On Techniques for Handling Soft Errors in Digital Circuits

Warin Sootkaneung and Kewal K. Saluja
Department of Electrical and Computer Engineering
University of Wisconsin-Madison
Madison, WI 53706, USA
sootkaneung@wisc.edu, saluja@ece.wisc.edu

Abstract

Dealing with soft errors due to particle strikes is the next major challenge in implementing digital systems. This study thoroughly investigates the effect of device size on circuit soft error rate and identifies methods to reduce soft error rate in combinational circuits. In particular, we propose three novel methods that upsize only selected gates and /or transistor networks. In order to obtain the most appropriate technique for soft error rate reduction in small technology node circuits, we conduct experiments and compare the results for several upsizing techniques including all gates, selected gates and transistor networks based on their fault sensitivities, and parallel networks with soft error rate saturation consideration. Consequently, it is discovered that some upsizing scenarios perform large improvement whereas others do not or even increase the soft error rate. The use of fault sensitivity analysis approach with parallel transistor network upsizing based on the contribution of each sensitive gate can reasonably reduce overall circuit sensitivity. Experimental results show an average reduction in soft error rate about 20% with a very small area overhead of 2% for benchmark circuits using our technique.

1.   Introduction

Soft errors are transient errors that cause incorrect operation of a digital circuit. Two major sources of soft errors are alpha particles emitted from radioactive impurities in device packages and cosmic rays carrying neutrons [1], [2], [3]. These particles may strike sensitive parts of the circuit and induce transient glitches at primary outputs. In the past, high energy alpha particles were predominant source of soft errors. However, due to improvements in packaging technology materials, soft errors caused by alpha particle strikes can almost be neglected. Nonetheless, as technology nodes are continuously being scaled, digital circuits are turning more vulnerable even to weak particle or neutron strikes. As a result, soft errors are becoming one of the most problematic issues. There is an urgent need to develop novel techniques to combat this problem.

Addressing soft error problems is even more crucial to life-critical computing systems, such as avionic, biomedical, and spacecraft systems. Many techniques for improving soft error immunity in digital circuits on either sequential or combinational parts have been proposed. A register file (RF) is normally in the critical path of a processor and more importantly, it is the hottest part, so adding protection scheme as in main memory parts is too costly and unacceptable. Pure software and compiler optimization approaches to reduce soft errors in the RF were introduced in [4], [5], [6]. Conventional memory structures such as caches and main memories are well protected using the Error Correcting Code (ECC) [1]. In embedded systems which require high performance operation for given applications, hardware based techniques for soft error protection in memory were proposed in [7], [8]. In addition, Biswas et al. [9] conducted measurements of soft error rate using high-energy proton beams and performed a detailed simulation on write-back caches to investigate architectural root cause of a super-linear increase in detected unrecoverable errors (DUEs) when the cache size is increased. In combinational circuits, it is more challenging to combat soft error effects since neither coding protection nor software based approaches can easily be applied on these parts. Additional hardware is required for almost all techniques for soft error resilience in combinational units. Those techniques include inserting cross-coupled pairs [10], use of low threshold voltage devices [10], flip-flop selection [11], [12], filtering [13], and resizing [10], [11], [14], [15]. Apart from reliability improvement of the target circuit, a satisfactory power, delay, and area overheads are of concern. Unfortunately, overhead (area, performance, power) used for this task is still relatively high. The study of gate-level soft error mitigation techniques for neutron-induced errors in [10] reveals that upsizing technique provides an impressive power and delay performance for reliability enhancement per unit area. The above advantage of the upsizing technique also motivated us to investigate redesign methods through sizing which can boost the circuit reliability with low area overhead. Redesigning sensitive parts of a circuit is a viable alternative. On the other hand, soft error rate simulators have gradually been developed to be used as dependable tools for circuit reliability improvement task.

In this paper, we investigate methods for soft error reduction in experimental benchmark circuits from ISCAS'85, ISCAS'89 (combinational parts) and ITC suit. We employ our previously proposed simulator [15] which

can take into account the effect of transistor positions and gate input patterns on soft error rate estimation. The gate error rate acquired from the simulator provides us information regarding gate sensitivity. This information can be used to distribute the overhead budget to the selected candidate devices for enhancing the reliability. In our experiment, we also evaluate our proposed resizing techniques that upsize selected gates or transistors in the complementary network of each sensitive gate based on the *probability of failure* (*POF*). In addition, we assess our novel method that takes the saturation of gate *POF*, which may occur when some transistor networks are over upsized, into account. Our experiments on various submicron benchmark circuits show that with a proper sensitive gate selection and upsizing weight distribution, upsizing transistor networks with *POF* saturation consideration yields larger soft error immunity improvement compared to the techniques which upsize all devices in a circuit and in sensitive gates.

This paper is organized as follows. Section 2 provides an overview of the previously proposed soft error modeling and mitigation techniques. Section 3 gives the details of techniques for mitigating soft errors. In Section 4, the experimental results on a number of benchmark circuits are presented. Section 5 contains discussions and some remarks and we conclude this research in Section 6.

2. Related Work

A soft error caused by a particle strike can be modeled as a current source injecting into the drain of a transistor as shown in Figure 1. The injection current source model is based on two types of particles: alpha and neutron. For alpha-particle strikes, a double exponential current source proposed in [16] is widely used. In this study, since we consider the soft error caused by a neutron strike, the current source model consists of a single exponential term as shown in equation (1) below [17], [18].

$$I(t) = \frac{Q}{T}\sqrt{\frac{t}{T}}e^{-t/T},\qquad(1)$$

In this equation, Q and T are the amount of charge deposition and the time constant for the charge collection, respectively. The direction of the current depends on whether the strike happens on a PMOS or NMOS device [19]. A gate is assumed to be failed if the amount of charge Q in (1) is equal to or larger than the *critical charge* (Q_{crit}) because this charge will alter the output voltage above or below $V_{DD}/2$.

We performed extensive SPICE simulation to determine Q_{crit} under variation in device size, similar to the experiment conducted in [20]. However, since the amount of Q_{crit} of each transistor in a gate is dependent on the input patterns [15], we consider all possible input combinations to a gate for the determination of its Q_{crit}. The upsizing method directly relies on the increase of Q_{crit} and it does not substantially affect the delay performance of the circuit. Hence, many sizing techniques have been

proposed that use upsizing [10], [11], [14], [15], [18], [19], [21], [22], [23], [24], [25]. However, their major drawback is that they offer relatively low reliability improvement per unit area overhead.

In any case, after Q_{crit} of each device is obtained, we use the energy transfer model for silicon from [26] to trace back to the energy of the strike which is subsequently mapped to the neutron flux. In other words, we can determine the neutron flux which contains sufficient energy to produce Q_{crit} and potentially causes a circuit to fail. The neutron flux information is available from the Joint Electron Device Engineering Council (JEDEC), Solid State Technology Association standard (JEDEC89A) [27]. Our previously proposed sizing technique [15] provided high accuracy in soft error rate estimation using an *input/position-dependent* gate level model for soft errors. We define *Tr POF$_{i(t, j)}$* as the *POF* of transistor t for an input vector j of gate i. This can be calculated as in (2) below.

$$Tr\ POF_{i(t,j)} = \frac{1}{k} * \emptyset_{i(t,j)} * Ad_{i(t)} * w_{i(t)} * E_{i(j)}\qquad(2)$$

where k is the total number of simulated input vectors, $\emptyset_{i(t,j)}$ is the total probability of strike in term of total neutron flux [27] producing the charge deposition above Q_{crit} of transistor t and input vector j of a gate i, $Ad_{i(t)}$ is the drain area of transistor t, $w_{i(t)}$ is the weight factor defined as the ratio of the active area to the circuit area, and $E_{i(j)}$ is the error count which is accumulated by +1 if input j of gate i suffers from a neutron strike and the bit flip at the gate output can propagate to the circuit primary output. The error count indicates the logical masking probability of a circuit for given workload. In this study, random inputs to the circuit are provided in order to estimate the error count. However, it is possible to use workload traces without any additional computation effort in the algorithm presented in this paper if such traces are available. In any case, assuming that only one transistor fails at a time, the *POF* of a gate (a circuit) can be achieved by summing the *POF* of each transistor in the gate (the circuit).
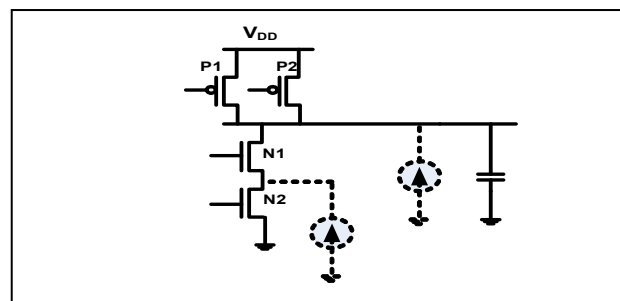

Figure 1   Neutron strike models at target transistor drains in a two-input NAND gate.

Deogun, Sylvester, and Blaauw [10] presented a reliability improvement technique using a combination of upsizing, inserting a cross-coupled pair, and reducing threshold

voltage. Their study carried out an analysis of soft error rate in a 10-stage 2-input NAND chain and concluded that the upsizing protection scheme still has small degree of improvement.

The study of *fault sensitivity* in [14] used SPICE level modeling to investigate *POF* of various blocks in analog-to-digital converters. Only a small number of blocks which are more sensitive, or have higher *POF*, are resized; as a result, this local redesign can efficiently enhance the circuit soft error immunity with low area overhead. However, use of SPICE at block and circuit level is far too expensive in computation time and such an approach cannot be applied to large digital circuits.

Sizing techniques with mathematical optimization formulation can simultaneously minimize soft error rate with any other constraints related to the circuit performance [24], [25]. However, in these two papers, the soft error rate of a gate was assumed to be a linear function of either masking probability or gate size. This simplistic assumption may lead to less accurate solutions.

Our work in [15] introduced a sizing approach based on fault-sensitivity concept. That technique equally widens parallel circuits in selected sensitive gates. Although the technique proposed in [15] can significantly boost circuit reliability with very small area overhead, it does not allocate the additional area fairly. It provides nearly equal area factor to all sensitive gates. Thus, candidates which have low *POF* to begin with and do not need as large additional area as other more sensitive gates with higher *POF* are all treated the same.

3. Methodologies to Improve Soft Error Rate

This section briefly explains the deficiency of previous soft error immunity improvement techniques in small technology node circuits and introduces the solutions for this problem. First of all, we discuss an anomaly in using traditional upsizing method in submicron circuits. Next, two main algorithms are presented in order to handle soft errors in combinational circuits with reasonable reliability gain per unit area cost.
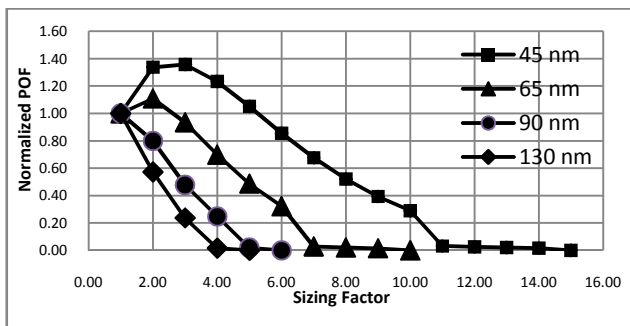
Figure 2   Normalized *POF* as a function of sizing factor of the circuit c6288 mapped with various technology nodes.

3.1 Limits of the Traditional Upsizing Method

An upsizing method which distributes the additional area to all devices in a circuit is inefficient because it provides poor reliability improvement per unit area as technology nodes are scaled. Our study in [28] revealed that in some small size circuits, increasing the size of all devices may degrade the circuit reliability due to soft errors. Figure 2 shows the normalized *POF* and sizing factor of the minimum sized layout of the benchmark circuit c6288 for different technology nodes. It can be seen in Figure 2 that at some points for smaller nodes (45nm and 65 nm traces), the *POF* increases when all devices in the circuit receive the same additional sizing factor. The reason is that for small technology nodes, the increase in the probability of neutron strike which is area dependent may dominate the increase in electrical masking probability due to Q_{crit} improvement. This in turn can result into the degradation of soft error tolerance. As a result, such a conventional sizing technique in which all transistors are upsized is a poor choice to improve reliability of submicron combinational circuits.
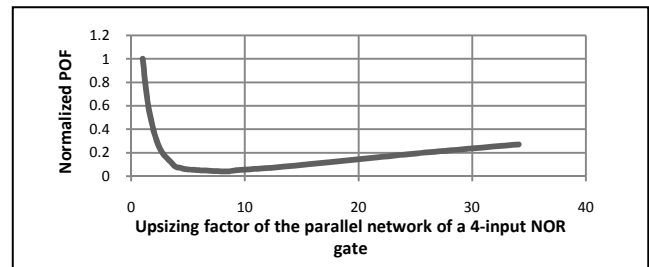
Figure 3   Normalized *POF* of a 4-input NOR in 32 nm predictive technology when the parallel network is upsized with different upsizing factors.

On the other hand, in almost all of the gates, sensitivity of a serial network is approximately 10^2-10^3 times the sensitivity of a parallel network. It is discovered that giving an extra area to the parallel network can significantly reduce Q_{crit} of serial transistors and vice versa [15]. Due to the fact that this approach causes a sharp increase in Q_{crit} of the most sensitive transistor networks with very small amount of area cost; as a result, the electrical masking probability significantly improves and dominates the increase in the probability of strike. Therefore, it offers large reduction in the circuit *POF* even when each gate receives relatively small additional area. However, unlike the overall-upsize method in which the circuit *POF* approaches zero when a circuit receives superfluous area (both series and parallel network are upsized), excessive area assigned to the parallel network does not cause much decrease in Q_{crit} of the parallel transistors. Thus, increasing the area of parallel transistor network increases the probability of neutron flux hitting the parallel network; consequently, it may even increase the corresponding gate *POF*. This observation is evident from Figure 3. In this figure, the relationship between

parallel network upsizing factor and *POF* of a 4-input NOR gate is shown. It can also be seen that circuit *POF* of the 4-input NOR decreases significantly when the upsizing factor of its parallel devices is less than 4, but increases gradually when they receive more area.

For the reasons above, we intend to find a better solution for reliability improvement on nanometer node circuits. Our objective is to study and determine appropriate upsizing methods for small technology node circuits. This includes evaluating our proposed approaches in which the overhead area budget is fairly distributed to sensitive gates based on soft error rate of each gate, and the share of area assigned to the transistors must not be so large that it offers no reduction in *POF* when applying transistor network upsizing method.

3.2   Upsizing Method with Weighted Area Overhead

The main problems of traditional sizing technique are discussed in Section 3.1. Below, we provide a quick outline of the kernel of the new idea and also show how it evolved from studying the problems with previous approaches.

This upsizing scenario takes into consideration the weight of desired area overhead dedicated to sensitive gates. We use the argument that the higher sensitivity of a gate, the larger additional area it deserves, to maximize the return on the investment by way of reduced *POF*. Thus, the upsizing factor of each selected gate should be determined based on its contribution to overall circuit *POF*.

To distribute additional area to sensitive gates, we first reckon the original *POF* of all gates. This is done using the simulator we developed [15]. After we simulate for the gate *POF*, each gate *POF* is sorted in descending order. The area budget is distributed to the most s sensitive gates based on the gate *POF*. The value s can be chosen by the users either as a percentage of all gates or based on a threshold of *POF*. Let POF_{i0} be the original *POF* of gate i, and POF_{s0} be the original *POF* of gate $i = s$ which has the smallest *POF* value among the candidate gates. To improve reliability of gate i, an additional gate area Δa_i is given to the gate. Clearly, for a gate i, the additional gate area, Δa_i, and the change in *POF* of gate i, ΔPOF_i, are related by a function, Γ. Therefore,

$$\Delta a_i = \Gamma(\Delta POF_i) \tag{3}$$

Now, since we would like to decrease the *POF* of each sensitive gate based on its relative sensitivity with respect to the smallest original *POF* value among the candidate gates, we then set the following equation:

$$\Delta POF_i \propto \frac{POF_{i0}}{POF_{s0}} \tag{4}$$

The additional area from (3) becomes as follow:

$$\Delta a_i = \alpha \, \Gamma(\frac{POF_{i0}}{POF_{s0}}) \tag{5}$$

where α is a constant which is assumed to be same for all sensitive gates. Basically, the operation Γ relies on the factors such as circuit topology, workloads, and the number of selected sensitive gates. Next, we sum all extra area terms in (5) together and this summation is equal to the desired circuit area overhead. The constant, α can be obtained by (6) below:

$$\alpha = \frac{desired\ circuit\ area\ overhead}{\Sigma_{i=1}^{s}\ \Gamma(\frac{POF_{i0}}{POF_{s0}})} \tag{6}$$

Note that the term in (5), which is the extra area given to gate i, becomes a function of original *POF* of all the selected candidate gates and the desired circuit area overhead. In this study, we simply assign Γ as a simplistic operator such that $\Gamma(x) = x^r$, where r is any positive real number. To be exact, we allow r to vary and we select the best value of r that offers the maximum yield (decrease in *POF*) for each experimental circuit. This upsizing method gives larger additional area to more vulnerable or sensitive gates than those which are less vulnerable. Hence, this method is expected to offer higher reliability enhancement than upsizing all sensitive gates or transistor networks without weighting. Figure 4 shows the flowchart that summarizes the upsizing method with weighted area overhead step by step.
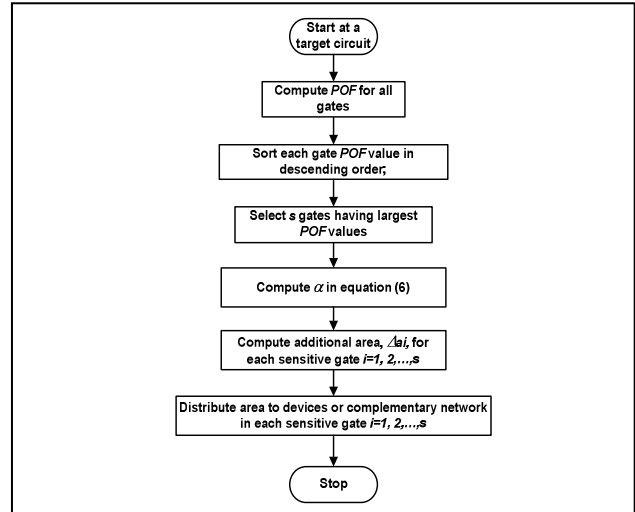

Figure 4   Flow chart of upsizing method with weighted area overhead.

3.3   Upsizing Method with Gate *POF* Saturation Consideration

Soft error rate of the most sensitive transistors (normally the stack of transistors in a gate) can be suppressed considerably by upsizing the parallel network part of a CMOS gate. However, excessively upsizing the parallel transistors may increase the number of neutrons striking on the extended area of parallel transistors as argued before in Section 3.1. This can potentially cause an increase in the soft error rate. In order to avoid allocating excessive or undesirably large area to sensitive gates, we set the

maximum upsizing factor of each gate type in the cell library to a precomputed value. The maximum upsizing factor is selected such that beyond this, the *POF* of the gate does not improve (saturation point). This method predefines an upper bound on the upsizing factor of the sensitive gates and it is implemented in conjunction with the previous method. For each selected gate i, the extra area from (5) is added to the original gate area, a_{i0}. The new area, $a_i = a_{i0} + \Delta a_i$, is then compared to the maximum value, and the result is one of the following two cases:

1) The new area of a gate i (a_i) is greater than the maximum value (a_{i_max}): in this case, the excess area of the gate beyond the allowed limit is returned to area pool for redistribution to the other sensitive gates.

2) The new area of a gate i is less than the maximum value: the area from the pool is given to the gate as long as the area in the pool is still available.

Note that in this algorithm, a gate that is more sensitive than any other gates has higher priority to receive the extra area from the pool until it reaches the maximum value. In some cases, some less sensitive gates, which are not in the set of the selected candidates, may also receive the additional area from the pool. The flowchart of this method is illustrated in Figure 5.
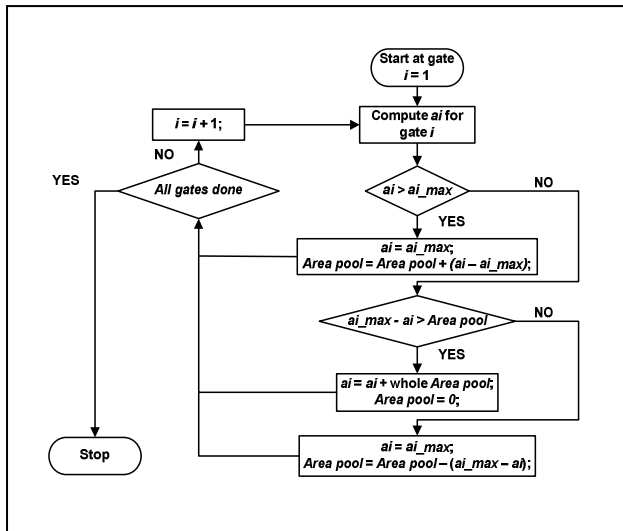

Figure 5   Flow chart of upsizing method considering *POF* saturation.

4.   Experimental Results

In this section, we determine the relative performance of our three methods proposed in this paper and compare them to a baseline method in which all gates or transistors are upsized. We evaluate various ISCAS'85, ISCAS'89 (combinational parts), and ITC benchmark circuits at operating temperature of 25ºC using 65 nm and 90 nm predictive technology nodes from [29]. The cell library

consists of 2-, 3-, and 4-input NAND, NOR gates, and Inverters. Furthermore, all reported *POF* values are normalized with respect to the base case of circuit layout that meets the original design objective of minimum area and equal rise/fall delay property in order to provide comparative performance of different methods.

4.1   Upsizing Method with Weighted Area Overhead

This part of our work evaluates proposed methodology in which the additional area spent for improving soft error rate of a circuit is distributed based on the vulnerability of sensitive gates.

In this experiment, we follow the flowchart in Figure 4 to obtain the weighted area for each sensitive gate. The experiment is set to investigate two approaches of area distribution. First, we provide the weighted area to all transistors in sensitive gates. Second, the area is shared by only parallel transistor network in each sensitive gate. We dispense only 2% area overhead to each benchmark circuit. Any gate that has the sensitivity greater than the threshold *POF* is defined as the *sensitive gate*, and these gates are added to the set of candidates to be upsized. In this study, we set the threshold *POF* between 20%-90% of the most sensitive gate.

Tables I and II show the comparison of normalized values of *POF* of selected benchmark circuits for 65 nm and 90 nm technology nodes when we perform five different upsizing scenarios. The baseline of this experiment is the traditional upsizing method in which all devices in the circuit are upsized. For fault-sensitivity based techniques, the most sensitive gates are selected to be upsized using non-weighted and weighted area overhead. The last experiment analyzes the circuit *POF* when parallel transistors in each sensitive gate are upsized with and without weighted area considerations.

From Tables I and II, it is seen that although the fault sensitivity concept is applied to enhance the upsizing method, we may face some difficulties in soft error rate reduction. It is clear from these tables that on average, upsizing sensitive gates decreases circuit reliability even more than upsizing the whole circuit does. On the other hand, the results from upsizing parallel transistors reveal that most circuits have reduced *POF* (average reduction at 15% for 65 nm-circuits and 20% for 90 nm-circuits). Additionally, when the size of sensitive transistors is weighted based on (5) and (6), this technique performs the largest reliability improvement. The experimental results from our upsizing method with weighted area overhead on parallel transistors show that this method, on average, yields slightly better results than the method without weighted area overhead. However, the *POF* of all experimental circuits upsized with weighted area overhead never goes higher than non-weighted area case and it decreases as much as 16% compared to upsizing method

without weighted area overhead in some circuits. Note in this table that smaller a value, better it is.

We notice from Tables I and II that as the technology nodes become smaller, the relative gains also reduce. However, our sensitivity based approaches causes some circuits to distract from this trend since the variations on the number of selected gates affect the area distribution which subsequently causes an increase in reliability.

Table I   Normalized *POF* of 65 nm Circuits with 2% Total Area Overhead.

Circuit	Whole Circuit Upsize	Sensitive Gate Upsize		Parallel Transistor Upsize	
		Non weighted	*Weighted*	*Non weighted*	*Weighted*
C432	1.0059	1.0504	1.0559	0.9064	0.8992
C499	1.0052	1.0071	1.0071	0.9166	0.9166
C1196	1.0042	1.0140	1.0104	0.9087	0.9040
C1908	1.0046	1.0141	1.0141	0.8450	0.8450
C6288	1.0076	1.0174	1.0216	0.9224	0.9201
i1	1.0026	1.0320	1.0320	0.8069	0.8069
i2	1.0030	0.7501	0.7501	0.7600	0.7600
i3	1.0022	1.1478	1.1478	0.4688	0.4688
i4	1.0036	1.0574	1.0499	0.8189	0.7810
i5	1.0060	1.0275	1.0257	0.9627	0.9516
i6	1.0048	1.0117	1.0114	0.9866	0.9859
i7	1.0032	1.0153	1.0153	0.8481	0.8481
i8	1.0039	1.0357	1.0378	0.8413	0.8408
S13207	1.0064	0.9909	0.9809	0.9614	0.9599
S15850	1.0067	1.0505	1.0569	0.8916	0.8891
Ave.	1.0047	1.0148	1.0145	0.8564	0.8518

Table II   Normalized *POF* of 90 nm Circuits with 2% Total Area Overhead.

Circuit	Whole Circuit Upsize	Sensitive Gate Upsize		Parallel Transistor Upsize	
		Non weighted	*Weighted*	*Non weighted*	*Weighted*
C432	0.9997	1.0303	1.0342	0.8705	0.8695
C499	0.9987	0.9875	0.9875	0.8794	0.8794
C1196	0.9970	1.0119	1.0153	0.8072	0.8069
C1908	0.9983	0.9974	1.0020	0.8341	0.8322
C6288	1.0010	1.0084	1.0089	0.8939	0.8917
i1	0.9952	1.0116	1.0116	0.7531	0.7531
i2	0.9950	0.9454	0.7588	0.6257	0.5228
i3	0.9950	1.0720	1.0720	0.3825	0.3825
i4	0.9963	1.0213	1.0128	0.7955	0.7727
i5	0.9993	1.0149	1.0115	0.9408	0.9313
i6	0.9980	0.9983	0.9942	0.9390	0.9322
i7	0.9963	0.9971	0.9977	0.7907	0.7843
i8	0.9967	1.0146	1.0154	0.7924	0.7915
S13207	0.9993	1.0270	1.0337	0.8498	0.8422
S15850	0.9996	1.0281	1.0326	0.8610	0.8585
Ave.	0.9977	1.0110	0.9992	0.8010	0.7900

Table III   Normalized *POF* of 65 nm Circuit with Parallel Transistor Network Weighted Upsize and *POF* Saturation Consideration.

Circuit	2% Area Overhead	5% Area Overhead
C432	0.8992	0.8305
C499	0.9166	0.8280
C1196	0.7921	0.7157
C1908	0.8450	0.7508
C6288	0.9201	0.8360
i1	0.8069	0.6662
i2	0.6330	0.6244
i3	0.4688	0.2929
i4	0.7810	0.7497
i5	0.9516	0.9193
i6	0.9859	0.8982
i7	0.8481	0.7608
i8	0.8293	0.6888
S13207	0.8756	0.8246
S15850	0.8891	0.8490
Ave.	0.8295	0.7490

Table IV   Normalized *POF* of 90 nm Circuit with Parallel Transistor Network Weighted Upsize and *POF* Saturation Consideration

Circuit	2% Area Overhead	5% Area Overhead
C432	0.8695	0.7848
C499	0.8794	0.8004
C1196	0.7735	0.7131
C1908	0.8322	0.7235
C6288	0.8917	0.7646
i1	0.7531	0.6164
i2	0.5228	0.5203
i3	0.3825	0.2484
i4	0.7727	0.7432
i5	0.9313	0.8841
i6	0.9322	0.8539
i7	0.7843	0.6529
i8	0.7915	0.7434
S13207	0.8422	0.8149
S15850	0.8585	0.8391
Ave.	0.7878	0.7135

4.2 Upsizing Method with Gate *POF* Saturation Consideration

Next, we conducted extensive experiments to determine the impact of area overhead on the reliability improvement. We found that many gates in many of the circuits are sensitive to saturation effect; as a result, the circuit *POF* does not reduce beyond a point and in some cases it even increases when the area overhead increases.

Tables III and IV show the comparison of normalized values of *POF* of all experimental benchmark circuits for 65 nm and 90 nm, respectively, when gate *POF* saturation is taken into consideration. The experiment is set with 2% and 5% area overheads and the gates that have the sensitivity greater than the same threshold *POF* as identified in the previous experiment are added to the set of candidates to be upsized. We also initialize the size of parallel transistor networks with the weighted area overhead from (5) and (6).

It is evident from Tables III and IV that when the saturation takes place, some circuits that derive benefit of this method receive larger reliability improvement than upsizing sensitive transistors without bounded size as reported in the previous experiment. Our method with gate *POF* saturation consideration prevents the gates from over upsizing and as a result, the relative circuit *POF* reduces.

5. Discussions

Although we believe that our proposed methods efficiently handle soft error problems in submicron circuits, it is not guaranteed that the improvement is optimized by these methods for a given area budget. The algorithm for distributing the area as discussed in section 3.2 is one of the most important factors that can justify the use of this heuristics. Specially, the relationship between the area provided to a sensitive gate and the corresponding decrease in the gate *POF* as given in (3) is dependent on circuit types, technology nodes, and circuit inputs. The function in (3) can only be identified heuristically based on the behavior of a target circuit for a particular application and technology node. Further, the threshold *POF* which is specified differently among all experimental circuits is also a crucial factor that can affect the yield of our proposed method. Since the threshold *POF* defines the number of sensitive gates we select to upsize, a change in this value will cause variations in the area distribution and subsequently vary the reliability achievement. In this study, we allow the threshold *POF* of each circuit under test to vary between 20%-90% of the largest gate *POF* thereby, providing a range of variation on total circuit *POF*. In addition, it is also found that the proper threshold *POF* of a circuit relies on circuit types and technologies.

With gate *POF* saturation consideration and upsizing the parallel transistor network, only circuits which contain sensitive gates falling into saturation condition benefit from this approach. This method assures that the *POF* will not increase when a transistor network is destined to receive much more exaggerated area, yet it may not give the optimal results. Furthermore, it is not easy to compute the maximum area that each parallel network can accept without increasing the gate *POF*. Since as mentioned before, the *POF* values are input dependent, the same type of gate in different part of a circuit may have different *POF* and hence, different maximum area for the parallel network. We chose the simplest case in which each gate type receives equal input weight to predetermine its maximum area. Nevertheless, this assumption provides acceptable results.

Methods for distributing the area budget in an optimal manner are being investigated. We believe that a mathematical optimization formulation is a proper approach to share the overhead to sensitive parts of a circuit fairly. However, our initial efforts have led us to formulate this as a nonlinear optimization problem. As a result, the process of solving this optimization problem requires massive amount of time and cannot be used for large circuits with large number of variables.

In this paper, all gate sensitivities were determined using random inputs to the circuit. We assume that all inputs are equally likely. In practical circuit, this often is not true. Possible solutions to this problem is either to use workload traces or use weighted inputs using weight sets based on workload trace. This will require collection of traces for a given circuit but the method proposed in this paper can be used without any changes in the algorithms.

This study has been limited to combinational circuits. One may ask if this can be extended to handle soft errors in sequential circuits. We believe that it can be and a possible way to extend this is as follows. In case of a sequential circuit, we can use the concept of "latching window" to determine the temporal masking probability which takes into consideration the pulse-width of the soft error glitches and the frequency of the circuit. This information can be used to estimate the circuit *POF*. We believe that we may be able to use the flip-flop selection approach [11], [12] in conjunction with our method to provide large soft error rate reduction in sequential circuits.

6. Conclusions

This paper discusses techniques for mitigating soft errors in digital circuits for small technology nodes with tight area budget. From our detailed investigations, we can conclude that the three techniques proposed by us in this paper are superior to the traditional circuit upsizing techniques in which all devices are upsized which can cause reliability degradation in smaller sized geometries. The first technique, upsizing parallel networks based on fault-sensitivity without weighted area overhead, equally distributes the additional area to parallel networks in sensitive gates. This method offers up to 20% reduction of soft error rate. The second approach which upsizes parallel networks with weighted area based on the original gate *POF* provides even more impressive results than the first

method. The last technique is developed to limit saturation effect on the gate *POF*. The experimental results reveal that the technique with gate *POF* saturation consideration provides the largest improvement, especially, on those circuits in which very small numbers of gates are selected and these gates are also relatively extremely sensitive. We also discover that the sensitive-gate upsizing method, which upsizes all devices in each sensitive gate, results in poor soft error rate reduction even though this method considers the weighted area distribution. Therefore, the sizing technique featuring upsizing all transistors either in a circuit or in selected sensitive gates may not be a good choice for handling soft errors in nanometer circuits.

7.   Acknowledgements

8.   References

[1] S. Mitra, N. Seifert, M. Zhang, Q. Shi, and K. S. Kim, "Robust System Design with Built-In Soft-Error Resilience," *Computer*, vol. 38, no. 2, pp. 43-52, Feburary 2005.

[2] F. Wang and V. D. Agrawal, "Single Event Upset: An Embedded Tutorial," in *the 21st International Conference on VLSI Design*, Hyderabad, India, 2008, pp. 429-434.

[3] Y. Tosaka et al., "Cosmic ray neutron-induced soft errors in sub-half micron CMOS circuits," *IEEE Electron Device Letters*, pp. 99-101, March 1997.

[4] N. Oh, P. P. Shirvani, and E. J. McCluskey, "Error Detection by Duplicated Instructions in Super-Scalar Processors," *IEEE Transaction on Reliability*, vol. 51, no. 1, pp. 63-75, March 2002.

[5] G. A. Reis, J. Chang, N. Vachharajani, R. Rangan, and D. August, "SWIFT: Software Implemented Fault Tolerance," in *the International Symposium on Code Generation and Optimization*, Washington, DC, 2005, pp. 243-254.

[6] J. Lee and A. Shrivastava, "A Compiler Optimization to Reduce Soft Errors in Register Files," in *the 2009 Conference on Languages, Compilers, and Tools for Embedded Systems*, Dublin, Ireland, 2009, pp. 41-49.

[7] H. R. Zarandi and S. G. Miremadi, "Soft Error Mitigation in Cache Memories of Embedded Systems by Means of a Protected Scheme," *Lecture Notes in Computer Science, Springer Berlin / Heidelberg*, vol. 3747, pp. 121-130, 2005.

[8] V. Gherman, S. Evain, M. Cartron, N. Seymour, and Y. Bonhomme, "System-Level Hardware-Based Protection of Memories against Soft-Errors," in *the DATE 2009*, Nice, France, 2009, pp. 1222-1225.

[9] A. Biswas et al., "Explaining Cache SER Anomaly Using DUE AVF Measurement," in *the 16th IEEE International Symposium on High-Performance Computer Architecture (HPCA-16)*, Bangalore, India, 2010, pp. 1-12.

[10] H. S. Deogun, D. Sylvester, and D. Blaauw, "Gate-Level Mitigation Techniques for Neutron-Induced Soft Error Rate," in *the 6th International Symposium on Quality of Electronic Design (ISQED2005)*, San Jose, CA, 2005, pp. 175-180.

[11] R. R. Rao, D. Blaauw, and D. Sylvester, "Soft Error Reduction in Combinational Logic Using Gate Resizing and Flipflop Selection," in *the 2006 IEEE/ACM international Conference on Computer-Aided Design*, San Jose, CA, 2006, pp. 502-509.

[12] E. L. Hill, M. H. Lipasti, and K. K. Saluja, "An Accurate Flip-Flop Selection Technique for Reducing Logic SER," in *the International Conference on Dependable Systems and Networks (DSN 2008)*, Anchorage, AK, 2008, pp. 128-136.

[13] J. W. Choi, B. Shim, A. C. Singer, and N. I. Cho, "Low-Power Filtering via Minimum Power Soft Error Cancellation," *IEEE Transactions on Signal Processing*, vol. 55, no. 10, pp. 5084-5096, October 2007.

[14] M. Singh and I. Koren, "Fault-Sensitivity Analysis and Reliability Enhancement of Analog-to-Digital Converters," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 11, no. 5, pp. 839-852, October 2003.

[15] W. Sootkaneung and K. K. Saluja, "Sizing Techniques for Improving Soft Error Immunity in Digital Circuits," in *the International Conference on VLSI Design and Communication Systems (ICVLSICOM-10)*, Chennai, India, 2010, pp. 87-92.

[16] G. C. Messenger, "Collection of Charge on Junction Nodes from Ion Tracks," *IEEE Transactions on Nuclear Science*, vol. 29, no. 6, pp. 2024–2031, 1982.

[17] P. Hazucha and C. Svensson, "Impact of CMOS Technology Scaling on the Atmospheric Neutron Soft Error Rate," *IEEE Transactions on Nuclear Science*, vol. 47, no. 6, pp. 2586-2594, December 2000.

[18] P. Shivakumar, M. Kistler, S. W. Keckler, D. Burger, and L. Alvisi, "Modeling the Effect of Technology Trends on the Soft Error Rate of Combinational Logic," in *the International Conference on Dependable Systems and Networks (DSN 2002)*, Bethesda, MD, 2002, pp. 389-398.

[19] C. Hungse, E. M. Rudnick, J. H. Patel, R. K. Iyer, and G. S. Choi, "A Gate-Level Simulation Environment for Alpha-Particle-Induced Transient Faults," *IEEE Transactions on Computers*, vol. 45, no. 11, pp. 1248-1256, November 1996.

[20] D. Rossi, J. M. Cazeaux, M. Omana, C. Metra, and A. Chatterjee, "Accurate Linear Model for SET Critical Charge Estimation," *IEEE Transactions on VLSI Systems*, vol. 17, no. 8, pp. 1161–1166, August 2009.

[21] Q. Zhou and K. Mohanram, "Gate Sizing to Radiation Harden Combinational Logic," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and System*, vol. 25, no. 1, pp. 155-166, January 2006.

[22] N. Miskov-Zivanov and D. Marculescu, "MARS-C: Modeling and Reduction of Soft Errors in Combinational Circuits," in *the 43rd Annual Design Automation Conference (DAC'06)*, San Francisco, CA, 2006, pp. 767-772.

[23] R. R. Rao, K. Chopra, D. T. Blaauw, and D. M. Sylvester, "Computing the Soft Error Rate of a Combinational Logic Circuit Using Parameterized Descriptors," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and System*, vol. 26, no. 3, pp. 468-479, March 2007.

[24] W. Sheng, L. Xiao, and Z. Mao, "Soft Error Optimization of Standard Cell Circuits Based on Gate Sizing and Multi-Objective Genetic Algorithm," in *the 46th Annual Design Automation Conference (DAC'09)*, San Francisco, CA, 2009, pp. 502-507.

[25] K. Bhattacharya and N. Ranganathan, "Reliability-centric Gate Sizing with Simultaneous Optimization of Soft Error Rate, Delay and Power," in *the International Symposium on Low Power Electronics and Design (ISLPED 08)*, Bangalore, India, 2008, pp. 99-104.

[26] D. G. Mavis and P. H. Eaton, "Soft Error Rate Mitigation Techniques for Modern Microcircuits," in *the 40th International Reliability Physics Symposium*, Dallas, Texas, 2002, pp. 216-225.

[27] JEDEC89A Standard, "Measurement and Reporting of Alpha Particles and Terrestrial Cosmic Ray-Induced Soft Errors in Semiconductor Devices," Joint Electron Device Engineering Council, Solid State Technology Association, 2006.

[28] W. Sootkaneung and K. K. Saluja, "Gate Input Reconfiguration for Combating Soft Errors in Combinational Circuits," in *the 4th Workshop on Dependable and Secure Nanocomputing*, Chicago, IL, 2010, pp. 107-112.

[29] HSPICE PTM website. [Online]. Available: http://www.eas.asu.edu/~ptm.