Method for Qcrit Measurement in Bulk CMOS Using a Switched Capacitor Circuit

John Keane[†‡] jkeane@ece.umn.edu
Alan Drake[†] adrake@us.ibm.com
AJ KleinOsowski[†] ajko@us.ibm.com
Ethan H. Cannon[*] cannon1@us.ibm.com
Fadi Gebara[†] fhgebara@us.ibm.com
Chris Kim[‡] chriskim@umn.edu

[†]IBM Austin Research Laboratory
Austin, TX USA

[*]IBM Systems and Technology Group
Essex Junction, Vermont USA

[‡]Dept. of ECE, University of Minnesota
Minneapolis, MN USA

***Abstract*—As the CMOS IC industry continues to follow an aggressive scaling roadmap, device sizes and voltage margins are rapidly shrinking. These factors, along with the construction of larger and more complex systems, increasing clock frequencies, and other complications lead to a growing sensitivity to radiation events. Of particular concern is the sensitivity of memory cells to soft errors, in which a radiated particle strike generates charge that is collected by a reverse-biased junction in these circuits, causing the stored bit to be flipped. In this work, we propose a new method to measure the robustness of storage circuits to radiation events using a switched capacitor circuit. This method would significantly reduce testing time and costs compared to existing methods, while providing circuit designers with reliable hardware measurements. Simulations demonstrate the ability of this method to find the critical charge needed to flip a stored bit within an acceptable error range of generally less than 25% when compared with the widely used current source modeling method.**

I. INTRODUCTION

Rapidly shrinking device sizes and aggressive voltage scaling in CMOS circuits have exacerbated the threat posed by radiation events. These events occur when energetic particles—usually alpha particles emitted by impurities in IC packaging materials, or neutrons resulting from cosmic ray interactions with earth's atmosphere—strike silicon devices, causing localized bursts of charge generation. The resulting current and voltage transients can flip the logic state of a storage cell if the strike occurs sufficiently close to a charge sensitive reverse-biased junction. This mechanism is called a soft error, or a Single Event Upset (SEU), since no permanent damage is done to the circuit. The flipped cell can simply be rewritten to carry on with normal operation. Additionally, particle strikes in combinational logic can cause an incorrect value to propagate to a latch boundary where it is clocked in, which is called a single event transient (SET). The present work will focus on upsets in storage cells.

Although radiation events in silicon have been studied since the 1970s, this reliability mechanism has recently attracted wide-spread attention in the semiconductor industry. Smaller nodal capacitances and voltage margins result in an increased sensitivity to radiated particle strikes, meaning that the critical amount of charge needed to flip a stored bit (Qcrit) has decreased with technology scaling. Smaller device sizes have also reduced the sensitive volume in each silicon device capable of collecting the charge created by a strike, but the growing number of transistors on a chip which are vulnerable to soft errors has offset this benefit, so the overall Soft Error Rate (SER) of silicon chips continues to increase with scaling. Therefore, great care must be taken to mitigate the impact of these soft errors in many critical systems. In the early days of radiation event research, robustness to soft errors was only given significant consideration in military and space applications. However, today this issue must be dealt with in many commercial and consumer products with low tolerances for failure, such as large servers and medical devices.

While Qcrit is one parameter required to quantify the sensitivity of storage cells to energetic particle strikes, additional information is needed to calculate the SER for these circuits, and in turn, an entire system. In particular the charge collection efficiency, or the amount of charge that a node can collect after a radiation event, must be understood. This parameter is a function of the transistor architecture [1, 2], and requires a detailed knowledge of device physics and the particle interactions that take place during and after a radiated particle strike. Much research has been devoted to understanding these issues through hardware experiments and modeling.

Leading semiconductor companies devote large amounts of resources, including money and a great deal of valuable time, to soft error testing. Existing methods include exposing large numbers of chips to natural radiation, or taking accelerated measurements by bombarding ICs with radiation from particle beams. The former method requires extensive testing periods (up to months) to gather reliable statistics. The latter necessitates specialized equipment, trained personnel, and careful experiment design. A fast hardware testing procedure requiring fewer resources and readily available test equipment would therefore be extremely valuable.

In order to facilitate the characterization of radiation events in silicon devices, several simulation schemes have been proposed at the device, circuit, and system levels. These simulation models are critical for pre-silicon analysis of the expected error rates in different circuit configurations and architectures. Expected failure rates are needed to

ensure that reliability requirements are met without expensive redesigns after the first tapeout. Methods employed to improve the robustness of circuits to soft errors include redundancy, the use of radiation-hardened latches, and error-correction coding (ECC) in memory systems.

At the device level, numerical device simulators such as that described in [3] are used to model the effects of particle strikes in great detail. This process provides information regarding the charge collection efficiency of particular points in the device under test, and the shape of the current wave pulse created by a particle strike. It can also aid in improving radiation immunity through process changes. A faster simulator used to accomplish some of the same goals was presented in [4]. System level simulation programs locate vulnerable portions of a chip where upsets are more likely to cause errors seen by the end user [5].

Information gathered from hardware experiments and detailed device simulations of radiation events are used to develop accurate circuit-level simulation techniques. At this level, a radiated particle strike is abstracted as a current pulse which is meant to be representative of a particle strike on one particular node in the circuit under test (CUT). A number of current pulse shapes have been proposed in prior work for use in circuit simulations. A roughly triangular current wave shape was used in [2], while the authors of [6] stated that the wave form varies but can generally 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. A double exponential current wave form is used in [7-9], and is described as

$$i(t) = \frac{Q_{total}}{\tau_f - \tau_r}\left\{e^{-t/\tau_f} - e^{-t/\tau_r}\right\}. \quad (1)$$

In this equation, Q_{total} corresponds to the total amount of charge generated by a particle strike, and τ_r and τ_f are the rising and falling time constants, respectively. In all cases, the majority of the charge collection takes place on the order of tens of picoseconds after a sharp current rise time, while the decaying current tail varies from tens to hundreds of picoseconds depending on operating conditions, particle type, and a number of other factors.

Our proposed test circuit extends this ideal current source modeling method to a hardware implementation by utilizing a switched capacitor design, which can inject charge carriers into a node of the CUT. Simulation results show that our proposed method compares favorably with results found using the ideal current pulse method, which has in turn been shown to provide accurate Qcrit values [7]. With proper component sizing, Qcrit results found with our method are in most cases within 25% of the ideal double exponential pulse results. This is an acceptable error when dealing with numbers in the femtocoulomb range and a dynamic parameter such as the critical charge. This point will be expanded upon in section III.

Information gathered from such a test would provide another source of data to be used in the development of accurate circuit-level modeling tools for radiation events. One advantage of this method is the fact that it provides us with a means of testing individual nodes of a circuit, which is not possible in typical hardware experiments involving exposure to radiation. Another benefit is the obvious point that this experiment would result in hardware measurement data (albeit, without an exact reproduction the extreme current and voltage transients that can result from a particle strike, but the closest we can safely come to that with standard circuit techniques). As the author of [10] stated, Qcrit is dependent on many device, circuit, and technology parameters, some of which are not accurately reflected in available device models. Therefore, simulations alone do not result in reliable failure statistics unless they are based on accurate hardware data, so there is a need for more experimental measurements. Implementing our scheme has the potential to provide us with that data without long test periods, specialized equipment, or large test arrays. In the present work, we will describe our design, simulation results, and implementation considerations.

II. CIRCUIT DESCRIPTION

A. Design Overview and Operation

The basic structure of our charge injection circuit is presented in Figure 1, within the context of one particular storage cell under test (a latch in this case). When testing the sensitivity of an NMOS device in the off state within a given circuit configuration, electrons are injected into the reverse-biased drain junction (i.e., current is drawn away from the node). This region is the most charge sensitive part of the circuit, so the specified setup will give us a worst-case sensitivity for particle strikes affecting this device.

The charge injection is accomplished by discharging the storage node capacitance (C_{SN}) through N1, then turning that device off, and finally switching on N2 for the minimum pulse width achieved by on-chip pulse generation circuitry. If no state change is detected, that process is repeated by first rewriting the cell state and then increasing the on-time for N2 until the state of the storage device is flipped (while rewriting the cell state for each incremental on-time increase). At that point, Qcrit for the present operating conditions and current waveform has been deposited on the node under test. Qcrit, which is used to quantify the sensitivity of circuits to radiation events, is calculated after measuring C_{SN}, as well as the voltage at that node before and after the charge injection event, by using the following equation:

$$Q_{crit} = C_{SN}\left(V_{final} - V_{initial}\right). \quad (2)$$

Hole injection into PMOS drains is similar, but current is injected into (rather than drawn out of) the reverse-biased

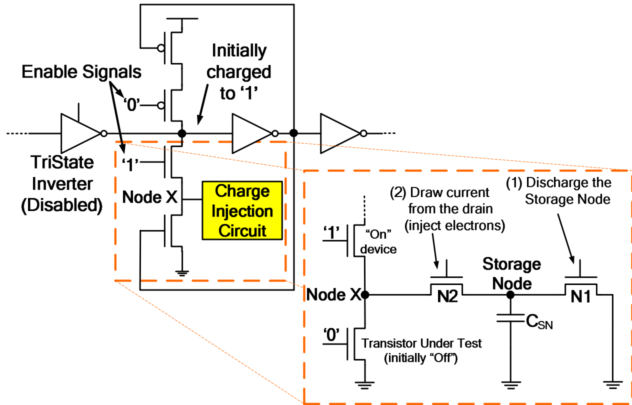

Figure 1. Switched capacitor charge injection implementation used here in the second stage of a standard flip-flop for NMOS drain injection.

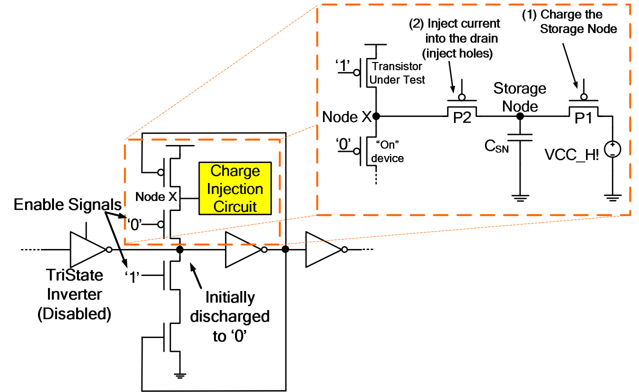

Figure 2. Switched capacitor charge injection implementation used here in the second stage of a standard flip-flop for PMOS drain injection.

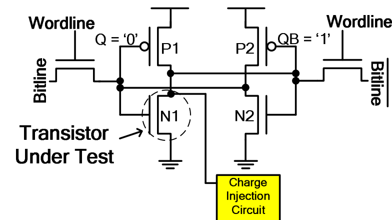

Figure 3. SRAM cell including the switched capacitor charge injection block from Figure 1.

PMOS drain node. Therefore, in this case C_{SN} is first charged to a level higher than the standard digital supply (VCC_H! in Figure 2), rather than discharging C_{SN}. A varied injection pulse width to turn on device P2, injecting current *into* Node X, is used as described above for the NMOS implementation. Alternatively, it is possible to use a constant pulse width and a variable voltage source (i.e., incrementally raise VCC_H! until an upset occurs) in order to measure the sensitivity of a particular PMOS device. In either case, thick oxide devices should be used in the charge injection circuitry with this increased voltage, and the PMOS switch body terminals tied to VCC_H!.

Figure 3 illustrates another example schematic, specifically an SRAM cell, including the charge injection block. This figure indicates that NMOS injection circuitry from Figure 1 is utilized to test for a high to low transition on node QB due to a strike at the drain of N1. However, the sensitivity of the cell to a particle strike at the drain of P1 when QB is low can be tested with the same configuration, and the PMOS charge injection circuit from Figure 2.

B. Design Considerations

Several factors must be taken into consideration when designing switched capacitor circuits such as those described above. For instance, while larger switches provide a low resistance current path for fast injection, they also lead to increased capacitive coupling, channel charge injection, and clock feedthrough. The latter is suppressed with a half-dummy switch, while the former two effects are kept to a minimum through careful device sizing.

Sizing the storage node capacitor requires a trade-off between area and accuracy, or in the worst case even functionality. In the NMOS injection setup for example, C_{SN} must be sufficiently larger than the capacitance of Node X (see Figure 1) in order for the injection circuit to sink enough current to upset the state of the CUT. Although in this fast transient case, charge sharing equations based on DC conditions or simple linear currents will not apply, the need for careful capacitor sizing can be derived intuitively.

Simply put, with a small C_{SN}, the source-drain voltage across switch N2 will quickly drop, causing this device to fall out of saturation. Note that a MOS capacitor can not be used in this case, as the capacitance should not vary with voltage. Metal capacitors are a better choice here.

Figure 4.a shows the current waveforms through switch N2 for an example latch (SC#2 in Figure 5) in 65nm PTM technology [11, 12] operating at 25C with a 1V supply. The size of the storage node capacitor is varied as indicated. The normalized minimum on-time values for N2 which results in a logic state upset for those capacitor sizes are plotted in Figure 4.b. These pulses were generated through chains of inverters rather than using ideal input patterns to produce more realistic results. In all test cases here and throughout this work, the minimum pulse widths required were on the order of tens of picoseconds (generally 10ps-40ps), which can be created by on-chip equipment, and correspond closely to the widths of current pulses created by radiation events. Simulations results indicated that for each storage cell configuration tested, a C_{SN} size of 30fF to 70fF was sufficient to reliably cause a data upset with a supply voltage range of 0.8V to 1.2V.

III. SIMULATION RESULTS

Simulations were performed on a number of latches found in IBM's standard cell library, along with other typical storage cell designs. All experiments utilized 65nm PTM technology at 25C. (Similar results were observed at temperatures ranging to 85C.) Qcrit values were

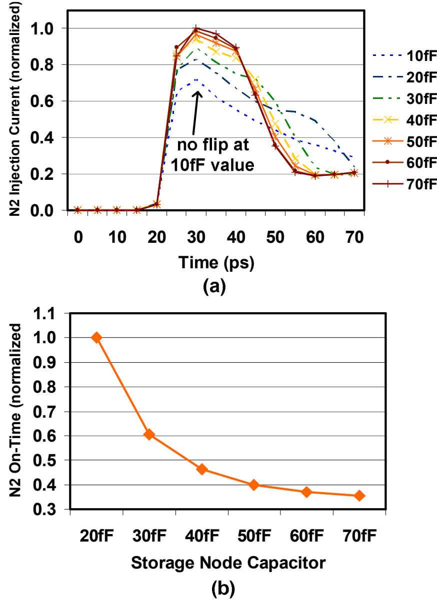

Figure 4. Current injection waveforms for different C_{SN} capacitor sizes (a), and the normalized minimum pulse times required to cause a logic state upset for those sizes.

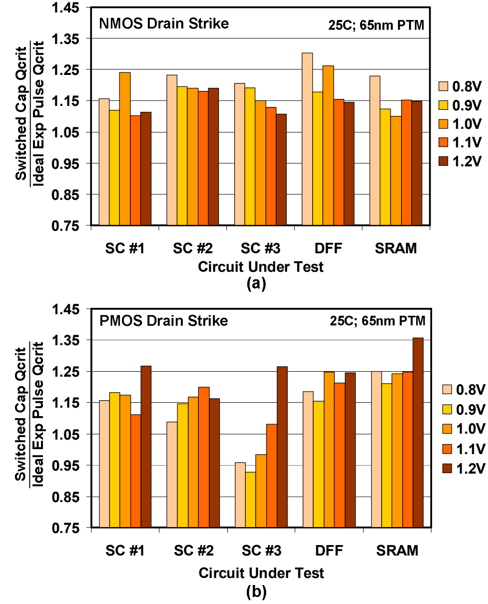

Figure 5. Simulation results for NMOS (a) and PMOS (b) charge injection experiments. Qcrit values found with the injection experiment are normalized to those found with an ideal current source driving a double exponential pulse.

calculated as described in equation (2), using the capacitance value for the storage node found in the HSpice output capacitance table, and the voltages measured on the storage node one nanosecond before the start of the current injection event and one nanosecond after its completion. Note that the only timing requirement for those voltage measurements is that they take place soon before and after the charge injection when the voltages are settled.

Qcrit values found with the switched capacitor charge injection circuit were compared to those found with the typical current source modeling method. For this comparison, a double exponential current source with a rising time constant of 50fs and a falling time constant of 5ps was used, for a total pulse time of roughly 25ps. This shape is representative of those described in literature (some of which were summarized in section I). In order to find the critical charge with this current source, the magnitude of the current spike was varied while the rise and fall times remained constant.

Figure 5 displays Qcrit values found with the switched capacitor injection circuit normalized to those found with the ideal current source just described. The error across a range of supply voltage values for both PMOS and NMOS drain strikes is less than 25% with few exceptions. The ideal current source modeling method has been shown to provide exceptionally accurate data compared to hardware measurements [7], so we surmise that when properly tuned and implemented, our proposed injection circuit should lead to accurate Qcrit measurements.

As covered in [9], the shape of the current pulse is a function of a number of factors including substrate doping, operating voltage, and the position of the particle strike relative to the junction where the charge is collected. In that work, a constant fast rise time is used to model the range of possible waveforms, while the falling time constant was treated as a variable. In order to demonstrate the accuracy of our method as compared to the double exponential pulse model using a range of falling time constants, we swept from the low value of 5ps used as our base case up through 50ps for the three standard cell latches used earlier. The latter value gives a total current pulse time of roughly 300ps. The results for a range of supply voltages are plotted in Figure 6. We see that each measurement stays within the 25% error range up to a 20ps falling time constant (with a corresponding total pulse time of ~120ps), and error is less than 50% even up to the 300ps waveform for most cases examined here. These error values can be improved upon in certain instances with capacitor and transistor sizing in the charge injection circuit. That improved accuracy will come at a cost of increased area and/or degraded sampling precision.

To those not familiar with Qcrit and SER measurements, a 25% error would seem to be unacceptable. However, this is not the case when measuring such dynamic values. The critical charge for any given circuit is not a constant—it is a variable parameter that depends on radiation pulse characteristics, the dynamic response of the circuit under consideration, temperature, and device parameter variations [13, 14]. For these reasons, and a host of other complications, the critical charge and the corresponding error rate values are acceptable within an unusually large range of error. The authors of [1] and [2] cite SER errors

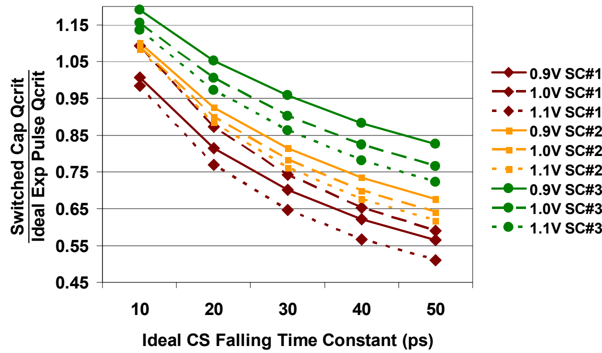

Figure 6. Qcrit results normalized to ideal double exponential current source results with a varied falling time constant.

between modeled and measured values of over 75% and 2X, respectively. The author of [14] states that variations in the SER of a bipolar SRAM array can range up to 60X, and specifically plots large Qcrit error ranges (although this paper is somewhat dated). Based on this and other literature dealing with radiation events in silicon, an error of up to 25% and even more can still provide designers with useful information regarding the robustness of certain circuit configurations.

IV. TEST CHIP INTERFACE

A high level diagram of the experimental test setup for the proposed charge injection circuit is presented in Figure 7. We would plan to implement as much of the control and measurement logic on the test chip as possible. One important step done with off-chip equipment is the capacitance measurement on the storage node (C_{SN}), which could be completed with a high speed network analyzer.

In the illustrated setup, data is scanned in through an interface such as a JTAG port, or simply a pad on the test chip. The FSM controls the test procedure timing by synchronizing the measurement steps outlined in section II.A. The pulse generator would be a programmable circuit capable of incrementing the size of pulses by small steps (<5ps increments). One such generator was described in [15]. The storage node voltage measurements would preferably be done by on on-chip voltmeter capable of storing multiple voltage measurements for a later scan-out, but this measurement could be taken off chip if so desired.

V. CONCLUSIONS

We have proposed a switched capacitor circuit for measuring the critical charge of CMOS storage cells. This test circuit was demonstrated to provide results within a 25% error margin in most cases, with respect to the widely used ideal current source modeling method. Given the large range of acceptable Qcrit and SER error, the proposed circuit is a promising option for obtaining hardware measurement data. This data could contribute to the development of the device models circuit designers require to implement robust systems.

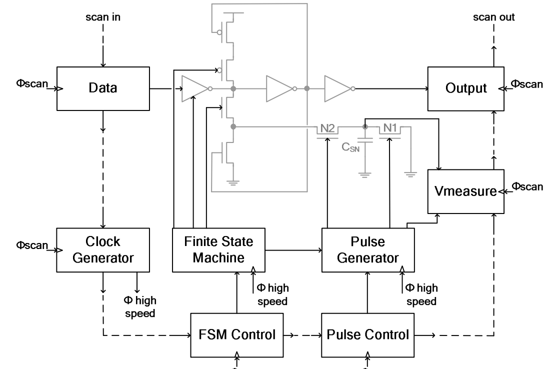

Figure 7. Proposed experimental test setup. Portions of this logic could be shared between multiple storage cells under test.

VI. ACKNOWLEDGEMENT

This work was supported in part by DARPA contract NBCH30390004.

VII. REFERENCES

[1] T. Heijman, "Analytical Semi-Empirical Model for SER Sensitivity Estimation of Deep-Submicron CMOS Circuits," IEEE Int. On-Line Testing Symp., pp.3-8, July 2005.

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

[3] E. M. Buturla, P. E. Cottrell, B. M. Grossman, and K. A. Salsburg, "Finite-Element Analysis of Semiconductor Devices: The FIELDAY Program," IBM J. Res. and Develop., Vol. 25, No. 4, pp.218-231, July 1981.

[4] H. H. K. Tang and E. H. Cannon, "SEMM-2: A Modeling System for Single Even Upset Analysis," IEEE Trans. on Nuclear Science, pp.3342-3348, December 2004.

[5] X. Li, et al., "SoftArch: An Architecture-Level Tool for Modeling and Analyzing Soft Errors," Int. Conf. on Dependable Systems and Networks, pp.496-505, June 2005.

[6] C. Dai, et al., "Alpha-SER Modeling & Simulation for Sub-0.25µm CMOS Technology," IEEE Symp. on VLSI Tech., pp.81-82, June 1999.

[7] AJ KleinOsowski, et al., "Modeling Single-Event Upsets in 65-nm Silicon-on-Insulator Semiconductor Devices," IEEE Trans. on Nuclear Science, Vol. 53, No. 6, pp.3321-3328, December 2006.

[8] L. W. Massengill, et al., "Single-Event Charge Enhancement in SOI Devices," IEEE Electron Device Letters, Vol. 11, No. 2, pp.98-99, February 1990.

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

[10] G. R. Srinivasan, "Modeling the Cosmic-Ray-Induced Soft-Error Rate in Integrated Circuits: An Overview," IBM J. Res. and Develop., Vol. 40, No. 1, pp.77-89, January 1996.

[11] Predictive Technology Model, online: http://www.eas.asu.edu/~ptm/

[12] W. Zhao and Y. Cao, "New generation of Predictive Technology Model for sub-45nm Design Exploration," Int. Symp. On Quality Electronic Design, pp. 585-590, 2006.

[13] R. C. Baumann, "Radiation-Induced Soft Errors in Advanced Semiconductor Technologies," IEEE Transactions on Device and Materials Reliability, Vol. 5, No. 3, pp. 305-316, September 2005.

[14] L. B. Freeman, "Critical Charge Calculations for Bipolar SRAM Array," IBM J. Res. and Develop., Vol. 40, No. 1, pp.119-129, January 1996.

[15] V. Ramakrishnan and P. T. Balsara, "Very High Precision Venier Delay Line Based CMOS Pulse Generator," IEEE Dallas/CAS Workshop, pp.225-228, October 2005.