SEU-Aware Low-Power Memories using a Multiple Supply Voltage Array Architecture

Seokjoong Kim and Matthew R. Guthaus

University of California Santa Cruz, Santa Cruz, CA US
{seokjkim, mrg}@soe.ucsc.edu

Abstract. Electric devices should be resilient because reliability issues are increasingly problematic as technology scales down and the supply voltage is lowered. Specifically, the Soft-Error Rate (SER) increases due to the reduced feature size and the reduced charge. This paper describes an adaptive method to lower memory power using a dual V_{dd} in a column-based V_{dd} memory with Built-In Current Sensors (BICS). Using our method, we reduce the memory power by about 40% and increase the error immunity of the memory without the significant power overhead as in previous methods.

1 Introduction

Single Event Upsets (SEUs) are caused by alpha particles or cosmic rays which create temporary electron-hole pairs upon collision with a silicon surface. In the past, these were common only in high-altitude (space) applications, but they are becoming more significant as process geometries and supply voltages shrink. Figure 1(a) shows the case when a particle hits the channel of a transistor. Depending on the energy and incident angle of the particle, an amount of electron hole pairs are created which can affect certain characteristics of the transistor, such as the drain current I_{ds}.

Many previous works have proposed methods to analyze transient errors induced by radiation [14, 18, 20, 27]. Basically, these methods have used a simulated pulse to emulate the spike induced by SEUs to simulate the effect at transistor/gate level as shown in Figure 1(b). In memory devices, this pulse targets the most sensitive storage node in the memory cell. The Soft-Error Rate (SER) [16,29], however, depends on the location, altitude and surrounding energy level [12] in which the circuit is operating. Researchers often use an empirical model for SER based on the critical charge Q_{crit} [4,7], but both the environmental and critical charge parameters of this model are challenging to estimate due to technology scaling and process variation. Other prior works [13] have proposed to use real measured data from radiation chambers to increase the accuracy of the prior models. This method improves the error rate accuracy, but it is costly in terms of resources and time to properly calibrate the model for the each chip designed.

To reduce the soft error rate, many previous works employ architectural techniques such as Error Correcting Codes (ECC) [5]. Error Correction Codes (ECC) add additional parity bits to original data bits to detect/correct errors. ECC can detect soft errors
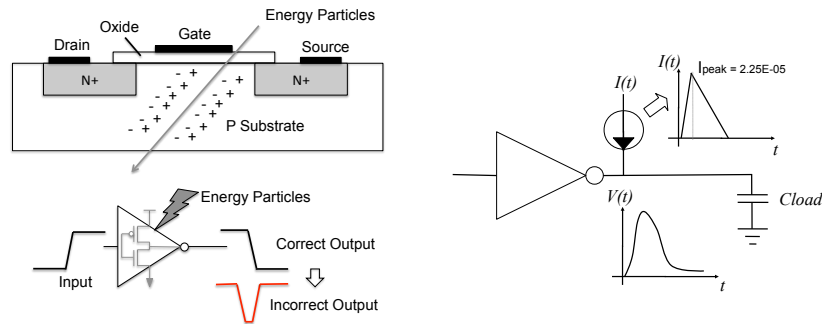
(a) Single Event Upsets (SEU) cause electron hole pairs in the transistor channel and incorrect output.

(b) Gate-level soft error simulation of the impact requires temporal modeling of the charged particles.

Fig. 1. Single Event Upset (SEU) example and gate-level SEU simulation methodology are used to analyze circuit robustness.

depending on the the number of parity bits. Single Error Correction Double Error Detection (SECDED) scheme is normally used for ECC due to its simple architecture, but Double Error Correction (DEC) can be implemented using more logics and gates and increases power. Also circuit sizing methods were also proposed [2]. Circuit level techniques can increase the soft error immunity using hardened memory cells. The basic idea of hardened memory cell is increasing a capacitance of stored node to increase the critical charge Q_{crit} level. This method improves the soft error tolerance but it affects to the memory performance due to the increased capacitance.

The major issue with the prior approaches is that they can't dynamically react to immediate changes in the flux energy level. Built-In Current Sensors (BICS) have been proposed that detect transient errors in real time [15, 19] so that the errors may be immediately detected and corrected by Error Correction Codes (ECC). This enables the SER to be controlled within a recoverable range while the memory operates. Although it keeps the SER within recoverable margins, the additional BICS and ECC may increase the cost and power consumption of the chip.

Dynamic Noise Margin (DNM) has been previously introduced to quantify the transient response of SRAM cells in the presence of noise [3]. DNM quantifies a memory cell's fault tolerance to a transient voltage instead of static voltage. This means that DNM can more accurately quantify the tolerance of a memory cell to realistic external noise since SEUs from alpha and neutron particles have both temporal and voltage level components. Previous researchers have proposed many different analysis methods to compute DNM [8, 22, 23, 25].

In this work, we propose a SEU-tolerant SRAM architecture using BICS to detect SEUs and then improve the dynamic noise immunity using a dual-supply Dynamic Voltage Scaling (DVS) scheme. Since most memory designs perform DVS by selecting from pre-defined V_{dd}, we propose the methods to determine the optimal supply voltage

levels considering both error tolerance and power reduction based on column-based V_{dd} array architecture with BICS.

Our major contributions are as follows:

– We are the first to propose an adaptive architecture using BICS in column-based V_{dd} memory architecture.
– We are the first to quantify the optimal voltage considering power and SEU tolerance through a new Monte Carlo framework.
– We analyze the impact of peak current variation and explicitly consider the Dynamic Noise Margin (DNM) of the memory cells.
– We also show the SER improvement by increasing transistor size in memory cell.

The rest of this paper proceeds as follows: Section 2 describes the overview of our BICS architecture, Section 3 introduces our MC framework and calculates the optimal voltage levels, Section 4 describes our power model using the dual V_{dd}, Section 5 shows our experimental setup and results, and Section 6 concludes the paper.

2 Background

This section describes previous works that systematically detect transient errors in memory arrays and recent research into dual-supply voltage column based memories. The two components are integral to our approach which is presented next in Section .

2.1 Dynamic Transient Error Detection

Researchers have proposed built-in sensors to detect transient errors dynamically [15, 19, 21]. Figure 2(a), for example, shows a Built-In Current Sensor (BICS) implemented alongside a representative 6T SRAM cell. The BICS connects to each column at the bottom of the array. When a particle strikes an internal node of any memory cell in the column, the voltage of the internal node fluctuates due to the electron-hole pairs and immediately decreases the virtual V_{dd} (*VVDD*) of the BICS. This fluctuation turns on the PMOS transistor in pull-up path of the BICS which asserts the *UPSET* signal to indicate the presence of a transient particle.

2.2 Column-based Supply V_{dd} Array Architecture

Column-based V_{dd} memories have been recently proposed to reduce memory array power consumption [6, 26]. Figure 2(b) shows the memory array structure with each memory cell's V_{dd} is connected to the global V_{dd} in each column. Since SRAM read operations need higher V_{dd} for improved noise margins compared to write operations, a dual supply voltage saves power without performance or reliability degradation. When a column is read, the supply voltage is set to V_{high} and when a column is written, it is set to V_{low}. This approach reduces power by minimizing the supply voltage depending on the read/write operating pattern.
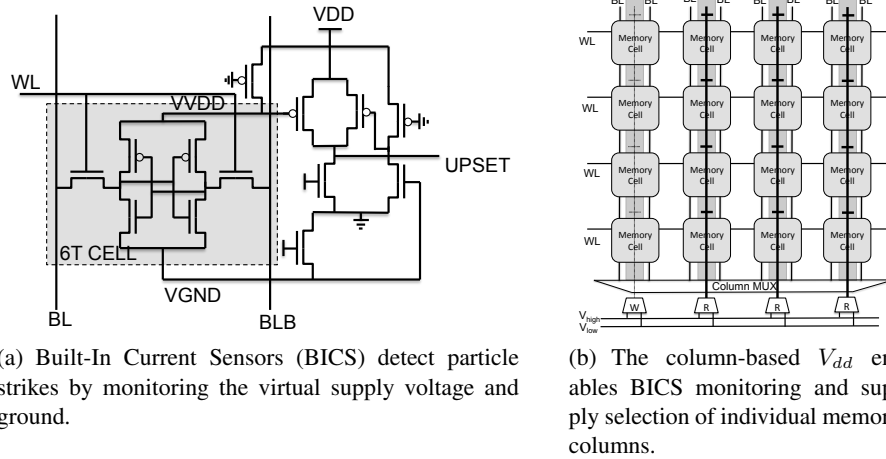
(a) Built-In Current Sensors (BICS) detect particle strikes by monitoring the virtual supply voltage and ground.

(b) The column-based V_{dd} enables BICS monitoring and supply selection of individual memory columns.

Fig. 2. Previous works have separately used Built-In Current Sensor (BICS) for error detection and column-based V_{dd} arrays for dynamic power savings depending on the operation (read or write).

3 Proposed Work

While column-based V_{dd} memory architectures have been used for power reduction in the previous section, our approach instead assumes the same voltage level for both operations (read and write). We alternatively combine the BICS and column-based V_{dd} array to dynamically select the minimum supply voltage to retain data values according to the present external noise conditions as shown in Figure 3. The combination of these two techniques lowers power consumption by dynamically adjusting what would otherwise be a conservative worst-case static guardband voltage while maintaining fault tolerance.

3.1 SEU-Aware Low-Power Memory Array

Our method uses the column-based V_{dd} memory architecture and BICS to detect transient errors and dynamically compensate for the noise in the memory cell using a high supply voltage. Our common-case strategy is to use a low voltage V_{dd} in normal standby operation and adapt with a high voltage, V_{high}, for active operation and to improve fault recovery time response. Because most memory cells spend most time in a standby mode, the low V_{dd} voltage efficiently reduces the stand-by leakage power. However, a low V_{dd} also reduces the robustness by directly increasing the memory cell recovery time due to transient errors. Figure 4 shows an example that illustrates how the recovery time depends on the supply voltage. The recovery time is faster with the high voltage than the low voltage.

In our approach, the supply voltage of a column is adjusted to V_{high} when a SEU is detected in memory cells in non-accessed columns. The low V_{dd} could be the Data
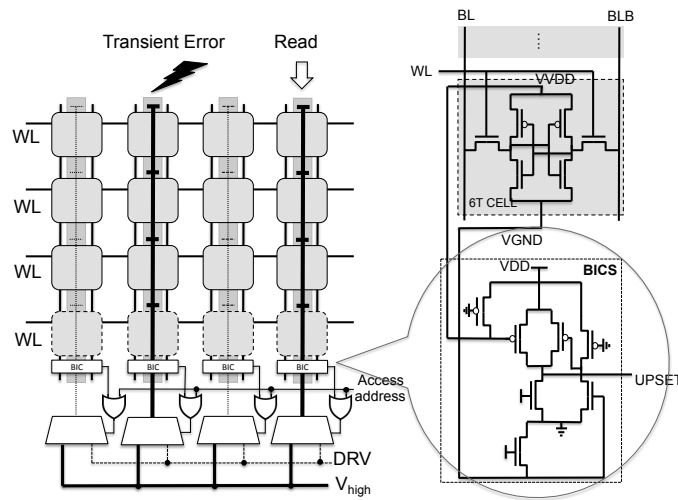
Fig. 3. Our approach uses Built-In Current Sensor (BICS) together with a Column-based V_{dd} Array to detect SEUs at a column granularity.

Retention Voltage (DRV), for example, but we need a method to improve DRV robustness.
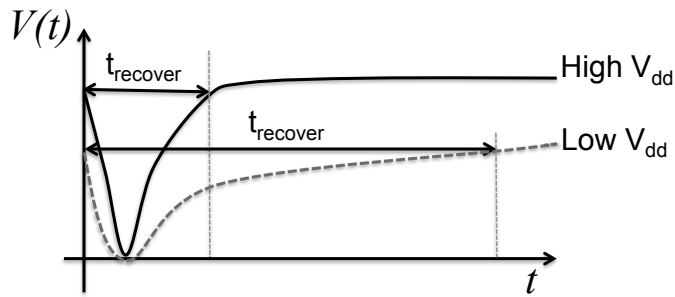
Fig. 4. Low V_{dd} reduces the cell recovery time from transient error.

Figure 3 shows the architecture using V_{high} and DRV. V_{high} is only enabled through the supply mux when (1) the SEU occurs or (2) the column address (read/write) is addressed. To do this, we add a logical *OR* operation to the bottom of each column and connect *UPSET* signal and Column Selection (*CS*) signal as inputs. For example, if the SEU occurs in column 2 and the *CS* signal for read operation accesses column 4, only

two columns are connected to V_{high}. The rest of columns are still connected to the low V_{dd}.

Our method adjusts V_{dd} of each column depending on whether a SEU is located in a column according to the BICS. This can happen in the background during idle periods. Therefore, the power consumption is reduced by only using the high supply voltage when necessary. We calculate the V_{high} supply voltages by analyzing the memory access delay constraint of a read operation and calculate the DRV using Monte-carlo SNM analysis [9]. The write operation is not directly considered, because the read operation has less noise margin and is more critical than the write operation [26].

3.2 Memory Characterization Framework
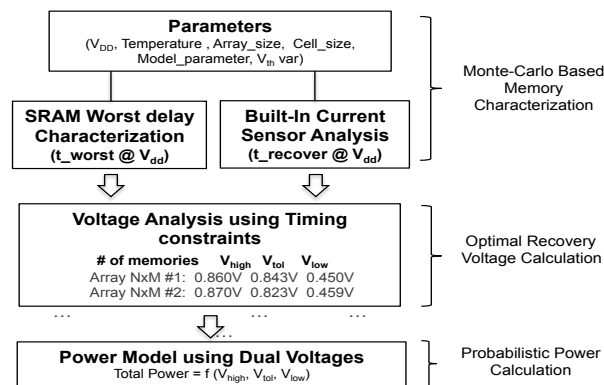
Fig. 5. A Monte Carlo framework is used to analyze the timing and power of the low and high supply voltage levels.

Figure 5 shows our Monte Carlo framework that is used to analyze the impact of SEUs on memory timing. It uses several configuration parameters to specify the supply voltage, memory size, device parameters, and transistor variation. Among the parameters, we consider V_{th} variation only for simplicity. It then executes two independent processes. One process performs worst case delay characterization during normal memory operation while the other analyzes the recovery time when performing an access with V_{high} during a SEU. Both modules internally perform a voltage sweep to study the impact of V_{dd}.

The worst case delay is a quadratic function of the supply voltage with the coefficient depending on the array size,

$$t_{worst}(V_{dd}) = f(V_{dd}, M, N). \tag{1}$$

Figure 6(a) shows this using simulation data (V_{dd} and array size $N \times M$). Similarly, the recovery time from a SEU using the BICS architecture is measured as the time required

for a memory node voltage to fully recover (99.9% of V_{dd}) using the dual voltage. This is a function of the memory column height due to the bit-line and supply rail capacitance and the supply voltage due to the memory cell drive strength,

$$t_{recover}(V_{dd}) = f(V_{dd}, DRV, N). \tag{2}$$

Figure 6(b) shows the recovery time $t_{recover}$ depending on column height N. As expected, large column height N increases $t_{recover}$ in both cases (I_{peak}=2.25E-05 and I_{peak}=6.25E-05) due to the linear increase in capacitance.
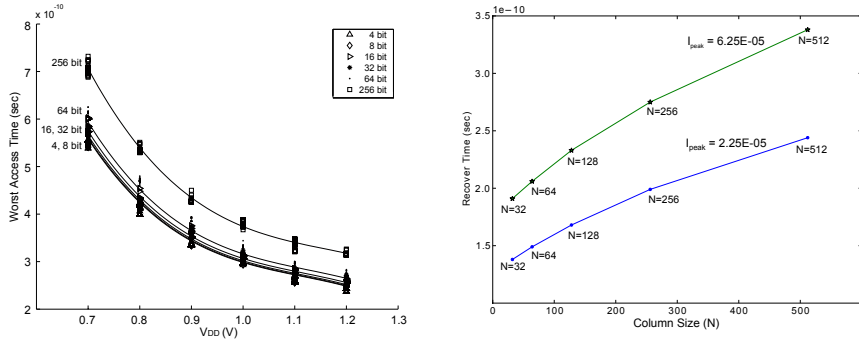
(a) Worst case delay is fit to a non-linear model for various array sizes.

(b) Plot (Column size N vs. $t_{recover}$) in different I_{peak}.

Fig. 6. The worst delay t_{worst} and the recovery time $t_{recover}$ are characterized independently in our Monte-Carlo Based framework.

Once the memory characterization step is finished, the timing constraints are used to calculate the dual supply voltages as described in Section 3.3 and then they are used to calculate the memory power as described in Section 4.

3.3 Optimal Recovery Voltage V_{high} Analysis

V_{high} must be large enough to prevent transient errors, but it should be set at a low value to preserve power. Granting that low V_{high} can reduce the power, making V_{high} too low will reduce the transistor's driving strength so that it causes read violation errors. Our method considers the recover time $t_{recover}$ of a memory cell and the worst case delay t_{worst} without a SEU as a constraint to find a proper value of V_{high}. In our feedback architecture, the *UPSET* signal is fed to a mux to adjust the voltage to V_{high}, t_{MUX} is the time required to select the V_{dd} through the mux so that node voltage can be eventually recovered when the SEU occurs. Even after the supply voltage is adjusted to V_{high}, additional time is required to increase the voltage of memory cell internal nodes. The total recovery time is

$$t_{recover} = t_{BICS} + t_{MUX} + t_{cell}. \tag{3}$$

Two of the sub-components (t_{BICS}, t_{MUX}) depend on the column height N while t_{cell} is largely determined by the supply voltage and cell driving strength.

The timing relation between t_{worst} and $t_{recover}$ is established as:

Criterion 1 *If a memory cell has a recovery time ($t_{recover}$) larger than the worst delay (t_{worst}), the memory cell can not recover from SEU.*

A proper V_{high} lower-bound must be calculated using two delay parameters ($t_{recover}$ and t_{worst}) at a given V_{dd}. In other words, the condition ($t_{recover} > t_{worst}$) will cause transient errors. Therefore, we can formulate the condition to avoid transient errors as:

$$t_{recover}(V_{dd}) \leq t_{worst}(V_{dd}). \tag{4}$$

V_{high} is the lowest V_{dd} that satisfies Equation (4) for an given I_{peak}. We can expect that different I_{peak} can change V_{high}. This will be discussed in Section 5.1.
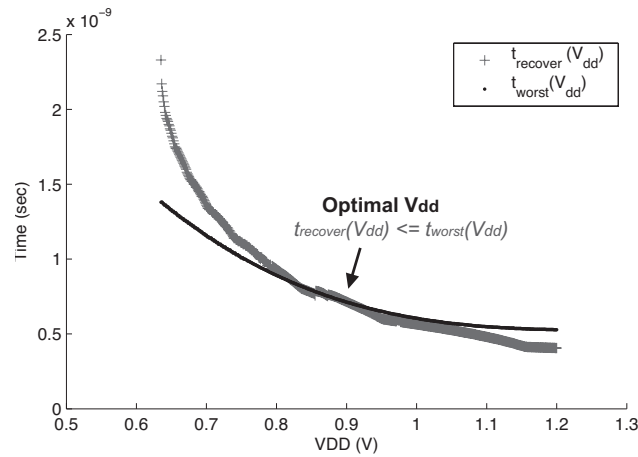
Fig. 7. Calculation of V_{high} lower-bound using t_{worst} model and $t_{recover}$ simulation with I_{peak}=*3.25E-05* shows that the criterion is satisfied around $0.9V$ in 1024K SRAM

For example, Figure 7 shows the plot of SRAM cell $t_{recover}$ and t_{worst} at various V_{high} supply voltages. Using this plot, the V_{high} lower bound condition is satisfied near V_{dd}=$0.9V$. It is interesting to note that the quadratic coefficient of the recovery time is much less than the worst case memory. This is because the higher supply voltage enables the memory cell to recover more quickly from a SEU.

4 Power Calculation

Our architecture employs a dual voltage (V_{high} and DRV) selectively depending on the SEU occurrence and active operation frequency. This means that the V_{high} duration time differs depending on the circumstances (e.g. altitude and location) due to the flux of SEUs. This can be modeled probabilistically to estimate overall memory power.

4.1 Probabilistic Power Model using V_{high} and DRV

There are several components that must be considered to compute the power of our proposed approach. First, the column-based architecture needs an additional mux in each column to select the proper supply voltage level. Also, the BICS operates independently from read/write operations to detect transient errors. The total memory power considering these issues is estimated as

$$P_{memory} = P_{array} + P_{MUX} + P_{BICS}. \tag{5}$$

where P_{array} is the $N{\times}M$ array power and denoted as $P_{array}(V_{high}, DRV)$ using a cell power P_{cell} and a ratio p and *(1-p)*. $p \in [0,1]$ means the ratio of V_{high} duration time over total transient time. Inversely, *(1-p)* means the ratio of DRV duration over total transient time.

P_{array} is calculated using one of the following approaches: In one approach, we can see the dual V_{dd} effect in a traditional row-based array, applying V_{high} and DRV to an entire array and estimate the power as:

$$P_{array} = p \cdot \sum_{i=1}^{N}\sum_{j=1}^{M} P_{cell(i,j)}(V_{high}) + (1-p) \cdot \sum_{i=1}^{N}\sum_{j=1}^{M} P_{cell(i,j)}(DRV). \tag{6}$$

In another approach, we can apply V_{high} and DRV to columns selectively and estimate the power as:

$$P_{array} = p \cdot \{P_{col}(V_{high}) + (M-1) \cdot P_{col}(DRV)\} + (1-p)M \cdot P_{col}(DRV) \tag{7}$$

In Equation (7), P_{col} shows the power consumption of a column according to

$$P_{col} = \sum_{i=1}^{N} P_{cell}(i,j) \tag{8}$$

assuming a one bit word size. Since the memory array consists of multiple bit words, Equation (7) uses the word size W to estimate the array power according to:

$$P_{array} = p \cdot \{P_{col}(V_{high}) \cdot W + (M-W) \cdot P_{col}(DRV)\} \\ + (1-p) \cdot M \cdot P_{col}(DRV). \tag{9}$$

In order to consider the power overhead of P_{MUX} and P_{BICS}, we simulate each component using the dual voltage stimulus with probabilities p and $(1-p)$ of SEUs occurring and sum up the respective power based on the corresponding memory column size M to calculate the overall power.

5 Experimental Results

All simulations use the 45nm PTM technology models [1] with a temperature of 25°C. We assume that transistors have independent $\pm 15\%/3\sigma$ variation of the nominal V_{th}.

The pull-up/pull-down SRAM transistor width size ratio is 0.5 and $\frac{PR}{CR} = \frac{90nm/45nm}{180nm/45nm}$ with identical gate lengths [11, 17]. The maximum particle flux is set to the typical ground-level total neutron (N_{flux}=56.5$m^{-2}s^{-1}$ [28]) while the cross-sectional area is assumed to be CS=0.296μm^2 [24]. We generate memories ranging from 1K-256K using a memory compiler and then calculate the worst access delay based on bit-cell location using Hspice simulation. The worst case delay model t_{worst} is fit using the Matlab command *nlinfit* due to the large t_{worst} simulation time on large memory arrays.

Our results are compared to a typical guardbanded approach. The transient error tolerant voltage V_{tol} [10] is selected such that no transient errors are expected with the given maximum particle flux.

5.1 Various peak current I_{peak} impact on V_{high}

Previous works modeled atomic spike pulse as an artificial current sources. The current sources are modeled as triangular model for simplicity [14, 18, 20, 27]. Without loss of generality, the energy particle injection occurs during very small time periods (less than ps). In reality, however, the induced peak current I_{peak} can be various depending on location, altitude and circumstance energy level.

We analyzed the various peak current I_{peak} (1.315E-5A to 3.215E-5A) impact on V_{high}. Figure 8 shows that I_{peak} has linear impact on V_{high} according to measurements on a 1K SRAM. We observed that the data can be modeled as linear equation

$$V_{high} = a \cdot I_{peak} + b \tag{10}$$

where a and b coefficients calculated from the curve fitting. Equation (10) implies
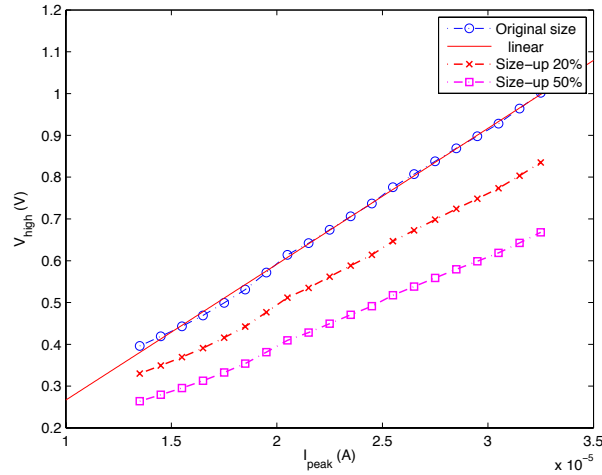
Fig. 8. Peak current's amplitude (I_{peak}) vs. V_{high} in 1K SRAM

that high supply voltage V_{high} is necessary for low SER condition (when higher I_{peak} exists) to error tolerant. If circuits are supposed to operate with low power and designed to be tolerable, the situation that V_{high} exceeds the maximum voltage limit of the design at certain I_{peak} would be a problem, because transient errors still occur even though V_{high} is applied. So we applied some techniques to avoid this situation in Section 5.2.

5.2 Transistor sizing impact on V_{high}

We also analyzed the impact of transistor sizing on reducing V_{high} in the case that a transient error happens even with V_{high} when V_{high} may exceed the voltage budget of design. In this case, we increased width(W) and length(L) size of the SRAM cell transistors while keeping the original W/L ratio of PMOS(W/L=90nm/45nm) and NMOS(W/L=180nm/45nm) to not affect t_{worst}. We observed that transistor sizing can reduce V_{high} effectively. We increased PMOS and NMOS size by 20% and compared to to the original I_{peak} and V_{high} plot as shown in Figure 8. Sizing up transistors by 50% also show a similar trend.

For example, when the original sized SRAM cell failed at V_{high}=0.9V with given peak current I_{peak}=2.915E-5 under the maximum budget V_{dd}=0.8V, sizing up transistors by 20% can satisfy the voltage budget V_{dd}=0.8V. In the 20% sized-up SRAM plots in Figure 8, the V_{high} that satisfies the SEU tolerance under I_{peak}=2.915E-5 is about V_{high}=0.75V.

5.3 Dynamic Noise Margin (DNM) for SEU Analysis

We first analyze the Dynamic Noise Margin (DNM) during an SEU. Figure 9 shows a plot with the x-axis representing the induced peak current I_{peak} and the y-axis as recovery time $t_{recover}$. Figure 9 shows three cases using dual V_{dd} ($0.4V/0.9V$, $0.4V/1.2V$, and $0.4V/1.5V$). The vertical lines are failure points. The lines show the maximum induced noise that can be tolerated given a recovery time constraint. Using this data, we can study the DNM when dual V_{dd} can aid recovery from SEUs. and we can find the optimal V_{dd} at given I_{peak}.

The DNM analysis describes whether a SEU creates a transient error or not at given I_{peak} condition. This means that we can know how dual V_{dd} schemes are tolerant to a given I_{peak}. For example, all three dual V_{dd} strategies can recover from a SEU at the condition $I_{peak} = 2.25E - 05$ although $t_{recover}$ in the case of $V_{dd} = 0.4V/0.9V$ is doubled compared to $t_{recover}$ of $V_{dd} = 0.4V/1.2V$. However, at the condition $I_{peak} = 3.75E - 05$, the $V_{dd} = 0.4V/0.9V$ case fails to recover. This means that the DNM of the memory cell determines the maximum peak noise tolerance as $I_{peak} = 3.75E - 05$ in the case $V_{dd} = 0.4V/0.9V$. Similarly, $I_{peak} = 5.45E - 05$ is the maximum peak current tolerated with $V_{dd} = 0.4V/1.2V$ and $V_{dd} = 0.4V/1.5V$.

The DNM analysis can also be used to determine the optimal V_{dd} that can tolerate a given noise I_{peak}. As expected, higher V_{dds} enable a faster recovery time. The recovery time $t_{recover}$ of the memory cell using V_{dd}=0.4V and 1.5V's is faster than the other cases at same I_{peak}. The higher V_{dd} increases the power unnecessarily although it enables the memory cell to recover more quickly. For example, both $V_{high} = 1.2V$ and $V_{high} = 1.5V$ have the same tolerance, however, the lower voltage should be selected
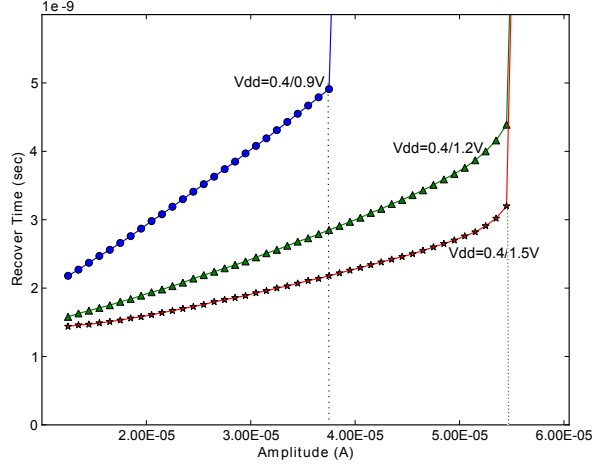
Fig. 9. Peak current's amplitude (I_{peak}) vs. $t_{recover}$ in different dual V_{dd} combinations (1K SRAM). V_{high} determines the memory tolerance to a given I_{peak} amplitude and it should be calculated to optimal V_{dd} level to reduce the power.

to save power. For this reason, the power-optimal V_{dd} should be near $V_{high} = 1.2V$ not $1.5V$.

5.4 Power Reduction

We now analyze the optimal supply voltages depending on the peak current I_{peak} that a flux generates [28]. The optimal voltages are calculated as $V_{high} = 0.948V$, $V_{tol} = 0.607V$ at a flux $N_{flux} = 56.5m^{-2}s^{-1}$ and $DRV = 0.186V$. Table 1 shows the comparison of our two strategies: 1) our proposed method with V_{high} and DRV applied to the entire array (column 3-column 4), b) our proposed method with V_{high} and DRV applied to the selected columns (column 5-column 8). The baseline is a traditional SRAM with a guard-banded error-tolerant supply voltage V_{tol} (column 2).

Table 1 and Table 2 compare proposed methods when energy particles strike the memory with probabilities $p = 0.1$ and $p = 0.2$, respectively. We assume two cases since the p value is not fixed and depends on the environment where the memory operates. It can be a large number when the radiation particles strike frequently. According to Table 1, simply applying V_{high} and DRV to the entire array can reduce the power consumption by an average of 16.38% compared to an SRAM with a guard-banded supply voltage, V_{tol}. Applying V_{high} to the column with SEU and active columns selectively reduces the power consumption by an average of 55.09% compared to the guard-banded supply voltage SRAM, V_{tol}. When particles hit the memory more frequently as shown in Table 2, the power reduction decreases to 7.03% and 49.87% com-

pared to each case in Table 1 since V_{high} is needed two times more than Table 1 to avoid errors.

Table 1. Power Reduction Results when Radiation strikes memory Once ($p = 0.1$)

Size	SRAM with V_{tol} only	Our Proposed I (V_{high},DRV to array)		Our Proposed II (V_{high},DRV to column)			
		Word size = 32		Word size = 32		Word size = 8	
	Power (W)	Power (W)	Improv.(%)	Power (W)	Improv.(%)	Power (W)	Improv.(%)
1K	3.336E-06	3.430E-06	-2.81%	2.767E-06	17.06%	2.431E-06	27.14%
4K	1.321E-05	1.115E-05	15.65%	6.963E-06	47.31%	6.293E-06	52.37%
16K	5.286E-05	3.883E-05	26.54%	1.944E-05	63.23%	1.810E-05	65.76%
64K	2.114E-04	1.363E-04	35.55%	5.836E-05	72.40%	5.572E-05	73.65%
256K	8.457E-04	5.448E-04	35.58%	2.075E-04	75.46%	2.022E-04	76.09%
	Avg. Improvement(%)		16.38%		55.09%		59.00%

Table 2. Power Reduction Results when Radiation strikes memory Twice ($p = 0.2$)

Size	SRAM with V_{tol} only	Our Proposed I (V_{high},DRV to array)		Our Proposed II (V_{high},DRV to column)			
		Word size = 32		Word size = 32		Word size = 8	
	Power (W)	Power (W)	Improv.(%)	Power (W)	Improv.(%)	Power (W)	Improv.(%)
1K	3.336E-06	3.744E-06	-12.23%	3.216E-06	3.61%	2.543E-06	23.78%
4K	1.321E-05	1.299E-05	1.68%	7.856E-06	40.55%	6.517E-06	50.69%
16K	5.286E-05	4.744E-05	10.25%	2.122E-05	59.86%	1.854E-05	64.92%
64K	2.114E-04	1.735E-04	17.92%	6.189E-05	70.73%	5.660E-05	73.23%
256K	8.457E-04	6.975E-04	17.53%	2.146E-04	74.62%	2.040E-04	75.88%
	Avg. Improvement(%)		7.03%		49.87%		57.70%

We also observe that our proposed architecture increases the power consumption in the case of small memories such as 1K, due to the additional circuitry to implement the column-based V_{dd}. The additional circuitry power overwhelms the small memory array power consumption, but in large memories (4K-256K) this cost is amortized and our architecture reduces the overall power more effectively.

In both tables, we use a 32-bit word size, but we have also performed analysis with an 8-bit word size. Smaller word sizes improve the power consumption, because our architecture enables fewer columns during active read/write operations. The background recover power of memory cells in stand-by mode is not affected by the word size.

6 Conclusions

We presented a soft-error tolerant low-power memory architecture that employs BICS in column-based V_{dd} SRAM to adaptively select from dual supply voltages. We then used a Monte Carlo framework to calculate the optimal dual supply voltages and demonstrated that our architecture can significantly reduce power compared to traditional

guard-banded static supply voltage architectures. On average, our architecture is able to reduce the power by an average of 39.5% without sacrificing error tolerance for an range of memory array sizes.

Acknowledgments

References

1. ASU. Predictive Technology Model (PTM). `http://ptm.asu.edu`.
2. K. Bhattacharya and N. Ranganathan. RADJAM: A novel approach for reduction of soft errors in logic circuits. In *VLSI Design*, pages 453 –458, Jan. 2009.
3. L. Ding and P. Mazumder. Dynamic noise margin: definitions and model. In *VLSI Design, 2004. Proceedings. 17th International Conference on*, pages 1001 – 1006, 2004.
4. L. B. Freeman. Critical charge calculations for a bipolar SRAM array. *IBM Journal of Research and Development*, 40:119–129, January 1996.
5. R. W. Hamming. Error detecting and error correcting codes. *Bell System Technical Journal*, 26(2):147–160, 1950.
6. F. Hamzaoglu, Y. Wang, and et al. Bit cell optimizations and circuit techniques for nanoscale sram design. *Design & Test of Computers, IEEE*, 28(1):22 – 31, 2011.
7. P. Hazucha and C. Svensson. Impact of CMOS technology scaling on the atmospheric neutron soft error rate. *Nuclear Science, IEEE Transactions on*, 47(6, Part 3):2586 – 2594, Dec 2000.
8. G. Huang, W. Dong, and et al. Tracing SRAM separatrix for dynamic noise margin analysis under device mismatch. In *Behavioral Modeling and Simulation Workshop, 2007. BMAS 2007. IEEE International*, pages 6 –10, sept. 2007.
9. S. Kim and M. Guthaus. Leakage-aware redundancy for reliable sub-threshold memories. In *Design Automation Conference (DAC), 2011 48th ACM/EDAC/IEEE*, pages 435 –440, june 2011.
10. S. Kim and M. Guthaus. Low-power multiple-bit upset tolerant memory optimization. In *Computer-Aided Design (ICCAD), 2011 IEEE/ACM International Conference on*, pages 577 –581, nov. 2011.
11. S. Kim and M. Guthaus. SNM-aware power reduction and reliability improvement in 45nm SRAMs. In *VLSI and System-on-Chip (VLSI-SoC), 2011 IEEE/IFIP 19th International Conference on*, pages 204 –207, oct. 2011.
12. A. Lesea, S. Drimer, and et al. The rosetta experiment: atmospheric soft error rate testing in differing technology fpgas. *Device and Materials Reliability, IEEE Transactions on*, 5(3):317 – 328, sept. 2005.
13. S. Michalak, K. Harris, and et al. Predicting the number of fatal soft errors in los alamos national laboratory's ASC Q supercomputer. *Device and Materials Reliability, IEEE Transactions on*, 5(3):329 – 335, sept. 2005.
14. P. C. Murley and G. R. Srinivasan. Soft-error monte carlo modeling program, SEMM. *IBM Journal of Research and Development*, 40(1):109 –118, Jan. 1996.
15. E. Neto, I. Ribeiro, and et al. Using bulk built-in current sensors to detect soft errors. *Micro, IEEE*, 26(5):10 –18, Sept. 2006.

16. E. Normand. Single event upset at ground level. *Nuclear Science, IEEE Transactions on*, 43(6):2742 –2750, Dec 1996.

17. A. Pavlov and M. Sachdev. CMOS SRAM circuit design and parametric test in nano-scaled technologies: process-aware SRAM design. *Springer*, Jan 2008.

18. R. Rajaraman, J. S. Kim, and et al. SEAT-LA: A soft error analysis tool for combinational logic. In *VLSI Design*, 2006.

19. P. Reviriego, J. A. Maestro, and et al. Reliability analysis of memories protected with BICS and a per-word parity bit. *ACM Trans. Des. Autom. Electron. Syst.*, 15:18:1–18:15, March 2010.

20. P. Shivakumar, M. Kistler, and et al. Modeling the effect of technology trends on the soft error rate of combinational logic. In *Proceedings of International Conference on Dependable Systems and Networks (DSN)*, pages 389 – 398, 2002.

21. F. Vargas and M. Nicolaidis. SEU-tolerant SRAM design based on current monitoring. In *Proceedings of the 24th International Symposium on Fault-Tolerant Computing (FTCS)*, pages 106 – 115, Jun 1994.

22. E. Vatajelu, A. Go andmez Pau, and et al. Transient noise failures in SRAM cells: Dynamic noise margin metric. In *Test Symposium (ATS), 2011 20th Asian*, pages 413 –418, nov. 2011.

23. J. Wang, S. Nalam, and B. Calhoun. Analyzing static and dynamic write margin for nanometer SRAMs. In *Low Power Electronics and Design. ISLPED, 2008 ACM/IEEE International Symposium on*, pages 129 –134, aug. 2008.

24. F.-L. Yang, C.-C. Huang, and et al. 45nm node planar-SOI technology with 0.296 μm^2 6T-SRAM cell. In *VLSI Technology, 2004. Digest of Technical Papers. 2004 Symposium on*, pages 8 – 9, June 2004.

25. B. Zhang, A. Arapostathis, and et al. Analytical modeling of SRAM dynamic stability. In *Computer-Aided Design, (ICCAD). IEEE/ACM International Conference on*, pages 315 – 322, nov. 2006.

26. K. Zhang, U. Bhattacharya, and et al. A 3-ghz 70-mb SRAM in 65-nm CMOS technology with integrated column-based dynamic power supply. *Solid-State Circuits, IEEE Journal of*, 51(1):146 – 151, Jan 2006.

27. M. Zhang and N. Shanbhag. Soft-error-rate-analysis (SERA) methodology. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 25(10):2140 –2155, Oct. 2006.

28. J. F. Ziegler. Terrestrial cosmic rays. *IBM Journal of Research and Development*, 40:19–39, January 1996.

29. J. F. Ziegler and W. A. Lanford. Effect of cosmic rays on computer memories. *Science*, 206(4420):776–788, 1979.