Impact of Technology and Voltage Scaling on the Soft Error Susceptibility in Nanoscale CMOS

Vikas Chandra
ARM R&D
vikas.chandra@arm.com

Robert Aitken
ARM R&D
rob.aitken@arm.com

Abstract

With each technology node shrink, a silicon chip becomes more susceptible to soft errors. The susceptibility further increases as the voltage is scaled down to save energy. Based on analysis on cells from commercial libraries, we have quantified the increase in the soft error probability across 65nm and 45nm technology nodes at different supply voltages using the Qcrit based simulation methodology. The Qcrit for both bit cells and latches decreases by ∼30% as the designs are scaled from 65nm to 45nm. This decrease is expected to continue with further technology scaling as well. The results show that at nominal voltage, the Qcrit for a latch is just ∼20% more than that of the bit cell in sub-65nm technology nodes. Further, as the voltage is scaled from 1V to 0.4V, Qcrit decreases by ∼5X which substantially increases the probability of an upset if a particle strike happens. This work shows that in sub-65nm technology nodes with aggressive voltage scaling, it is equally critical to solve the soft error problems in logic (latches, flip-flops) as it is in SRAMs.

1 Introduction

Rapidly shrinking technology node and aggressive scaling of voltage have increased the probability of soft errors. Soft errors are radiation induced faults which happen due to a particle hit, either by an alpha particle from impurities in packaging material or a neutron from cosmic rays [1, 10, 16]. When particles strike the silicon substrate they create hole-electron pairs which are then collected by *pn* junctions via drift and diffusion mechanisms. This collected charge creates a transient current pulse and if it is large enough, it can flip the value stored in the state saving element (bit cell, latch etc.). These upsets are called Single Event Upsets (SEU). When particle strike happens in combinational circuit, the result is a glitch which can then propagate to a latch where it could be clocked in and incorrect data can be latched.

Embedded SRAMs are especially vulnerable to SEU due to small size of bit cell and its small node capacitances. However, soft error will just not impact SRAMs but latches/flip-flops and combinational logic as well. At the chip level, the contribution to the soft error rate (SER) from latches and flip-flops is growing [7, 13]. As technology node scales, the susceptibility of soft error increases due to a decrease in the node capacitances. The second facet to this problem is due to aggressive voltage scaling to save power. The dynamic power consumption is reduced quadratically with decreasing supply voltage and linearly with decreasing frequency. To decrease the dynamic power consumption, the supply voltage is scaled down at the cost of increased delay. Decreasing the supply voltage impacts soft error susceptibility as the charge needed to upset a node is a function of the voltage level.

There are many techniques which are currently used to protect SRAMs, namely parity and error correction codes (ECC). However, due to spatial distribution of latches/flip-flops on chip, it is expensive to apply similar techniques in logic. Also, it has always been assumed that latches/flip-flops are much more robust than SRAM bit cells due to their larger size devices and larger node capacitances. This might have been true in technology nodes greater than 90nm, but this certainly is not true in sub-65nm technology nodes [11]. In this work we have shown that with voltage scaling the latches become as critical as bit cells in terms of soft error susceptibility. We have analyzed the susceptibility of bit cells and latches in commercial 65nm and 45nm technology nodes.

The rest of the paper is organized as follows. Section 2 describes the background of soft errors, related work and the simulation methodology for measuring soft error susceptibility of a storage element. Section 3 analyzes the process and voltage scaling characteristics of SRAM bit cell and latch with respect to soft error susceptibility.

1550-5774/08 $25.00 © 2008 IEEE
DOI 10.1109/DFT.2008.50

Section 4 explains the design techniques used to reduce soft error probability. Section 5 summarizes the work and concludes.

2 Background

Figure 1 shows a scenario where an α particle strikes a PMOS transistor. The particle hit could be a glancing blow or a penetrating strike. When particles hit silicon, they generate hole-electron pairs which can then be swept to a diffusion junction if an electric field exists across the device. This has the effect of causing a short duration pulse of current. The effect of the strike on the diffusion can be modeled as a current source as shown in Figure 1. There are other mechanisms of soft errors besides α particles like neutrons, heavy-ions etc. Figure 1 is meant to be representative and not exhaustive.

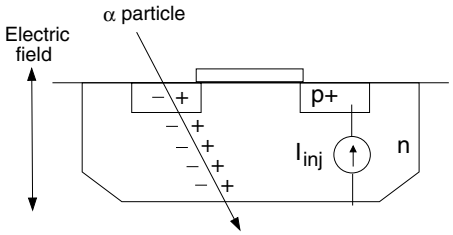


Figure 1: α particle hit on a PMOS transistor

A soft error will happen if the collected charge at a junction is equal to the critical charge (*Qcrit*). *Qcrit* is defined as the minimum charge which is needed to flip the bit stored in the storage cell. To estimate a circuit's sensitivity to soft error, the *Qcrit* values are calculated. The higher the value of *Qcrit*, the more difficult will be to flip the cell and hence it will be more robust. *Qcrit* is proportional to node capacitance and supply voltage. With technology scaling, usually the node capacitance and supply voltage also scale down thus decreasing the value of *Qcrit*.

2.1 Simulation methodology

Qcrit is an important parameter to characterize the sensitivity of a storage cell to a particle strike. A particle strike creates hole-electron pairs which when recombined forms a current spike. Hence, a particle strike at a node can be modeled as a current pulse. For a given current pulse waveform $i(t)$, *Qcrit* is defined as the minimum time integral on $i(t)dt$ that results in the cell flip (Equation 1).

$$Qcrit = \int_0^t i(t)dt \tag{1}$$

There has been a lot of work on choosing the right shape of the current pulse. In [17], the authors use a triangular pulse and in [4], the authors conclude that the shape of the pulse will vary but it can be represented by a piecewise linear function with a peak corresponding to the funneling charge collection and a more slowly decaying tail for the diffusion charge collection. In [9, 15], the authors describe a double exponential current pulse, as described by Equation 2.

$$i(t) = \frac{Q_{total}}{\tau_f - \tau_r} \cdot \left(e^{-t/\tau_f} - e^{-t/\tau_r} \right) \tag{2}$$

In this equation, Q_{total} denotes the total amount of charge generated by the strike, τ_r and τ_f are the rise and fall time constants respectively. The value of τ_r is much smaller than τ_f and hence most of the charge collection happens right after the steep current rise. Another commonly used current pulse is given by Equation 3 [2, 14].

$$i(t) = \frac{Q_{total}}{\tau} \cdot \sqrt{\frac{t}{\tau}} \cdot e^{-t/\tau} \tag{3}$$

In Equation 3, Q_{total} is the amount of charge collected during the event and τ is is the time constant of the collection process. In reality, the peak value of the current pulse and the rise/decay time constant will depend on particle type, particle energy, angle of incidence etc. In order to avoid the complexities associated with choosing values for the models described by Equations 2 and 3, we elected to use a simple triangular model for the current pulse $i(t)$. We believe that this model is accurate for a relative comparison of $Qcrit$ across technology nodes and supply voltages. The rise time constant (τ_r) is chosen as 50fs and the decay time constant (τ_f) is 5ps. These values are representative of particle strike characteristics described in literature [8]. Based on these values, the current pulse is shown in Figure 2.

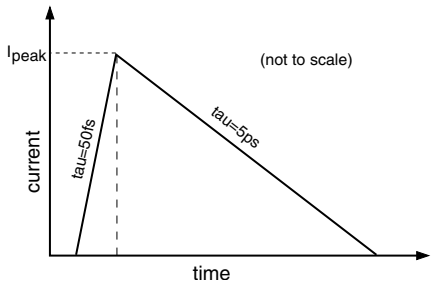


Figure 2: Current pulse shape showing rise and decay time constants

The area under the curve shown in Figure 2 denotes the charge dumped on (or taken away from) a node where the strike happens. For a more robust storage cell, the area under the curve will be larger. Since we fixed τ_r and τ_f, we change I_{peak} to vary the area under the curve. For a given set of operating conditions, I_{peak} is chosen so that the storage cell just flips. In other words, I_{peak} is proportional to the minimum amount of charge ($Qcrit$) needed to flip the cell.

3 Susceptibility analysis

As discussed in Section 2.1, $Qcrit$ is used as a metric to evaluate the susceptibility to soft errors. We discuss the susceptibility of two different classes of circuits:

- 6T bit cell
- Latch

These two classes of circuits represent the majority of storage cells in current VLSI implementations. A 6T bit cell is a dense, symmetrical storage cell consisting of two identical back-to-back inverters and is a basic building block of a high density SRAM. A latch is an asymmetrical storage cell which usually consists of an inverter and a clock (also called clk) controlled tri-state inverter. All the logic storage cells are built out of latches and flip-flops. A flip-flop consists of two latches which operate on different phases of clk. Sections 3.1 and 3.2 analyze these two circuits in terms of their soft error susceptibility. Both these circuits are from commercial libraries in 65nm and 45nm technology nodes. Also, the analysis in the following sections are based on detailed extracted netlists which include all the parasitic capacitances. This is important as parasitics change the value of $Qcrit$ substantially [6].

3.1 SRAM 6T bit cell analysis

A 6 transistor bit cell is shown in Figure 3. The storage inverters are connected to bit lines (bl and blb) through NMOS transistors. These NMOS transistors are driven by wordline (wl). During write operation, wl is asserted and one of the bit lines is pulled low to write a "1" or a "0". During read, the wordline is asserted but the bit lines are not driven. Depending on the value stored in the cell, one of the bit lines is pulled low and the voltage differential is sensed by a sense amplifier to read out the value. Usually a large number of bit cells connect to a bit line so the capacitance of the bit line is quite large. The factors which impact the critical charge of a bit cell are described in [6]. The two critical nodes in Figure 3 are shown as nodes $g1$ and $g2$. The criticality of the

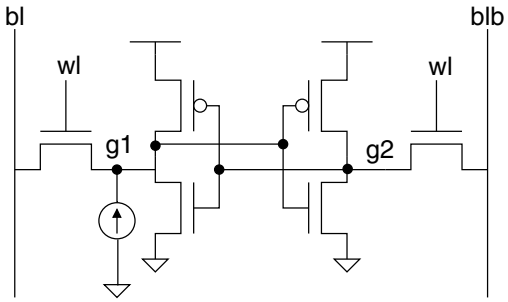
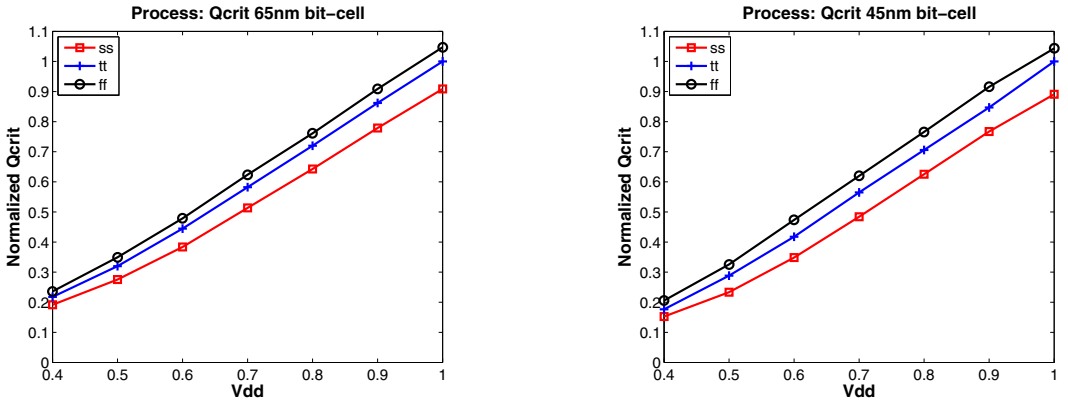


Figure 3: Particle strike modeling in a 6T bit cell

nodes results from the small size of the bit cells and their associated tiny capacitances (10^{-16} to 10^{-17} F). As described in Section 2.1, a particle strike is modeled as a current source. The current source shown in Figure 3 represents the charge generated at node $g1$ when it gets struck by a particle. Since the bit cell is symmetrical, the analysis of particle strike on node $g2$ will be identical to that of node $g1$.

Figures 4(a) and 4(b) show the impact of three different process corners on the *Qcrit* of 65nm and 45nm bit cells respectively. Process corners denote the performance spread of transistors in silicon. The corners are given as fast (f), typical (t) and slow (s). The graph's data are individually normalized *wrt* the *Qcrit* value at



(a) (b)

Figure 4: Impact of process corners on *Qcrit* for a bit cell (a) 65nm process (b) 45nm process

the tt corner (PMOS typical, NMOS typical) and 1V supply. It is interesting to note that the normalized data look very similar for the two technology nodes. The *Qcrit* decreases by a factor of 5X as the voltage is scaled from 1V to 0.4V. As expected, the ff corner is the most robust owing to faster transistors. In addition to the variation observed in *Qcrit* within process generations, it is also interesting to compare the relative ratios of bit cell *Qcrit* values at 45nm and 65nm nodes, as shown in Figure 5. At nominal voltage (1V), the value of *Qcrit* decreases by a factor of 0.74 which is close to the *scaling factor* of the technology shrink. However, at lower voltage (0.4V) the factor is close to 0.6. So, voltage scaling increases the soft error upset probability by ∼20%.

3.2 Latch analysis

A simple latch design is shown in Figure 6. When *clk* is "1", the latch is transparent and *Q* follows *Data*. When *clk* goes low, the input switch turns off and the latch closes with the value stored. During this close phase, the latch becomes vulnerable to soft error as a particle strike can flip the value stored in the latch. Unlike the bit cell described in Section 3.1, the latch is not symmetrical. In other words, nodes $n1$ and $n2$ behave differently

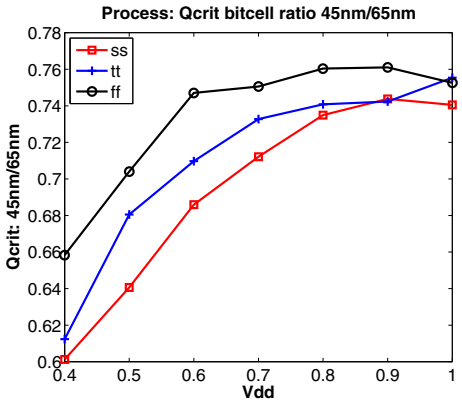


Figure 5: Ratio of bit cell $Qcrit_{45nm}/Qcrit_{65nm}$

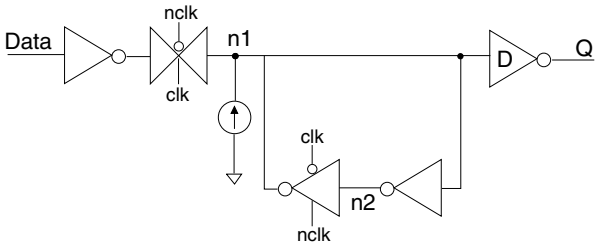


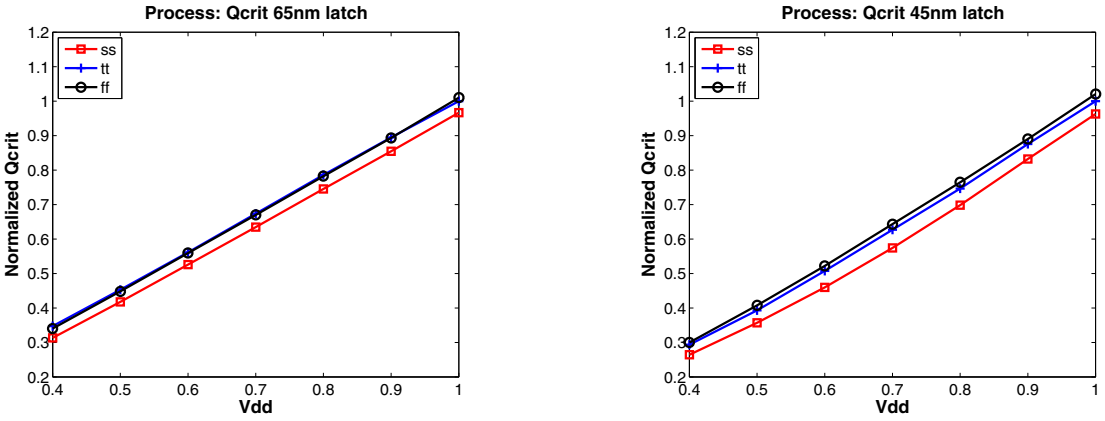
Figure 6: Particle strike modeling in a latch

when a particle strike happens. The two factors which cause the asymmetric behavior of nodes are driving gate strength and node capacitance. As can be inferred from Figure 6, both the factors will be different for nodes $n1$ and $n2$. For the latch we analyzed, node $n1$ has smaller $Qcrit$ than node $n2$. This is primarily due to weaker drive strength of the tri-state inverter (a factor of 2X) as compared to the drive strength of the inverter driving $n2$. Hence, to account for the worst case, we have done the particle strike analysis on the node $n1$. Since a flip-flop consists of two latches, the soft error analysis of a latch extends to a flip-flop as well.

Figures 7(a) and 7(b) show the impact of process variation on voltage scaling for the latch in 65nm and 45nm technology nodes respectively. The decrease in the value of $Qcrit$ is linear with respect to voltage. This can be attributed to the relation of charge to voltage ($Q = CV$). But as the voltage is reduced to the sub-threshold regime, the value of gate capacitance also begins to change. It is not shown in Figures 4 and 7, but the $Qcrit$ variation becomes non-linear *wrt* voltage near 0.2V or so. The data in each of the graphs in Figure 7 have been individually normalized *wrt* the $Qcrit$ value at the tt corner and 1V supply. If we compare the relative scaling of $Qcrit$ for bit cells and latches *wrt* voltage at 65nm and 45nm (Figures 4 and 7), we can see that the latches are more robust at lower voltage. The normalized $Qcrit$ for the bit cell is approximately 0.2 when the voltage is 0.4V, whereas for latches the relative $Qcrit$ only scales to 0.3 at 0.4V. Figure 8 shows the ratio of 45nm $Qcrit$ to 65nm $Qcrit$ for the latch. With voltage scaling this ratio decreases from 0.7 to 0.62. This behavior is very similar to that of the bit cell.

3.2.1 Latch drive strength variants

A standard cell library has typically many drive strength variants of a cell. We considered three drive strengths of the latch shown in Figure 6, namely A, B and C. For the library we analyzed, C is about 50% bigger than B which is twice the size of A. Drive strength of the latch shown in Figure 6 is controlled by the size of the inverter D. An increase in drive strength affects $Qcrit$ due to increase in the node capacitance of node $n1$. Figure 9 shows the increase of $Qcrit$ for drive strengths B and C *wrt* A. So, for a latch based design, increasing the drive strength will help to reduce soft error upset probability. For this work, we used a latch with a drive strength of A as it has a smaller $Qcrit$ than a latch with B or C drive strength.



(a) (b)

Figure 7: Impact of process corners on *Qcrit* for a latch (a) 65nm process (b) 45nm process

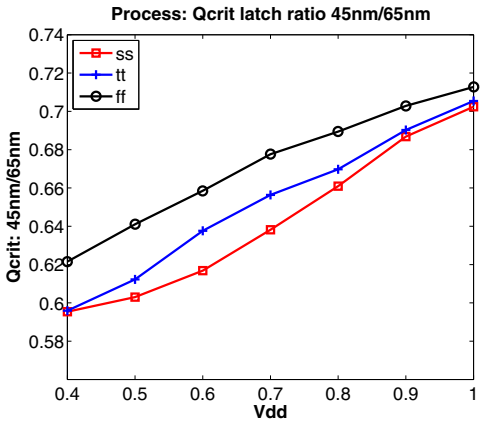


Figure 8: Ratio of latch $Qcrit_{45nm}/Qcrit_{65nm}$

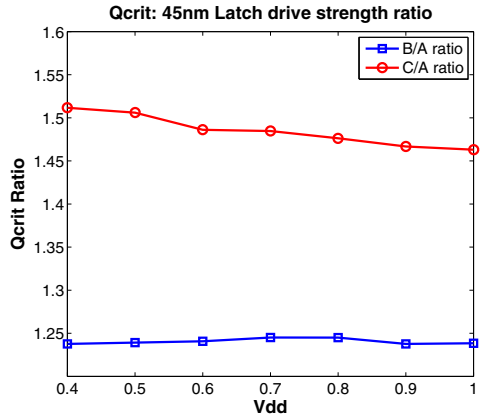
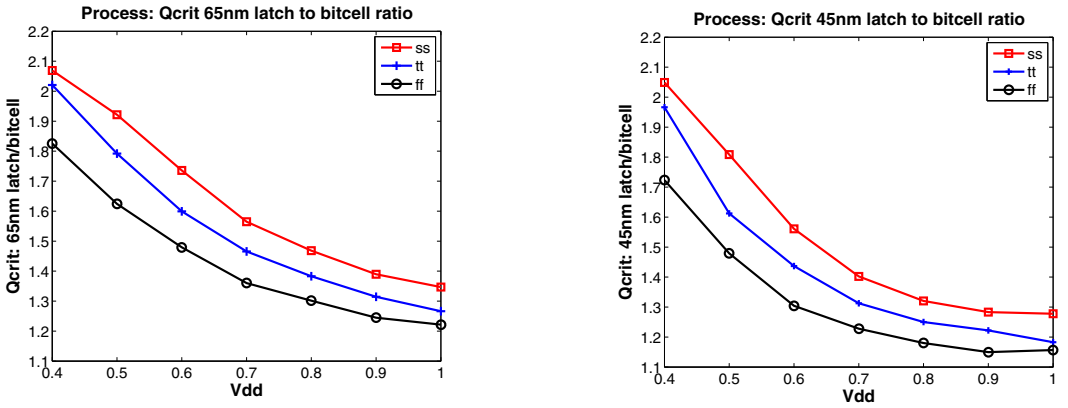


Figure 9: Effect of drive strength on *Qcrit* in a latch

3.3 Bit cell and latch comparison

In a new technology node, a lot of time is spent on designing a bit cell. The important criteria being yield, density, leakage, performance, Vmin behavior and dynamic power. Typically a bit cell is designed to be as small as lithographically printable. On the other hand, latches (and other standard cells) do not use pushed lithography rules and are typically larger than bit cells. Hence there is a mismatch in the way these two circuits scale. Figures 10(a) and 10(b) show the ratio of latch to bit cell *Qcrit* for 65nm and 45nm processes respectively. At 1V and tt corner, the latch is around 27% (ratio of 1.27) better than the bit cell for 65nm. For 45nm, the value drops down to 19% (ratio of 1.19). In other words, at nominal operating voltage the latch is marginally better than bit cell with regards to soft errors. As the voltage is scaled down, the latch starts to look better than the bit cell. For 65nm, at 0.4V, the best case ratio is 2.08 (ss corner) and the worst case ratio is 1.82 (ff corner). For 45nm, at 0.4V, the best case ratio is 2.05 and the worst case ratio is 1.72. Even in the low voltage regime, a latch is only twice as robust as a bit cell!



(a) (b)

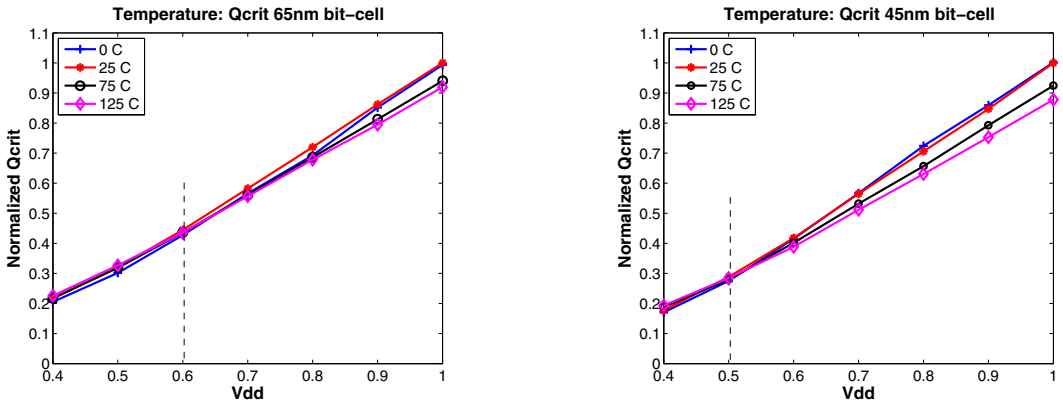
Figure 10: Ratio of latch to bit cell *Qcrit* for (a) 65nm process (b) 45nm process

This observation tells us that soft errors in latches have already become as important as those in bit cells. With technology node scaling to 32nm and below, the soft error susceptibility of latches and bit cells is likely to become very similar. There are inexpensive error correcting schemes to correct soft errors in SRAMs. These schemes are not feasible for latches due to their spatially distributed nature. The techniques commonly employed to make the latch more robust are discussed in Section 4.

3.4 Temperature and *Qcrit* scaling

Temperature also affects the susceptibility of a design to soft error. We analyzed the bit cells in 65nm and 45nm technology nodes for their sensitivity to temperature. Figures 11(a) and 11(b) show the effect of temperature on *Qcrit* with voltage scaling for a bit cell. From the figures, it is clear that temperature does not change *Qcrit* significantly (10% in 65nm, 15% in 45nm), although at higher voltage the *Qcrit* difference caused by temperature is higher than at lower voltage. The spread at 1V in the 45nm node is higher than that in the 65nm node, although they behave more similarly at 0.4V.

In [9], the authors analyzed the impact of temperature on the SER of a 65nm SOI SRAM. It was observed that the *Qcrit* decreases with increase in temperature. This trend was observed across all voltages. Our analysis shows a similar trend at higher voltages. However, we observed that at lower voltages the high temperature *Qcrit* was higher than the low temperature *Qcrit*. At high voltage, 0C case has the best *Qcrit* and 125C case has the worst *Qcrit*. At low voltage, 0C has the worst *Qcrit* and 125C has the best *Qcrit*. This *temperature inversion* effect happens for both 65nm and 45nm technology nodes (the inversion voltage is shown by the dotted lines in Figure 11). However, the voltage at which this inversion occurs reduces with technology scaling. For 65nm, the inversion voltage is 0.6V whereas for 45nm, the inversion voltage is 0.5V. The same effect was



(a) (b)

Figure 11: Impact of temperature on *Qcrit* for bit cell in (a) 65nm process (b) 45nm process

observed for latches as well. In our opinion, this effect is related to the temperature inversion effect already being observed in nanometer bulk devices and which probably is not that prominent in SOI devices.

4 Soft error mitigation

Based on our analysis in Section 3, it is clear that as technology shrinks and voltage scales, soft error becomes an increasingly important reliability bottleneck. The techniques to reduce the probability of soft error can be employed at various levels of abstraction, including process, circuit and architecture.

- **Process:** Fully depleted SOI devices are much more robust against soft error due to lack of the bulk. Recently published data show that even partially depleted SOI based SRAMs are 5X more robust than bulk based SRAMs [3]. However, the SOI process is very different from the bulk process and volume manufacturing of fully depleted SOI devices is still challenging.
- **Circuit:** There are various techniques which can make a circuit more robust against soft errors (e.g. *rad-hard* circuits). Common approaches include increasing the node capacitance by resizing [19] and changing the architecture of the storage cell [5, 12]. Another common technique is to add redundancy to latch/flip-flop which can detect and correct soft errors [18]. It is useful to note that redundancy is a kind of error correction, but not a very efficient one.
- **Architecture:** The architectural techniques used for SRAMs typically employ parity or ECC to detect errors. One common ECC implementation is SECDED (single error correct double error detect), which is used for processor caches.

All the techniques described above come at a cost. SOI has higher manufacturing cost than bulk, *rad-hard* cells usually have large area and power overhead and ECC techniques can have performance penalty during cache access. Nevertheless, these techniques, or others like them, will be needed to combat the substantial increase in soft error susceptibility due to technology and voltage scaling. These techniques must be an integral part of the design process, not an afterthought.

5 Conclusions

Technology scaling combined with voltage scaling exacerbates soft error upset susceptibility. The value of *Qcrit* decreases by ∼30% between the 65nm and 45nm technology nodes. This decrease will continue as the technology scales further down. With every technology shrink, the soft error susceptibility of a latch gets closer to that of a bit cell. For a 45nm process, the *Qcrit* for the latch is just 19% higher than the bit cell at 1V. As the voltage

is scaled down, the latch starts to look better but it is only ∼2x better at 0.4V. Further, with aggressive voltage scaling, the susceptibility to soft errors increases substantially. When voltage is scaled from 1V to 0.4V, the value of *Qcrit* decreases by a factor of 5X. In summary, shrinking transistor geometries combined with voltage scaling make soft errors one of the most crucial reliability concerns. With the soft error susceptibility of latches approaching that of bit cells, increasing design and architecture effort is required in order to protect against soft errors.

References

[1] R. C. Baumann, "Soft errors in advanced semiconductor devices - Part I: The three radiation sources," *IEEE Transactions Device and Materials Reliability*, vol. 1, no. 1 pp. 17-22, Mar 2001.

[2] M. Baze and S. Buchner, "Attenuation of Single Event Induced Pulses in CMOS Combinational Logic," *IEEE Transactions on Nuclear Science*, Dec. 1997.

[3] E. H. Cannon *et al*, "SRAM SER in 90, 130 and 180 nm Bulk and SOI Technologies," *Intl. Reliability Physics Symposium*, pp. 300-304, 2004.

[4] C. Dai *et al*, "Alpha-SER Modeling and Simulation for Sub-0.25μm CMOS Technology," IEEE Symposium on VLSI Technology, pp. 81-82, June 1999.

[5] P. Hazucha *et al*, "Measurements and Analysis of SER-Tolerant Latch in a 90-nm Dual-V_T CMOS Process," *IEEE Journal of Solid-State Circuits*, Vol. 39, No. 9, Sep. 2004.

[6] T. Heijmen, "Factors that Impact the Critical Charge of Memory Elements," *IEEE Intl. On-Line Testing Symposium (IOLTS)*, 2006.

[7] T. Karnik, P. Hazucha and J. Patel, "Characterization of soft errors caused by single event upsets in CMOS process," *IEEE Transactions on Dependable and Secure Computing*, pp. 128-143, 2004.

[8] J. Keane *et al*, "Method for Qcrit Measurement in Bulk CMOS Using a Switched Capacitor Circuit," *NASA Symposium on VLSI Design*, June 2007.

[9] A. J. KleinOsowski *et al*, "Modeling Single-Event Upsets in 65-nm Silicon-on-Insulator Semiconductor Devices," *IEEE Transactions on Nuclear Science*, Vol. 53. No. 6, pp. 3321-3328, Dec. 2006.

[10] T. C. May and M. H. Woods, "Alpha-particle induced soft errors in dynamic memories," *IEEE Transactions on Electron Devices*, pp. 2-9, Jan 1979.

[11] S. Mitra, T. Karnik, N. Seifert and M. Zhang, "Logic Soft Errors in Sub-65nm Technologies: Design and CAD Challenges," *IEEE Design Automation Conference*, June 2005.

[12] R. Naseer and J. Draper, "DF-DICE: A Scalable Solution for Soft Error Tolerant Circuit Design," *ISCAS*, pp. 3890-3893, 2006.

[13] N. Seifert, D. Moyer, N. Leland and R. Hokinson, "Historical trend in alpha-particle induced soft error rates of the Alpha microprocessor," *IEEE International Reliability Physics Symposium (IRPS)*, pp. 259-265, 2001.

[14] P. Shivakumar, M. Kistler, S. Keckler, D. Burger and L. Alvisi, "Modeling the effect of technology trends on the soft error rate of combinational logic," *International Conference on Dependable Systems and Networks*, pp. 389-398, June 2002.

[15] G. R. Srinivasan *et al*, "Accurate, Predictive Modeling of Soft Error Rate Due to Cosmic Rays and Chip Alpha Radiation," *IEEE International Reliability Physics Symposium*, pp. 12-16, 1994.

[16] Y. Tosaka *et al*, "Impact of cosmic ray neutron induced soft errors on advanced CMOS circuits," *IEEE Symposium on VLSI Technology*, pp. 148-149, 1996.

[17] S. Walstra and C. Dai, "Circuit-Level Modeling of Soft Errors in Integrated Circuits," *IEEE Transactions on Device and Materials Reliability*, Vol. 5, No. 3, pp. 358-364, Sep. 2005.

[18] M. Zhang *et al*, "Sequential Element Design with Built-In Soft Error Resilience," *IEEE Transactions on VLSI*, Vol. 14, No. 12, pp. 1368-1378, Dec. 2006.

[19] Q. Zhou and K. Mohanram, "Transistor Sizing for Radiation Hardening," *Intl. Reliability Physics Symposium*, pp. 310-315, 2004.