

A Method to Cope With Soft Errors

Michael N Skoufis, Haibo Wang, and Spyros Tragoudas
 Department of Electrical and Computer Engineering
 Southern Illinois University Carbondale
 Carbondale, IL 62901 USA

Abstract: - Radiation effects account for increasing soft error occurrences in deep submicron circuits. A method for calculating an upper bound for the width of an erroneous pulse that may reach a flip flop or a primary output is presented. This is required in order to determine pulse removal (PR) circuit characteristics for error correction at each flip flop or primary output side.

Key-Words: Soft errors, SET, fault tolerance

1 Introduction

Technological evolution currently allows us to integrate a very large number of components in integrated circuits. However this extraordinary improvement does not come without a cost. As a consequence, internal nodes are more vulnerable to hazardous intruding glitches with a high probability that such an error can be latched in due to a faster clock.

Such hazards indeed result from atmospheric neutrons generated by cosmic rays as well as alpha particles found in packaging and die materials [2]. The outcome of a study indicates that roughly 20 neutrons/sq cm/hr make their way to Earth being at an energy level greater than 10MeV [3]. This imposes a realistic threat in fault-free operation for technologies under 100 nm due to these electrical phenomena. A current pulse at some node entails a voltage pulse called single event transient (SET). The latter might cause a logic state inversion at latches or flip flops yielding a single event upset (SEUs) or soft error.

Previously, soft errors associated with memories have been widely studied. Techniques, such as error correcting codes (ECC), have been proposed to address soft errors in memory circuits. The lack of concern for soft errors in combinational circuits in the past is partially due to a natural masking attitude of combinational logic [6], which is characterized by three mechanisms: logical, electrical and latching-window masking. The first one does not let the error propagate through the logic. Electrical masking eliminates the error naturally because of gate delays

on the propagating path. Lastly, latching-window masking justifies the risk-free operation even in the case of the error reaching a primary output, if the latch is closed. The effectiveness of the above mechanisms quickly degrades with the continuous technology scaling down. Consequently, soft error rate for combinational circuits has been estimated to increase and equal the soft error rate of unprotected memory by 2011 [1].

In relation to the aforementioned window-latching masking, the nature of soft errors requires resolution methods other than shifting the latching window of the flip flops. Even if the clock signal is delayed, an error might still be latched in due to the randomness of SETs. The focus of this research is the development of a fault tolerant technique to assure fault-free operation at all times. An approach is devised for estimating a worst case propagation of a single event transient through combinational logic. It is based on the bounded delay model of a gate and it returns the computed value of the grown error width at a primary output or a flip flop input line. This value is essential in the design of single circuits for pulse removal (PR). Such a circuit is presented in Fig. 1 below.

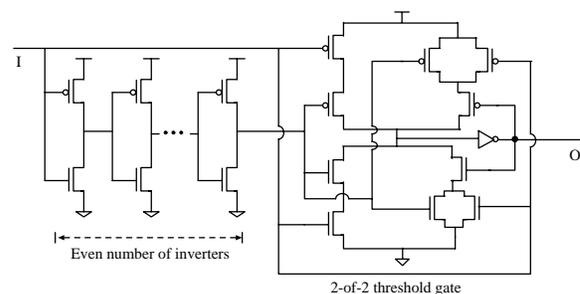


Figure 1: PR circuit.

It consists of a chain of inverter circuits and a 2-of-2 threshold gate [6]. The inverter chain has an even number of inverters and functions as a delay component. The output of the 2-of-2 threshold gate follows its input signals only when both inputs have the same logic value; otherwise, its output remains at the previous output value. If the width of a pulse (signal glitch) occurred at the input of the PR circuit is smaller than the propagation delay of the inverter chain, the short-period signal transitions will not be observed by both inputs of the threshold gate at the same time. Thus, the glitch is not propagated to the output of the PR circuit. The delay of the inverter chain corresponds to the maximum width of pulses that can be filtered out, which is also referred to as the threshold width of the glitch-removing circuit.

Such a circuit can be inserted at primary outputs or before flip flops, in order to avoid latching in the propagated SET. In order to obtain small threshold width resolutions, inverter circuits with small propagation delays are preferred. Practically, such inverters can be implemented using minimum-size transistors.

The rest of the paper is organized as follows: Section 2 explains the approach used for the computation of the longest time-duration transient that might be latched in. Experimental results are presented in Section 3 and the paper is concluded in Section 4.

2 Computing time-duration of an SET

In this section we present an approach for estimating an upper bound for the duration of a transient that may result from the appearance of an SET in the circuit. Assuming that the particle strikes on some line at time $t = 0$, we estimate the range in which an error might occur by monitoring the time. The proposed method is a vector independent analysis according to which all gate inputs not carrying SETs are considered to be at non controlling values. Thus, input masking is avoided and a worst case scenario is satisfied.

A '0-type' pulse and a '1-type' pulse are defined as low logic and high logic pulses contained in a specific '0-type' and '1-type' time range respectively. Both types of pulses trigger logic

changes at the output of all gates. This is because the side inputs are set to non controlling values.

A 'time range' is the time interval in which the contained pulse can manifest itself. Outside such a range, the hazard does not exist. These propagated transients are assumed to be generated by a single initial SET. They will travel along different sensitized paths and most likely will meet at a 're-convergency' node on one or more gates.

We define 'merging' as the union of at least two such overlapping pulses of the same polarity at two inputs of a gate. The width of the resulting pulse after merging may vary but here is assumed to be the largest in duration or 'worst case' that could occur, under the given time ranges and widths of the merged pulses.

Assume two consecutive pulses on the same or two different gate input lines. Let the term 'tolerance' be the maximal allowable distance between the two corresponding time ranges for which 'merging' of the individual inputs is justified by the slow response of the gate. In other words, the gate output does not have enough time to settle due to the proximity of the input pulses. When this tolerance measure is exceeded then merging cannot be done and a sequence of independent consecutive pulses at the gate output takes form. This sequence is called a 'pulse sequence'.

Consider an output line of gate G that carries at least one single pulse in a specific time range. This range is determined by the bounded delay of the gate G as well as the time range of its input pulses. In Fig. 2, a generic example is given. The inputs of the AND gate each carry a single pulse in two different but overlapping time ranges. The resulting time range at the output of the AND gate will carry the worst case potential pulse of width 2. This may occur in the time range determined by the earliest to the latest time instant for which the gate could generate a possible transition at its output, which may propagate to the following logic levels.

Such a pulse propagates forward through different sensitized paths and exploits circuit re-convergencies. Hence, due to these re-convergent paths, a wider hazard is produced at the output of a gate.

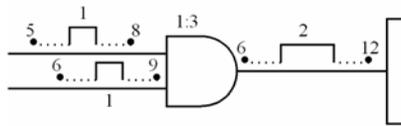


Figure 2: Generic case for error propagation

Thus, the width of a pulse or SET transient at a gate output might grow. This will depend on the number of re-convergencies ending in this gate as well as the time range lineup for the input transients.

When pulses meet at the input lines of a gate, the resulting output depends largely on the time ranges during which the input transients are active, as well as on the actual pulse duration. Both of these factors contribute in merging and propagating progressively wider pulses that may utterly have a serious impact on the functionality of a circuit.

The time periods for which SETs are defined allow margin for overlaps that could potentially create pulses of larger widths at the output of a gate. The proposed pulse propagation algorithm looks for possible ways in which a pulse of the longest possible width can be generated. This is achieved by examining when these multiple errors may occur with respect to one another, so that these stimuli can line up and synchronize in a way to collectively apply the widest probable error at the inputs.

In Fig. 3, two input pulses far enough from each other will be evaluated separately yielding a train of pulses at the output. Assuming that the pulses were closer to each other so that the tolerance condition is satisfied, then as shown in Fig. 4 the gate will respond uninterruptedly to the incoming transitions without having time to settle in-between.

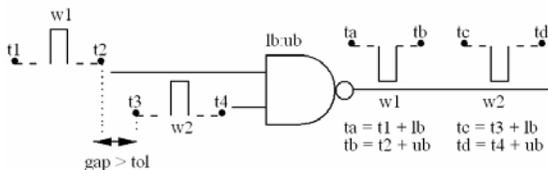


Figure 3: Pulse sequence formation

When two input pulses in a gate (as the ones previously discussed) get even closer so that there is some form of generic overlap, then merging is

highly likely depending on the timing definition of those errors. As illustrated in Fig. 5, two overlapping pulses will possibly merge to generate a wider error at the output.

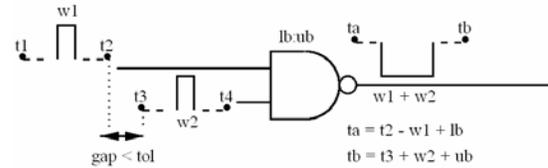


Figure 4: Merging for close pulses

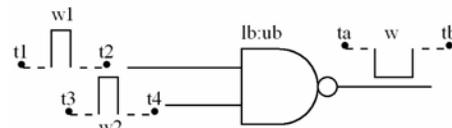


Figure 5: Merging for overlapping pulses

The maximum justified output pulse width is the summation of the individual incoming pulse widths, $w_{max} = w_1 + w_2$, subject to the time ranges for each input pulse. On the other hand, the minimum possible width is the maximum of all, $w_{min} = \max(w_1, w_2)$. Such a case is the one described in Fig. 6 below.

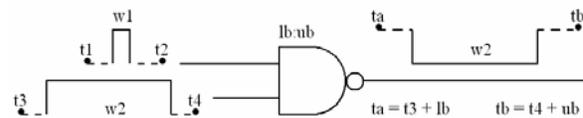


Figure 6: Masking of a narrow pulse

The described approach for propagating an SET is a simple topological traversal of a logic circuit. It is an input vector independent approach that computes the worst case pulse width that can be generated in a circuit due to a soft error injected at some line. A generic circuit line carries a sequence of 0 - type and 1 - type pulses generated by the initial random SET.

The proposed method will maintain *only one* pulse within any time range at any gate G. This is certainly true at the gate where the SET is initially generated. However, this approach maintains the same property for any gate G in the circuit.

3 Experimental Results

In this section, we give an estimate of the worst case error that can occur in the combinational part of some ISCAS89 benchmark circuits. Based on this estimated error, correction can be performed and the smallest suitable PR sensor sizes are suggested. The pulse propagation algorithm of section 3 was implemented in the C programming language. We present results on four ISCAS 89 benchmarks (s1238, s1423, s9234, and s38584) on Table 1.

To perform the analysis, the following gate delays were used: upper bound of inverter delay = lower bound of inverter delay = 25 ps, upper bound of AND/OR delay = lower bound of AND/OR delay = 70ps, upper bound of NAND/NOR delay = lower bound of NAND/NOR delay = 45 ps.

‘SET’ is the width of the initial transient and ‘Tol’ is the tolerance value for the gate. ‘Max Error’ is the worst case error we can obtain at the input of a flip flop or at a primary output. This max error indicates the minimum sensor size we need for complete error eradication. ‘Critical Delay’ is the critical path delay of the circuit for the specific setup. The ‘Total Sensor Delay’ is the overall actual delay for a sensor of given size. ‘% Delay Increase’ is the percentage delay overhead on the critical path after the insertion of the minimum size PR.

We observe that the maximum generated width was in the range of 20% of the delay of the critical path. Also, our results show that the 0-type and 1-type ranges in a pulse train for any line was smaller than 50. For the experiments, a Sun Fire V440 server with Ultra-SPARC IIIi processors was used. CPU execution time never exceeded 10 minutes.

Table 1: Upper bound for SET and PR sensor sizes at FFs.

(a)

SET=10ps	Tol=10ps			
	S1238	S1423	S9234	S38584
Max Error (ps)	80	190	100	330
Critical Path Delay (ps)	1280	2960	2365	1785
Total Sensor Delay (ps)	226	341	247	470
% Delay Increase	17.6%	11.5%	10.4%	18.5%

(b)

SET=10ps	Tol=5ps			
	S1238	S1423	S9234	S38584
Max Error (ps)	30	20	20	110
Critical Path Delay (ps)	1280	2960	2365	1785
Total Sensor Delay (ps)	173	163	163	257
% Delay Increase	13.5%	5.5%	6.9%	14.4%

References:

- [1] Dhillon, Y.S.; Diril, A.U.; Chatterjee, “A Soft-error tolerance analysis and optimization of nanometer circuits”, Design, Automation and Test in Europe, 2005. Proceedings 2005 Page(s):288 - 293 Vol. 1
- [2] Baumann, R.C.,”Soft-Errors in Advanced Computer Systems”, Device and Materials Reliability, IEEE Transactions on Volume 1, Issue 1, March 2001 Page(s):17 – 22
- [3] Nicolaidis, M.,”Design for soft-error robustness to rescue deep submicron scaling”, Test Conference, 1998. Proceedings.International, 18-23 Oct. 1998 Page(s):1140
- [4] Nicolaidis, M.,”Design for soft error mitigation”, Device and Materials Reliability, IEEE Transactions on Volume 5, Issue 3, Sept. 2005 Page(s):405 – 418
- [5] Kumar, J.; Tahoori, M.B,” A low power soft error suppression technique for dynamic logic”, Defect and Fault Tolerance in VLSI Systems, 2005. DFT 2005. 20th IEEE International Symposium on 3-5 Oct. 2005 Page(s):454 - 462
- [6] Dodd, P.E.; Massengill, L.W ,”Basic mechanisms and modeling of single-event upset in digital microelectronics”, Nuclear Science, IEEE Transactions on Volume 50, Issue 3, Part 3, June 2003 Page(s):583 – 602
- [7] G. E. Sobelman and K. Fant, “CMOS Circuit Design of Treshold Gates with Hysteresis”, Proc. ISCAS, Vol. 2, pages: 61 - 65