. Alternately, operation with the same gap will give resolution increased by the square root of the wavelength ratio.

In order to progress to below 2 micrometers it appears that limited field, step-and-repeat (i.e., direct-step-on-the-wafer or DSW) projection optics with automation alignment will be required. Many new IC facilities now under construction are strongly dependent on optical step-and-repeat lithography. Linewidths of 1 to 1.25 micrometers and registration of ±0.4 to ±0.6 micrometers should be possible near the end of the decade.

Beyond one micrometer, a strong competition is developing between direct write electron-beam and projection X-ray. Although the outcome will not be decided for at least 5 years, both techniques still have serious deficiencies. Electron-beam lithography is costly, complicated, and in need of more

robust and more sensitive resists. Electron-beam machines have, however, become widely accepted as mask makers for optical projection and, at IBM, are also used in production for the customizing of final level metal patterns in bipolar and FET logic arrays. X-ray lithography is less well developed and needs more energetic sources, stable masks, more sensitive resists, and an automatic alignment scheme. An interesting proposal in the lithography field is to develop a relatively smaller electron storage ring for the generation of intense X-rays (17). Such a ring would service several (e.g., up to 10) X-ray lithographic stations.

In the resist area, an important activity concerns conformal multilayer masking techniques (18, 19). This is a means for circumventing low resist sensitivity and depth of field constraints, and for improving resolution (14). As shown in Figure 5, a very high resolution pattern can be formed in a thin layer of resist which would be too thin to be used for etching. But this high resolution pattern can be transmitted down into a thicker layer of working resist, without much loss in resolution, by a blanket exposure. IBM calls this portable conformal masking because the multilayer resist structure is transported from one exposure station to another and because the thicker resist conforms to the topography below it.

Over the past ten years, photolithographic linewidths have halved every 6 years (see Figure 2). The combined progress in lithographic machinery, resist materials, and etching techniques can double the chip density about every 4 years at best (see Figure 1). It is expected that improvements in optical projection lithography can sustain this rate of improvement for at least the next five years.

The transition to E-beam or X-ray lithography in production should take place in the latter half of the decade as optical wavelengths constrain photolithography to about one micrometer dimensions (see Figure 4). This transition will be slowed by the severe technical and economic difficulties inherent in introducing any new lithographic technology, by the problems associated with even larger chip and wafer sizes, and by the staying power associated with an immense investment in the highly utilitarian optical technology.

4.0 SUBTRACTIVE OPERATIONS

The transfer into the substrate of the exposed and developed mask pattern in the photoresist layer is accomplished by a subtractive etching operation. Historically, such operations were carried out using wet chemical etchants such as hydrofluoric, sulphuric, or phosphoric acids. The attraction of wet etchants is that they are highly selective, generally attacking only one layer species. Unfortunately, they are

frequently isotropic or non-directional in nature and hence tend to etch under the masking layer. Wet etches are temperature and concentration dependent and overetching is often needed to insure complete material removal.

Important improvements have been made in recent years with etching in RF generated plasmas (20, 21, 22) which is sometimes referred to as dry etching. Figure 6 illustrates the difference between sputter, plasma, and reactive ion etching. In sputter etching, a non-reactive gas such as Ar is used and the sample is simply bombarded with directional, energetic ions. The etching, however, is indiscriminant, i.e., non-selective. Thus, there is no physical mechanism that stops the etching process when a second layer is revealed.

With plasma etching (20, 22) a reactive gas species like CF_4 or CCl_4 is used with the wafers at plasma or ground potential and a pressure of 0.5 to 2 Torr. The high gas pressure leads to a random incidence of etching species. These conditions and the wafer positioning lead to an isotropic (i.e., non-directional) but highly selective etching which is widely used in the industry today, for example for ashing (stripping) exposed resist layers and for etching thin Si_3N_4 layers. As contrasted to the barrel assembly, the use of a parallel plate reactor improves the directionality of plasma etching, but with a loss in selectivity. Selectivity ratios of 10 or 20 to 1 are possible although 5 to 1 is more typical. The transition from wet etching to dry (plasma) etching is required for linewidth control for features in the 2 to 3 micrometer regime (14).

With reactive ion etching (RIE), a reactive gas like CF_4, $CClF_3$, C_2F_6, CCl_4, etc., is used at lower pressures (.020 to .040 Torr) and with the wafers placed on the cathode (e.g., out of the plasma discharge region) (23). A directional and selective etching condition results. The etching behavior is sensitive to various parameters such as RF power, gas pressure, and the choice of etching gas and cathode material. A strong attraction of RIE is that the etching gas composition can be manipulated to obtain different high etching rates (24). Also, by increasing the gas pressure or altering the sample position, combinations of reactive ion and plasma etching can be utilized (22).

The drawbacks to reactive ion etching are primarily technical (batch sizes, etch rate uniformity, understanding of the etching chemistry) and hence RIE will become more widely accepted as techniques improve. The combination of selectivity and directionality afforded by reactive ion etching is essential to the development of one micrometer processes. A

particular need is for robust photoresists that are highly resistant to plasma or reactive ion etching. Another need is the ability to provide a controlled slope on the edge of an etched line to relieve line coverage problems.

5.0 PROCESS MODELING

Computer aided design (CAD) tools are of increasing benefit to chip and circuit design for wire routing, cell placement, timing analysis, and design rule checking. Device modeling is also valuable in predicting electrical parameters, cutoff frequencies, threshold voltages, and so on. Only recently has process modeling begun to play an important role in integrated circuit fabrication (25).

Integrated circuit modeling activities can be roughly subdivided into process, device, circuit, and chip architecture areas. Device modeling using Poissons' equation and the continuity equations has been active for over twenty years, however process modeling dates back less than a decade. The primary purpose of process modeling is to provide a description of the device structure that can be utilized in a device analysis model. This structural description may include impurity distributions, insulator thicknesses, and device dimensions such as channel lengths and widths. This activity can be referred to as process profile modeling. The device analysis model utilizes this profile information in predicting the device parameters such as current-voltage or voltage-voltage relationships. These electrical relationships can then be incorporated into a circuit simulation model to predict switching rates or signal propagation times. The result of such an analysis will be a description of the nominal circuit performance. Before attempting to fabricate a circuit however, one should also determine the statistical range of circuit performance.

Variations in the fabrication process are unavoidable. Some of these are natural in origin such as uncertainties in substrate resistivity, others are equipment related arising from variations in furnace temperatures and implantation doses, and yet others are due to operator differences in etching times or other procedural variables. The cumulative result of many small statistical variations in fabrication can induce a significant resultant error in the final electrical parameters. It is here that process control modeling can be of value. Deviations from the base line process conditions can be deliberately introduced into the model and a sensitivity analysis performed (26). This can help to determine if one of the process steps is close to a critical point. In one instance in the author's experience, process models showed that the chosen implanter energy was on a steep sensitivity slope. Reducing the energy and increasing the dose gave a safer

fabrication condition which had less effect on the range of the resultant electrical device parameters.

Relative to CAD for circuit development, process modeling is still in its infancy. The process work began as one-dimensional profile analysis, by either analytical or numerical models, and progressed to two and three-dimensional forms. The initial work concerned primarily the additive materials operations. More recently, the selective operations of resist exposure and development have been addressed with the intent of defining linewidth parameters (27). More work still remains to be done in modeling of oxidation and diffusion which are now well developed and suitable for computer modeling. Plasma and reactive ion etching, however, are less well understood which complicates the description of their behavior. In the area of profile modeling, two successfull efforts concern descriptions of channel stoppers for FET isolation (28), and double-diffused regions for bipolar emitters (29).

Until very recently process modeling has been mostly an "after-the-fact" activity. Typically, process modeling was brought into play only after difficulties were detected in a new process, or in a modified older process. The VLSI era promises to be different. First, strictly from an economic point of view, the cost of experimental pilot line facilities is so high that brute force trial-and-error process development is too expensive. Thus tightly coupled process and device modeling activities will be required to optimize the process development cycle. Second, more highly integrated circuits promise to be more sensitive to statistical variations whether natural in origin or introduced during fabrication. Statistical fluctuations in doping, high densities of small defects, layer thickness control, and a host of problems may plague VLSI fabrication in the micrometer and submicrometer lithographic regime. Consequently, two and even three dimensional process control models will be required to help identify, understand, and avoid costly fabrication problems. Third, novel process steps can now be investigated with a computer model. This gives the process engineer a new tool for innovation.

6.0 AUTOMATED PROCESSING

Process automation is slowly and steadily being incorporated into integrated circuit fabrication. Automation can be applied in several ways (30). One area is robotics, or automated wafer handling in which the wafers are moved on air tracks (rather than by operators) and are mechanically inserted into open mouthed diffusion and evaporation stations. Two such facilities exist within IBM. A second area is on-line inventory control, also used at IBM, in which the position of various wafer lots in the line is determined and monitored by

computer. This is especially important for high throughput manufacturing. A third area of automation is local process control in which, for example, microprocessors are used to ramp furnaces or insert wafers at a particular station.

An important adjunct to process automation is on-line monitoring of gas purity, furnace temperature, etc. This activity is highly transducer dependent and much progress can still be made here. An interesting hypothesis is that real time process information could be used to modify a step further along in the process. For example, if the monitor showed that the gate insulator thickness was above nominal, the subsequent channel implantation dose could be reduced accordingly. Such on-line process tailoring might be required for the most sophisticated VLSI fabrication.

The objective of process automation is improved process quality or process throughput, or both. Probably the simplest and most direct way to increase productivity, however, is to increase the wafer size. Over the past decade wafer diameters have increased from one to five inches. By 1990, we can expect wafer diameters of seven to nine inches. Silicon strips or ribbons may also be developed. Each move to a larger wafer diameter is tramatic due not only to equipment changes, but due to size related problems as well. As wafers increase, thermally induced wafer stresses (warpage), stress-induced dislocations, global and local distortions, and other problems arise. Improved lithographic dimension is an alternative productivity enhancer, and it is lithographic machinery development that experiences the most demanding requirements with each quantum jump in wafer diameter. The difficulties associated with increasing wafer size tend to constrain progress in lithographic dimension. Compared to full field exposure, direct step-on-the-wafer, whether optical, E-beam, or X-ray, is relatively less sensitive to increased wafer size due to the limited field exposure area.

7.0 DEVICE STRUCTURES

The intended result of new processing techniques is an integrated transistor device structure that is smaller, cheaper, faster, lower in power, and more reliable than its predecessor. The basic structural elements of an integrated circuit are electrical isolation (between devices), the active device itself, contact vias (layer intraconnection points), wiring (interconnection lines), and passivation.

7.1 MOSFET TECHNOLOGY

Figure 7 shows a high density N-channel MOSFET or NMOS structure typical of the industry today (31). The structure is characterized by semi-recessed oxide isolation, a polysilicon-

gate FET, an etched and rediffused contact area, polysilicon and diffusion and aluminum wiring paths, and phosphorous-doped glass passivation.

7.1.1 ISOLATION

In MOSFET technology, dielectric oxide isolation between devices has progressed from planar (global thick oxide with etched holes) to locally grown (semi or fully recessed). Semi-recessed oxide is the mainstay of polysilicon-gate MOSFET technology due to its simplicity of fabrication, higher density, and self-aligned parasitic channel stopper. The locally grown or recessed oxide techniques require an oxidation-resistant silicon nitride layer. Oxide growth gives rise to a lateral oxidation wedge under this layer shaped like a bird's beak which reduces the active device area and complicates the structure (32). Although lacking high surface planarity, a simple approach is to use a common channel and field boron doping thereby eliminating nitride from the isolation step (33). This technique becomes more attractive at smaller dimensions with thinner isolation layers. Ideally, one would eliminate silicon nitride from the process and yet provide a deep, narrow, oxide region flush with the substrate surface and possessing a doped channel stopper region.

Complementary (CMOS) FET technology has a special problem in that the isolation must help reduce PNPN (silicon-controlled rectifier) latchup. A common approach is to widely separate devices and use the inherent diffusion isolation of the structure. Sapphire substrates provide complete dielectric isolation, which greatly improves the density, but the cost and processing difficulties of sapphire (e.g., outdiffusion, epi control, etc.) have led to a declining interest in this substrate material. The idealized deep dielectric isolation discussed above could greatly benefit the bulk CMOS technology.

7.1.2 DEVICE STRUCTURES

Two of the most interesting developments in MOSFET devices are the double-diffused or DMOS device (34), and the lightly-doped drain MOSFET (35). Figure 8A illustrates the DMOS device. The idea is to mask the drain and to laterally diffuse in a narrow p-type channel region on the source side of the FET. By this method, for example, a 0.5 micrometer channel length can be fabricated with 2.0 micrometer lithography. The advantages are higher gain, faster switching, and better channel length control. The penalties are added cost (one additional masking operation) and unilateral device operation. Static and dynamic RAMs and uncommitted logic arrays have been made using DMOS devices, but it is yet to be established that the performance improvement warrants the additional processing complexity.

In the lightly-doped drain structure shown in Figure 8B, the electric field at the drain is reduced by grading the diffusion profile there (35). Consequently, the drain voltage may be increased, thereby increasing the switching speed and/or current carrying capability. Extra masking steps are not required as the device fabrication is based on a controlled oxide overhang.

7.1.3 CONTACT VIAS

Contact resistance promises to be one of the first major problems to be encountered with higher integration because the resistance increases even more than linearly as contact area decreases (36). This antiscaling behavior is combatted, to first order, by deeply rediffusing the contact hole (37) (see Figure 7). Widely used in production, this technique is referred to as a borderless contact, but actually the diffused area is expanded under the isolation oxide by a rediffusion step. Another approach to optimizing metal to diffusion or to polysilicon contact areas is the self-registering contact in which an oxide layer is locally grown up around a nitride protected contact area (38).

Direct polysilicon to diffusion contacts buried under thick oxide (i.e., "buried contacts") become more difficult to fabricate as processes are scaled down because thermal drive times are reduced and a thinner polysilicon layer contains less N-type dopant. New approaches will be needed here.

7.1.4 WIRING

The most important advances in MOSFET wiring are double level metal and silicide-on-polysilicon (polycide) interconnections. N+ diffused regions can also be used for wiring, however, with scaling the higher sheet resistance of shallower diffused lines discourages it. Although it requires at least two extra masks for contact via and wiring patterns, double level metal is finding acceptance in FET production. The advantages include density improvements in 4 Kbit quasi-static RAMs, lower resistance wiring in CMOS microprocessors, and increased yield with redundancy wiring for 64 Kbit dynamic RAMs.

The refractive silicide on polysilicon technique offers 5 to 10 times lower line resistance without additional masking operations (13, 14). The processing difficulties with incorporating a reproducible W, Ta or Mo silicide layer into the existing FET process are formitable, however, this technique should be available in the near future. The reduced interconnection line resistance will lead to higher speed operation, especially in static RAMs and in microprocessors.

The use of an intermetallic silicide layer to reduce the sheet resistance of polysilicon or diffused region wiring represents one of the most important processing trends today. Although the polycide wiring approach has been demonstrated with both bipolar and FET test vehicles, thus far it has not been incorporated into mass manufacturing due to fabrication difficulties. When used on top of polysilicon, the silicide layer must be essentially transparent to the process, that is it must be patterned and oxidized along with the underlying polysilicon layer. This is not a trivial requirement. Generally silicide layers are more resistance to plasma and wet etches, and can become brittle when oxidized. Also, the silicide may fail to adhere to the polysilicon layer.

The most popular approach is to use a silicide layer over a polysilicon layer, the idea being to simultaneously retain the favorable properties of polysilicon either as a gate electrode material in FETs or as a diffusion source and contact for bipolar emitter or base regions, and to incorporate into it the low sheet resistance of the intermetallic silicide. A silicide layer alone would be easier to pattern but has poor oxidation properties and cannot serve as a controlled diffusion source. The most popular silicides being investigated are the high temperature ones like WSi_2, $TaSi_2$, $MoSi_2$, $NbSi_2$, and $TiSi_2$. The primary deposition techniques are sputtering, or co-evaporation by electron beam. Following the deposition, the silicide molecule must be established, or formed, by heating the composite at an elevated temperature (e.g., 900°C for 30 minutes). The patterning may be done either before or after the forming step.

Two interesting new techniques are the use of ion implantation (39) and laser annealing (40) to form the silicide. In the former case the deposited silicide layer is bombarded by an arsenic beam, the energy dissipation of which causes the silicide layer to form, thus eliminating the high temperature heating step.

It is interesting to speculate that the ion beam annealing technique for silicides might also be used to anneal, for example, implanted source and drain regions, possibly with an argon beam. Of course, lasers or electron beams can be incorporated into ion implanters. Overall, there is a constant desire to combine processing steps in situ, although thus far little of this has occurred. There still is a great deal of wafer handling involved in an IC process which may require as many as 150 sequential operations.

7.1.5 MOSFET

Figure 9 shows a hypothetical IC MOSFET of 1990. It is a polysilicon-gate bulk CMOS structure with one layer of polysilicon with silicide over it. Very shallow and lightly doped source and drain regions are used with laterally diffused regions for threshold control. Two layers of metal wiring are employed. The topography of the structure is highly planarized due to the fully recessed, deep dielectric, field isolation and the planarizing passivation layers. High conductivity silicide layers over the diffused and polysilicon regions greatly improve the electrical conductivity. Contact vias to connect metal lines to metal, diffused, and polysilicon lines are refilled with conductive material. The vias are self-registering to the lines they contact. The channel length of the device is 1.0 ± 0.25 micrometers and the threshold voltage is 0.5 volts. The gate insulator thickness is 250 Angstroms.

7.2 BIPOLAR TECHNOLOGY

Recently, the major emphasis in MOSFET development has been on dimensional reduction rather than on structural innovation. In contrast, bipolar technology is going through a minor renaissance in structural improvement. Just as the emergence of polysilicon gate electrodes spurred advances in FETs, the invention of integrated-injection logic has inspired innovation in bipolar device structures. Additionally, fabrication advances in FETs have influenced bipolar structures which now include, for example, self-aligned regions, polysilicon doping, and polycide interconnections.

7.2.1 ISOLATION

Bipolar isolation progressed directly from diffusion isolation to fully recessed oxide. Figure 10 shows a cross-section of an IBM masterslice bipolar logic structure (41). Planarity is a key requirement as the substrate must support three levels of metal wiring for which the linewidth control is affected by surface topography. The idealized, deep, narrow dielectric isolation described for MOSFETs would, of course, also benefit bipolars. Historically, novel isolation schemes have been more readily accepted into bipolar processes.

7.2.2 DEVICE STRUCTURE

A high degree of novelty is being incorporated into new bipolar structures. Among the most important of these are collector regions doped from and contacted by polysilicon, self-aligned collector-base contact edges, collector regions that abut the isolation oxide areas, and metal-interconnected base regions (42). A major thrust is toward 0.1 micrometer base widths facilitated by limited outdiffusion. Another is toward reducing parasitic capacitances by reducing collector area (butted regions) and by moving the base contact closer to the active base region (self-alignment).

7.2.3 CONTACT VIAS

To date, conventional etched contact vias have been used to emitter, subcollector, and extrinsic base regions. The use of polysilicon or silicide-on-polysilicon diffusion sources allows aluminum lines to contact polysilicon regions over the isolation regions. This relieves aluminum spiking problems and reduces the overall device area.

7.2.4 WIRING

Bipolar logic structures today use three levels of metal wiring (41) and this might increase to four or five levels in the future. The use of polycide layers, however, may alter this trend. The attraction of multilevel metal is high conductivity and low temperature processing, the drawbacks being extra masking steps and increasingly larger groundrules for the upper levels. Conductive metal refill techniques for contact vias between metal levels are needed, as are self-alignment techniques to register holes to lines.

7.2.5 FUTURE STRUCTURE

Figure 11 shows a hypothetical bipolar transistor of 1990. It is a T^2L structure with silicide-on-polysilicon for the emitter, base and collector doping. Device regions butt up against the ideally deep dielectric isolation. Three layers of metal wiring are employed. Contact vias between metal layers are refilled with metal and self-registering contact techniques are used. The base width of the transistor is 0.1 micrometers. Like the future FET, the bipolar structure is highly planarized to relieve line coverage problems and reduce the linewidth of upper metal layers.

8.0 STRUCTURAL PROBLEMS

8.1 ISOLATION AREA

Over the past decade, reduction in isolation area relative to device area has led to significant density improvements. This is exemplified by the transition from diffusion isolation to fully-recessed oxide isolation in bipolars, and from thick oxide with non-registered channel stoppers to semi-recessed oxide with self-aligned channel stoppers in N-channel MOSFETs. Nevertheless, today about 50% of an IC chip area is still devoted to isolation. A major improvement in isolation is needed for density enhancement. The idealized deep and narrow dielectric isolation, which we hypothesized earlier, might reduce isolation area to 25%. The isolation dielectric could be oxide, nitride, polysilicon, or combinations of these materials, and the refill should be planar with respect to the substrate surface.

8.2. CONTACT RESISTANCE

Contact resistance is one of the parameters that defies scaling and promises to present a major difficulty for VLSI (1). As contact areas are reduced, contact resistances increase linearly or superlinearly with contact diameter. One micrometer contact diameters with 10 to 100 ohms contact resistance can be expected (36). Techniques to insure uniformity such as rediffused (37) or laser annealed (40) contact holes help reduce contact resistance which is determined by the area, thickness, and resistivity of the contacted region. In this regard, an aluminum to silicide-on-polysilicon contact becomes particularly attractive.

8.3 LINE COVERAGE

The ability to cross one conductive line with another is impaired if the edges of the lines are not sloped or if the intervening insulating layer is not conformal. The problem is manifested in presentday products in aluminum lines crossing dry etched polysilicon lines in which a reflowed phosphorus-doped glass insulator is not used. Line thinning or breaks can occur at the crossing points. This will present an additional burden to the plasma etching techniques which will have to provide controlled slopes during delineation.

8.4 RADIATION

Radiation introduced during fabrication can detrimentally affect integrated circuits. Deposition, etching, and lithographic equipment is particularly suspect, especially for thin MOSFET gate insulators (43). Although it appears that any radiation damage that might be introduced by chemical vapor deposition steps or plasma etching can be annealed out at subsequent processing operations at over 800°C, metal deposition presents a different situation. Aluminum, for example, is deposited in vacuum by RF heating, by sputtering, or by electron-beam heating. The latter technique involves considerable radiation damage which cannot be annealed out at high temperatures as aluminum melts near 500°C. Electron-beam (or X-ray) lithography also introduces radiation damage, which, if used for the final metal definition may be hard to remove (44). Plasma etching can also introduce radiation damage (45).

A deleterious effect of process radiation is to produce damage states (traps) in the thin gate insulator which may charge up during the operating life on the device. Thus a device parameter such as the threshold voltage may slowly drift out of specification during the life of the device. One trend is to try to improve low temperature (below 500°C) annealing techniques by RF annealing or by thermal or plasma annealing in

pure hydrogen. The area of process induced radiation damage is a new one that promises to be of much greater importance for VLSI fabrication.

9.0 SUMMARY

A variety of new fabrication techniques and structural elements have been reviewed thus far in this work. We have speculated as to how these trends in silicon processing might come together to produce the integrated circuit device structures of the future. In particular we have tried to imagine what the silicon MOSFET and bipolar transistors will look like ten years hence and what techniques will be required to fabricate them.

9.1 IC FABRICATION OF THE FUTURE

Based on present projections, by 1990, one micrometer lithography with mask-to-mask alignment capability of ±0.25 micrometers will be practiced in mass production. Most likely this will be achieved with optical direct-step-on-wafer (DSW) projection systems. MOSFET memory chips with one million components will be available yielding 128 Kbit static RAMs and 256 Kbit dynamic RAMs in production with developmental chips of twice that capacity just emerging. Furthermore, 64 bit FET microprocessors with 250,000 components will also be available with the equivalent computing power of 50,000 logic circuits. High performance (less than 10 nanoseconds) bipolar cache memory chips of 32 Kbits will be available. Bipolar masterslice logic will have 10,000 circuits (about 50,000 components) and denser bipolar circuitry (PLAs and semi-custom chips) will reach 100,000 components. Chip power dissipation will play a major role in determining that the leading MOSFET technology will likely be polysilicon-gate bulk CMOS while the bipolar technology will be low power T^2L, or I^2L. Key improvements in packaging technology will also occur.

Ten years from now, the IC fabrication process will use ion implantation for all doping steps and dry etching for all material removal steps. Polycide layers will be used for gate electrodes, wiring, contacts, and controlled doping of shallow regions. Combinations of plasma and reactive ion etching will be used to achieve the required degree of selectivity, directionality, and line shape. Multilayer resists will be available to withstand the plasma etching. Seven inch wafer diameters and chips as large as 100 mm^2 will be processed. Low pressure and plasma assisted CVD will be commonplace, as will high pressure oxidation. Scanned laser beams will be used for various annealing and forming steps. Process control and monitoring will be highly automated so that operators will be employed primarily for maintaining and repairing equipment or moving containers of wafers from one station to another.

Monitoring information will be processed in real time so that later process steps may be customized to accommodate variations in earlier steps. The process line of tomorrow will look and operate somewhat like the computer center of today.

Technological prediction is an unreliable and highly imprecise art. In 1970, device and process researchers could not have predicted the HMOS polysilicon-gate FET or I^2L bipolar structures of today, and fabrication techniques like plasma etching and laser annealing were then unknown. In a ten year timeframe, lithographic dimensions decreased from over 6 to about 2.5 micrometers, wafer diameters increased from 1 to 5 inches, and device structures underwent revolutionary changes. In the next decade groundrules will decrease from 2.5 to 1 micrometer, wafer diameters will increase from 5 to 7 inches, and device structures will undoubtedly again undergo revolutionary changes. About the only certain future characteristic of integrated circuit technology is its unpredictability.

9.2 THE TWO CULTURES

The advances in very large scale integrated circuit design will be accompanied by substantial progress in microcircuit fabrication. This paper has concentrated almost exclusively on trends in silicon wafer fabrication techniques, only mildly considering lithographic progress, and ignoring completely requirements in packaging and circuit design. Obviously progress on all fronts will be necessary.

A challenging aspect of VLSI is the range of its impact upon the electronics industry. For example, digital technology ranges from solid state physical effects, through processing, devices, circuits, and chip architecture to system design. Two camps or cultures may be identified: chip fabricators that work in the semiconductor "foundry" handling silicon wafers and system designers that work in the CAD "foundry" handling terminal keyboards. The device and circuit designers work in the intermediate region between these two extremes. Clearly the coming of VLSI has required people with widely differing skills to work together.

The interface between the two cultures is often a difficult one. This is partly because the system tends to work from the top down with the design ideas driving the fabrication capability. Consider for example that a fabricator needs a ten million dollar laboratory and fifty associates to do a one micron feasibility study, while the designer can simulate an entire chip right on his own computer terminal. Apparently the role of the lone innovator has shifted from the laboratory to the office. Certainly a very innovative and excited atmosphere exists today in the design world.

In the future fabricators and designers will have to work closely together and there is a tremendous range of technology to span. One small way of improving the fabrication-design interface and encouraging process innovation is to promote process modeling and automation. VLSI brings the two cultures closer together which in itself may be one of the most important future trends in integrated circuits.

10.0 ACKNOWLEDGEMENTS

The author is partially indebted to S.C.Su. of the Hughes Research Center for making available information on low temperature processing. Discussion and helpful suggestions were also received from the author's collegues at IBM including: J. M. Aitken, E. Bassous, J. M. Blum, L. M. Ephrath, W. D. Grobman, R. D. Isaac, B. J. Lin, L. M. Terman, and M. Y. Tsai.

11.0 REFERENCES

1. V. L. Rideout, "Limits to Improvements of Silicon Integrated Circuits," Joint Internat. Seminar on Teaching of Computer Science-VLSI, Newcastle upon Tyne (Sept. 9, 1981).

2. A. B. Glaser and G. E. Subak-Sharpe, Integrated Circuit Engineering, Addison-Wesley Pub., Reading, Mass. (1977) or R. A. Colclaser, Microlectronics Processing and Device, John Wiley Pub., New York (1980).

3. Courtesy of S. C. Su, Hughes Research Center, Newport Beach, Calif.

4. R. J. Robinson, "CVD Process Trends," Semiconductor Internat., pp. 27-37 (March, 1979).

5. P. E. Luscher, W. S. Knodle, and Y. Chai, "Automated Molecular Beams Grow Thin Semiconductor Films," Electronics, pp. 160-168 (August 28, 1980).

6. Work reported by Hughes Research Laboratories, Malibu, Calif.

7. P. S. Burggraaf, "Plasma Depostion Production Trends," Semiconductor Internat., pp. 23-34 (March, 1980).

8. R. H. Dennard, V. L. Rideout, F. H. Gaensslen, and H. N. Yu, "Uses of Ion Implantation in Advanced MOSFETS," Electrochem. Soc. Ext. Abstracts, pp. 326-329, Dallas (October 1974).

9. Y. Wada, S. Nishimatsu, and N. Hashimoto, "Arsenic Ion Channeling Through Single Crystal Silicon," J. Electrochem. Soc., Vol. 127, pp. 206-210 (January, 1980).

10. R. L. Seliger and P. A. Sullivan, "Ion Beams Promise Practical Systems for Submicrometer Wafer Lithography," Electronics, pp. 142-146 (March 27, 1980).

11. V. L. Rideout, "A Review of the Theory, Technology, and Applications of Metal-Semiconductor Rectifiers," Thin Solid Films, Vol. 48, pp. 261-291 (1978).

12. V. L. Rideout, "Reducing the Sheet Resistance of Polysilicon Lines in Integrated Circuits," IBM Tech. Disc. Bul., Vol. 17, p. 1831 (1974).

13. B. L. Crowder and S. Zirinsky, "1μm MOSFET VLSI Technology: Part VII -- Metal Silicide Interconnection Technology -- A Future Perspective," IEEE Trans. Electron Dev., Vol. ED-26, pp. 369-371 (April, 1979).

14. J. Lyman, "Scaling the Barriers to VLSI's Fine Lines," Electronics, pp. 115-126 (June 19, 1980).

15. T. Takahashi, S. Wakamatsu, and K. Kimura, "A High Speed Multiplier Using Subnanosecond Bipolar VLSI Technologies," Eur. Solid-State Cir. Conf. Tech. Dig., pp. 110-112, Southampton (September, 1979).

16. V. L. Rideout, "Development of One-Device Random Access Memory Cells: A Tutorial," IEEE Trans. Electron Dev., Vol. ED-26, pp. 839-852 (June, 1979).

17. W. D. Grobmann, "Synchrotron Radiation X-ray Lithography," to be published.

18. B. J. Lin, "Portable Conformable Mask -- A Hybrid Near-Ultraviolet and Deep-Ultraviolet Patterning Technique," SPIE, Vol. 174, Develop. in Semicond. Microlitho. IV, pp. 114-121 (1979).

19. B. J. Lin and T. H. P. Chang, "Hybrid E-beam/Deep UV Expsoure Using Portable Conformable Masking (PCM) Technique," J. Vac. Soc. Technol., Vol. 16, pp. 1669-1771 (November/December, 1979).

20. P. S. Burggraaf, "Plasma Etching Technology, "Semiconductor Internat., pp. 49-58 (December, 1979).

21. C. J. Mogab and W. R. Harshbarger, "Plasma Processes Set to Etch Finer Lines with Less Undercutting," Electronics, pp. 117-121 (August 31, 1978).

22. L. M. Ephrath, "Dry Etching Review," Silicon Symposium, Electrochem. Soc. Spring Meeting, Minneapolis (May, 1980).

23. L. M. Ephrath, "Reactive Ion Etching for VLSI," IEEE Internat. Electron Dev. Meeting Tech. Digest, pp. 402-404, Washington, D.C. (December 8-10, 1980).

24. L. M. Ephrath, "Selective Etching of Silicon Dioxide using Reactive Ion Etching with CF_4-H_2," J. Electrochem. Soc., Vol. 126, pp. 1419-1421 (August, 1979).

25. J. Meindl, et al, Process Models for Integrated Circuits, Stanford Univ., to be published.

26. D. A. Antoniadis and R. W. Dutton, "Models for Computer Simulation of Complete IC Fabrication Process," IEEE Trans. Electron Dev., Vol. ED-26, pp. 490-500 (April, 1979).

27. A. R. Neureuther, D. F. Kyser, and C. H. Ting, "Electron-beam Resist Edge Profile Simulation," IEEE Trans. Electron Dev., Vol. ED-26, pp. 686-692 (April, 1979).

28. V. L. Rideout, B. L. Crowder, and F. F. Morehead, "Implanted Boron Channel Stoppers for MOSFET Integrated Circuits," talk presented at IEEE Semicond. Interface Specialists Conf., New Orleans (December, 1979).

29. R. Reif, R. W. Dutton, and D. A. Antoniadis, "Computer Simulation in Silicon Epitaxy," J. Electrochem. Soc., Ext. Abstracts, Vol. 79-1, pp. 352-355 (May 6-11, 1979).

30. See special session on Process Monitoring and Automation at the Electrochem. Soc. Meeting in St. Louis (May, 1981).

31. M. Eklund, "IC Technology in the Eighties," Semicond. Internat., Vol. 3, pp. 29-38 (January, 1980).

32. E. Bassous, H. N. Yu, and V. Maniscalco, "Topology of Silicon Structures with Recessed Silicon Dioxide," J. Electrochem. Soc., Vol. 123, pp. 1729-1737 (November, 1976).

33. R. H. Dennard, and V. L. Rideout, "Method of Fabrication for Field Effect Transistors Having a Common Channel Stopper," U. S. Patent 4,090,289 (May 23, 1978).

34. Y. Tarui, et al, "Diffusion Self-aligned MOST -- A New Approach for High Speed Devices," in Proc. First Conf. Solid-State Devices (Suppl. to J. Japan Soc. Appl. Phys., Vol. 39, pp. 105-110, 1970).

35. S. Ogura, P. J. Tsang, W. W. Walker, D. L. Critchlow, and J. F. Shepard, "Design and Characteristics of the Lightly Doped Drain-Source (LDD) Insulated Gate Field-Effect Transistor," IEEE Trans. Electron Dev., Vol. ED-27, pp. 1359-1367 (August 1980).

36. H. Nozawa, S. Nishimura, Y. Horiike, K. Okumura, H. Jizuka, and S. Kohyama, "High Density CMOS Processing for a 16 Kbit RAM," IEEE Internat. Electron Dev. Meet. Tech. Digest, pp. 366-369, Washington, D.C. (December, 1979).

37. W. G. Watrous, "MOSFET Transistor and Method of Fabrication," U. S. Patent 3,986,903 (October 19, 1976).

38. V. L. Rideout, J. J. Walker, and A. Cramer, "A One-device Cell Using a Single Layer of Polysilicon and a Self-Registering Metal-to-Polysilicon Contact," IBM J. Res. Develop., Vol. 24, pp. 339-347 (May, 1980).

39. M. Y. Tsai, C. S. Peterson, F. M. d'Heurle, and V. Maniscalco, "Refractory Metal Silicide Formation Induced by As+ Implantation," Appl. Phys. Lett., Vol. 37, pp. 295-298 (August, 1980).

40. C. J. Doherty, T. E. Seidel, H. J. Leamy, and G. K. Celler, "Formation of p-n Junctions and Ohmic Contacts at Laser Processed Pt-Si Surface Layers," J. Appl. Phys., Vol. 51, pp. 2718-2721 (May, 1980).

41. H. W. Curtis, "Integrated Circuit Design, Production, and Packaging for System/38," IBM S/38 Technology Development Report, pub. by GSD Tech. Comm., Atlanta, Georgia (1978).

42. D. D. Tang, T. N. Ning, R. D. Isaac, G. C. Feth, S. K. Wiedmann, and H. N. Yu, "Sub-nanosecond Self-aligned I^2L/MTL Circuits," IEEE Internat. Electron Device Meeting., Tech. Digest, pp. 201-203, Washington, D.C., (December, 1979), also to be published in IEEE Trans. Electron Dev. (August, 1980).

43. R. A. Gdula, "The Effects of Processing on Radiation Damage in SiO_2," IEEE Trans. Electron Dev., Vol. ED-26, pp. 644-646 (April, 1979).

44. J. M. Aitken, "1μm MOSFET VLSI Technology: Part VIII -- Radiation Effects," IEEE Trans. Electron Dev., Vol. ED-26, pp. 372-378 (April, 1979).

45. D. J. DiMaria, L. M. Ephrath, and D. R. Young, "Radiation Damage in Silicon Dioxide Films Exposed to Reactive Ion Etching," J. Appl. Phys., Vol. 50, pp. 4015-4021 (June, 1979).

12.0 FIGURES

1. Yearly rate of improvement in components per chip (after ref. 1).

2. Progress in lithographic linewidth (after ref. 1).

3. Wet oxidation growth curves for 1 and 10 atmospheres (after ref. 3).

4. Development of lithographic exposure techniques.

5. Conformal masking technique (after ref. 18).

6. Comparison of sputter, plasma, and reactive ion etching (after ref. 22).

7. Present MOSFET structure (after ref. 31).

8. DMOS device (A) and lightly-doped drain FET (B) (after refs. 34 and 35, respectively).

9. Future FET structure.

10. Present bipolar structure (after ref. 41).

11. Future bipolar structure.

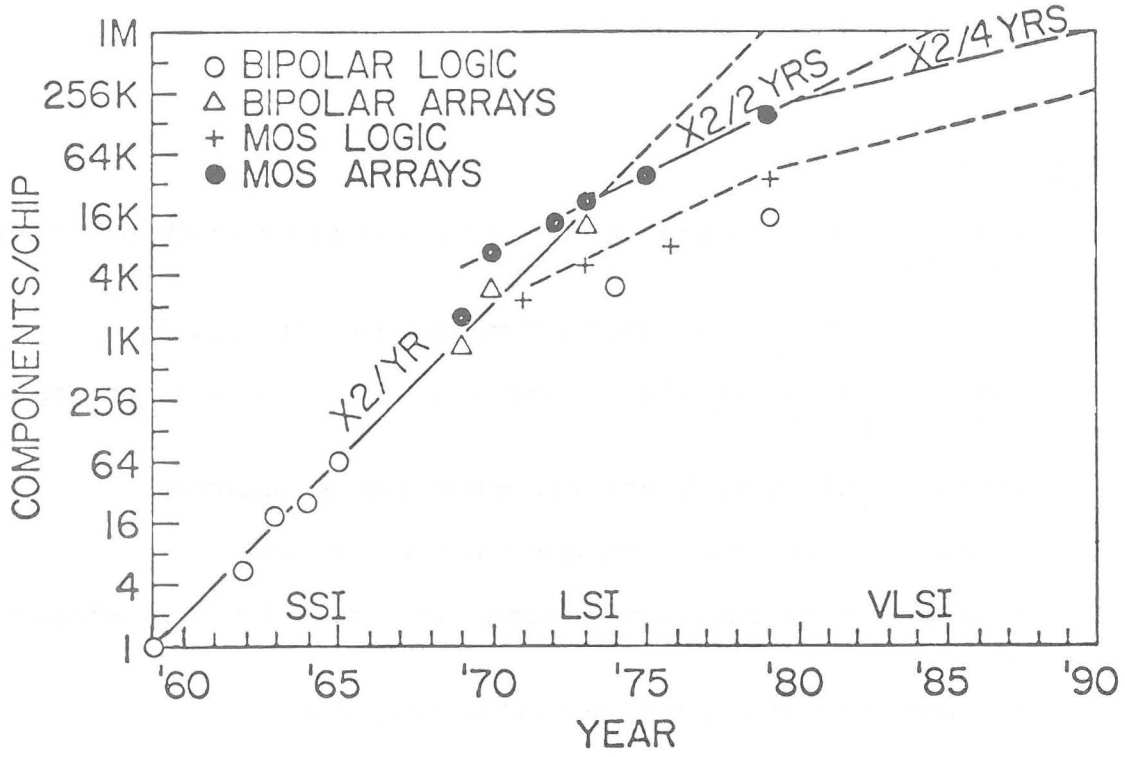

FIGURE 1: YEARLY RATE OF IMPROVEMENT IN COMPONENTS PER CHIP.

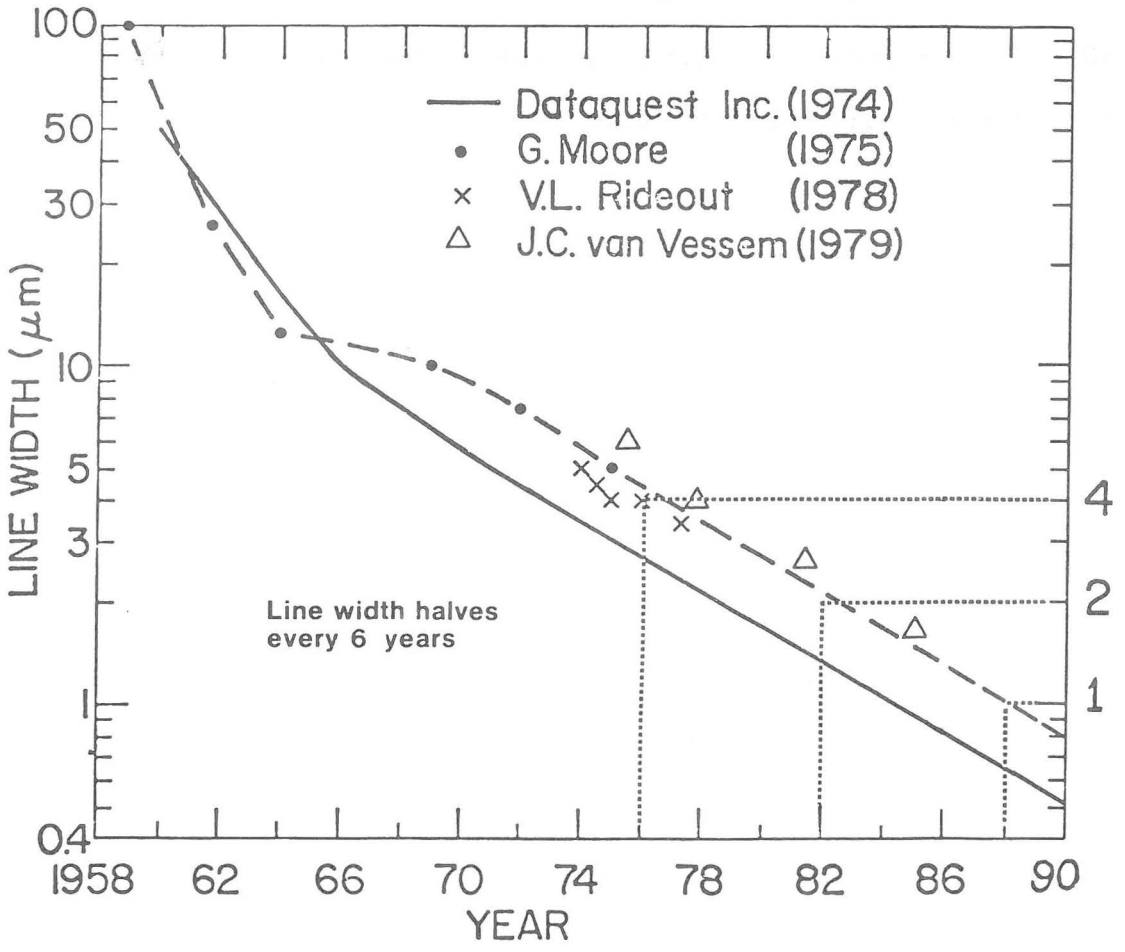

FIGURE 2: PROGRESS IN MINIMUM LINEWIDTH FOR PRODUCTION ICs.

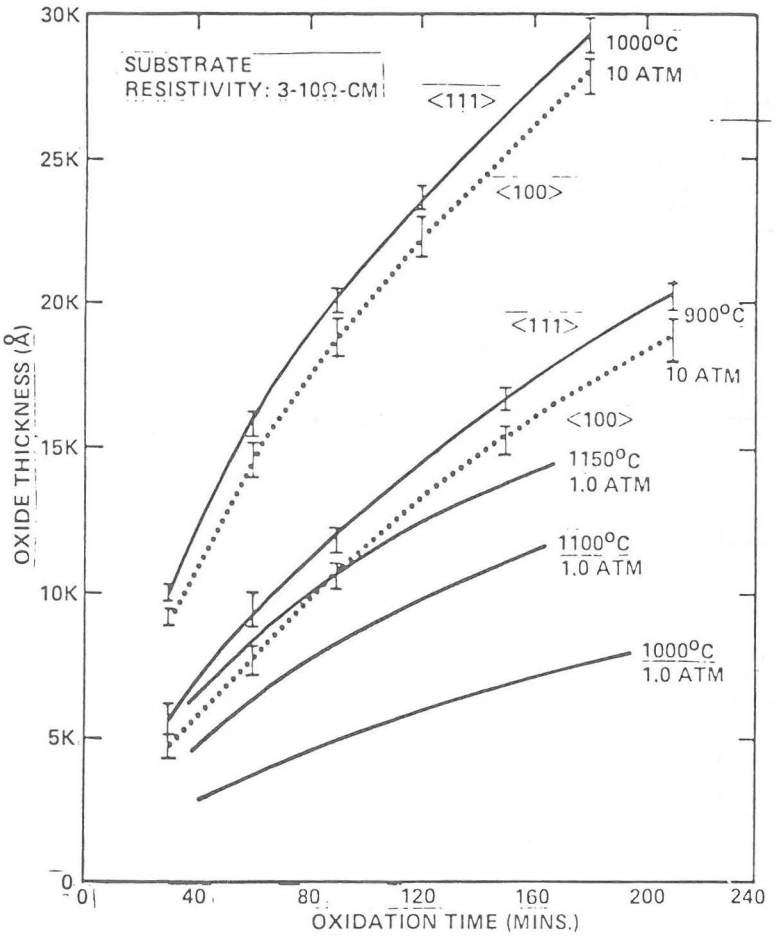

FIGURE 3:

WET OXIDATION GROWTH CURVES FOR 1 AND 10 ATMOSPHERES.

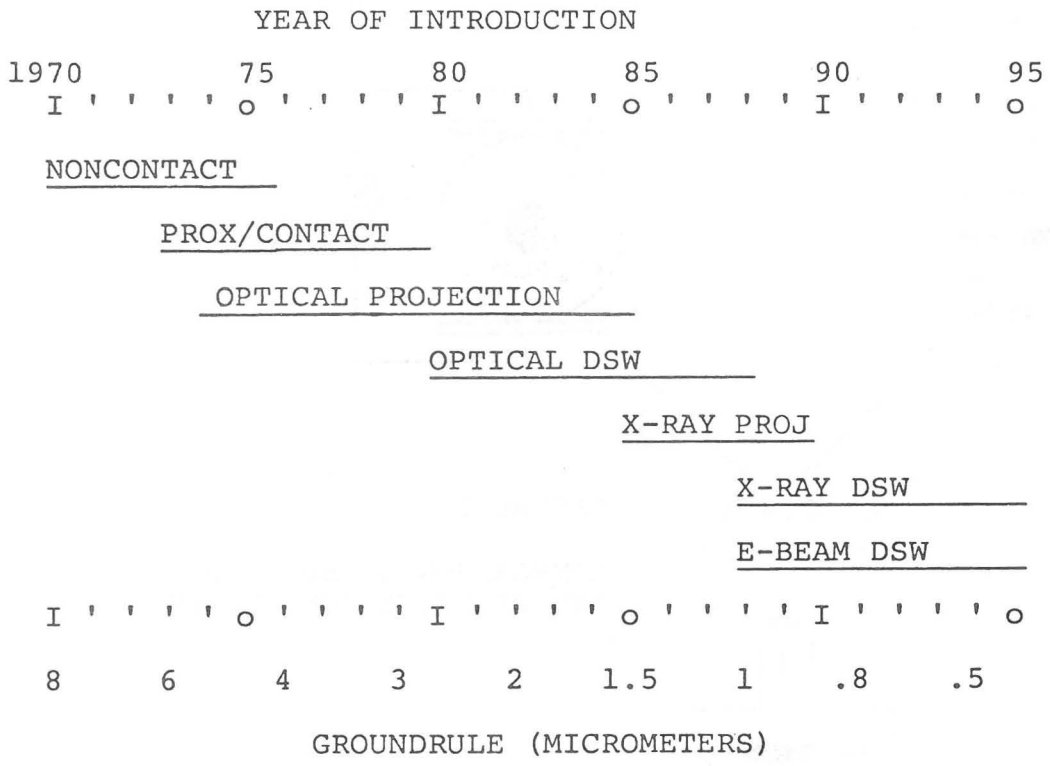

FIGURE 4: DEVELOPMENT OF LITHOGRAPHIC EXPOSURE TECHNIQUES.

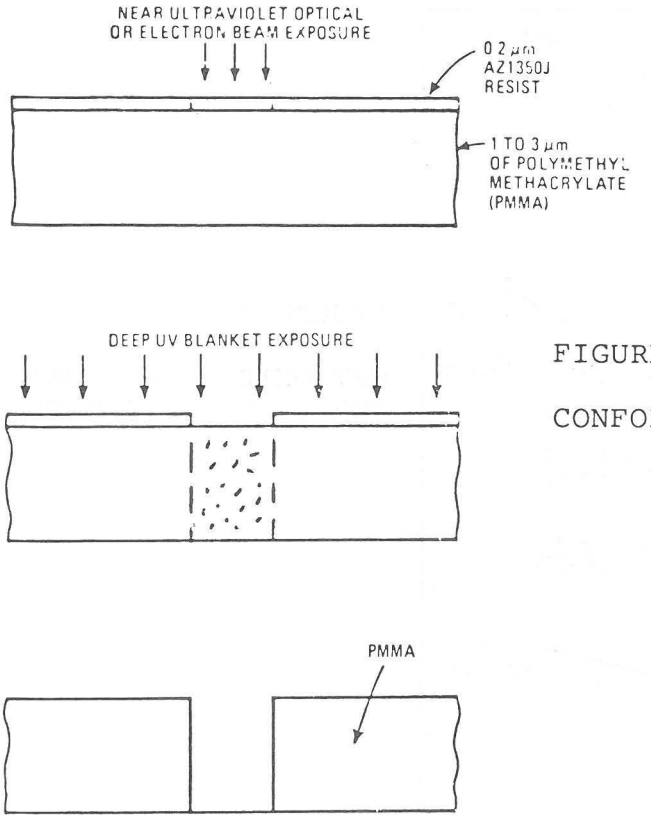

FIGURE 5:

CONFORMAL MASKING TECHNIQUE

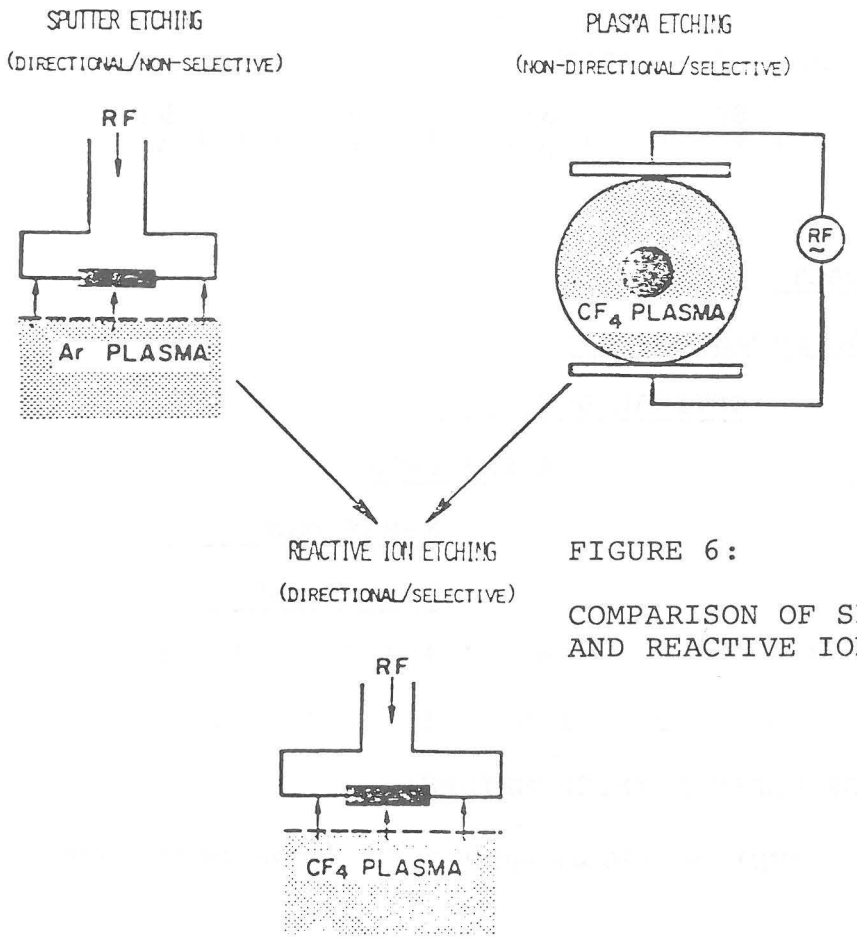

FIGURE 6:

COMPARISON OF SPUTTER, PLASMA, AND REACTIVE ION ETCHING.

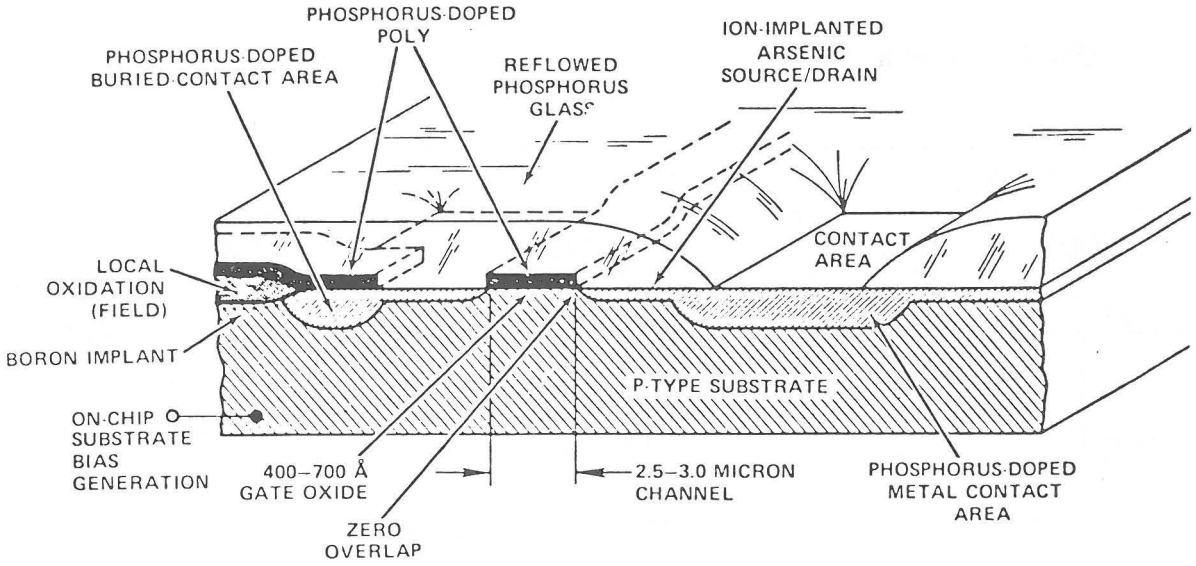

FIGURE 7: PRESENT MOSFET STRUCTURE.

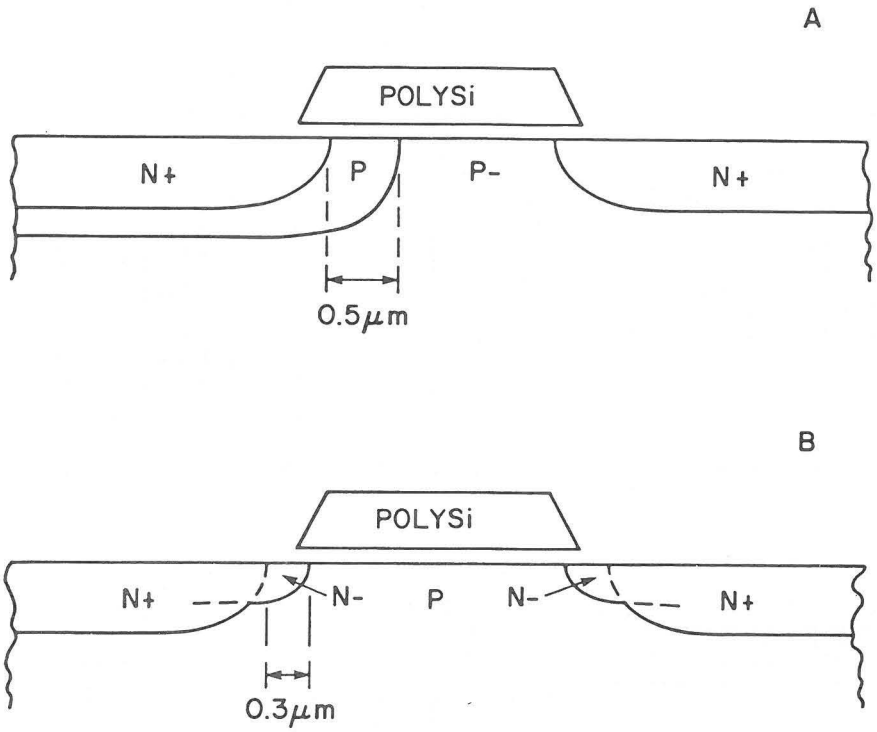

FIGURE 8: DMOS DEVICE (A) AND LIGHTLY-DOPED DRAIN FET (B).

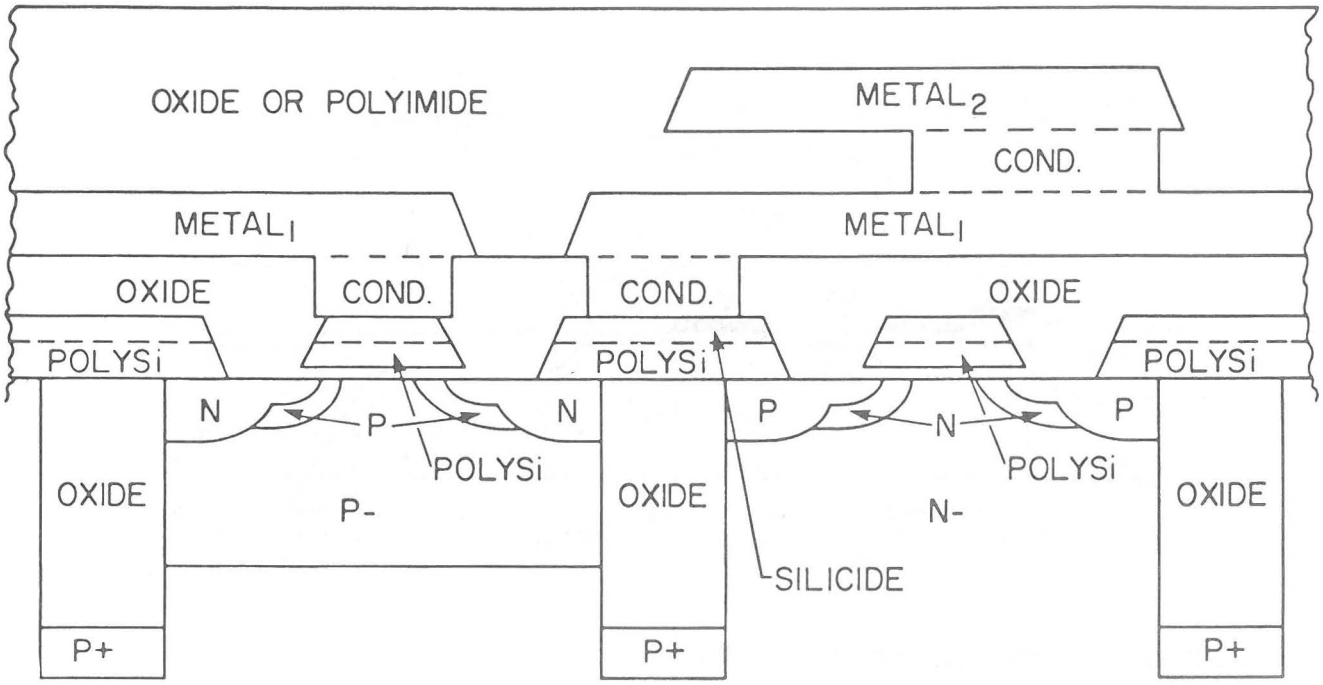

FIGURE 9: FUTURE FET STRUCTURE

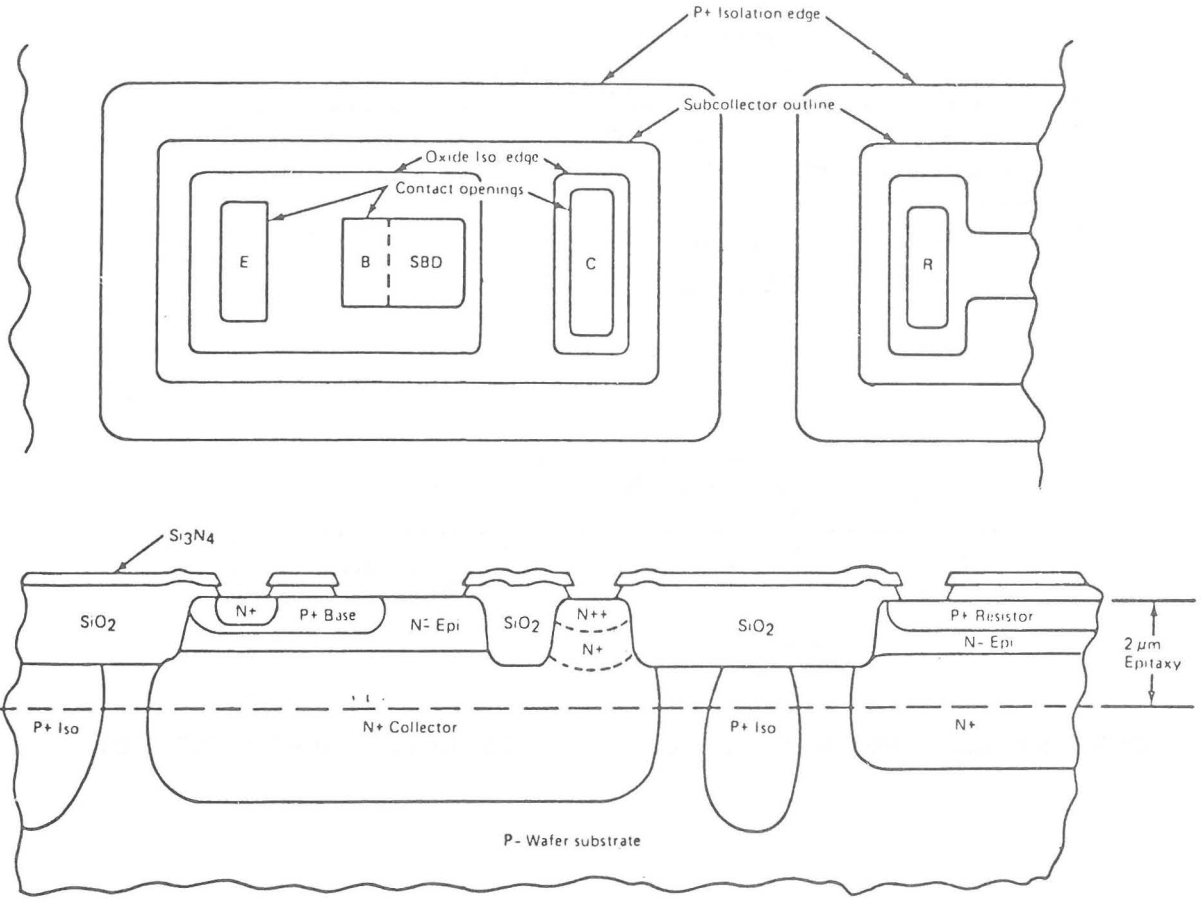

FIGURE 10: PRESENT BIPOLAR STRUCTURE

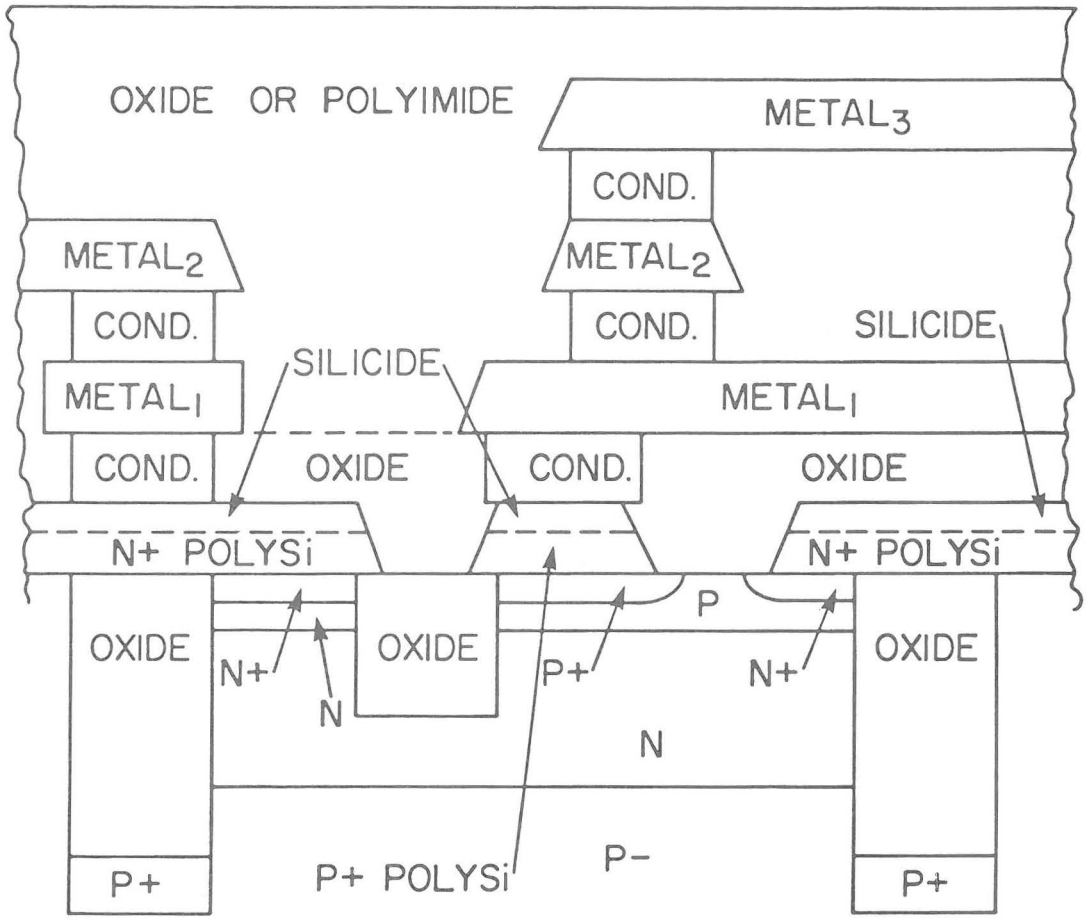

FIGURE 11: FUTURE BIPOLAR STRUCTURE.