

Single-Event Effects in Advanced CMOS Devices

A. H. Johnston
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA USA 91109

Abstract - Despite the decrease in critical charge that occurs for highly scaled CMOS devices, laboratory test data has shown that single-event upset rates are actually somewhat better for scaled devices compared to older devices with larger feature size. This paper discusses single-event upset in memories and microprocessors, along with the mechanisms that cause SEU effects. Predictions for future devices are made using the Semiconductor Industry Roadmap along with modeling and radiation test results.

I. INTRODUCTION

Single-event upset on microelectronics in space was first reported by Binder, Smith and Holman in 1975 [1]. Improvements in semiconductor device design and performance in the ensuing 27 years have resulted in mainstream devices with feature size approaching 0.1 μm , and more than 10^8 transistors per chip. No one could have predicted such a dramatic evolution in this field. Along with these increases in density and performance, there has been increasing concern about single-event upset in space, as well as in terrestrial applications.

The first efforts to predict how device scaling would impact single-event upset in space were done by Petersen, et al. in 1982 [2]. Figure 1, taken from that work, has been widely used, and has turned out to be a surprisingly good predictor of critical charge over several generations of devices.

Initially the issues for space applications were overcome by designing special hardened circuits. For example, adding resistors to SRAMs slows down the response, improving the SEU hardness by large factors [3]. Although a limited number of hardened circuits is available, their performance lags commercial technology by a considerable amount, which is a serious restriction for spacecraft designers.

The research in this paper was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration (NASA).

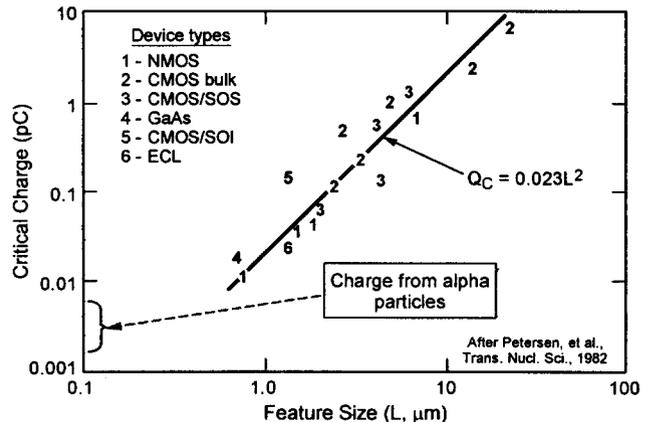


Figure 1. Critical charge vs. feature size (after Petersen, et al. [2].)

Consequently there is considerable interest in the application of state-of-the-art commercial integrated circuits in space, even though spurious responses due to SEU effects have to be taken into account in overall system design.

Single-event upset from alpha particles that are produced from radioactive decay became an issue for commercial manufacturers in 1979 [4]. As devices have been scaled to even smaller feature sizes, the semiconductor industry has been forced to deal with upset from neutrons, produced by the interaction of cosmic rays in the upper atmosphere, that are present in significant numbers at the earth's surface. Neutrons and protons with energies above 30 MeV cause very similar reactions with silicon. Thus, the concern – and moderate hardening of commercial semiconductors to atmospheric neutrons – is directly applicable to upset from energetic protons in space. However, upset from highly energetic cosmic rays is not considered by the commercial semiconductor industry. The charge deposited by cosmic rays is considerably higher than charge produced by either alpha particles or neutrons, and thus all high-density integrated circuits are susceptible to upset from heavy ions in space.

This paper discusses the effects of device scaling on upset sensitivity in typical space applications. The first section discusses basic scaling considerations,

based on design requirements for various circuit applications, along with predicted improvements and changes from the Semiconductor Industry roadmap [5]. The next section discusses charge generation and collection, along with basic considerations for creating upsets in highly scaled devices. Section IV includes recent experimental results for state-of-the-art commercial devices, benchmarking these results for devices with feature sizes between 0.13 and 0.35 μm . The last section puts these results into perspective, and makes predictions about the way in which future changes in technology and scaling are likely to affect single-event susceptibility in space.

Although total dose damage is potentially an issue for some space applications, charge trapping is so small in the thin gate oxides used in highly scaled devices that gate threshold shifts are no longer of much concern. Total dose effects in trench isolation regions is potentially important, but tests of advanced microprocessors have shown little effect at levels below 100 krad(Si). For these reasons, total dose effects have not been included in this paper.

II. SCALING RELATIONSHIPS

A. General Issues for Conventional CMOS

Device scaling for CMOS is a complex problem, which requires tradeoff of many different parameters [6-8]. Earlier scaling predictions were done with constant scaling factors for dimensions, and reduced voltage (constant field scaling, [9]). However, threshold voltage and the subthreshold slope of the gate-voltage/drain current characteristics dictate a different approach, where the electric field in the channel region and the field across the gate oxide are allowed to increase.

Analysis of scaling effects is quite complex, and more recent work has subdivided scaling into three basic circuit applications: (1) high-performance devices, such as microprocessors, where the main overall criterion is speed; (2) low-power devices, where speed and power dissipation are both involved in establishing design tradeoffs; and (3) memories, which require a very different set of tradeoffs. These are discussed in more detail below.

High Performance Devices. For high-performance devices, power dissipation is a key issue, but the design criteria allow design of circuits that can dissipate large amounts of power (i.e., the fan-cooled packages in high-speed microprocessors). Because the device normally operates at high frequency,

standby current and gate leakage currents can be much higher than for other scaling scenarios. For space applications, more realistic limits need to be placed on power requirements, so the appropriate scaling for high-performance devices roughly follows scaling for devices used in desktop computers, which is less aggressive than applications in servers where very high power dissipation can be tolerated.

Gate leakage has become a major issue for scaled devices because of tunneling through the very thin oxides. Current projections are based on assigning 10 to 20% of the total power to gate leakage, which is a major departure from earlier scaling assumptions. Note that gate leakage does not scale with frequency, so that a significant fraction of the total power is effectively standby power..

Low Power. Low power devices place more stringent demands on standby power, and may also implement a reduced voltage mode to further reduce power during periods when maximum performance is not required. One of the main differences is the requirement for thicker gate oxides compared to high-performance devices in order to reduce gate leakage effects. Initial scaling projections for low-power applications used much lower power supply voltage compared to high-performance scaling, but more recent scaling projections use nearly the same voltages for both scenarios. The reason for this is the need for high on/off ratios for low-power applications.

Memories. Memories involve very different tradeoffs compared to speed-driven technologies, such as microprocessors. The on/off current ratio for memories needs to be about two orders of magnitude higher for memories than for logic. This requires thicker gate oxides compared to logic applications. Changes in device architecture and (for DRAMs) the design of the storage capacitor are key factors in memory evolution. Advanced DRAMs use trench capacitors with very high aspect ratios in order to minimize cell area. As discussed later, this has a major impact on single-event upset sensitivity.

Leakage currents in DRAMs must be very low in order to meet refresh rate requirements. Variations in threshold voltage occur within a large DRAM for several reasons, including statistical fluctuations in the number of dopant atoms that are present in each device. The threshold voltage variations cause large differences in leakage current, requiring considerably

lower mean leakage currents in the DRAM array in order to keep leakage currents in the more extreme part of the distribution within tolerance. Measurements of refresh rates over the array provide direct evidence of this effect.

Mainstream commercial memories are also extremely cost sensitive. Consequently, they use more conservative design approaches, and are nearly always made with bulk rather than epitaxial substrates. This makes them more susceptible to latchup than most logic-based circuits. Some advanced DRAMs are susceptible to catastrophic latchup from heavy ions.

B. Silicon-on-Insulator

Silicon-on-insulator CMOS has many potential advantages. Much has been made of the potential advantage of the very shallow silicon film thickness - typically 120 – 200 nm for partially depleted SOI. In principle this should result in far less charge collection compared to bulk/epitaxial processes where the charge collection depth is approximately 2000 nm, resulting in much lower soft error rates as well as improved SEU hardness in space. However, in partially depleted structures excess charge collection can occur because of the parasitic bipolar transistor (unless body ties are used in the process), largely negating the advantage of the reduced charge collection region [10]. This will be discussed further in Section IV.

C. The Semiconductor Roadmap

The semiconductor industry has established a “roadmap” in order to facilitate overall planning for equipment manufacturers. The roadmap is updated periodically. Table 1 shows some key properties of predictions for high-performance devices, based on the 1999 SIA Roadmap. Several things should be noted. First, there is a steady decrease in power supply voltage, which is required in order to keep power dissipation within reasonable bounds. This means that internal switching voltages for high-performance logic will continue to decrease, affecting noise margin as well as the allowable number of gates through which high-speed pulses can propagate (each transition reduces the switching amplitude somewhat, and eventually the pulses will no longer propagate). Second, the on/off ratio is rapidly decreasing, and actually changes more than any of the other parameters. It is predicted to decrease by a factor of ten during the current decade. Although this decrease

can be tolerated by high-performance devices, it is clearly at odds with the requirements for low-power devices and memories.

Table 1
Key Properties for High-Performance Devices

Year	2000	2001	2002	2003	2004	2005	2008	2011
T_{ox} (nm)	2.5	1.9	1.9	1.9	1.5	1.5	1.2	0.8
V_{dd} (V)	1.8	1.5	1.5	1.5	1.2	1.2	0.9	0.6
Gate Length (nm)	120	100	85	80	70	65	45	32
On/Off Ratio	107	94	75	58	47	38	19	9.4

III. THE UPSET PROCESS

A. Charge Generation

The charge generated from a heavy ion is usually assumed to be proportional to linear energy transfer. For particles incident on the device surface, the LET remains constant until the particle has traveled at least a factor of ten beyond the depletion region of the MOSFET drain region. This concept has worked well for older devices, and continues to be a reasonable approach for alpha particles, which have relatively small track diameters ($\approx 0.1 \mu\text{m}$). A great deal of the work done in the device community has been done using the track diameter and charge density of 5-MeV alpha particles.

However, particles in space have a very different charge-deposition structure because they have such high energies. For example, Figure 2 (after Dodd, et al. [11]) shows how the track diameter of a 1-GeV silver ion changes as it traverses a silicon region. The track diameter is so large at the surface – $8 \mu\text{m}$ – that it extends over many different devices, but produces much less charge underneath the region where it would affect a single device than one would expect using LET as a figure of merit. In contrast, the energy of a 100-MeV silver ion has a diameter near the surface of about $1.6 \mu\text{m}$. The strong difference in track structure for high-energy particles is even more important in space, where galactic cosmic rays have even higher energies with much larger track diameter.

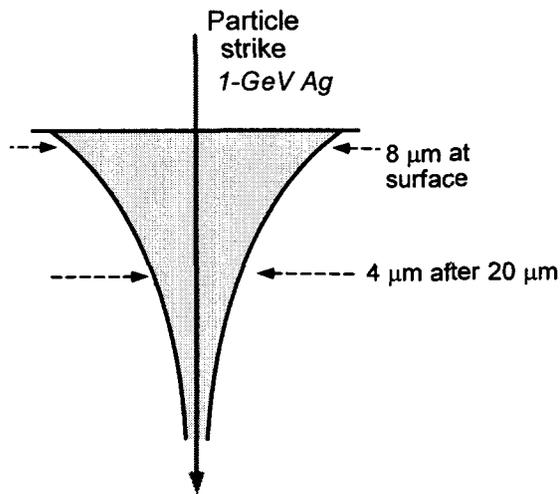


Figure 2. Decrease of track diameter of a 1-GeV silver ion as it traversed a silicon layer.

The diagram in Figure 2 shows the approximate diameter of the particle track just after the particle has traversed the region. The very high density of charge surrounding the track creates a plasma that causes any electric field that is applied across the region to collapse. The holes and electrons within the plasma separate because of the difference in mobility. They diffuse at different rates, changing the lateral and vertical extent of the deposited charge [12]. This affects charge collection and recovery, as discussed below.

B. Charge Collection

Initial treatments of charge collection have been done for structures with much larger areas than the highly scaled devices that are available today. Initial work by Hsieh et al. showed that the collapse of the charge track along the path of an alpha particle allowed direct collection of prompt charge in a p-n junction at distances well beyond the depletion region (charge funneling [13]). The concept of funneling was further discussed by McLean and Oldham in 1992 and corroborated with charge-collection measurements on p-n junctions [12]. Since then, device analysis codes have been used to do more detailed studies of charge collection at time scales that provide better insight into charge funneling and device size effects.

An example of charge collection modeling is shown in Figure 3, after Shin [14]. A 3-D modeling code was used to show how charge collected from an alpha particle strike was affected by the location of the ion strike within the structure, as well as the power supply voltage and the total junction area.

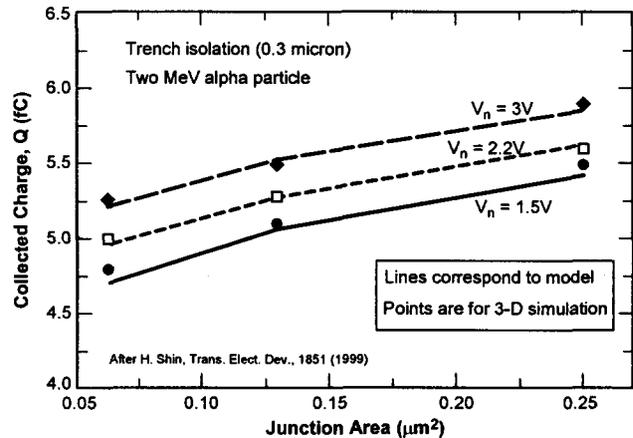


Figure 3. Effect of ion strike position and power supply voltage on collected charge from a 2-MeV alpha particle strike.

These results show that the charge collected from the alpha particle strike is about 25% lower as junction area is reduced, along with slight reductions in power supply voltage. It demonstrates one reason that upset sensitivity does not necessarily increase for highly scaled devices. The effect of small device geometry on charge collection may be larger for energetic ions in space because of the larger track radius compared to the alpha particles that are important for upset from packages, which is widely studied by semiconductor manufacturers.

C. Circuit Effects

Circuit design also influences single-event upset. Basic 6-T memory cells can be upset when the charge collected at the drain of the “off” transistor exceeds the critical charge for upset. This critical charge depends on the switching margin (related to the threshold voltage and the voltage at which the circuit will actually switch); and total capacitance, which includes the capacitance from drain to substrate, gate capacitance, and stray capacitance in the isolation region. In the past, charge collected from the ion strike has been considered instantaneous, but newer device structures respond so quickly that the detailed time response of the charge collection process is also important. Estimates of critical charge can be determined with the SPICE program, but more precise calculations using 2-D or 3-D analysis codes are required for accurate analysis.

Critical charge is affected by design details, such as the channel width and loading, as well as differences in circuit implementation that involve multiple transistors in AND, NAND or NOR configurations.

Clock rates and switching chains also can affect critical charge; a chain of gates operating near the

frequency limit will result in reduced logic swing as the switching pulse progresses through the chain. This reduces the circuit margin, making the device more sensitive to single-event upset.

IV. RADIATION DATA

A. DRAMs

Even though they are among the most sensitive devices to single-event upset, commercial DRAMs have been used during the last 15 years for solid-state recorder applications on many spacecraft. Older DRAMs had very simple response modes, and it was possible to use elementary error-detection-and-correction algorithms to accommodate the errors at the system level. For example, the Clementine spacecraft used an array of 4-Mb DRAMs in a 2.4-Gbit recorder, with a mean error rate of 71 ± 2 errors per day. Moon mapping with this storage approach was completely successful; no missing pixels occurred during the six months that the recorder was used.

Unfortunately newer DRAMs are not so easy to use in space because they are far more complex than earlier devices. Upsets in the internal architecture or registers (for SDRAMs) can alter the functionality of the memory, creating large numbers of errors that cannot be dealt with in a straightforward manner. Multiple-bit upset from a single ion strike can also occur, with up to several hundred errors for ions with large LETs. However, if one ignores those complexities, the upset sensitivity of modern DRAMs has actually improved as they have been scaled from the 16-Mb to the 256-Mb generation. Figure 4 compares proton upset results for various devices, normalized to the error rate per bit. If the error rate were constant, then the cross section should decrease with the cell size. However, it is clear from this figure that there is a large change in the slope of this curve for more advanced DRAMs. The reason for this is changes in the way that the storage capacitor is designed in more advanced DRAMs. Trench capacitors with very large aspect ratios are used, reducing charge collected in the capacitor compared to earlier DRAM technologies.

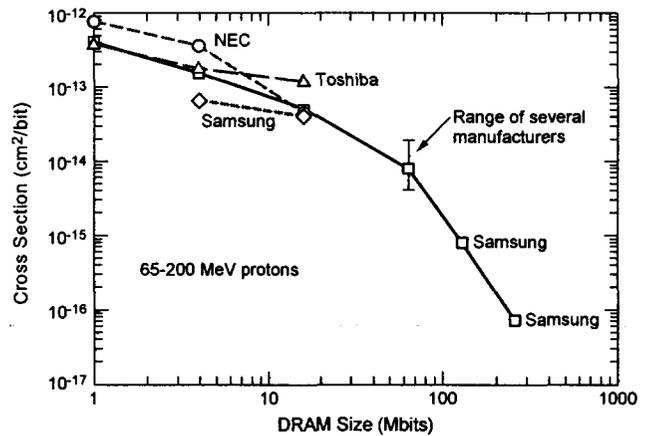


Figure 4. Cross section for proton upset as DRAMs are scaled to smaller feature size.

Another important issue for advanced DRAMs is stuck bits. There are several possible mechanisms for stuck bits, but it is usually assumed that the mechanism is localized ionization damage (microdose) caused by the ion, which affects only a single device. Microdose damage causes a shift in threshold voltage, increasing leakage current. For DRAMs the leakage current must remain very low in order to avoid refresh errors, particularly at moderately high temperatures ($50 - 70^\circ\text{C}$) which is the typical operating temperature in DRAM arrays.

Figure 5 shows how the cross section for single-event upset and hard errors depends on linear energy transfer for a 64-Mb SDRAM. The ratio is only about 10,000 to 1, which affects EDAC in typical space applications. Not all DRAMs show such large cross sections for hard errors, but the sensitivity of DRAMs to hard errors has steadily increased with scaling.

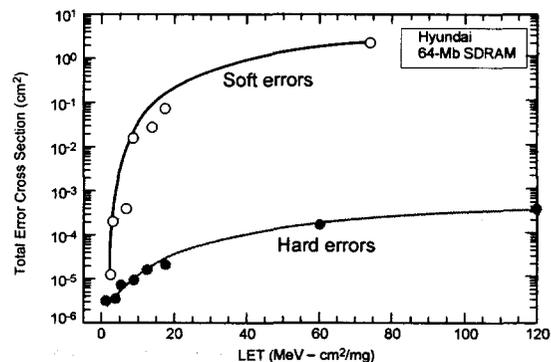


Figure 5. Upset and hard error cross section dependence on LET for an advanced DRAM.

B. Microprocessors

Fabrication techniques for microprocessors have advanced more rapidly than for other mainstream semiconductor devices because of the need for extremely high speed in processor applications. Thus,

SEU effects in processors provide a good measure of how device scaling affects real devices.

Microprocessor testing is not very straightforward because the devices are so complex, requiring extensive support circuitry (at very high operating frequency) as well as an operating system in order to exercise the device. The operating system is critically important. If it is too complex, then elementary malfunctions of the processor under test will be hidden by malfunctions in the operating system. Recent tests of microprocessors have been based on board-level development systems, designed by the manufacturer, with very simple operating systems. Software can be designed to test device operation at various levels. For example, register-intensive tests can be used that continually evaluate internal registers to isolate SEU effects in various regions of the device. The cache section of modern processors can also be evaluated separately using software techniques. These methods, along with tests at different clock frequencies to determine the effects of internal transients on upset rates, allow basic comparisons to be made in the SEU sensitivity of different types of processors, although they do not answer the more practical question of how many upsets will actually occur in a real software application. Figure 6 shows an example of register-level tests for floating-point registers in a Power PC750 microprocessor [15]. In this case there is a significant difference in the cross section for “1 to 0” compared to “0 to 1” transitions. Note that the threshold LET for upsets in the processor is about the same as that of the DRAM in Figure 5, even though the internal design and scaling rules for the microprocessor are quite different.

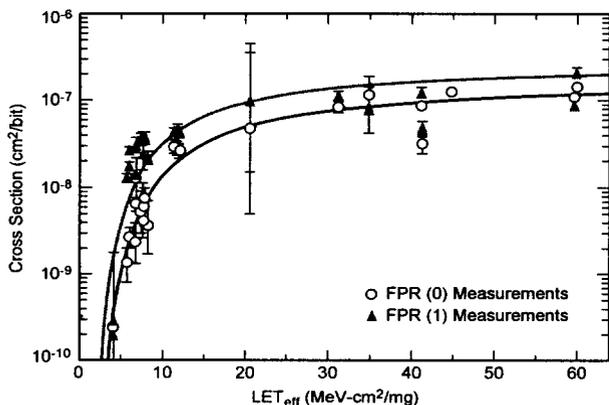


Figure 6. Upset cross section for floating-point registers in the Motorola Power PC750 microprocessor.

Although the threshold LET for upset in these processors is very low, the number of upsets that will occur in a high-inclination earth is not that high.

Correctable errors will occur roughly once every 24 hours, and uncorrectable errors or “crashes” are predicted every few weeks. (Note however, that the number of correctable errors depends on the specific software that the processor is running). These error rates are low enough to consider these unhardened devices in applications that can tolerate occasional operational malfunctions.

V. DISCUSSION

The extreme SEU sensitivity that has often been predicted for highly scaled devices has not developed in practice, as can be seen in the earlier results of Figure 4 for DRAMs. Upset in microprocessors involves very different circuitry, without storage capacitors. However, scaling trends for microprocessors also show some reduction in single-event upset sensitivity. Figure 7 compares upset results for three different generations of Power PC microprocessors [16]. The tests were done at maximum frequency, which increased as devices evolved. First note that the upset cross section in the Power PC750 with a feature size of 0.29 μm is slightly greater than the cross section of the G4, with a feature size of 0.2 μm . The core voltage for the two processors are 2.5 and 1.8 V, respectively.

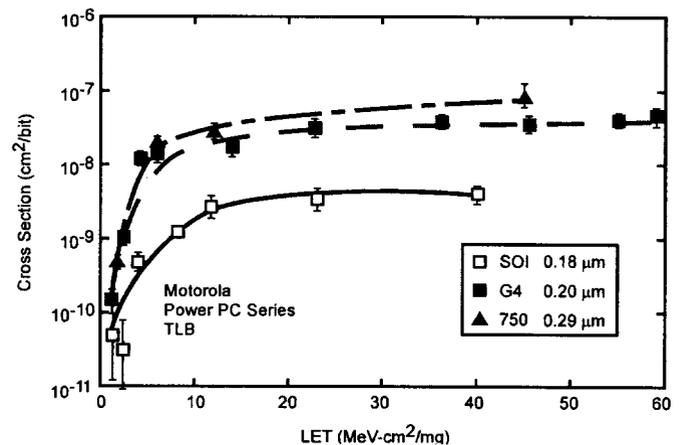


Figure 7. Register error rates for three generations of Power PC microprocessors from Motorola.

Results for the more advanced SOI version of the processor are somewhat different. That device has a feature size of 0.18 μm and core voltage of 1.6 V. The cross section of the SOI device is about an order of magnitude lower than that of the bulk processors, which is expected because the SOI structure has much smaller area than the bulk devices, and cannot collect charge from the substrate. However, the threshold LET is nearly the same as that of the bulk devices, which is inconsistent with earlier scaling projections

for SOI devices. This is probably due to excess charge collection because of the parasitic bipolar transistor within the compact SOI MOSFET (the processor is fabricated with partially depleted technology, with a film thickness of about $0.2 \mu\text{m}$). Very similar results were obtained for an equivalent processor that was manufactured by IBM, but used a lower core voltage with a feature size of $0.13 \mu\text{m}$. Thus, this SOI result appears to be consistent between two manufacturers.

Charge collection in SOI transistors is very complicated, and is heavily influenced by specific processing techniques. Figure 8 shows how the parasitic transistor gain affects measured error rates using an alpha emitter [17] (this is a standard test that is used by semiconductor manufacturers). All of the devices were made with the same feature size, but special ion implant steps were used for the second SOI process in order to reduce the bipolar transistor gain. The error rate is clearly much lower for the device with lower gain. Note that without the special implant, the error rate of the SOI device is significantly higher than an equivalent device, fabricated on bulk technology.

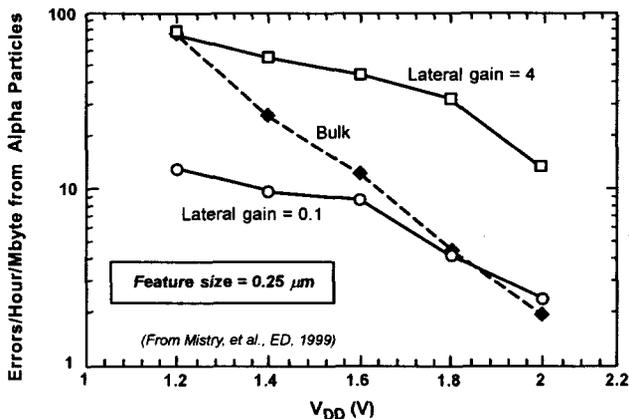
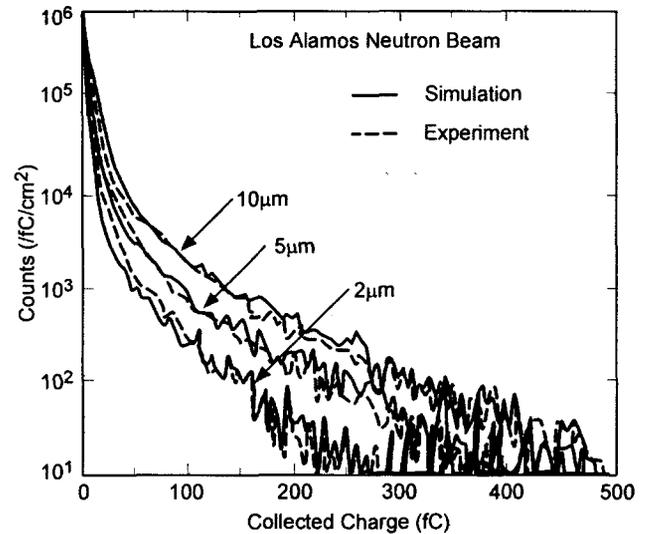


Figure 8. Effect of parasitic bipolar transistor gain on alpha-particle induced upset rates for SOI circuits.

In addition to alpha particle upset effects, semiconductor manufacturers are also concerned with upset effects from neutrons at ground level. Figure 9 shows an example of test data for SOI structures from an advanced process that were tested at the Los Alamos neutron facility [18]. They show the expected dependence on collection volume, and also illustrate good agreement between special modeling done by the manufacturer with test data. A great deal of work is being done on atmospheric neutron upset effects by the device community, which can be used to bound upset rates in the space environment. However, it is

important to realize that these commercial products are not hardened in the conventional sense, and will be subject to significant numbers of upsets in space.



After Tosaka, et al., IEEE Trans. Nucl. Sci., June, 1999

Figure 9. Neutron upset results for commercial memory devices.

This paper has discussed some basic concepts and recent trends for single-event upset in highly scaled devices. Most of the results were obtained during the last two years, some from specific radiation tests and others from modeling studies. The results show that the radiation susceptibility has actually improved somewhat for device generations that have advanced to the $0.18 \mu\text{m}$ level, which contradicts earlier predictions. This trend is encouraging, but it may not necessarily continue for devices that are scaled below $0.1 \mu\text{m}$. The cross section of partially depleted SOI technologies is substantially lower than for bulk technologies, but the threshold LET is nearly the same. If fully depleted SOI circuits become available, there may be significant improvement in threshold LET as well as in cross section that will be a major advantage for space applications. However, fully depleted devices will likely be produced on processes with even lower voltages and critical charge.

Bulk devices will continue to be scaled to smaller dimensions as well. Fundamental considerations of switching energy and noise margin suggest that sudden changes in SEU sensitivity are unlikely, and that the relatively flat dependence of SEU effects on scaling is likely to continue. However, circuit architecture and functional errors will likely become more important as devices are scaled to even smaller dimensions.

REFERENCES

- [1] D. Binder, E. C. Smith and A. B. Holman, "Satellite Anomalies from Cosmic Rays," *IEEE Trans. Nucl. Sci.*, **22**(6), 2675 (1975).
- [2] E. L. Petersen, et al., "Calculation of Cosmic Ray Induced Soft Upsets and Scaling," *IEEE Trans. Nucl. Sci.*, **(29)6**, 2055 (1982).
- [3] S. E. Diehl, et al., "Error Analysis and Prevention of Cosmic-Ion Induced Soft Errors in Static CMOS RAMs," *IEEE Trans. Nucl. Sci.*, **29**(6), 2032 (1982).
- [4] T. C. May and M. H. Woods, "Alpha-Particle Induced Soft Errors in Dynamic Memories," *IEEE Trans. Elect. Dev.*, **26**, 1 (1979).
- [5] SIA Industry Roadmap. 1999.
- [6] B. Davari, R. H. Dennard and G. G. Shahidi, "CMOS Scaling for High-Performance and Low Power – The Next Ten Years," *Proc. IEEE*, **89**, 595 (1995).
- [7] Y. Taur, et al., "CMOS Scaling into the Nanometer Region," *Proc. IEEE*, **85**, 486 (1997).
- [8] D. J. Frank, et al., "Device Scaling Limits of Si MOSFETs and Their Application Dependencies," *Proc. IEEE*, **89**(3), pp. 259-288 (2001).
- [9] R. H. Dennard, et al., "Design of Ion-Implanted MOSFET's with Very Small Dimensions," *IEEE J. Solid State. Circuits*, **9**, 256 (1974).
- [10] O. Musseau, et al., "Charge Collection in Submicron MOS/SOI Technology," *IEEE Trans. Nucl. Sci.*, **44**(6), 2124 (1997).
- [11] P. E. Dodd, "Device Simulation of Charge Collection and Single-Event Upset," *IEEE Trans. Nucl. Sci.*, **43**(2), 561 (1996).
- [12] C. M. Hsieh, et al., "A Field Funneling Effect on the Collection of Charge from Alpha Particle Tracks in Silicon," *IEEE Elect. Dev. Lett.*, **2**(4), 103 (1981).
- [13] F. B. McLean and T. R. Oldham, "Charge Funneling in n- and p-Type Substrates," *IEEE Trans. Nucl. Sci.*, **(29)6**, 2018 (1982).
- [14] H. Shin, "Modeling of Alpha Particle Soft Error Rate in DRAM," *IEEE Trans. Elect. Dev.*, **46**(19), 1850 (1999).
- [15] G. W. Swift, et al., "Single-Event Upset in the Power PC750 Microprocessor," *IEEE Trans. Nucl. Sci.* **48**(6), 1822 (2001).
- [16] F. Irom, et al., "Single-Event Upset in Commercial Silicon-on-Insulator PowerPC Microprocessors," presented at the 2002 Nuclear and Space Radiation Effects Conference, Phoenix, AZ, July 15-19, 2002.
- [17] K. R. Mistry et al., "Parasitic Gain Reduction and the Optimization of 0.25 um Partially Depleted SOI MOSFETs," *IEEE Trans. Elect. Dev.*, **46**(11), 2201 (1999).
- [18] Y. Tosaka, et al., "Simulation Technologies for Cosmic Ray Neutron-Induced Soft-Errors: Models and Simulation Systems," *IEEE Trans. Nucl. Sci.*, **46**(3), 774 (1999).