Impact of Dynamic Voltage Scaling and Thermal Factors on FinFET-based SRAM Reliability

F. R. Rosa[1], R. M. Brum[2], G. Wirth[2], L.Ost[3], and R. Reis[1]

[1]PGMICRO\PPGC Instituto de Informtica, Universidade Federal do Rio Grande do Sul
{frdarosa,reis}@inf.ufrgs.br
[2]Departamento de Engenharia Eletrica, Universidade Federal do Rio Grande do Sul
{brum,wirth}@inf.ufrgs.br
[3]Department of Engineering, University of Leicester, Leicester, United Kingdom
luciano.ost@leicester.ac.uk

***Abstract*—FinFET technology appears as an alternative solution to mitigate short-channel effects in traditional CMOS down-scaled technology. Emerging embedded systems are likely to employ FinFET and dynamic voltage scaling (DVS), aiming to improve system performance and energy-efficiency. This paper claims that the use of DVS increases the susceptibility of FinFET-based SRAM cells to soft errors under radiation effects. To investigate that, a methodology that allows determining the critical charge according to the dynamic behaviour of the temperature as a function of the voltage scaling is used. Obtained results support our claim by showing that both temperature and voltage scaling can increase up to five times the susceptibility of FinFET-based SRAM cells to the occurrence of soft errors.**

I. Introduction

The scaling of traditional CMOS technology is facing reliability and power issues due to the short channel and the leakage effects [1], [2]. In this context, semiconductor foundries, such as Intel, TSMC and Global Foundries, have been adopting the multi-gate, nonplanar transistor technology known as Fin Field-Effect Transistor, or FinFET [3], [4]. The FinFET is pointed as a promise alternative to improve the performance and energy tradeoffs, while keeping the compatibility with CMOS process [5].

Emerging embedded System-on-Chip (SoC) will employ FinFET and dynamic voltage scaling (DVS), aiming to improve system performance and energy-efficiency. DVS reduces energy consumption by lowering the supply voltage and system operating frequency. However, the use of DVS has a direct influence on the sensitivity of both CMOS-based [6]–[8] and FinFET-based SRAM cells to Single Event Upsets (SEUs).

Researches have been analyzed radiation effects as SEUs in bulk CMOS, and more recently, in FinFET, showing the susceptibility of both technologies to neutron or alpha particles strikes [9]–[11]. Excluding the work proposed in [12], other approaches consider only static variations of supply voltage and temperatures.

This paper contributes by adapting the methodology proposed in [12] to inspect the FinFET-based SRAM cells vulnerability considering critical charge under dynamic temperature and DVS variations. Summarizing, this paper contributes in the following aspects:

- inclusion of FinFET-based SRAM cell and temperature-power model calibration,
- methodology validation using 20nm PTM (Predictive Technology Model) FinFET technology, considering four voltage scaling scenarios,
- automatization and optimization of proposed methodology for CMOS and FinFET technologies.

Section II describes the proposed methodology. Obtained results considering the 20nm technology for four different voltage scaling scenarios are presented in Section III. Afterwards, conclusions and perspectives are discussed in Section IV.

II. Proposed Methodology to Evaluate critical charge ($Qcrit$) of FinFET-based SRAM cells

We develop a methodology to follow the SEU sensibility of a FinFET SRAM bit cell over a time interval. The SEU sensibilty take into account factors as temperature and supply voltage. In this context, we define SEU sensibility as the minimum amount of collected charge in the transistor junction needed to cause a bit-flip, designated critical charge ($Qcrit$).

Our proposed methodology consists on pre-characterizing the FinFET SRAM bit cell $Qcrit$ under a broad range of supply voltage and temperature conditions, thus creating a two-dimensional matrix. This profile matrix contains the $Qcrit$ information for every operating point (i.e. voltage and temperature), respecting a minimum granularity. Afterward, we correlate a temporal SRAM cell simulation in SPICE level with its SEU sensibility by accessing this profile matrix for every time step as the flow displayed by Fig. 1 (a).

The promoted methodology has five steps: The first step consists in embed a source current in the FinFET SRAM description to model the SEU fault injection by injecting a transient current pulse in the cell node, as illustrated in Fig. 1 (b). We use a double exponential current model for the fault injection, as defined in equation 1 and display in Fig. 1 (c).

To acquire the $Qcrit$ value we search for the smallest value of $\int I_p(t)dt$ by slightly changing the curve parameters,

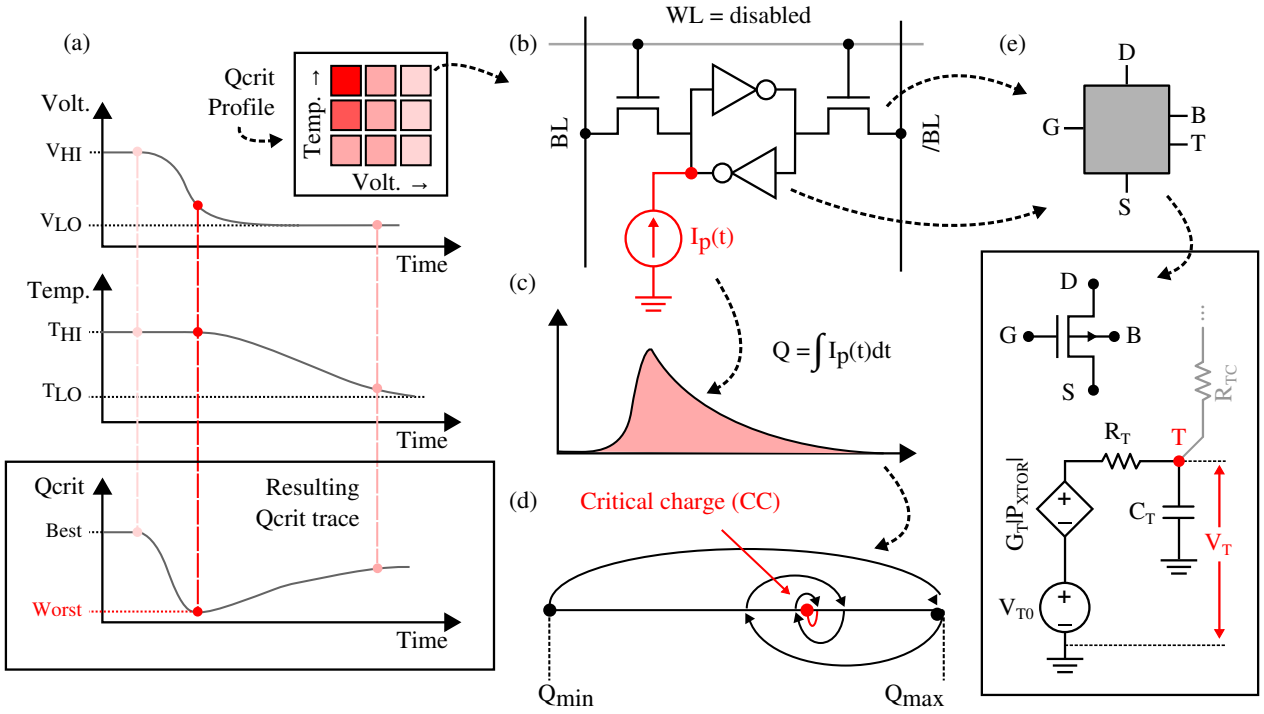

Fig. 1. Proposed methodology for profiling the critical charge: (a) Deriving time-dependent critical-charge curves using a pre-characterizing profile matix; (b) Bit-cell simulation with SEU emulated by current injection; (c) Relationship between critical charge and the injected current; (d) Binary search for the critical charge; (e) Self-heating transistor model.

consequently the total injected charge as highlighted in red(see Fig. 1 (c)). This approach enables find out the smallest amount of collected charge that triggers an SRAM bit-flip for a defined temperature and voltage supply.

$$I_p(t) = I_0(e^{\frac{t}{\tau_{fall}}} - e^{\frac{t}{\tau_{rise}}}) \quad (1)$$

Next step, we deploy the search methodology to find the $Qcrit$ for a broad range of temperature and voltage conditions, generating a $Qcrit$ profile matrix. The $Qcrit$ search employs a bisection search method. This method works by halving the search interval at each iteration until it reaches a small error from the desired value, pictured in Fig. 1 (d). This step is then repeated for each supply voltage and temperature condition, allowing us to construct a profile similar to the one represented by Fig. 1 (a), where a darker shades represents the smaller $Qcrit$ and, consequently, a more sensible cell. While the example shown in this picture is symmetrical with respect to the main diagonal, temperature and voltage may affect the $Qcrit$ differently, as shown in the next section. The worst case will be located in the upper-left corner, as higher temperatures and lower voltages are expected to reduce the $Qcrit$. However, seldom chips will work in such critical conditions, at least not for a long time. To estimate the timeframe, during which the circuit is exposed to this situation, we may combine the $Qcrit$ profile with time-dependent supply voltage and temperature curves.

Afterward in the third step, we introduce a simple self-heating model for the transistors, similar to the one proposed by Bielefeld et al. [13] to mimic a temperature behavior following the supply voltage wave, its schematic is shown in Fig. 1 (e). The passive elements in this model are related to the transistor's physical layers, the resistive elements related to the thermal conductivity, whereas capacitors are related to the specific heat and mass of these materials, and active elements will be related to the power-temperature relationship of the transistor. To obtain an accurate numerical result, a proper calibration of these elements would be required. For the purpose of proving that DVS impacts on the circuit's reliability, we are only interested in the ratio between the time constants associated with the thermal circuit (RCtemperature) and the voltage supply circuit. By using this model, the associated effect on the bit cell temperature can be obtained and crossed with the $Qcrit$ profile. To account for the power supply capacitance, we introduce a lumped capacitor in the power line.

Fourth step: We simulate the SRAM cell (attached with the power model) for a supply voltage wave to generate a temporal voltage and temperature trace. This supply voltage has two distinct phase: In the first phase, a high throughput memory access in overdrive Vdd, driving more current, and, therefore, increasing the SRAM cell temperature. Our methodology uses a 0.5 activity factor, a good average approximation (i.e. neither conservative nor pessimistic) for the least significant bits in a cache memory [14]. In the second phase, the memory

voltage supply is reduced to a low power level, which is defined according to the adopted technology (e.g. 0.5V for 20nm). The high temperature (T_{HI} in Fig. 1 (a)) during the high memory usage (V_{HI} in Fig. 1 (a)) is conducted through the chip package. At this point, the SRAM Cell has a low voltage supply and a high temperature, therefore, rising by several times its vulnerability to SEU. The period where the $Qcrit$ is smaller than expect bounds the worst case, which is defined according to the confluence of temperature and low Vdd. The energy dissipation occurs according to the $RC_{temperature}$ and its ratio with the $RC_{electric}$, which is determined by the technology construction. The last step, these two traces are analyzed at each timestep, consulting the profile matrix to acquire the $Qcrit$ at this particular temperature and voltage point. Thus creating a trace for the SEU sensibility of a FinFET SRAM bit cell as shown in Fig. 1 (a). In the next section, we present some of the obtained results for the 20nm Predictive Technology Model (PTM) from Arizona State University [15].

III. Experimental Setup and Results

In order to exploit the proposed methodology, we will present results for four different DVS scenarios. Our experimental setup uses the Cadence SPECTRE simulator to perform all experiments in an Intel Xeon L5520 2.27 GHz with 32 GB RAM. The 6T-SRAM FinFET cell use two inverters, and two passing gates as the case study. The PTM defines the height of the Fin (28nm), the thickness of the Fin (15nm), and gate length (24nm). Our cell uses a 48nm width with minimum length. Fig. 2 shows the $Qcrit$ matrix profile obtained for the 20nm PTM technology, with a resolution of 0.001 V and 0.1 C, covering temperatures ranging from -40C to 125C and supply voltages ranging from 0.45 V to 1.25 V.

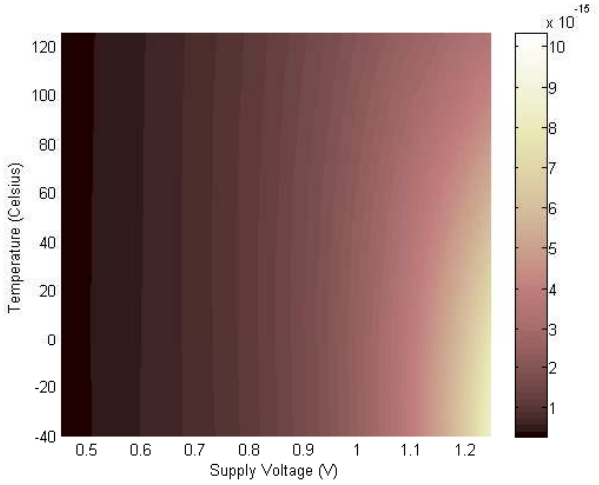

Fig. 2. Critical charge profile obtained for the FinFET 6T-SRAM cell built for the PTM 20nm technology.

From this profile matrix, it becomes apparent that the supply voltage has a stronger influence than the temperature over the considered operational ranges. Notwithstanding, while operating at higher voltages, the circuit sensibility is firmly influenced by the temperature: moving from room temperature to 125C can almost halve the $Qcrit$ of the bit cell. Indeed, higher supply voltage enhances radiation resilience as improves the noise margin, reaching five times between extreme points. Nevertheless, dynamic and static power dissipation effects are directly related to the supply voltage and therefore designers often deploy DVS to reduce its power consumption. When applying DVS, both effects of temperature and supply voltage are seen in a dynamic way where there are periods when the voltage is constant, and the temperature behaves dynamically and periods that the supply voltage and temperature are changing to adjust to the DVS.

The Fig. 3 displays (from top to down) the voltage wave, the resulting temperature from this context according to our model, $Qcrit$ value, and $Qcrit$ zoom for ratios from 1 to 100K. The first scenario (a) maintains a constant overdrive voltage of 1.2V until 10μs, consequently heating the cell, then lowering to 0.5V until the simulation ends. Where it releases the accumulated thermal energy through the chip package. As this $\frac{RC_{temperature}}{RC_{electric}}$ ratio increases the heat dissipates faster. A rapid heat dissipation reduces the coincidence between the low voltage supply and high temperature. In this scenario, shortly after the supply voltage transition has its SEU sensibility increased four times, additionally, the confluence of low supply voltage and higher temperature creates an even greater SEU sensitivity.

The second scenario (b) (see Fig. 3) adds a second overdrive period from 20μs to 30μs, consequently the SRAM experiences a reheating interval. At this moment, with lower temperatures and higher supply voltage, the SRAM cell has an overprotection (i.e. seven times) against particles strikes compared with the previous moments. This supply voltage rapidly leads to a cell warming, decreasing again the $Qcrit$. A shorter low Vdd interval will increase the minimal temperature reach while in low supply voltage. Additionally, the overdrive interval leads to a significant reheating. A longer high Vdd period reduces the cell vulnerability, besides the temperature increase as the result of the higher importance of supply voltage over temperature to the SEU sensitivity.

The third scenario (C) proposes a two-step voltage scaling, first passing to the nominal voltage (i.e. 0.9V) in 20μs and then an overdrive voltage from 30μs until 40μs. The supply voltage effect in the SEU sensitivity is not linear, as in the first transition from 0.5V to 0.9V the $Qcrit$ increases from 0.43 fC to 1.5 fC while transitioning to 1.2V its value grows three times. In the fourth scenario (d) a constant square wave with 20μs period excites the SRAM cell, nevertheless in this scenario the maximum supply voltage is the nominal Vdd. Higher $\frac{RC_{temperature}}{RC_{electric}}$ ratio increases the temperature swing, increasing SEU sensitivity unpredictability.

IV. Conclusion

In this work, we use a methodology to determine the $Qcrit$ of FinFET SRAM bit cells under voltage scaling, considering

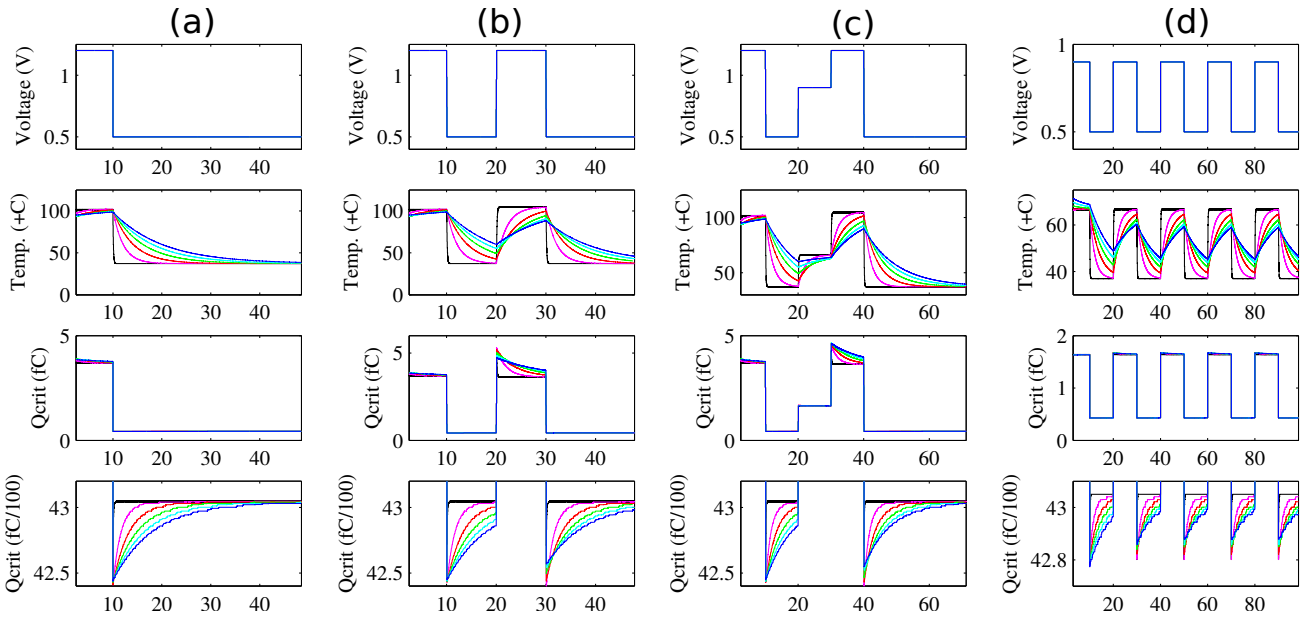

Fig. 3. (from top to down) voltage waveform, temperature, critical charge, and a critical charge zoom for proposed DVS scenarios using our method.

transient temperature effects. By introducing a temperature-power model, we demonstrated that FinFET-based SRAM cells are more sensitive to radiation when the supply voltage is transitioning from a higher to a lower voltage, reducing the $Qcrit$ few times. SEU sensitivity is highly dependable on parameters as temperature, supply voltage and their temporal confluence, as see in Fig. 3. Embedded systems dynamically adjust the supply voltage according to these parameters, normally controlled by off-chip management units, which have little or no configuration options available to the user.

Although DVS impacts directly the SEU sensitivity, its deployment is imperative in nowadays embedded systems, which are mostly battery-driven devices. The use of DVS may expose particular code fragments to the occurrence of soft errors. In this context, the system engineer requires a greater understanding of the DVS impact on the soft error sensitivity for a given system. Solutions for this challenge may include either delaying supply voltage transition for radiation-hardened circuits, or using a gradual step-based supply voltage transition. In the future, we intend to study new approaches to radiation-aware voltage scaling.

REFERENCES

[1] J. Qin, S. Chen, and J. Chen, "3-D TCAD simulation study of the single event effect on 25 nm raised source-drain FinFET," *Science China Technological Sciences*, vol. 55, no. 6, pp. 1576–1580, Feb. 2012.

[2] L. Artola, G. Hubert, and M. Alioto, "Comparative soft error evaluation of layout cells in FinFET technology," *Microelectronics Reliability*, vol. 54, no. 910, pp. 2300–2305, Sep. 2014.

[3] B. Doyle, B. Boyanov, S. Datta, M. Doczy, S. Hareland, B. Jin, J. Kavalieros, T. Linton, R. Rios, and R. Chau, "Tri-Gate fully-depleted CMOS transistors: fabrication, design and layout," in *2003 Symposium on VLSI Technology, 2003. Digest of Technical Papers*, Jun. 2003, pp. 133–134.

[4] C. Russ, "ESD issues in advanced CMOS bulk and FinFET technologies: Processing, protection devices and circuit strategies," *Microelectronics Reliability*, vol. 48, no. 89, pp. 1403–1411, Aug. 2008.

[5] J. Kawa and A. Biddle, "FinFET: The Promises and the Challenges." [Online]. Available: http://www.synopsys.com/COMPANY/PUBLICATIONS/SYNOPSYSINSIGHT/Pages/Art2-finfet-challenges-ip-IssQ3-12.aspx

[6] R. Baumann, "Soft errors in advanced computer systems," *IEEE Design Test of Computers*, vol. 22, no. 3, pp. 258–266, May 2005.

[7] G. Tsiligiannis, E. Vatajelu, L. Dilillo, A. Bosio, P. Girard, S. Pravossoudovitch, A. Todri, A. Virazel, F. Wrobel, and F. Saigne, "SRAM soft error rate evaluation under atmospheric neutron radiation and PVT variations," in *On-Line Testing Symposium (IOLTS), 2013 IEEE 19th International*, Jul. 2013, pp. 145–150.

[8] E. Vatajelu and J. Figueras, "Robustness analysis of 6t SRAMs in memory retention mode under PVT variations," in *Design, Automation Test in Europe Conference Exhibition (DATE), 2011*, Mar. 2011, pp. 1–6.

[9] G. Hubert and L. Artola, "Single-Event Transient Modeling in a 65-nm Bulk CMOS Technology Based on Multi-Physical Approach and Electrical Simulations," *IEEE Transactions on Nuclear Science*, vol. 60, no. 6, pp. 4421–4429, Dec. 2013.

[10] G. Hubert, L. Artola, and D. Regis, "Impact of scaling on the soft error sensitivity of bulk, FDSOI and FinFET technologies due to atmospheric radiation," *Integration, the VLSI Journal*, vol. 50, pp. 39–47, Jun. 2015.

[11] L. Artola, G. Hubert, and R. Schrimpf, "Modeling of radiation-induced single event transients in SOI FinFETS," in *Reliability Physics Symposium (IRPS), 2013 IEEE International*, Apr. 2013, pp. SE.1.1–SE.1.6.

[12] F. R. Rosa, R. M. Brum, G. Wirth, F. Kastensmidt, L. Ost, and R. Reis, "Impact of dynamic voltage scaling and thermal factors on SRAM reliability," *Microelectronics Reliability*, vol. 55, no. 910, pp. 1486–1490, Aug. 2015.

[13] J. Bielefeld, G. Pelz, H. Abel, and G. Zimmer, "Dynamic SPICE-simulation of the electrothermal behavior of SOI MOSFET's," *IEEE Transactions on Electron Devices*, vol. 42, no. 11, pp. 1968–1974, Nov. 1995.

[14] H. Homayoun, M. Rahmatian, V. Kontorinis, S. Golshan, and D. Tullsen, "Hot peripheral thermal management to mitigate cache temperature variation," in *2012 13th International Symposium on Quality Electronic Design (ISQED)*, Mar. 2012, pp. 755–763.

[15] Y. K. Cao, "What is Predictive Technology Model (PTM)?" *SIGDA Newsl.*, vol. 39, no. 3, pp. 1–1, Mar. 2009.