Scaling Effects on Neutron-Induced Soft Error in SRAMs Down to 22nm Process
Eishi Ibe, Fellow, IEEE, Hitoshi Taniguchi, Yasuo Yahagi, Ken-ichi, Shimbo, and Tadanobu Toba

Abstract—Trends in terrestrial neutron-induced soft-error in SRAMs from 130nm to 22nm process are predicted by using the Monte-Carlo simulator CORIMS, which is validated to have less than 20% variations from experimental soft-error data on 250-130nm SRAMs in a wide variety of neutron fields like field testings at low and high altitudes and accelerator testings in LANSCE, TSL and CYRIC.

The following results are obtained:
(1) Soft-error rates per device in SRAMs will increase x7 from 130nm to 22nm process.
(2) As SRAM is scaled down to smaller size, soft error rate is dominated more significantly by low energy neutrons (<10MeV).
(3) The area affected by one nuclear reaction spreads over 1M bits and bit multiplicity of multi-cell upset become as high as 100 bits and more.

Index Terms—single event upset (SEU), multi-cell upset (MCU), SRAM, CORIMS, scaling, bit multiplicity, multi-node upset (MNU)

I. INTRODUCTION

Scaling down of semiconductor devices to sub-100nm technology encounters a wide variety of technical challenges like Vth variation[1], NBTI[2], short-channel effect[3], gate leakage[4] and so on. Terrestrial neutron-induced single event upset (SEU) are one of such key issues that can be a major setback in scaling. In particular, “multi-cell upsers (MCUs)”, which are defined as simultaneous errors in more than one memory cell induced by a single event, have been under close scrutiny[5-9]. The concept of the MCU, therefore, contains both upsets that can be corrected by error detection/correction code (EDAC/ECC) as well as those which cannot. The latter is called “multiple bit upset” or “multi-bit upset” (MBU) of memory cells in the same word, and can lead, for example, to hang-ups of computer systems. This means that an MBU is a subset of an MCU. Though MBUs can be avoided by a combination of ECC and the interleaving technique [9], MCUs that can be corrected by EDAC/ECC can still be problematic in high performance devices such as contents addressable memories (CAMs) used in network processors and routers. In the case of system design, it is therefore very important to evaluate MCUs as well as soft-error rates (SER) of the device in advance.

In addition, MCU can be a threat in mission-critical systems that are mainly protected by spatial or time redundancies. Typically redundancy circuits like TMR, DMR and DICE cannot be effective when relevant nodes are corrupted simultaneously by an MCU. Since such redundancy systems in electronic systems are strictly relevant to the international standard IEC61508 that defines the functional safety of electrical /electronic/ programmable electric safety related systems, protection technologies against MCUs may have to be consistent with the scope of the standard.

Historically, MCUs are understood as taking place as a result of collection of charges produced by a secondary ion from nuclear spallation reaction in a device. As device scaling down proceeds, novel MCU modes are being reported as “charge sharing among memory storage nodes in the vicinity [8,10] or bipolar effects in p-well [9,11,12]. Ibe et. al. have proposed multi-coupled bipolar interaction (MCBI) that is regarded as a parasitic thyristor effect triggered by a single event snapshot in the p-well and causes MCU bit multiplicity of more than 10 bits [9]. It is also reported that MCU physical address pattern differs depending on written data patterns typically ALLX (All”1” or All”0”) or Checkerboard (CB or its compliment CBc).

In this paper, the statistics in SEUs and MCUs in SRAMs are predicted down to 22nm process by using the Monte-Carlo simulator CORIMS [13].

In section II, the physical model, major algorithms and statistical parameters are reviewed. In section III, simulation results are presented and discussed in conjunction with impacts on logic devices. Section IV concludes the insights from the simulation results.

At present, simulation results are limited only relevant to the charge collection mode, but simulation results with simplified bipolar models typically MCBI shall be included in the final version.

II. MODEL DESCRIPTION

A. Nuclear Spallation Reaction Models

Monte Carlo single event simulator CORIMS (Cosmic Ray Impact Simulator) [13] is equipped with numerical solutions for nuclear spallation reactions of silicon, ion track analysis in an infinite layout of a semiconductor device, and charge collection to the diffusion layer of the device. The model of the nuclear spallation reaction is based on the intra-nuclear cascade (INC) model and the evaporation model by Weisskopf and Ewing. The INC model is applied to prompt collision process,

Eishi Ibe, Hitoshi Taniguchi, Yasuo Yahagi, Ken-ichi, Shimbo, and Tadanobu Toba are with Production Engineering Research Laboratory, Hitachi, Ltd., Yokohama, Kanagawa, Japan. (e-mail: hidefumi.ibe.hf@hitachi.com, hitoshi.taniguchi.dn@hitachi.com, yasuo.yahagi.sg@hitachi.com, kenichi.shimbo.tu@hitachi.com, tadanobu.toba.ee@hitachi.com)
where many-body collision among nucleons (neutron and proton) is treated numerically as a cascade of relativistic binary collisions between two nucleons in the target nucleus. The evaporation model of light particles from excited nucleus is also applied for delayed nuclear reaction process, where nucleons (n and p), deuterons (\(^2\)H or D), tritons (\(^3\)H or T), helium and residual nucleus are released. The inverse reaction cross section necessary for determination of an evaporation channel (a set of evaporated light particle and residual nucleus) is calculated based on the GEM model [14]. Nucleus type, energy, and direction of each secondary ion produced in a spallation reaction are thus determined and reaction locations are randomly set in the device model.

Accuracy of nuclear reaction model is validated through comparison of nuclear reaction data of high-energy proton and aluminum. SER in the device under any neutron spectra can be simulated. In the case of simulation at a specific location at ground level on the Earth, the terrestrial neutron spectrum at the location is corrected in accordance with the geomagnetic latitude and the altitude [15].

B. SRAM Device Model

The model layout of MOSFET SRAM cell is illustrated in Fig.1. Since the active regions are isolated by STI oxide in lateral direction and wells line up across the word lines, charge collection in the lateral direction is tightly limited in the present device. Bits in a word are aligned along a word line so that MBUs in this device is tightly limited eventually.

C. Cell Matrix Model

Naturally, a model with a number of fixed physical cell models may be applied to investigate MCU effects. Such a method, however, has an inherent limitations on the memory and speed of the simulations.

We have developed a dynamic cell-shift (DCS) method to overcome such limitations. Figure 2 shows the basic idea of the dynamic cell-shift method.

When an ion crosses a memory cell matrix along the line A-B-C-D, the track of the ion may be traced as long as the ion has a possibility to hit the sensitive components. This method requires a cell-matrix that is wide enough compared to the ion range. The proposed method does not need an actual cell matrix. Instead, it utilizes only one physical cell model. When an ion reaches the boundary at B, for example, the track is virtually shifted to B’-C’ by using a shift in the Y-coordinates and physical address of Y direction Yad is incremented by 1 bit. Similarly, when the actual ion crosses C-D, virtual track is shifted to C’-D’ and the physical addresses of X and Y direction are incremented by 1 bit from the original address. In this way, any ion track can be traced until it stops, regardless of the length of the ion’s range. In the present device, the condition of data “1” or “0” corresponds to the position (left or right) of “high” node in one bit of SRAM and the layout of the SRAM is symmetry to its center, all “1” and all “0” have the same feature and susceptibility to neutron impacts.
D. MCU Classification

As proposed in the MCBI experimental analyses[9], the following MCU classification rules are also applied in CORIMS.

(1) MCU pattern is classified into three basic categories like a single line along BL (category “b”), a single line along WL (category “w”) and cluster (an MCU that has two or more bits along with both BL and WL directions; category “c”).

(2) MCU code that can be almost uniquely relevant to physical address pattern in an MCU is given as:

\[ C_{N1}N2N3N4_P \]

- **C:** category (b/w/c);
- **N1:** MCU size (\(N3 \times N4\));
- **N2:** bit multiplicity in an MCU;
- **N3:** width in the BL direction (bits);
- **N4:** width in the WL direction (bits);
- **P:** parity (A1: initial data in an MCU bits are all “1”; A0: initial data are all “0”; MX: initial data are a mixture of “0” and “1”).

E. Recycle Simulation Method

In extreme cases, CPU time may exceed several days. This makes parametric survey study difficult in wide scope. To cope with this problem, CORIMS saves the virtual single events with extremely low critical charge with a certain input conditions, and re-runs later to recycle them for parametric survey on the effects of different critical charge, data pattern, and interleaving within one hour CPU time.

F. Roadmap

Table 1 summarizes the typical roadmap parameters in 20-130nm SRAM assumed based on ITRS2007. Lateral two-dimensional scaling is assumed to reduce area by a factor of 2 by each generation. Depth profile is assumed to be constant due to lack in the roadmap information. As parasitic capacitance is basically in proportion to device area, critical charge is also assumed to decrease by a factor of 2 by each generation. Although \(V_{dd}\) is assumed to be constant, the critical charge will decrease more rapidly if the \(V_{dd}\) is reduced by generation.

G. Validation of SRAM Model

SRAM models in CORIMS has been validated to have less than 20% variations from experimental data of 250-130nm SRAMs in a wide variety of neutron fields like field testings [16] and accelerator testings in LANSCE[17], TSL[18] and CYRIC [19]. Figure 4 shows such an example of justification of 130nm SRAM simulation with measured data in three different locations in Japan [16].

![Fig. 3 Data pattern implementation in CORIMS.](image)

![Fig. 4 Comparison of Measured and Simulated SER in 180nm SRAM at three different location in Japan [16]](image)
any ultra-high-speed automatic classification tool. In neutron accelerated testing for 45-22 nm SRAM would become very painful or almost impossible task unless there is perfect soft-error immune. Heavy ions cause high density charge deposition but their frequencies are relatively low. Lighter particles (proton and alpha particle) play major roles for the deposition density below 10fC/μm.

The frequencies of charge deposited at the boundary of the storage node by secondary ions are shown in Fig.5 for proton, alpha particle, heavier particles (atomic number is 10 or more) and total particles. Basically, there are no differences in the shape of spectra with generation with maximum deposition density of 110fC/μm (not shown in Fig.4). This means that any device which can tolerate the density 110fC/μm can be perfectly soft-error immune. Heavy ions cause high density (10-110fC/μm) charge deposition but their frequencies are relatively low. Lighter particles (proton and alpha particle) play major roles for the deposition density below 10fC/μm.

By contrast, the soft error susceptibility improves only slightly when the critical charge increases from 5fC to 10fC for 130nm, but the change in the critical charge of 1 → 2 → 4 → 10fC improves the susceptibility by one order of magnitude for each step for 22nm SRAM. Proton and alpha particle play major role when the critical charge is low. This range in the deposited charge becomes narrower as scaling proceeds.

D. Failed Bit Map (FRM)

Figure 7 shows the distribution of failed bit map in the BL (perpendicular axis) and WL (vertical axis) address space when nuclear reaction takes place in the four bits near the origin for the data pattern CB. It is seen that the area densely affected drastically increases from 130nm (about 50x50 bits) to 22nm (about 500x500). This implies that making a statistics for MCU in neutron accelerated testing for 45-22 nm SRAM would become very painful or almost impossible task unless there is any ultra-high-speed automatic classification tool.

B. Charge Deposition Density for Secondary Ions

The frequencies of charge deposited at the boundary of the storage node by secondary ions are shown in Fig.5 for proton, alpha particle, heavier particles (atomic number is 10 or more) and total particles. Basically, there are no differences in the shape of spectra with generation with maximum deposition density of 110fC/μm (not shown in Fig.4). This means that any device which can tolerate the density 110fC/μm can be perfectly soft-error immune. Heavy ions cause high density (10-110fC/μm) charge deposition but their frequencies are relatively low. Lighter particles (proton and alpha particle) play major roles for the deposition density below 10fC/μm.

By contrast, the soft error susceptibility improves only slightly when the critical charge increases from 5fC to 10fC for 130nm, but the change in the critical charge of 1 → 2 → 4 → 10fC improves the susceptibility by one order of magnitude for each step for 22nm SRAM. Proton and alpha particle play major role when the critical charge is low. This range in the deposited charge becomes narrower as scaling proceeds.

C. Total Charge Deposited to Storage Node

Figure 6 shows the spectra of the total charge deposited to the storage nodes for 22nm and 130nm SRAMs. When the deposited charge exceeds the critical charge, SER takes places. In contrast to the charge deposition density, there are differences among different generations. Maximum deposited charge decreases from 130nm SRAM (36fC) to 22nm SRAM (20fC), but the difference is not so large (16fC).

By contrast, the soft error susceptibility improves only slightly when the critical charge increases from 5fC to 10fC for 130nm, but the change in the critical charge of 1 → 2 → 4 → 10fC improves the susceptibility by one order of magnitude for each step for 22nm SRAM. Proton and alpha particle play major role when the critical charge is low. This range in the deposited charge becomes narrower as scaling proceeds.

D. Failed Bit Map (FRM)

Figure 7 shows the distribution of failed bit map in the BL (perpendicular axis) and WL (vertical axis) address space when nuclear reaction takes place in the four bits near the origin for the data pattern CB. It is seen that the area densely affected drastically increases from 130nm (about 50x50 bits) to 22nm (about 500x500). This implies that making a statistics for MCU in neutron accelerated testing for 45-22 nm SRAM would become very painful or almost impossible task unless there is any ultra-high-speed automatic classification tool.

E. Energy Dependency of SEU/MCU Cross Section

SEU and MCU cross sections for each generation are shown as a function of neutron energy in Figs.8 and 9, respectively. As scaling proceeds, the contribution of neutrons with energy lower than 10MeV drastically increases due to increase in contribution of lighter particles as scaling proceeds. This implies that two essential changes may be needed in the standard methods including JESD89A to estimate SER from accelerator-based testing. Namely,

1) To include the contribution of neutrons with energy lower than 10MeV to avoid large error in SER estimation when the spallation neutron sources are used, and
2) The ordinary excitation function with saturated cross section should be modified to have a sharp peak at low neutron energy when the (quasi-) monoenergetic neutron sources are used.
By contrast, there are not essential changes in MCU cross section shapes. This can be understood from the fact that the contribution of lighter particle to MCU is relatively low.

F. Trends in MCU Ratio

Figure 10 shows the trends in MCU ratio to total SEU. The ratio generally increases as neutron energy increases and scaling proceeds. The maximum ratio exceeds as high as 0.5 for 22 nm SRAM.

G. Trends in MCU Multiplicity

Figure 11 shows the changes in multiplicity in MCUs. It is seen that the multiplicity shifts to larger number of bits as scaling proceeds. The maximum multiplicity is well beyond tens of bits as mentioned before.

H. Validity of Simulated Results

In the present model, the depth profile of impurities and the maximum funneling length are fixed for all generation. But in reality, depth profile will be shallower. Funneling length will also be shorter as concentration of impurities become higher. These effects would work for suppressing SER. On the other hand, operation voltage will be reduced in reality as scaling proceeds. This works for worsening SER.

Change in the material in the device would make wider variation in the prediction. If the high-\(k\) material is used for gate oxide like HfO, the critical charge is increased to result in lower SER. Meanwhile, if the low-\(k\) material used for interlayer oxide, parasitic capacitance is reduced to result in lower critical charge and higher SER.

The bipolar effects, which are not implemented in CORIMS at present, are somewhat in the trade-off relationship with the charge correction mode. When the operation voltage is reduced, the bipolar mode would decrease. When p-well size is shrunk, charge collection mode would be minor but bipolar mode would be activated due to shrinkage of distance of p-n junctions. The bipolar effects are more realistically implemented in the final version of this manuscript but it is noteworthy that charge correction mode only does have serious impacts on future silicon industry.

I. Insights in the Impact on Logic Devices

It is reported by a number of researcher that SER in logic devices will become serious including impacts on PLL[24] or clock line[8]. As mentioned before, MNU can be major setback in reliability design of logic devices and systems. Interactive countermeasures among device/component/system layers have to be established to cope with the real threats in the future. Because perfect countermeasures in each isolated layer may not be realized, interactive countermeasure is very important. Such countermeasures should be consistent with the scope of relevant international standards such as IEC61508 for the functional safety of electrical/electronic/programmable electronic safety-related systems.

IV. CONCLUSIONS

Trends in terrestrial neutron-induced soft-error in SRAMs down to 22nm process are predicted by using the Monte-Carlo simulator CORIMS, which is validated to have less than 20%
variations from experimental data in a wide variety of neutron fields like the low and high altitude field testings and the accelerator testings in LANSCE, TSL and CYRIC.

The following results are obtained:

1. Soft-error rates per device in SRAMs will increase x7 from 130nm to 22nm process.

2. As SRAM is scaled down to smaller size, SEU is dominated more significantly by low energy neutrons (<10MeV). But MCU does not change drastically.

3. The area affected by one nuclear reaction spreads well beyond 1M bits and bit multiplicity of multi-cell upset become as high as 100 bits and more.

The discussions are extended to the MNU of logic devices and systems. Development of interactive countermeasures among device/component/system layers, which will be consistent with IEC61508, is proposed to cope with the real threats in the future.

V. ACKNOWLEDGMENT

We gratefully acknowledge Profs. Emeritus T. Nakamura and M. Baba of Tohoku Univ. for helpful discussions and support for the database on nuclear reactions.

VI. REFERENCES


