Nanometer Reliability

Dr. Danny Rittman
danny@tayden.com

Abstract

The great success of semiconductor industry has been driven by the advancement in transistor technology in its early era. The industry could improve the performance of their products by shrinking the transistor dimensions and integrating more transistors. However, this strategy is becoming less effective, as the transistors demanded substantial interconnections between them, and the speed of integrated circuit products are being dominated by interconnections. Innovations are necessary in the interconnection technology to overcome the barrier. As we step into deep nanometer arena major reliability issues arise. Among them are, hot electron degradation, Electromigration & Self-Heating, Oxide Breakdown (TDDB) , P transistor degradation (NBTI), latchup, ESD, Voltage Drop, Soft Error and packaging issues.

"Within high-k gate dielectrics, metal gate, copper/low-k interconnects, the introduction of new materials, processes, and devices presents challenges. Bulk material and interface properties usually define the intrinsic reliability characteristics while defects establish the extrinsic reliability characteristics. Process integration flow, techniques, and process tools often create first order reliability effects (both intrinsic and extrinsic). The importance of characterizing these materials and processes for reliability as well as for performance during the early development stage cannot be overstated. System-on-chip (SOC) products that typically integrate new function and often include large memories (SRAM, DRAM, and Flash) bring about unique design, integration, and test challenges. Microsystems require consideration of a wider range of failure modes than microelectronics alone and introduce new failure modes because of the interaction of diverse technologies that would not be present if each technology were manufactured on a separate chip. In addition, optical, chemical, and biometric sensors and micromachines (MEMs) require the development of new accelerated tests and failure mechanism models. Electrostatic discharge (ESD), latchup, and packaging in the nanometer regime also raise reliability concerns. Even though ESD and latchup effects have been well characterized for many years, scaling brings about new issues and concerns. Similarly, the increased complexity and performance requirements for packaging these products act as an exponential multiplier for many of the failure mechanisms besides introducing new ones. Finally, two critical crosscut issues are related to design and test. These may be the more difficult challenges as the work needed to reach solutions is typically dispersed across many organizations, sites, and partners. Integration efforts tend to be less focused than material and device issues. Although the challenges may be clear, the paths to find solutions tend to be fragmented and obscure; consequently, these items require special research focus. This document is neither a complete nor exhaustive list of reliability challenges for the ITRS. Certainly any area of technology advancement includes its own set of potential reliability problems and new challenges. Instead, those broadest or most critical challenges are highlighted." **The International Technology Roadmap for Semiconductors (ITRS) - Critical Reliability Challenges for the International Technology Roadmap for Semiconductors (ITRS); March 2004**

System Reliability has been a practiced for decades now. However, as we move into the deep nanometer arena, reliability estimation methods need to be looked into in the new light of today's realities. In this paper we'll discuss the nature of reliability issues for nanometer design. We discuss about various failure phenomenon that are detrimental to the reliability of ICs. We then describe the industry way to approach this subject and EDA solutions.

1 Introduction

With feature sizes also moving into deep nanometer scale, it becomes imperative for the technical community to critically review the issues that the new age has brought with itself. High speed and low power are not the only targets that designers have to design for. Specifications now include additional functionality and consistent performance. With devices becoming smaller, performance variation caused by a process variation is much larger. The development of semiconductor technology in the next decade will bring a broad set of reliability challenges at a pace that has not been seen in the last 20 years. Many aspects of semiconductor design and manufacturing will undergo dramatic changes that threaten the nearly unlimited lifetime and high level of reliability that customers have come to expect even as product complexity and performance have increased. The introduction of new materials, processes, and novel devices along with voltage scaling limitations, increasing power, die size, and package complexity will impose many new reliability challenges. With structures on the chips becoming smaller and the number of such structures increasing exponentially, even if reliability and manufacturability of the structures remains same, total reliability of the chip goes down. More number of structures increases the probability of failure on a chip. Two trends are forcing a dramatic change in the approach and methods for assuring IC's reliability. First, the gap between normal operating and accelerated test conditions is continuing to narrow, reducing the acceleration factors. Second, increased device complexity is making it impossible or prohibitively expensive to exercise or stimulate the product to obtain sufficient fault coverage in accelerated life tests. As a result, the efficiency and even the ability to meaningfully test reliability at the product level are rapidly diminishing.

As the market demand continues to push product performance to its technological limits, the tradeoff between performance and lifetime must be tailored to the needs of different market segments. No longer can a single product satisfy all applications with significant reliability and performance margins. This in turn requires that accurate reliability models and tools for lifetime estimation must be available during the product design stage. A failure mechanism-driven approach must be employed, identifying the potential failures and evaluating their kinetics and impact based on the specific application conditions and requirements of each market segment. A much improved understanding of materials properties and failure mechanisms and models is required. If decisions rely on standards-based tests, then performance may be artificially limited and/or development costs increased while driving reliability to levels beyond the product or application needs. Together these trends demand that reliability be modeled much more precisely during the product design cycle to make the correct performance vs. reliability tradeoffs. The introduction of new materials with more limited operating margins further accelerates this shift and

requires that the potential failure mechanisms, the required test structures, and the corresponding models be identified and developed well in advance of the technology qualification. Similarly, the introduction of novel devices and new components for SOC integration will have profound reliability impacts as they bring new failure modes and mechanisms. Many of these devices are now in their infancy, and practical devices are still too far away for reliability characterization. If history is a guide, it is likely that work on the reliability of these novel devices will be late and sub-critical. Bringing reliability issues upstream in the development process will result in a better assessment of the readiness of these technologies for volume production.

2. Nanometer Reliability Issues

A. Hot-electron effects and oxide degradation *

The term **'hot electrons**' refers to **electrons** (`**hot carriers**` can refer to either holes or electrons) that have gained very high kinetic energy after being accelerated by a strong electric field in areas of high field intensities within a semiconductor (especially MOS) device. Because of their high kinetic energy, hot electrons can get injected and trapped in areas of the device where they shouldn't be, forming a space charge that causes the device to degrade or become unstable. The term **'hot electrons effects'**, therefore, refers to device degradation or instability caused by hot electron injection.

There are four commonly encountered hot carrier injection mechanisms. [According to the 5th Edition Hitachi Semiconductor Device Reliability Handbook] These are:
1) Drain avalanche hot carrier injection
2) Channel hot electron injection
3) Substrate hot electron injection
4) Secondary generated hot electron injection.

1 - Drain avalanche hot carrier (DAHC) injection (Figure 1) - Produces the worst device degradation under normal operating temperature range. This occurs when a high voltage applied at the drain under non-saturated conditions (VD>VG) results in very high electric fields near the drain, which accelerate channel carriers into the drain's depletion region. Studies have shown that the worst effects occur when VD = 2VG.

The acceleration of the channel carriers causes them to collide with Si lattice atoms, creating dislodged electron-hole pairs in the process. This phenomenon is known as ***impact ionization***, with some of the displaced e-h pairs also gaining enough energy to overcome the electric potential barrier between the silicon substrate and the gate oxide.

Under the influence of drain-to-gate field, hot carriers that surmount the substrate-gate oxide barrier get injected into the gate oxide layer where they are sometimes trapped. This hot carrier injection process occurs mainly in a narrow injection zone at the drain end of the device where the lateral field is at its maximum.

Hot carriers can be trapped at the Si-SiO2 interface (hence referred to as 'interface states') or within the oxide itself, forming a space charge (volume charge) that

increases over time as more charges are trapped. These trapped charges shift some of the characteristics of the device, such as its threshold voltage (Vth) and its conveyed conductance (gm). The injected carriers that do not get trapped in the gate oxide become gate current. The majority of the holes from the e-h pairs generated by impact ionization, flow back to the substrate, comprising a large portion of the substrate's drift current. Excessive substrate current may therefore be an indication of hot carrier degradation. In gross cases, abnormally high substrate current can upset the balance of carrier flow and facilitate latch-up.

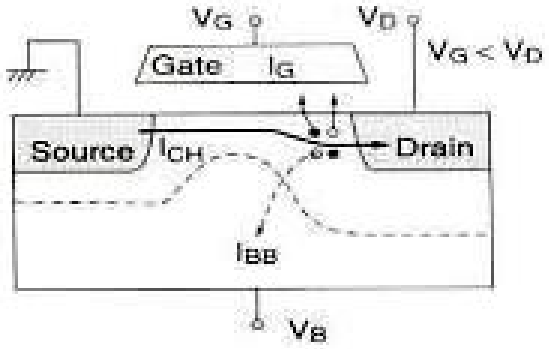


Figure 1 - DAHC injection involves impact ionization of carriers near the drain area
Image source: Hitachi Semiconductor Reliability Handbook

2 - Channel hot electron injection (CHE) (Figure 2) - This phenomenon occurs when both the gate voltage and the drain voltage are significantly higher than the source voltage, with VG≈VD. Channel carriers that travel from the source to the drain are sometimes driven towards the gate oxide even before they reach the drain because of the high gate voltage.

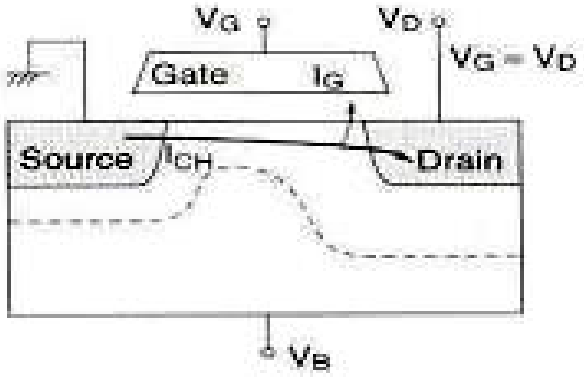


Figure 2 - CHE injection involves propelling of carriers in the channel toward the oxide even before they reach the drain area
Image source: Hitachi Semiconductor Reliability Handbook

3 - Substrate hot electron (SHE) injection (Figure 3) - Occurs when the substrate back bias is very positive or very negative, i.e., |VB|>> 0. Under this condition, carriers of one type in the substrate are driven by the substrate field toward the Si-SiO2 interface. As they move toward the substrate-oxide interface, they further gain kinetic energy from the high field in surface depletion region. They eventually overcome the surface energy barrier and get injected into the gate oxide, where some of them are trapped.

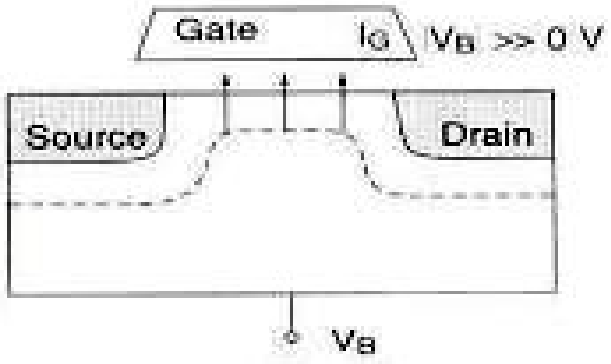


Figure 3 - SHE injection involves trapping of carriers from the substrate
Image source: Hitachi Semiconductor Reliability Handbook

4 - Secondary generated hot electron (SGHE) injection (Figure 4) - This phenomenon is based the generation of hot carriers from impact ionization involving a secondary carrier that was likewise created by an earlier incident of impact ionization. This occurs under conditions similar to DAHC, i.e., the applied voltage at the drain is high or VD>VG, which is the driving condition for impact ionization. The main difference, however, is the influence of the substrate's back bias in the hot carrier generation. This back bias results in a field that tends to drive the hot carriers generated by the secondary carriers toward the surface region, where they further gain kinetic energy to overcome the surface energy barrier.

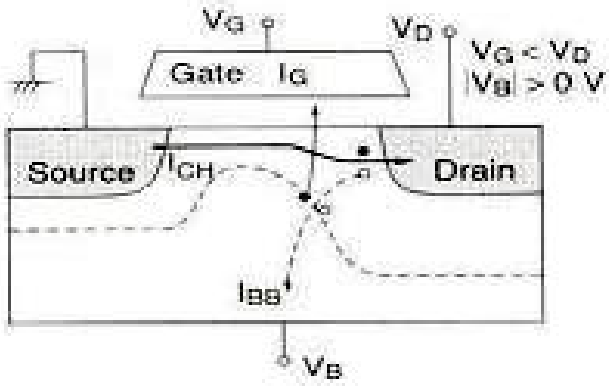


Figure 4 - SGHE injection involves hot carriers generated by secondary carriers
Image source: Hitachi Semiconductor Reliability Handbook

Hot carrier effects are brought about or aggravated by reductions in device dimensions without corresponding reductions in operating voltages, resulting in higher electric fields internal to the device. Problems due to hot carrier injection therefore constitute a major obstacle towards higher circuit densities. Recent studies have even shown that voltage reduction alone will not eliminate hot carrier effects, which were observed to manifest even at reduced drain voltages, e.g., 1.8 V.

Thus, optimum design of devices to minimize, if not prevent, hot carrier effects is the best solution for hot carrier problems. Common design techniques for preventing hot carrier effects include: 1) increase in channel lengths; 2) n+ / n- double diffusion of sources and drains; 3) use of graded drain junctions; 4) introduction of self-aligned n- regions between the channel and the n+ junctions to create an offset gate; and 5) use of buried p+ channels.

Hot carrier phenomena are accelerated by low temperature, mainly because this condition reduces charge de-trapping. A simple acceleration model for hot carrier effects is as follows:

$$AF = R2 / R1$$
$$AF = e^{([Ea/k]\,[1/T_1 - 1/T_2] + C\,[V_2 - V_1])}$$

where:
AF = acceleration factor of the mechanism;
R1 = rate at which the hot carrier effects occur under conditions V1 and T1;
R2 = rate at which the hot carrier effects occur under conditions V2 and T2;
V_1 and V_2 = applied voltages for R1 and R2, respectively;
T_1 and T_2 = applied temperatures (deg K) for R1 and R2, respectively;
Ea = -0.2 eV to -0.06 eV; and
C = a constant.

B. Electromigration & Self Heat

Nanometer designs contain millions of devices and operate at very high frequencies. The current densities (current per cross-sectional area) in the signal lines and power are consequently high and can result in either signal or power electromigration problems. The electron movement induced by the current in the metal power lines causes metal ions to migrate. That phenomenon of transport of mass in the path of a DC flow, as in the metal power lines in the design, is termed power electromigration. There are two types of electromigration. Uni-Directional, for example power and static signals and Bi-Directional, for example clocks and other switching signals. The most critical is the Uni-Directional electromigration type since the electron 'erosion' move constantly in one direction and can cause signal line failure. The power electromigration effect is harmful from the point of view of design reliability, since the transport of mass can cause open circuits, or shorts, to neighboring wires.

Electromigration is actually not a function of current, but a function of current density. It is also accelerated by elevated temperature. Thus, electromigration is easily observed in Al metal lines that are subjected to high current densities at high temperature over time. (Figures 4, 5)

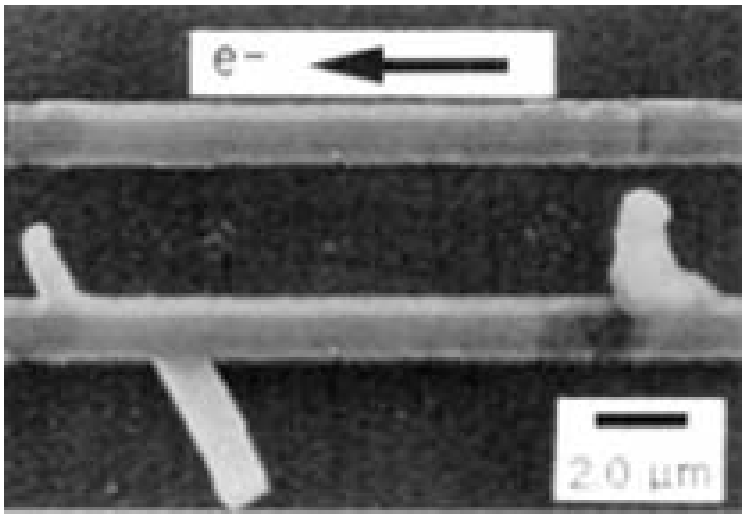


Figure 4 - Electromigration Effect – Short Circuit
Image: Computer Simulation Laboratory

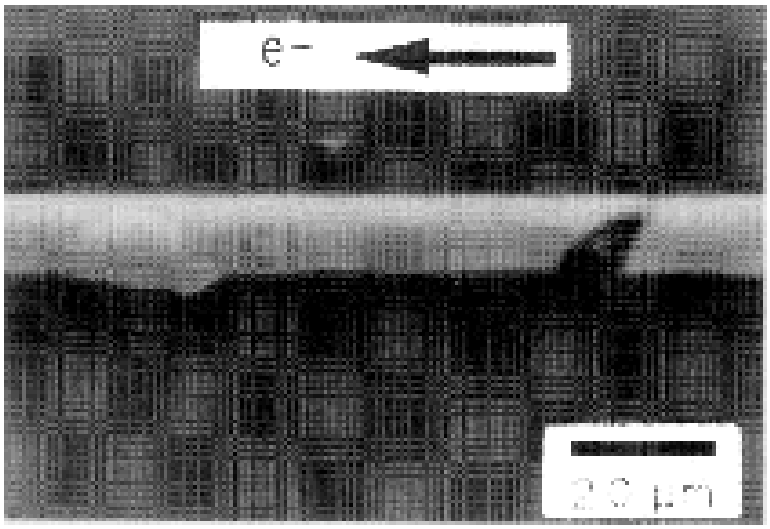


Figure 5- Electromigration Effect – Open Circuit
Image: Computer Simulation Laboratory

The higher current density around the void results in localized heating that further accelerates the growth of the void, which again increases the current density. The cycle continues until the void becomes large enough to cause the metal line to fuse open.

Electromigration may be modeled by the following equation, which is known as Black's Equation:

$$t_{50} = CJ^{-n}e^{(Ea/kT)}$$

where:
t50 = the median lifetime of the population of metal lines subjected to electromigration;
C = a constant based on metal line properties;
J = the current density;
n = integer constant from 1 to 7; many experts believe that n = 2;
T = temperature in deg K;
k = the Boltzmann constant; and
Ea = 0.5 - 0.7 eV for pure Al.

Electromigration failures take time to develop, and are therefore very difficult to detect until it happens. Thus, the best solution to electromigration problems is to prevent them from taking place.

Electromigration can be prevented by: 1) proper design of the device such that the current densities in all parts of the circuit are practically limited; 2) increasing of the grain sizes of the metal lines such that these become comparable to their widths (whereby bamboo structure is achieved); and 3) good selection and deposition of the passivation or thin films placed over the metal lines in order to limit extrusions caused by electromigration.

C. Oxide Breakdown *

Oxide Breakdown (Figure 6) is the destruction of an oxide layer (usually silicon dioxide or SiO2) in a semiconductor device. Oxide layers are used in many parts of the device: as gate oxide between the metal and the semiconductor in MOS transistors, as dielectric layer in capacitors, as inter-layer dielectric to isolate conductors from each other, etc. Oxide breakdown is also referred to as 'oxide rupture' or 'oxide punchthrough'.

Oxide breakdown has always been of serious reliability concern in the semiconductor industry because of the continuous trek towards smaller and smaller devices. As other features of the device are scaled down, so must oxide thickness be reduced. Oxides become more vulnerable to the voltages fed into the device as they get thinner. The thinnest oxide layers today are already less than 50 angstroms thick. An oxide layer can break down instantaneously at 8-11 MV per cm of thickness, or 0.8-1.1 V per angstrom of thickness.

Oxide breakdowns may be classified as one of the following:
1) EOS/ESD-induced dielectric breakdown
2) Early-life dielectric breakdown
3) Time-dependent dielectric breakdown (TDDB).

The first classification is self-explanatory, referring merely to oxide destruction due to the application of excessive voltage or current to the device (see ESD section).

Early-life and time-dependent dielectric breakdowns are technically the same failure mechanism, except that the former involves a breakdown that occurs early in the life of the device (say, within the first 2 years of normal operation), while the latter involves a breakdown that occurs after a longer time of use (mainly in the 'wear-out' stage). Both categories involve destruction of the dielectric while under normal bias or operation.

1 - Early life and time-dependent dielectric breakdowns are primarily due to the presence of weak spots within the dielectric layer arising from its poor processing or uneven growth. These weak spots or dielectric defects may be caused by:

1) The presence of mobile sodium (Na) ions in the oxide
2) Radiation damage
3) Contamination, wherein particles or impurities are trapped on the silicon prior to oxidation
4) Crystalline defects in the silicon such as stacking faults and dislocations.

The risk of dielectric breakdown generally increases with the area of the oxide layer, since a larger area means the presence of more defects and greater exposure to contaminants. The worse cases of oxide defects are the ones that result in early life dielectric breakdowns. It must be pointed out, however, that even very high quality oxides can suffer breakdown with time, especially in the 'wear-out' period of its lifetime. This latter case is the classic 'TDDB' mechanism.

2 – Time-Dependent Dielectric Breakdown (TDDB) Previous studies have shown that SiO_2 TDDB is a charge injection mechanism, the process of which may be divided into 2 stages - the build-up stage and the runaway stage.

During the build-up stage, charges invariably get trapped in various parts of the oxide as current flows in the oxide. The trapped charges increase in number with time, forming high electric fields (electric field = voltage/oxide thickness) and high current regions along the way. This process of electric field build-up continues until the runaway stage is reached.

During the runaway stage, the sum of the electric field built up by charge injection and the electric fields applied to the device exceeds the dielectric breakdown threshold in some of the weakest points of the dielectric. These points start conducting large currents that further heat up the dielectric, which further increases the current flow. This positive feedback loop eventually results in electrical and thermal runaway, destroying the oxide in the end. The runaway stage happens in a very short period of time.

The presence of defects in the dielectric greatly reduces the time needed to transition from the build-up to the runaway stage. These defects actually have the effect of 'thinning' down the oxide where they are located, since they are occupying space that should have been occupied by the dielectric. The effective electric field is higher in these thinned-out areas compared to defect-free areas for any given voltage. This is why it takes a lower voltage and shorter time to break down the dielectric at its defect points.

There are many lifetime equations used in the industry today to model the reliability of an oxide layer. One of the simplest, however, can be seen in www.semicon.toshiba.co.jp. According to this site, TDDB may be modelled by:

$$\mathbf{T_f = Ae^{(-BV)}}$$

where:
Tf = the time to failure;
A = a constant;
V = the voltage applied across the dielectric layer; and
B = a voltage acceleration constant that depends on the properties of the oxide.

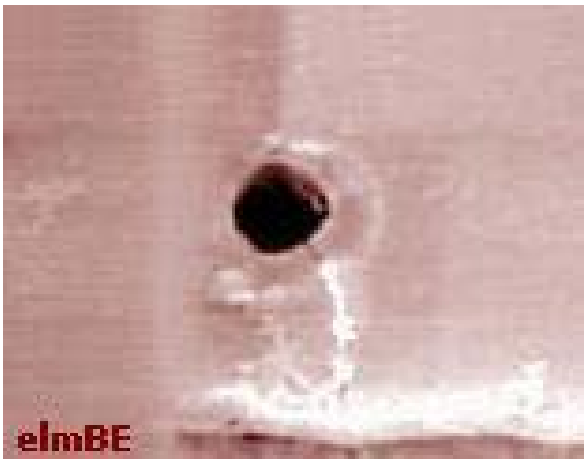


Figure 6 - Photo of an ESD-induced Oxide Breakdown
Image Source: Toshiba

D. P transistor degradation (NBTI - Negative Bias Temperature Instability)

With the continuous shrinking of the transistor dimensions, generation of interface traps during negative bias temperature instability (NBTI) stress in p-MOS transistors has become one of the most critical reliability issues that ultimately determine the lifetime of CMOS devices. It is important to categorize NBTI to two (2) effect types, static (SNBTI) and Dynamic (DNBTI). Static NBTI phenomenon can be described as a constant negative bias is applied to the gate electrode of a p-MOS transistor at high temperatures with S/D grounded. Dynamic NBTI occurs during the operation of a p-MOSFET in a CMOS inverter, when the applied gate bias is switching between "high" and "low" voltages, while the drain bias is alternating between "low" and "high" voltages, correspondingly. This creates dynamic stress conditions. The conventional static NBTI measurement has neglected the passivation effects of the interface traps during the operation of p-MOSFETs in digital CMOS circuits, and therefore overestimates the degradations of p-MOS devices. A large portion of the interface traps generated under the NBTI stressing, corresponding to p-MOSFET operating condition of the "high" output state in a CMOS inverter, are passivated electrically when the gate to drain voltage switches to positive corresponding to the p-MOSFET operating condition of the "low" output state in a CMOS inverter. As a result dynamic NBTI (DNBTI) effect greatly prolongs the lifetime of p-MOSFETs operating in a digital circuit, while the conventional static NBTI (SNBTI) measurement underestimates the p-MOSFET lifetime. Although electric passivation (EP) of interface traps has been reported before in MOSFETs during hot-carrier stress, its effects on NBTI and device lifetime have not been investigated. Due to EP effect, the lifetime of p-MOSFETs under DNBTI stress corresponding to a realistic operation condition in a digital circuit is approximately one order of magnitude longer than that under conventional SNBTI stress.

E. Latchup

Latchup is a known reliability issue in nanometer design arena. Latchup may be defined as the creation of a low-impedance path between power supply rails as a result of triggering a parasitic device. In this condition, excessive current flow is possible, and a potentially destructive situation exists. After even a very short period of time in this condition, the device in which it occurs can be destroyed or weakened; and potential damage can occur to other components in the system. Latchup may be caused by a number of triggering factors, to be discussed below—including overvoltage spikes or transients, exceeding maximum ratings, and incorrect power sequencing. (Figure - 7, 8)

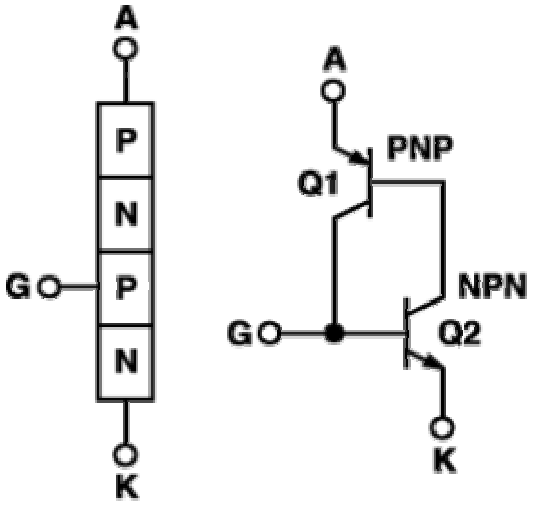


Figure 7 - Transistor equivalent of an SCR.

Image Source: Analog Devices

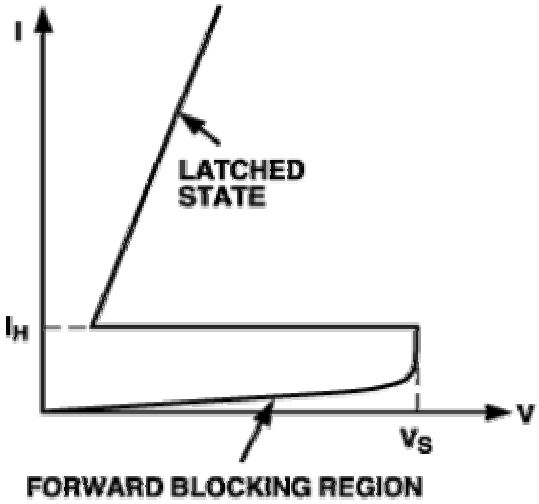


Figure 8 -Current- voltage characteristic of an SCR.

Image Source: Analog Devices

An SCR is a normally *off* device in a "blocking state", in which negligible current flows. Its behavior is similar to that of a forward-biased diode, but conducts from anode, A, to cathode, K, only if a control signal is applied to the gate, G. In its normally off state, the SCR presents a high impedance path between supplies. When triggered into its conducting state as a result of excitation applied to the gate, the SCR is said to be "latched". It enters this state as a result of current from the gate injected into the base of Q2, which causes current flow in the base-emitter junction of Q1. Q1 turns on causing further current to be injected into base of Q2. This positive-feedback condition ensures that both transistors saturate; and the current flowing through each transistor ensures that the other remains in saturation. **

When thus latched, and no longer dependent on the trigger source applied to the gate (G), a continual low-impedance path exists between anode and cathode. Since the triggering source does need not be constant, it could simply be a spike or a glitch; removing it will not turn off the SCR. As long as the current through the SCR is sufficiently large, it will remain in its latched state. If, however, the current can be reduced to a point where it falls below a holding-current value, I_H, the SCR switches off. Figure 1b shows the current-to-voltage transfer function for an SCR. In order to bring the device out of its conductive state, either the voltage applied across the SCR must be reduced to a value where each transistor turns off, or the current through the SCR must be reduced below its holding current. **

A CMOS switch channel effectively consists of PMOS and NMOS devices connected in parallel; control signals to turn it off and on are applied via drivers. Since all these MOS devices are located close together on the die, it is possible that, with appropriate excitation, parasitic SCR devices may conduct — a form of behavior possible with any CMOS circuit. Figure 2 illustrates a simplified cross section showing two CMOS structures, one PMOS and one NMOS; these could be connected together as an inverter or as the switch channel. The parasitic transistors responsible for latch-up behavior, Q1 (vertical PNP) and Q2 (lateral NPN) are also shown in figure 9. **

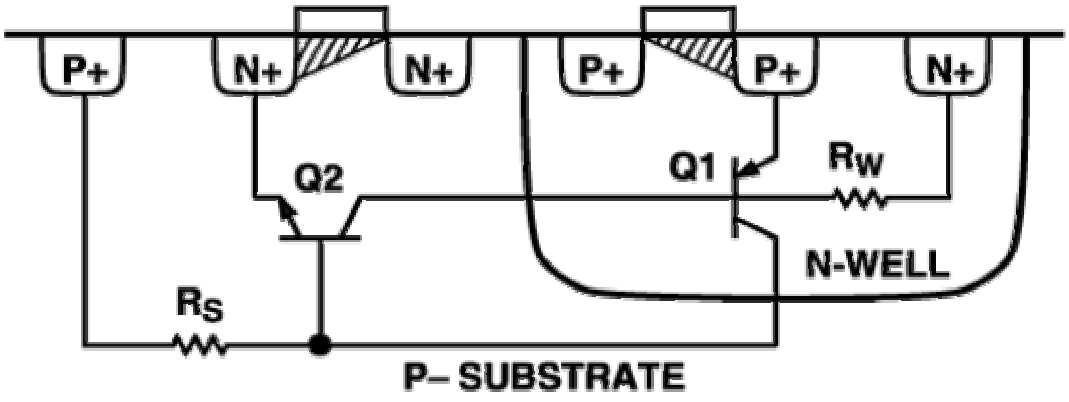


Figure – 9: The parasitic transistors responsible for latch-up behavior, Q1 (vertical PNP) and Q2 (lateral NPN)
Image Source: Analog Devices

Having described the architecture that makes latchup possible, we now discuss the events that can trigger such behavior. SCR latchup can occur through one of the following mechanisms.

- *Supply voltages exceeding the absolute maximum ratings.* These ratings in the data sheet are an indication of the maximum voltage that can safely be applied to the switch. Anything in excess may result in breakdown of an internal junction and hence damage to the device. In addition, operation of the switch under conditions close to the maximum ratings may degrade long-term reliability. It is important to note that these ratings apply at all times, including when the switch is being powered on and off. The triggering mode could result from transients on supply rails.
- *Input/output pin voltage exceeding either supply rail by more than a diode drop.* This could occur as a result of a fault on a channel or input—if a part of the system is powered on prior to the supplies being present at the switch (or similar CMOS components in the system). The powered part of the circuit would be sending signals to other devices in the design which may not be able to handle the voltage levels presented. The resulting voltage levels could exceed the maximum rating of the device, and possibly result in latchup. Again, this could occur as a result of spikes or glitches on input or output channels.
- *Poorly managed multiple power supplies.* Switches that have multiple power supplies tend to be more susceptible to latchup resulting from improper power-supply sequencing. Such switches usually have two analog supplies, V_{DD} and V_{SS}, and a digital supply, V_L. In some cases, when the digital supply is applied prior to the other supplies, it may be possible for maximum ratings to be exceeded and the device to enter a latchup state. In general, for those devices that require an external digital supply, V_L, we recommend that when power is being applied to and removed from the device, care should be taken to ensure the maximum ratings are not exceeded.

When any of the triggering mechanism described above occur, the parasitic SCR structure of Figure 1a may begin to conduct, producing a low impedance state between power supply rails. If there is no current limit mechanism on the supplies, excessive current will flow through this SCR structure and through the switch. This could destroy the switch and other components if allowed to persist. With high current levels, a device would not have to remain in a latch-up state for very long; even very brief latchup can result in permanent damage if current is not limited. **

Latchup can be classified into two generalized categories: internal and external. Internal latchup occurs when circuits are not connected to I/O pads, whereas external latchup occurs when circuits or injection sources are connected to pads. With the aggressive scaling of CMOS, SOI, and BICMOS technologies, the ground rules are being reduced to allow greater numbers of transistors in a given die size. The reduction in the ground rules leads to smaller N+(PWELL)/P+(NWELL) spacing, which in turn increases the parasitic NPN and PNP betas, lowering the trigger currents/voltages and the holding voltage. With the introduction of triple well bulk CMOS technologies, new NPNs and PNPs are

formed that will need to be considered beyond the classical NPNs and PNPs formed in a dual well bulk CMOS technology.

The reduction in the ground rules leads to smaller N+(PWELL)/P+(NWELL) spacing, which in turn increases the parasitic NPN and PNP betas, lowering the trigger currents/voltages and the holding voltage. With the introduction of triple well bulk CMOS technologies, new NPNs and PNPs are formed that will need to be considered beyond the classical NPNs and PNPs formed in a dual well bulk CMOS technology.

Latchup Prevention

Fab/Design Approaches (Figure 10)

1. Reduce the gain product b1 x b1
    - move n-well and n+ source/drain farther apart increases width of the base of Q2 and reduces gain beta2 > also reduces circuit density
    - buried n+ layer in well reduces gain of Q1
2. Reduce the well and substrate resistances, producing lower voltage drops
    - higher substrate doping level reduces Rsub
    - reduce Rwell by making low resistance contact to GND
    - guard rings around p- and/or n-well, with frequent contacts to the rings, reduces the parasitic resistances.

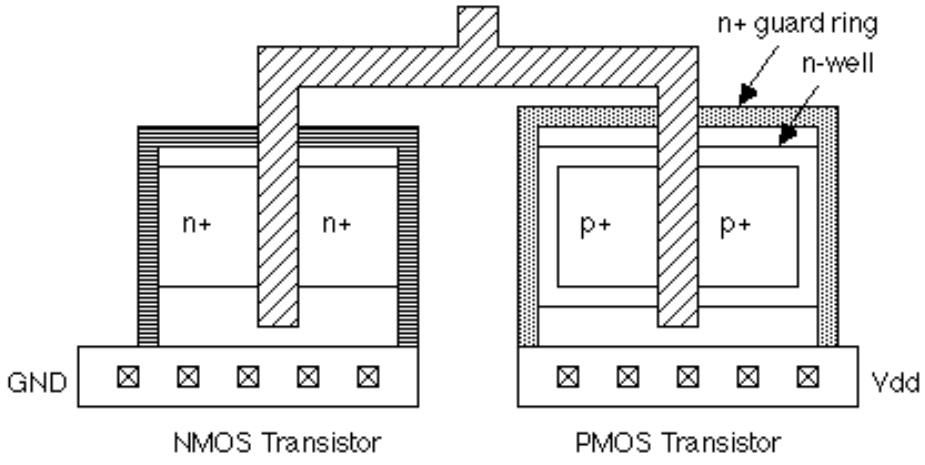


Figure 10 – Preventing Latchup using CMOS guard-rings

Systems Approaches

1. Make sure power supplies are off before plugging a board. A "hot plug in" of an un-powered circuit board or module may cause signal pins to see surge voltages greater than 0.7 V higher than Vdd, which rises more slowly to is peak value. When the chip comes up to full power, sections of it could be latched.
2. Carefully protect electrostatic protection devices associated with I/O pads with guard rings. Electrostatic discharge can trigger latchup. ESD enters the circuit through an I/O pad, where it is clamped to one of the rails by the ESD protection circuit. Devices in the protection circuit can inject minority carriers in the substrate or well, potentially triggering latchup.
3. Radiation, including x-rays, cosmic, or alpha rays, can generate electron-hole pairs as they penetrate the chip. These carriers can contribute to well or substrate currents.
4. Sudden transients on the power or ground bus, which may occur if large numbers of transistors switch simultaneously, can drive the circuit into latchup. Whether this is possible should be checked through simulation.

F. ESD (Electrostatic Discharge)

An ESD event is the transfer of energy between two bodies at different electrostatic potentials, either through contact or via an ionized ambient discharge (a spark). This transfer has been modeled in various standard circuit models for testing the compliance of device targets. The models typically use a capacitor charged to a given voltage, and then some form of current-limiting resistor (or ambient air condition) to transfer the energy pulse to the target. ESD protection devices attempt to divert this potentially damaging charge away from sensitive circuitry and protect the system from permanent damage, as shown in Figure - 11.

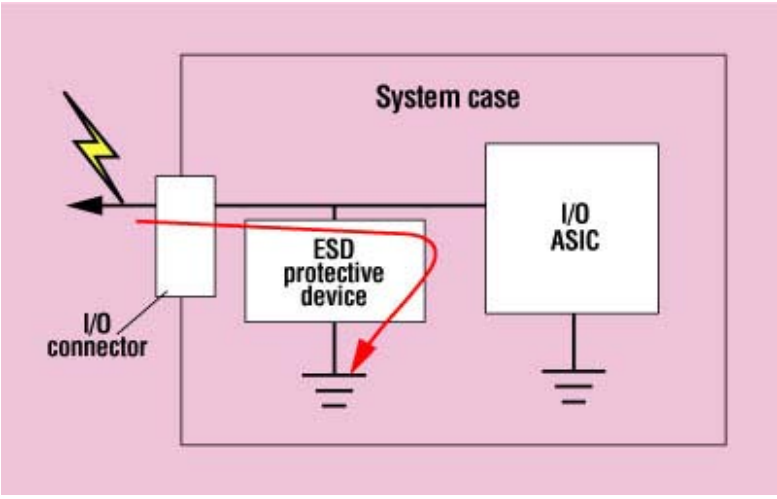
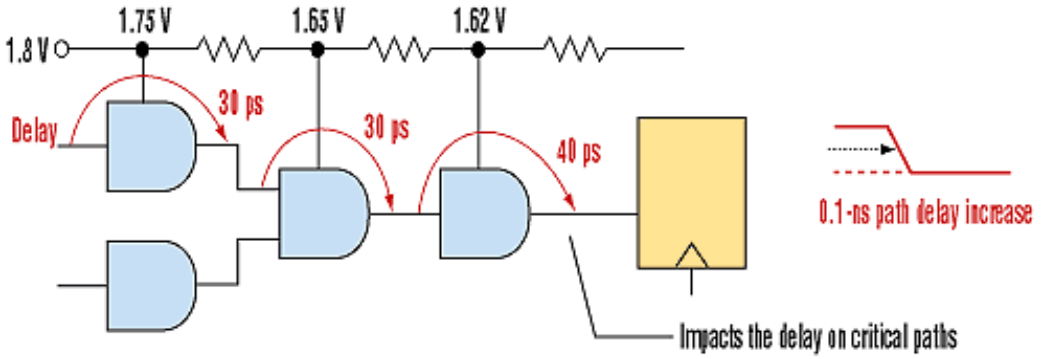


Figure – 11: ESD protection devices attempt to divert a potentially damaging charge away from sensitive circuitry and protect the system from permanent damage.

An integrated circuit (IC) connected to external ports is susceptible to damaging electrostatic discharge (ESD) pulses from the operating environment and peripherals. The same ever-shrinking IC process technology that enables such high-port interconnect data rates can also suffer from higher ESD susceptibility because of its smaller fabrication geometry. Additional external protection devices can violate stringent signaling requirements, leaving design engineers with the need to balance performance and reliability. Traditional methods of shunting ESD energy to protect ICs involves devices such as zener diodes, metal oxide varistors (MOVs), transient voltage suppression (TVS) diodes, and regular complementary metal oxide semiconductor (CMOS) or bipolar clamp diodes. However, at the much higher data rates of USB 2.0, IEEE 1394, and digital visual interface (DVI), the parasitic impedance of traditional protection devices can distort and deteriorate signal integrity.

G. Voltage Drop

One of the major issues of signal integrity and reliability is the IR drop effect. IR drop is a signal integrity effect caused by wire resistance and current drawn from the power and ground grids. If the wire resistance is too large or the cell current is higher than predicted, an undesirable voltage drop may happen. The voltage drop causes the voltage supplied to the affected cells to be lower than required, which leads to larger gate and signal delays (TPD), which in turn can cause timing discrepancies in the signal paths as well as clock skew. Voltage drop on power and ground grids can also affect the noise margins and compromises the signal integrity of the design. Therefore special attention should be taken to resolve the IR drop effects during post-layout phase.



Voltage drops through a chain of logic gates can severely influence SoC timing closure. Here, 0.1 ns of additional delay, which may not have been accounted for in static timing analysis, is imposed by an aggregate drop of less than 200 mV.

Figure -12: Voltage Drop Case
Image: Magma Design Automation

Nanometer designs are extremely susceptible to voltage drop because power and ground wire resistivity increases with decreasing geometries, while the overall power supply voltage decreases. Gate delays increase non-linearly as voltage at gates decrease. The result is poor performance and increased noise susceptibility. Furthermore, gates with different voltage levels communicating with each other across the chip can propagate erroneous data, causing a malfunction. The power grid must be robust enough to prevent reliability problems from EM effects without costly over-design. In nanometer design, it is essential to understand power issues early in the design cycle, and in detail, to minimize power consumption and to address considerations such as temperature, leakage, return path, etc.

H. Soft Errors

Soft errors are transient faults that occur in VLSI circuits due to external radiation and affect the logic states of sensitive nodes. They generally occur from nuclear decay of packaging materials or atmospheric particles accelerated towards the earth by cosmic rays. Neutron radiation interferes with charges held in sensitive nodes in circuits causing soft errors - or SEU (Single event upset) and they generally affect storage elements such as memory, latches and registers. Logic cores and FPGAs are known to be much less sensitive to soft errors than memories but the operating frequency increase, the geometry shrinking and the power supply reduction tend to drastically raise the soft error sensitivity of these devices.

Due to aggressive scaling down of the power supply voltage (Vdd), reduction in the minimum feature size, and the use of flip-chip packaging, the sensitivity of a circuit to single event upset increases. Vdd is the main sensitivity factor, decreasing the node charge to a critical level. High clock rates foster soft error vulnerability in logic parts: the probability to latch a single event upset is becoming more and more significant. The metric for soft errors is well defined. A FIT (failure in time) is one soft error for 10e9 hours / device. Recent radiation tests have shown that for a 1Gb memory in 0.25µm, the current average for Soft Error Rate (SER) is one error per week. With 0.13µm, usual memory specifications require a design under 1000 FIT: with 50 devices, this specification means 1 error per week again. Radiation tests must be performed to get accurate data for IC SER sensitivity.

Soft Errors Protection - To quantify the sensitivity of chips to soft errors, chip manufacturers can run radiation testing on their latest IC products. The next step would be to estimate the soft error rate during the design cycle in order to reduce soft error sensitivity before sending chips to production. But tests and estimates are not enough for chip designers; they will have to protect the designs against soft errors. We have to implement protection techniques like ECC and other efficient solutions to ensure an acceptable level of robustness against soft errors.

3. Industry's Approaches and Solutions

Reliability is a critical concern for the manufacturers and users of integrated circuits (ICs). Developing practical, affordable techniques to ensure reliability has always been challenging. It is even more challenging as problems with scaling require the introduction of new materials, new operating regions and the reduction of reliability margins.

Successful nanometer design requires reliability estimations built-in flow. Nanometer routers must be reliability aware, taking physical effects such as SI into consideration on-the-fly. They must also be manufacturing-aware, with capabilities such as variable-spacing and variable-width routes to support copper, CMP, and subwavelength processes. Silicon integrity and reliability have become first-priority effects for successful tapeout. For the past decade the EDA industry has provided extensive solutions for reliability phenomenon, yet significant improvement is needed in order to efficiently provide a unified solution. Since reliability and signal integrity issues are directly connected, it creates great difficulties to achieve a comprehensive solution within EDA tool.

For Example, Synopsys Galaxy™ Signal Integrity is a new and complete signal integrity solution within the Galaxy Design Platform that addresses crosstalk delay, noise (glitch), IR (voltage) drop and electromigration. Galaxy SI provides designers with comprehensive prevention, analysis and sign-off, speeding SI closure for 130-nanometer designs and below.

Cadence offers design for manufacturing (DFM) technologies enable customers to verify and optimize layouts in digital and custom IC designs, while providing a reliable way to achieve manufacturing sign-off before tape-out. Complex combinations of voltage drop, signal cross-coupling, and circuit parasitics interact to stretch design cycles and force re-spins. Process variations across the die, wafer, and batch affect yield, performance, and reliability. In addition, burgeoning volumes of parasitic data strain storage facilities and choke chip analysis software.

4. Conclusion

This article briefly describes major nanometer reliability issues. Nanometer design implementation places extraordinary demands on design teams. In order to provide an effective solution a built-in reliability technology has to be implemented within design tools as an integral part of the design flow. Current EDA tools provide partial solution for nanometer reliability issues. EDA technology is constantly evolving with developments in IC technology that continue to break the limits with each new technology, and producing chips that are faster, more powerful and smaller than previously imaginable. With each change of IC process technology, there is also the need for new design methodologies and tools. The main trick is to keep EDA technology constantly 'in sync' with physical process advancement.

References

* Taken from Analog Devices
** Taken from SemiCon Fareast

1. Mahapatra, A. M. Ionescu and K. Banerjee, "Quasi-Analytical Modeling of Drain Current and Conductances of Single Electron Transistors with MIB," *32nd European Solid-State Device Research Conference (ESSDERC)*, Florence, Italy, September 24-26, 2002. (to appear)
2. K. Banerjee and A. Mehrotra, "Analysis of On-Chip Inductance Effects for Distributed RLC Interconnects," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems,* August 2002. (to appear)
3. C. Ito, K. Banerjee and R. W. Dutton, "Analysis and Design of Distributed ESD Protection Circuits for High-Speed Mixed-Signal and RF ICs," *IEEE Transactions on Electron Devices*, August 2002. (to appear)
4. K. Banerjee and A. Mehrotra, "Power Dissipation Issues in Interconnect Performance Optimization for Sub-180 nm Designs," *IEEE Symposium on VLSI Circuits*, Honolulu, HI, June 13-15, 2002. (to appear)
5. A. M. Ionescu, M. J. Declercq, S. Mahapatra, K. Banerjee and J. Gautier, "Few Electron Devices: Towards Hybrid CMOS-SET Integrated Circuits," *39th ACM Design Automation Conference (DAC)*, New Orleans, LA, June 10-14, 2002. (to appear)
6. S. Mahapatra, A. M. Ionescu and K. Banerjee, "A Quasi-Analytical SET Model for Few Electron Circuit Simulation," *IEEE Electron Device Letters*, June 2002. (to appear)
7. S. Mahapatra, A. M. Ionescu, K. Banerjee and M. J. Declercq, "A SET Quantizer Circuit Aiming at Digital Communication System," *IEEE International Symposium on Circuits and Systems (ISCAS)*, Scottsdale, AZ, May 26-29, 2002. (to appear)
8. A. M. Ionescu, M. J. Declercq, K. Banerjee and S. Mahapatra, "Teaching Microelectronics in the Silicon ICs Showstopper Zone: A Course on Ultimate Devices and Circuits: Towards Quantum Electronics," *4th European Workshop on Microelectronics Education (EWME)*, Baiona, Mancomunidad de Vigo, Spain, May 23-24, 2002. (to appear)
9. S. Mahapatra, A. M. Ionescu, K. Banerjee and M. Declercq, "A SET based Quantizer Circuit for Digital Communications," *IEE Electronics Letters*, May 2002. (to appear)
10. K-H. Oh, C. Duvvury, K. Banerjee and R. W. Dutton, "Investigation of Gate to Contact Spacing Effect on ESD Robustness of Salicided Deep Submicron Single Finger NMOS Transistors," *40th IEEE Annual International Reliability Physics Symposium (IRPS)*, Dallas, TX, April 8-11, 2002, pp. 148-155.
11. S. Im, K. Banerjee and K. E. Goodson, "Modeling and Analysis of Via Hot Spots and Implications for ULSI Interconnect Reliability," *40thIEEE Annual International Reliability Physics Symposium (IRPS)*, Dallas, TX, April 8-11, 2002, pp. 336-345.
12. A. M. Ionescu, V. Pott, R. Fritschi, K. Banerjee, M. J. Declercq, Ph. Renaud, C. Hibert, Ph. Fluckiger and G-A. Racine, "Modeling and Design of a Low-Voltage SOI Suspended-Gate MOSFET (SG-MOSFET) with a Metal-Over-Gate-

Architecture," *IEEE International Symposium on Quality Electronic Design (ISQED)*, San Jose, CA, March 18-21, 2002, pp. 496-501.
13. K. Banerjee and A. Mehrotra, "Inductance Aware Interconnect Scaling," *IEEE International Symposium on Quality Electronic Design (ISQED)*, San Jose, CA, March 18-21, 2002, pp. 43-47.
14. P. Dainesi, A. M. Ionescu, L. Thevenaz, K. Banerjee, M. J. Declercq, Ph. Robert, Ph. Renaud, Ph. Fluckiger, C. Hibert and G-A. Racine, "3-D Integrable Optoelectronic Devices for Telecommunications ICs," *IEEE International Solid State Circuits Conference (ISSCC)*, San Francisco, CA, February, 4-6, 2002, pp. 360-361.
15. T-Y Chiang, K. Banerjee and K. C. Saraswat, "Analytical Thermal Model for Multilevel VLSI Interconnects Incorporating Via Effect," *IEEE Electron Device Letters*, Vol. 23, No. 1, 2002, pp. 31-33.
16. R. H. Dennard, F. H. Gaensslen, H.-N. Yu, V. L. Rideout, E. Bassous, and A. R. LeBlanc, "Design of Ion-Implanted MOSFETs with Very Small Physical Dimensions," *IEEE J. Solid State Circuits* **SC-9,** 256–268 (1974).
17. M. Lenzlinger and E. H. Snow, "Fowler–Nordheim Tunneling into Thermally Grown SiO_2," *J. Appl. Phys.* **40,** 278–283 (1969).
18. I. C. Chen, S. E. Holland, K. K. Young, C. Chang, and C. Hu, "Substrate Hole Current and Oxide Breakdown," *Appl. Phys. Lett.* **49,** 669–671 (1986).
19. J. H. Stathis and D. J. DiMaria, "Reliability Projection for Ultra-Thin Oxides at Low Voltage," *IEDM Tech. Digest*, pp. 167–170 (1998).
20. B. Hoeneisen and C. A. Mead, "Fundamental Limitations in Microelectronics—I. MOS Technology," *Solid-State Electron.* **15,** 819–829 (1972).
21. K. Nagai and Y. Hayashi, "Static Characteristics of 2.3-nm Gate-Oxide MOSFETs," *IEEE Trans. Electron Devices* **35,** 1145–1147 (1988).
22. H. S. Momose, M. Ono, T. Yoshitomi, T. Ohguro, S. Nakamura, M. Saito, and H. Iwai, "Tunneling Gate Oxide Approach to Ultra-High Current Drive in Small Geometry MOSFETs," *IEDM Tech. Digest*, pp. 593–596 (1994).
23. T. H. Ning, "Silicon Technology Directions in the New Millennium," *Proceedings of the International Reliability Physics Symposium*, 2000, pp. 1–6.
24. T. Ghani, K. Mistry, P. Packan, S. Thompson, M. Stettler, S. Tyagi, and M. Bohr, "Scaling Challenges and Device Design Requirements for High Performance Sub-50 nm Gate Length Planar CMOS Transistors," *Symposium on VLSI Technology*, *Digest of Technical Papers*, 2000, pp. 174–175.
25. M. Hirose, M. Koh, W. Mizubayashi, H. Murakami, K. Shibahara, and S. Miyazaki, "Fundamental Limit of Gate Oxide Thickness Scaling in Advanced MOSFETs," *Semicond. Sci. Technol.* **15,** 485–490 (2000).
26. D. A. Muller, T. Sorsch, S. Moccio, F. H. Baumann, K. Evans-Lutterodt, and G. Timp, "The Electronic Structure at the Atomic Scale of Ultrathin Gate Oxides," *Nature* **399,** 758–761 (1999).