

Modeling Soft-Error Susceptibility for IP Blocks

Robert Aitken, Betina Hold
ARM, Physical IP Division
141 Caspian Court
Sunnyvale, CA, USA 94089
rob.aitken@arm.com

Abstract

As device geometries continue to shrink, single event upsets are becoming of concern to a wider spectrum of system designers. These “soft errors” can be a nuisance or catastrophic, depending on the application, but they must be understood and their effects budgeted for. Ultimately, experimental measurement is needed to quantify soft error rates, but after-the-fact measurement is too late to make changes. This paper shows a methodology that can be used to estimate the soft error properties of individual IP blocks by using a combination of critical charge calculations and experimental data.

1 Background: Soft Errors

Single event upsets (SEU) result from a variety of physical phenomena, including alpha emissions from radioactive decay in solder, high-energy (>1MeV) neutrons from atmospheric cosmic ray collisions, and thermal neutrons and heavy ions from nearby radioactive events (cosmic ray or decay related). All but high-energy neutrons can be protected against through judicious selection of manufacturing materials. Particles have a variety of energies, and can cause disturbances ranging from temporary charge impulses to physical destruction. See [5-7] for a good overview of these phenomena and approaches used to mitigate their effects. Historically, concern about single event upsets was limited to high reliability applications such as aerospace, but three recent trends are changing this.

First, the increasing volume of memory means that previously rare occurrences are becoming more common. For example, an event that occurs on a single memory bit once every billion operating hours will happen once every seven minutes in a data center with a terabyte of memory. Common failure rates for embedded SRAM are on the order of a failure per bit every trillion operating hours, but this still leads to a data center error about once every five days. If the data center is in Denver, the failure rate will be about 5 times higher.

Second, there is a trend towards higher electronic automation on systems that were formerly mechanical. Antilock braking systems, for example, require highly reliable electronics and must operate in a failsafe manner in the presence of SEU.

Finally, as CMOS process technology shrinks with every generation, the critical charge required to disturb the contents of an SRAM cell is declining over time. The shrinking area of bit cells has compensated for this in recent generations, but shrinking geometries mean more bit cells will be present, for an overall increase in system soft error rate (SER) [5]. In fact, as several researchers have observed, if these trends continue, error tolerance will be needed on latches and even standard cell logic in an increasing number of applications [6].

2 Implications of SoC Business Model

Calculation of soft error rates is more than a technical problem. In today’s system-on-chip (SoC) environment, designs are comprised of commercially available IP blocks, from memories to high speed I/O interfaces to CPU cores, together with user-designed logic that typically makes use of standard cell library components. All in all, many different organizations may be involved in a chip design, and all must provide and exchange information to provide a complete SEU solution.

Accurate estimation of SER and mean time between failures (MTBF) is a complex task that requires experimental analysis in both hardware and software. JEDEC has guidelines on experiments that can be used to calculate SER given certain observations under accelerated conditions [4]. Actual product SER will depend on these factors, plus application-dependent issues including physical orientation, local environment (e.g. a handheld game could be used at sea level, or at 10,000m in an aircraft).

While full quantification of SER must wait until product deployment, this does not mean that SER cannot be addressed earlier. Indeed, SER must be considered throughout the design cycle in order to avoid surprises at the end. This paper shows a method to estimate product SER given data available for embedded memory. It is also shown how limited data for some designs can be used to estimate behavior of others in the same manufacturing process. This allows SER to be traded off among other design attributes, such as area, performance, manufacturability, and power without resorting to the time and expense needed to build and test actual silicon.

This ability to prejudge the SER susceptibility of design is critical because the semiconductor industry has substantially disaggregated, allowing chip designers the freedom to select from a number of IP providers, wafer foundries, packaging and test houses, among others. Some of these complex relationships are shown in Figure 3. In this disaggregated situation, information exchange on SER-critical data is not guaranteed, and in some cases may be hindered, based on economic interests of the various parties involved. To succeed in this context, an SER methodology must provide some form of value to all companies involved. This constrains the problem significantly from the original vertically integrated case (for a detailed discussion on these issues as they relate to test, see [1]).

As an example, single port memory bit cell designs often originate with a foundry, because of their tight coupling to both process design rules and yield. The foundry is thus in the best position to conduct the required experiments on soft error rates for the bit cell. However, SRAM SER is not solely dependent on bit cells. The architecture is also important, and this could come from an IP company. Dual port bit cells often originate with IP companies. These must be available early in the process life cycle in order to enable the first designs in a process. There is not time in this scenario to go through multiple iterations on bit cell design, so it is critical to get the design right from the start.

3 Calculating Qcrit

The amount of charge required to flip a memory cell is known as Qcrit. For a symmetric cell, this charge will be essentially identical for flipping from both 0 to 1 and 1 to 0. Qcrit can be calculated using SPICE simulation of a current pulse applied to a memory bit cell. For an arbitrary current waveform $I(t)$, Qcrit is defined as the minimum time integral of $I(t)dt$ that results in a bit cell flip.

$$Q_{crit} = \int_0^{t_{flip}} I(t)dt$$

The waveform is often characterized as a rapid rise in current to a peak value, followed by an exponential decay, characterized by the function below [2][6], where Q is the amount of charge collected during the event and T is the time constant of the collection process:

$$I(t) \propto \frac{Q}{T} \times \sqrt{\frac{t}{T}} \times \exp\left(-\frac{t}{T}\right)$$

In reality, the peak value and decay time will depend on particle energy, angle of incidence, and so on. For simulation purposes, though, some value must

be chosen. SPICE is able to integrate the current through a time range to calculate Qcrit. The choice of value matters, as shown by the graph below (Figure 1), which shows the change in relative values of Qcrit for a given bit cell design as the peak value of the current pulse scales from 20uA to 50uA. Intuitively, this makes some sense: a circuit is more likely to successfully fend off a small current for a long time than a large current for a short time.

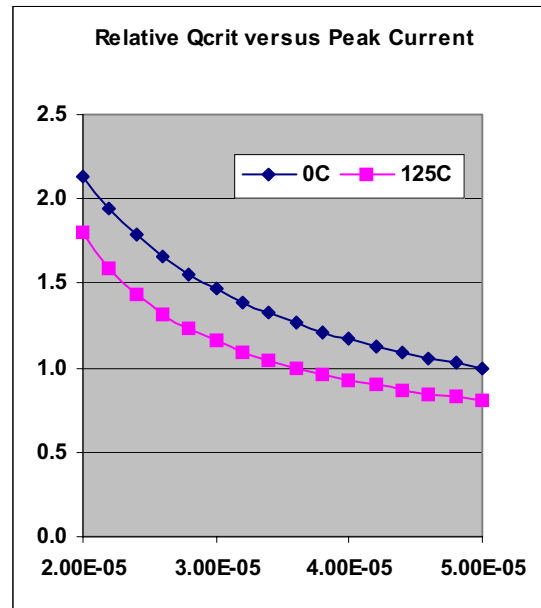


Figure 1 Qcrit change with peak pulse current

If an integration function is not available, it would be easier to approximate the exponential pulse with a triangular waveform with a short rise and longer decay. The duration of the waveform must be sufficient to cause a bit flip. Qcrit may be expressed in terms of the peak current of the triangle wave, as shown below:

$$Q_{crit} = 0.5 * I_{triangle} * t_{triangle}$$

Once Qcrit has been obtained, it becomes a useful measure for comparing bit cell designs. The variability of Qcrit calculations suggests that it is vital to settle on a methodology (maximum height of current pulse, process corner, temperature, voltage, etc.) before making comparisons. Once this has been done, comparison is straightforward.

Designs with similar electrical properties can have somewhat different Qcrits. Qcrit also varies with supply voltage, process corner and temperature. Low supply voltages result in lower values for Qcrit, as does high temperature (see Figure 2). The variation is particularly

apparent in some process corners (e.g. strong N devices and weak P devices). This means that designs targeting lower voltages in order to reduce dynamic current ([IEM ref]) must be more conscious of the potential impact on SER.

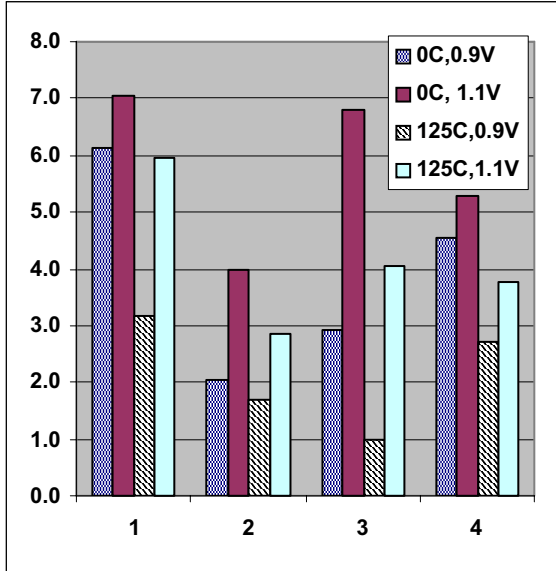


Figure 2 Change in Qcrit with temperature, voltage

Finally, there is a question about variability and SER. In deep submicron processes, bit cell parametric variation is substantial. Differences of 30% in read current between bit cells in the same array are commonplace. As a result, simulations performed on average case circuits may not be adequate to determine actual failure rate. Weaker cells will fail more rapidly than stronger ones. Intuitively, one might expect that the stronger cells would balance out the weaker ones, and this will be true if the variation in Qcrit were normally distributed. In general, SER due to neutron flux has been observed to vary linearly with sensitive device area (transistor drains for SRAM cells) and exponentially with Qcrit [3]:

$$SER \propto F \times A \times e^{-k \cdot Qcrit}$$

where F is the neutron flux in particles/cm²*s, A is the area of the design that is sensitive to particle strikes in cm² and k is a constant of appropriate units*. Note that SER (FIT rates) are the inverse of MTBF, which

* Note that k is actually the inverse of the charge collection efficiency of the device in fC, but this is assumed to be constant across comparable devices in the absence of data to the contrary.

results in a sign change for the constant if MTBF is used instead.

While Qcrit helps to compare designs, it does not help with estimating actual failure rates. For this, experimental corroboration is necessary. This corroboration is most often available for single port bit cells in an early release of a process. Alpha particle data is most frequently available, but neutron data may be available as well (the latter requires data collection that is only possible at a small number of facilities worldwide). Data is usually presented in terms of FIT/Mb of memory. By measuring the same circuit in the same flux at multiple voltages (which allows us to use multiple simulated Qcrit values), we can fit data to the equation below:

$$\frac{SER_1}{SER_2} = \exp(k \times (Qcrit_2 - Qcrit_1))$$

$$k = \frac{\ln(SER_1) - \ln(SER_2)}{Qcrit_2 - Qcrit_1}$$

For example, if a change in Qcrit from 2fC to 1.5fC results in a 3X change in SER, then k has a value of 2.197fC. With this value, a 10% increase in Qcrit leads to a 36% decrease in SER. Once k has been calculated, it may be used to estimate the SER of other related devices within a constant neutron flux, as shown below.

$$\frac{SER_{new}}{SER_{obs}} = \frac{A_{new}}{A_{obs}} \times \exp(k \times (Qcrit_{new} - Qcrit_{obs}))$$

The more a given circuit varies from the experimentally measured circuit, the less accurate the predicted SER is likely to be. Even so, the ability to estimate SER can be useful. Table 1 below shows how similar circuits will vary across differing values of area and charge. For example, it is clear from the table that a 20% increase in area is worthwhile for SER reduction if it gains 10% in critical charge (the net benefit is 23%). There will be some benefit even with a 5% change in Qcrit. This is not true if k is only 1.05, as shown in the sixth column. Experimental data to allow for estimating k is clearly highly beneficial to SER evaluation prior to committing designs to silicon.

A1	A2	Q1	Q2	delta(SER), k=2.2	delta(SER), k=1.05
1	1	2.00	1.00	8.998	2.858
1	0.8	2.00	1.00	7.198	2.286
1	1.2	2.00	1.00	10.798	3.429
1	1	2.00	1.80	1.552	1.234
1	0.8	2.00	1.80	1.241	0.987
1	1.2	2.00	1.80	1.862	1.480
1	1	2.00	2.00	1.000	1.000
1	0.8	2.00	2.00	0.800	0.800
1	1.2	2.00	2.00	1.200	1.200
1	1	2.00	2.20	0.644	0.811
1	0.8	2.00	2.20	0.516	0.648
1	1.2	2.00	2.20	0.773	0.973
1	1	2.00	2.50	0.333	0.592
1	0.8	2.00	2.50	0.267	0.473
1	1.2	2.00	2.50	0.400	0.710
1	1	2.00	3.00	0.111	0.350
1	0.8	2.00	3.00	0.089	0.280
1	1.2	2.00	3.00	0.133	0.420

Table 1: Delta SER for comparable devices

4 Conclusions

While experimental measurement is needed to precisely quantify soft error rates, we have shown that limited data may be generalized in a simple way in order to make predictions about likely soft error rates in practice. This allows for tradeoffs among possible circuit implementations that can be used, along with critical charge calculations, to estimate the soft error properties of individual IP blocks.

5 References

- [1] R. Aitken, "Applying Defect-Based Test to Embedded Memories in a COT Model.", Proc. IEEE Workshop on Mem. Tech. Design & Test, July 2003.
- [2] M. Baze and S. Buchner. Attenuation of Single Event Induced Pulses in CMOS Combinational Logic. *IEEE Transactions on Nuclear Science*, 44(6), December 1997.
- [3] P. Hazucha and C. Svensson. "Impact of CMOS Technology Scaling on the Atmospheric Neutron Soft Error Rate", *IEEE Transactions on Nuclear Science*, 47(6), pages 2586–2594, Dec. 2000
- [4] JEDEC standard JESD89, available at www.jedec.org
- [5] R. Baumann, "Single Event Effects in Advanced SRAM and Logic Components", Sematech TRC 2004, available at www.sematech.org
- [6] P. Shivakumar et al, "Modeling the Effect of Technology Trends on the Soft Error Rate of Combinational Logic", *Proc. Int. Conf. On Dependable Systems and Networks*, 2002.
- [7] J.H. Stathis, "Reliability limits for the gate insulator in CMOS technology", *IBM J. Res. & Dev.*, 44(2/3), May 2002.

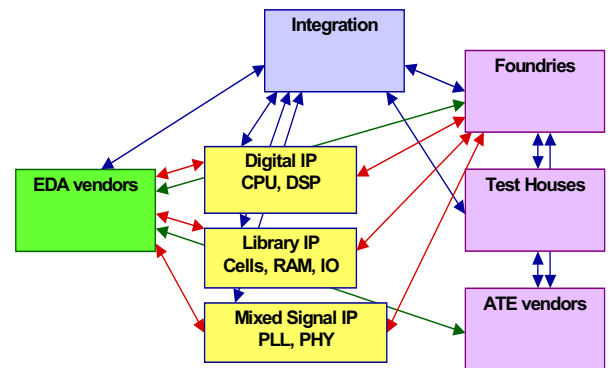


Figure 3 Example SoC supply chain