Sequential Circuit Temporal Hardening on an Advanced finFET Process

by

Clifford Samuel YoungSciortino

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Approved July 2022 by the
Graduate Supervisory Committee:

Lawrence T. Clark, Chair
Steven M. Guertin
Matthew J. Marinella

ARIZONA STATE UNIVERSITY

August 2022

ABSTRACT

Microelectronic circuits are prone to upsets in the natural and manmade radiation environments. As the scaling of these circuits continues, they have become more susceptible to these upsets. In highly scaled technologies even the terrestrial radiation environment is becoming increasing source of soft errors in integrated circuits. Simultaneously the means of protecting circuits via the process technology have become more and more limited. As a result, design techniques to mitigate the upsets are becoming a requirement in an ever-growing list of applications.

This work begins with an overview of radiation effects in integrated circuits. The phenomenology of upsets is discussed along with their basic mechanisms. How these effects are quantified in microelectronic circuits is then presented along with a summary of simulation methods. This is followed with a survey of the state of the field for radiation hardening by design techniques and a selection of radiation hardened flip flop designs.

Upsets within these sequential circuits like flip flops can lead to process failure or erroneous execution and thus much of the radiation hardening effort is focused on protecting them. This work applies a systematic approach to radiation hardening by design to a temporally hardened flip flop and implements it in a 14nm finFET process.

Forty-nine delay circuits are analyzed and compared on multiple performance metrics before a down select for integration. The resultant flip flop circuit is shown to have a minimum critical charge 3x higher than the baseline library flip flop. Physical design of the flip flop is outlined and nine configurations consisting of three delay lengths

and three levels if bit interleaving are accomplished. The circuits are integrated as shift registers in a radiation test chip and exposed to heavy ion testing.

Results of heavy ion testing demonstrate a threshold LET increase of approximately 6 MeV·cm2/mg with marginal increases in saturation cross section for the target LET range. A failure mode is detected while storing ones, that has both area and time dependence. Substrate charge collection is suggested as a cause and a new circuit design is presented to mitigate the error with minimal performance impact.

DEDICATION

To my wife Kaitlyn Rose YoungSciortino Ph.D., I love you and couldn't have done this without you. Your patience and understanding are boundless, you are my rock and my joy.

To my parents Mark and Shellie Short, thank you for all your love and support.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Single event effects (SEE) in advanced nodes are increasingly important as the reduced area and drive strength lower the amount of deposited charge needed to alter a node's state. This has direct applications in aerospace, autonomous driving, and large server farms, among others where reduced critical charge means the natural space environment and surface level radiation can produce increasing numbers of soft errors in integrated circuits (ICs). Designers are faced with the difficult problem of mitigating these errors since radiation hardening by process (RHBP) techniques have become increasingly difficult and costly to implement as feature sizes get smaller (Diggins, et al. 2013). As a result, identifying effective RHBD techniques in highly scaled nodes has become increasingly important.

This work presents the physical design and circuit selection process of a temporally hardened D-Flip-Flop (DFF) and the results of the subsequent heavy ion testing. A systematic approach to pre-silicon validation and design space exploration is presented, demonstrating a methodical system of design evaluation. Results of heavy ion testing are also analyzed. An error mode is identified, and an improved circuit is designed and simulated.

1.1 Radiation Effects

Radiation effects in ICs can be broadly divided into two categories, total ionizing dose (TID) and SEE. TID effects are cumulative degradation of device parameters due to

charge trapping in oxides and material breakdown of the device components (Barnaby 2006). SEE are discrete occurrences of charge collection by circuit nodes that alter the device state. This can lead to erroneous pulses in logic chains or data corruption in memory feedback elements. The mitigation of soft errors caused by SEE is the primary focus of this work.

Single event effects occur when a highly energetic charged particle passes through or near a sensitized node of a circuit. Sensitized nodes are those containing reverse biased PN junctions. Typically drain diffusions that are held at the voltage opposite to their well, i.e., n+ drain diffusion held at V_{dd} or p+ drain diffusion held at V_{ss}. As these charged particles travel through the device their electric field deposits charge within the device and alters its electrical state. Sources of charged particles include cosmic rays, solar wind, trapped ions in the earth magnetic field, radioactive decay of device materials, and manmade radiation sources (Dodd 1999).

There are three main categories of single event effects, single even latch-up, single even upsets, and single event transients. Single event latch-ups (SEL) are either destructive or nondestructive events in which a parasitic bipolar (p-n-p-n) short is created between power and ground via the source and well as shown in fig. 1.1 (Sexton 2003). This produces very high currents and can destroy the device. If the device is not damaged a power cycle is required to reset. Single event upsets (SEU) occur in memory cells when the deposited charge alters the state of one of its feedback nodes. This overwrites the stored data with its complement. Single event transients (SET) occur in logic chains when

2
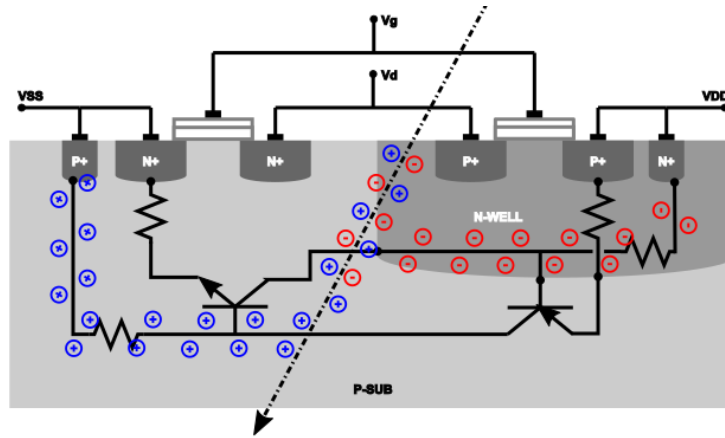
Figure 1.1 SEL parasitic BJTs

the deposited charge generates a pulse that propagates down the logic chain. If these are later captured by a latch, they will upset the circuit state.

The amount of charge that a particle deposits can be approximated by the linear energy transfer (LET) model (R. C. Baumann 2013). As the charged particle passes through the semiconductor material of the device it generates charged carrier pairs as it loses energy. LET describes this process of energy transfer in terms of dE/dx to define the energy transferred per unit path length through the semiconductor. After normalizing for the density of the target material (mg/cm^2) LET has units of $MeV \cdot cm^2/mg$ as shown by

$$LET = \frac{1}{\rho}\frac{dE}{dx} \quad (MeV \cdot \frac{cm^2}{mg}), \tag{1}$$

where ρ is the material density (2.42 g/cm^3 for silicon). This direct ionization is the primary mechanism for charge deposition for heavy ions, any ion larger than hydrogen, i.e. particles other than protons, neutrons, or electrons. Direct ionization from light particles generally does not deposit sufficient charge to cause an upset, however protons and neutrons do produce significant upset rates due to inelastic collisions with the target

nucleus. These collisions may produce alpha or gamma particles, and daughter nucleus or break the target into two fragments. Any of these products can then deposit energy along their recoil paths via direct ionization.

The amount of charge collected (Q_{coll}) by the struck node is dependent on device geometry, biasing of the circuit nodes, substrate structure, device doping, ion type, its energy, trajectory, path length, and the device electrical state. When the impinging particle passes through the device a high concentration of electron-hole pairs is generated around the ion track. When the generation path passes through the depletion region of a reverse biased junction, these carriers are rapidly collected by the electric field creating a large current/voltage transient at that node. As the ion continues into the substrate, the depletion region is extended and greatly enhances drift collection. These two collection mechanisms, shown in Fig. 1.2 (A), comprise the initial prompt collection of charge. Additional charge deposited in the deep substrate is later collected via diffusion currents over an extended period of time. In short channel devices, an additional source to drain PNP or NPN parasitic bipolar collection can occur as shown in Fig 1.2 (B).

The amount of charge necessary to upset a node is dependent primarily on the node capacitance, node voltage, restoration current, and transition time of the downstream connected logic gate (Clark 2010). The restoring current in CMOS circuits is provided by the complementary devices connected to the same circuit node as the sensitized device. For example, if the drain areas of the pull-down network are held high and thus sensitized and upset will pull them low. The pull up network will provide the
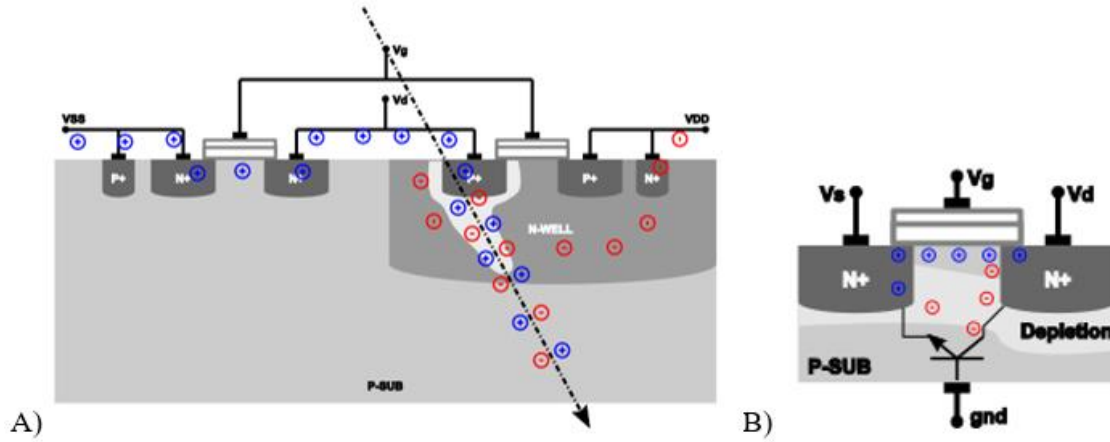
Figure 1.2 A) SEU/SET charge collection mechanisms, B) short channel collection

restoring current to pull them back up. This critical charge (Q_{crit}) can be expressed for

CMOS logic and memory cells as:

$$Q_{crit} = C_{node} \cdot V_{node} + I_{restore} \cdot t_{transition} \qquad (2)$$

In the case of the SET the transition time is equal to the propagation delay of the

downstream gate. For SEU the transition time is the propagation delay of the feedback

path. In either case the restoration current is provided by the drain current of the gate

driving the struck node as described above.

1.2   Quantifying Radiation Effects

Device sensitivity to radiation effects is a primary driver for soft error rates and

thus reliability. To quantify these metrics, in-situ and accelerated ground radiation testing

are performed. Measurements from these real silicon tests are then used to calibrate fault

models that are used in simulation to predict run time behavior in harsh radiation

environments (Mangeret 2018).

5

1.2.1   Cross Section

Radiation sensitivity is generally expressed in terms of cross-section (σ). This describes an effective area susceptible to upsets. This area is measured experimentally by placing the device into a radiation environment with a known flux for a given period and counting the number of errors induced. The equation for cross section is

$$\sigma = \frac{N_{error}}{\Phi} \; (cm^2) \tag{3}$$

Where N_{error} is the number of recorded errors and Φ is the particle fluence, the particle flux in particles per square cm per second integrated over the test time. Thus, the units of fluence are in number of particles per square cm and dividing the number of errors by this value gives a result in cm^2. Cross-section is typically defined in cm^2/bit or cm^2/device. For example, in a shift register this entails dividing the experimentally measured cross-section of the entire structure with the number of flip-flops in the register. A cross-section vs LET curve is typically used to characterize an IC's radiation performance. Figure 1.3 presents an ideal cross section curve. The minimum LET required to cause an upset is defined as the threshold LET and the maximum cross-section is defined as the saturation cross-section.

1.3   Simulation

Simulation of MOS technology-based circuits has a long history going back to at least the mid-60s (Van Lint, et al. 1967). The very first Simulation Program with
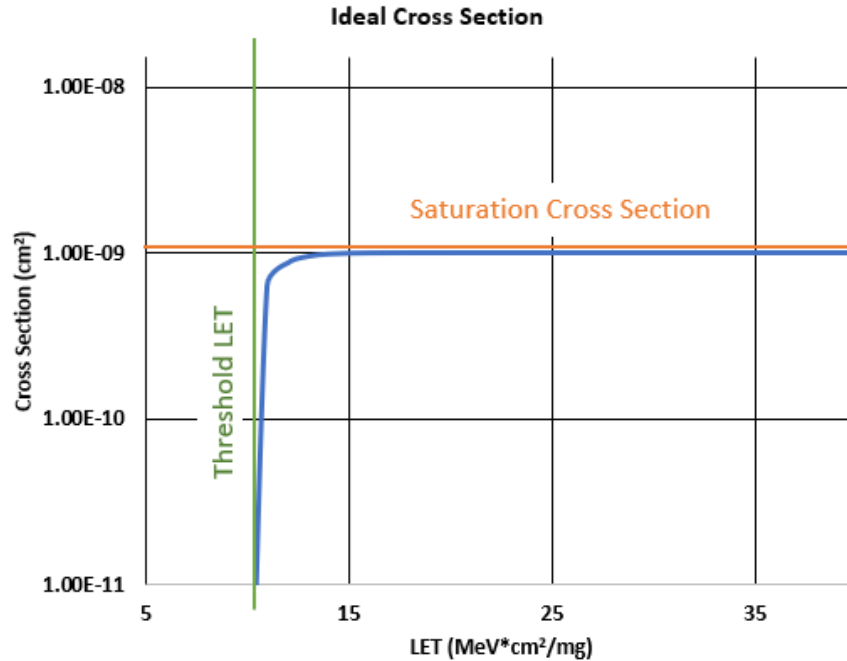
Figure 1.3 An idealized cross section curve showing the threshold LET and saturation cross section

Integrated Circuit Emphasis (SPICE) paper was released in 1973 (Nagel and Pederson 1973) as a follow up to earlier simulators BIAS and CANCER (Nagel and Rohrer 1971) out of the University of California at Berkeley. 3D-device simulation can be traced back to at least 1980 with IBMs FIELDAY (Buturla, et al. 1981). Since their introduction, circuit and device simulation have become an integral part of the design process. Qualitative and quantitative simulation results help guide early design decision making without the heavy cost of physical prototyping and testing. From the very beginning radiation effects in transistors was targeted by these simulators, with the first results from one-dimensional numerical modeling of drift-diffusion currents from radiation effects presented at the Nuclear and Space Radiation Effects Conference as early in 1967 (Gwyn, Scharfetter and Wirth 1967) and receiving the best paper award that year (IEEE 1967).

1.3.1 TCAD

Technology Computer Aided Design (TCAD) simulators use carrier transport models to solve for device behavior (Munteanu and Autran 2008). The most basic carrier transport model is the drift-diffusion model that solves the current continuity equation using the Poisson's equation. This basic model is best for long channel devices though and begins to break down in highly scaled nodes. For extremely small channel lengths Monte-Carlo solutions to the Boltzmann Transport Equation are used but these methods fail to capture quantum effects in nm scale devices. For nanoscale finFET technologies these quantum effects have to be included, such as quantum confinement, carrier gradient density, quantum tunneling, velocity saturation, etc. Once all the appropriate simulation models are identified it is required to calibrate the TCAD model to real silicon device test data. Accurate dimensions and doping gradients are necessary for good simulation matching to be achieved.

For SEE simulation in particular, another challenge presents itself for highly scaled devices (Artola, et al. 2015). The model of charge deposition becomes critical once the sensitized target region becomes sufficiently small (Raine, Guillaume, et al. 2011, Raine, Hubert, et al. 2011). Real ion tracks do not deposit charge in uniform cylinders and are instead both radially and linearly dependent. Calibrating these dependencies is a critical first step to reliably simulating an SEE response. Additionally, the surrounding circuit loading has huge impacts on the magnitude and duration of SEE and as such mixed mode simulation of the driving and load devices is required for accurate predictions.

1.3.1.1   Model Structure

The 3D-TCAD model of a 14-nm finFET, shown diagrammatically in Fig 1.4, was constructed in Victory Process and simulated in Slivaco's TCAD suite (Esposito, et al. 2021). The device dimensions and doping concentrations were extracted from literature (Synopsys 2013, 14nm Lithography Process n.d., Bazizi, et al. 2014, James 2016) to match the Process Design Kit data set. The resultant 2-fin inverter structure is representative of typical 14-nm finFET fabrication. 3D finite-element simulations were performed using bandgap narrowing, band-to band tunneling, Shokley-Read-Hall recombination using concentration-dependent low-field mobility, high field velocity saturation, and transverse-electric-field-dependent models. SEE Charge deposition was modeled using linear charge deposition (LCD) and a charge column radial dependency of 0.0015 um.   This value was experimentally determined by fitting the radial charge generation equation in TCAD to published radial dependency data for heavy ions in silicon (Fageeha, Howard and Block 1994).

1.3.2   SPICE

Many methods of accurately modeling SEE at the circuit level have been presented in the literature (Andjelkovic, et al. 2017). The simplest is driving a target node with a transient current source programmed with a double exponential model of an SEE. While this is simple to implement it has no physical meaning and requires the double exponential equation to be fit to experimental or TCAD simulation to be valid. Several models for the double exponential have been presented that attempt to tie the current

9
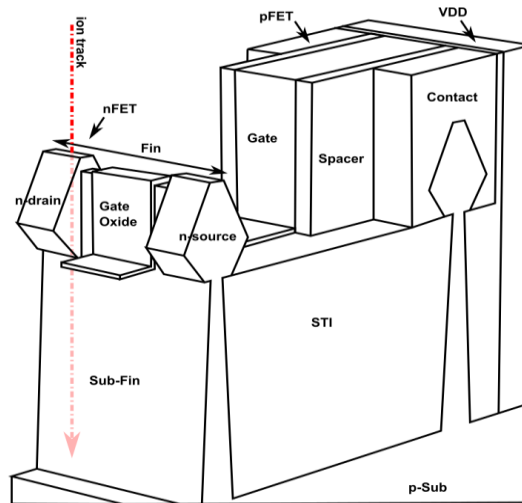
Figure 1.4 Single fin slice of 3D TCAD structure. Contacts and spacers removed on the nMOS for clarity. Actual spacing of active fins is wider and includes deep well trench separation that was excluded for brevity and space. Ion strike path show in red is typical of those simulated and is at normal incidence to the structure.

equation to device parameters but they are all limited by the same calibration requirements and accuracy limits. Other models utilize voltage dependent current sources in combination with a capacitor. The capacitor is programmed with the charge deposited by the strike and the current sources model the various collection and diffusion currents. These models have shown good correlation to TCAD SEE simulation results and are grounded in device physics. Other models include piecewise linear models, multiple independent sources, and switched resistor models. All these models can be made very accurate with proper tuning to TCAD but do not directly relate charge deposition to SEE parameters. This makes generalizing them across multiple nodes difficult.

For this work the model described by Privat (Privat and Clark 2015)was implemented, shown in Fig 1.5 (a), due to its ease of implementation, good numerical

10

properties, and use of total collected charge as a direct translation from charge deposition modeling in 3D-TCAD simulation of the same technology. This model is readily implemented in SPICE test benches as a standalone module that can be connected to any node of interest. The macros wave shaping parameters were programmed via the built-in optimizer provided with H-SPICE. The least squares method of error reduction was used to find the best fit across multiple LET by matching the upset node voltage and current waveforms and the output voltage swing of the mixed mode TCAD inverter chain to an identical SPICE only inverter chain biased using the same inputs. The results are shown in figure 1.5 (b). The uppermost pane shows the voltage transient of the struck node, the center pane shows the voltage transient of the downstream driven node, and the bottom pane shows the current transient on the struck node. Priority was given to the magnitude and duration of voltage transients since they are the primary parameters used to design and implement temporal filters.

1.4   Radiation Hardening by Design Techniques

Since soft errors represent erroneous data inputs, all non-destructive SEE can be mitigated with circuit design techniques (R. Baumann 2001). All these techniques rely on the fact that soft errors originate as discrete events in time and place. Therefore, redundancy in either time or space can be used to detect and filter out errors.

1.4.1   Spatial Redundancy

Spatially redundant systems utilize duplicate logical paths and memory locations to detect and correct errors. Fig 1.6 depicts basic redundancy schemes. In dual redundant
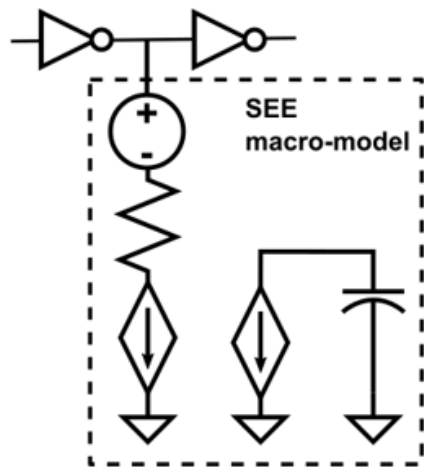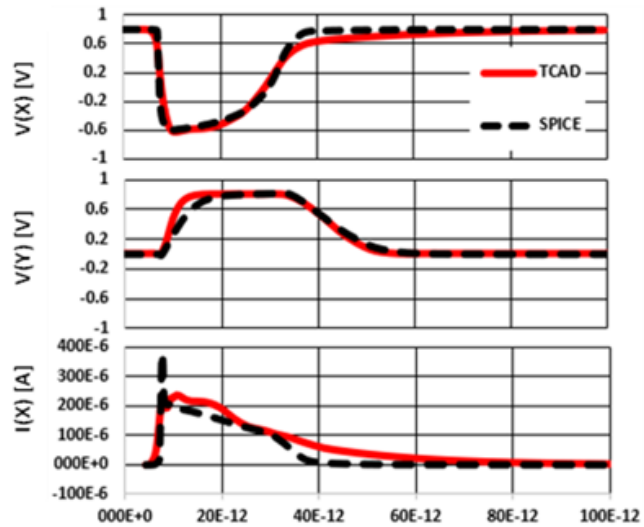
Figure 1.5 SPICE SEE (a) macro-model block diagram and (b) calibration to TCAD. Calibration curves shown include voltage swing of the struck node on the top, voltage swing of the output node in the middle, and the current response of the restoring device on the bottom.

systems the outputs of these parallel circuits are monitored and only allowed to propagate if they match. In triple redundant systems the outputs are voted, and the majority is allowed to pass (Triple Module Redundancy Design Techniques for Virtex FPGAs 2001). Only the latter allows for error correction. In dual redundant systems, errors are filtered and not allowed to propagate because the output of the c-element tri-states and maintains the previously propagated value until agreement is met. A clocked system must wait for the output to return to a normal state before sampling the data on the output or risk data loss. Triple redundant systems have constantly driven outputs and thus have higher throughput. The obvious penalty for these techniques is a 2 or 3x increase in area and power for the logical gates. Additional area and power requirements are needed to facilitate voters, filters, and to implement redundant control and clocking circuits etc.

12
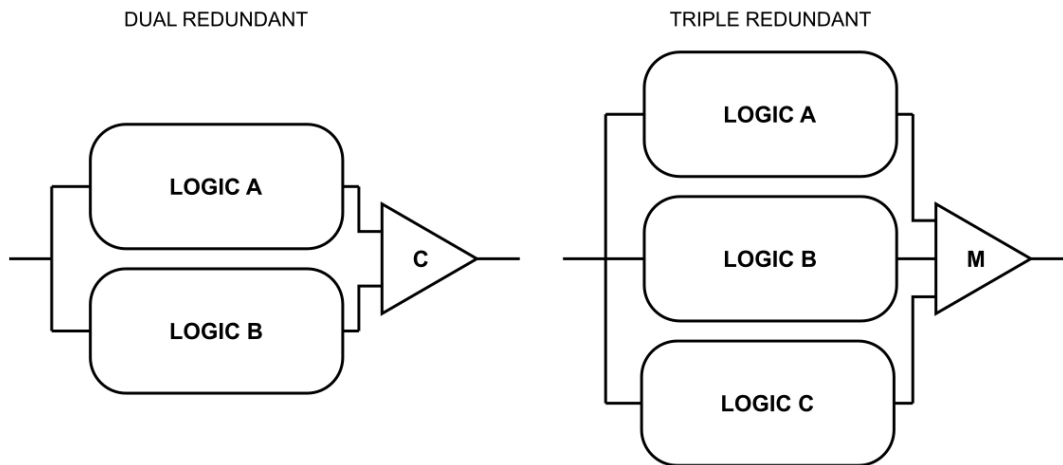
Figure 1.6 Spatial redundancy schemes. Note that C-elements tri-state in a mismatch condition and block the propagation of errors but do not correct them.

Additionally, path timing must be carefully matched lest extended settling time requirements hamper performance.

1.4.1.1   Interleaving and Node Separation

In highly scaled nodes, node-to-node proximity begins to confound spatial redundancy (Black, et al. 2008, Koga, et al. 1993). As feature sizes have become smaller than the area of carrier generation surrounding an ion-track, the likelihood of multiple node upsets has increased. This means that not only do the copies of the redundant circuits need to be separated from one another (modular redundancy), internal nodes that are co-sensitized need to be identified and separated as well. The most obvious being the internal nodes of feedback loops and the gates used to vote and filter redundant logic.

1.4.2   Temporal Redundancy

Temporally redundant systems rely on the fact that soft errors have a duration, and thus one can sample two points in time to detect and reject them with appropriate low

13

pass filtering. This is accomplished by passing the output of a logical path through two or more parallel paths with differing transmission delay (Mavis and Eaton 2002). The output of these paths are fed into a c-element which will only update its output if both paths agree or a majority voter that will filter the error as shown in Fig. 1.7. In this way any error that has a shorter duration than the difference in the path delays is ignored. As above triple redundancy is required for error correction with dual redundant systems suffering the same ambiguous output issues as in spatial redundant systems.
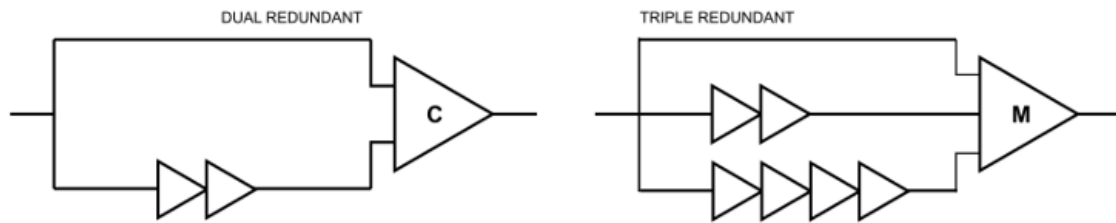
Figure 1.7 Temporal redundant circuits. Note that C-elements tri-state in a mismatch condition and block the propagation of errors but do not correct them.

The circuit effectively measures the circuit output across a time window and ensures that the data is stable for its entire length, and if not rejects the disturbance. To be effective, the difference in delay must be longer than the upset to be filtered. The primary hurdle for implementation in highly scaled nodes is the ability to produce a long delay at low power and area without creating a less resilient circuit than the unhardened one you are trying to protect. The obvious cost of these schemes is their direct impact on throughput. In practice the insertion of a temporal filter requires that at least twice the delay time be allocated to the total cycle time of a circuit. This includes the timing

14

margin required to cover both the pulse rejected and allow for the setup and hold timing

of sequential circuits latching data from the filter. Thus, for the optimal design

understanding the expected duration of an upset is required. This duration is process and

circuit dependent.

1.5   Radiation Hardened Flip-Flops

Many RHBD flip flops have been proposed in the literature with varying levels of

hardness and costs associated with them. Broadly, these designs can be broken down into

several families that have similar performance and tradeoffs (Hamed and Lee 2021). This

next section will provide a brief introduction to the various design families and address

the primary benefits and limitations of each.

1.5.1 Spatially Hardened Flip Flops

By duplicating the standard flip-flop logic one can readily implement spatial

redundant designs as shown in Fig 1.8. In dual redundant schemes, the output of both

flops can be passed to an XOR to produce an error signal that can be handled as a trap or

hazard depending on the system, limiting the risk of data loss. Unfortunately, a jam latch

as shown in Fig 1.8.b or similar mechanism may be required to hold the output when Q is

in tristate and prevent back writes or low threshold upsets (Zhang, et al. 2006). For error

correction a third flip flop is required to be implemented and the resulting triple modular

redundant flip-flop outputs are passed to a majority voter (Petrovic and Krstic 2015). This

voter can be integrated back into the constituent latches (Hindman, et al. 2011). Both

design approaches have a low performance cost since the timing impact is limited to a

15
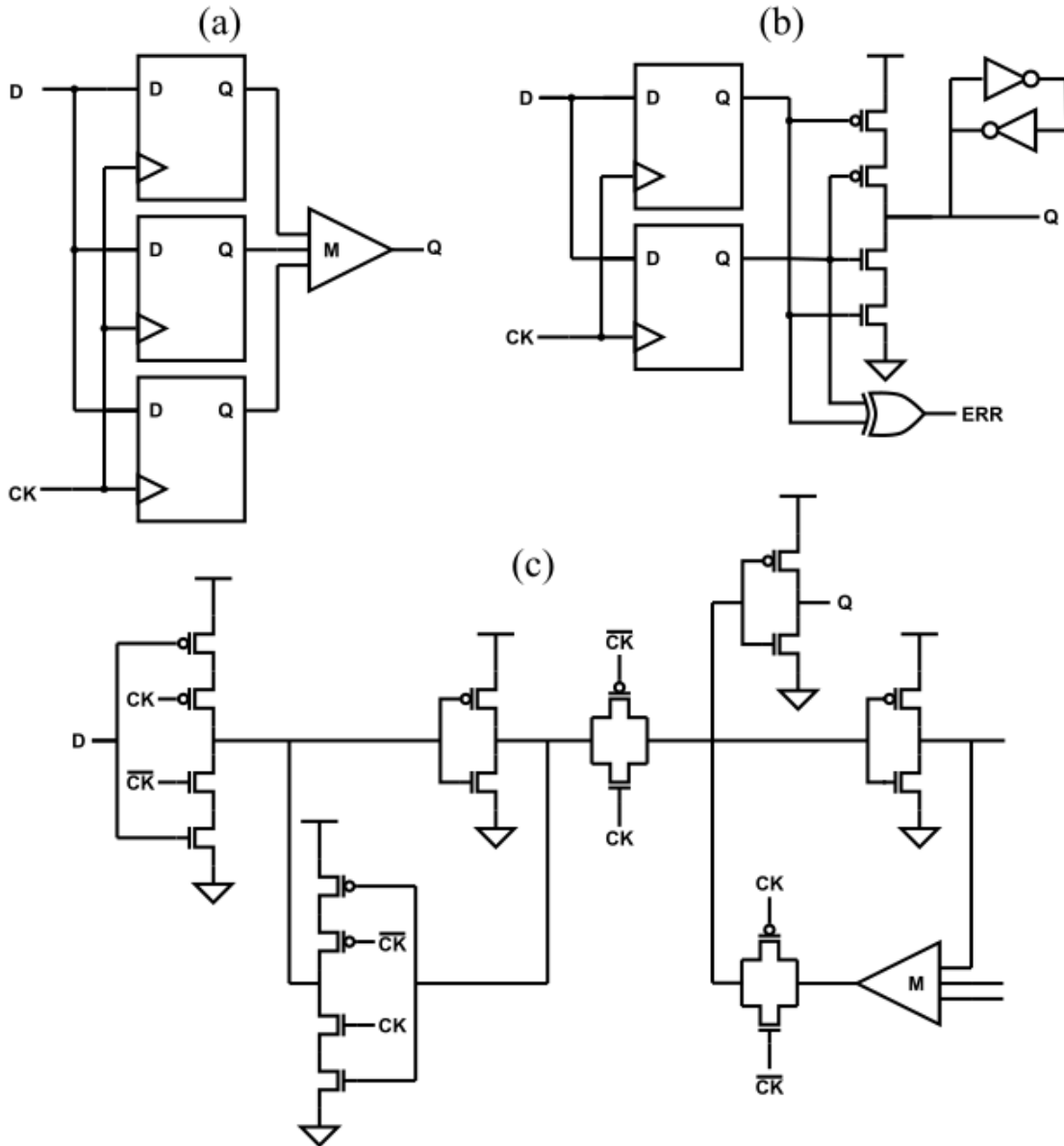
Figure 1.8 Spatially redundant Flip-Flops, (a) standard TMR FF as proposed by Petrovic and Krstic, (b) BISER dual redundant FF proposed by Zhang et al, (c) DFF with the voter integrated into the slave latch for TMR implementation proposed by Hindman et al.

single inversion stage added to the outputs. Unfortunately, to achieve SET immunity on

the data and clock inputs redundant logic and clock trees are required. Thus, the total area

and power cost of implementation can be more than 3x for systems using these designs.

16

1.5.1.1   DICE Flip Flop

The Dual Interlock storage Cell (DICE) flip-flop (Naseer and Draper 2006) is technically a spatially redundant flip flop but it differs in that the internal feedback nodes are duplicated and interlocked to be error-rejecting. The logic state of each internal node is controlled by the adjacent nodes, thus requiring two nodes to be struck in order to upset the latch. The DICE latch shown in Fig. 1.9 has the same SET issues as other spatially redundant flip-flops but has additional drawbacks in advanced nodes, the proximity of the redundant nodes makes it increasingly susceptible to multiple node charge collection (Warren, et al. 2009). This design like the others is susceptible to SET on the inputs. Errors on the clock or D port can cause erroneous data capture that the circuit cannot detect and clear. To remedy this issue Naseer and Draper proposed placing temporal filters at all inputs to the latch.
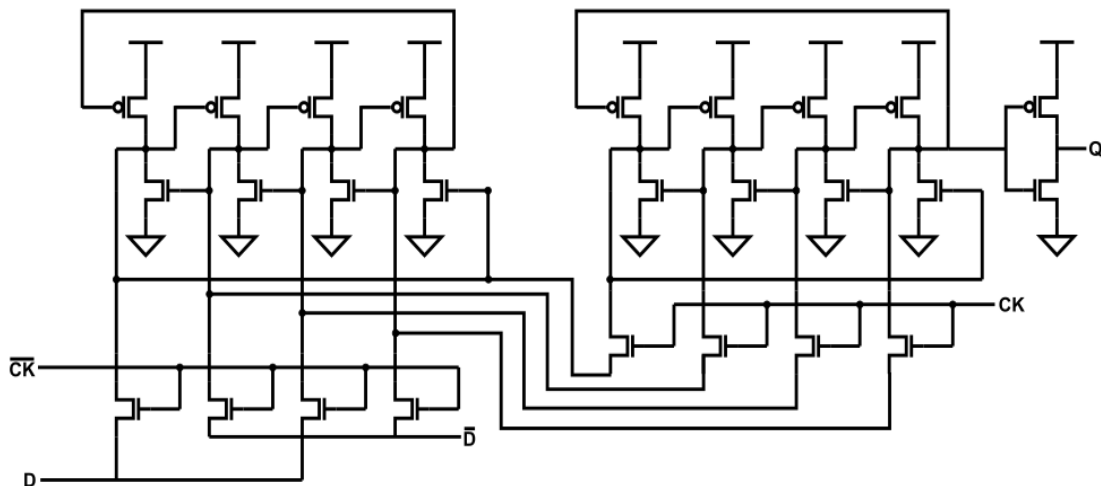
Figure 1.9 A DICE based FF as proposed by Naseer and Draper.

1.5.2   Temporally Hardened Flip Flops

The inputs to data, clock, and the data passing from setup back to hold within the latch can be delay filtered. Filtering the inputs to a latch will provide SET resilience up to the delay length for incoming pulses but will leave the internal nodes of the latch subject to SEU. Moving the delay filter into the latch as shown in Fig. 1.10 and protecting the feedback loop of the latch will provide both SET and SEU resiliency (Knudsen and Clark 2006). Upsets at the data or clock inputs will produce pulses on the hold node. As such they will be handled by the internal filter as if they had originated within the feedback loop. When placing the filter at a feedback node, care must be taken when using dual redundant schemes. SEU strikes at the output of the c-element can propagate around the loop and tri-state its output. This cuts off the restoring current and potentially latches the erroneous value. The insertion of delays in the feedback path can allow time for the circuit to clear the upset before it tri-states the C-element (Matush, et al. 2010).

As mentioned above the primary limitation of temporal hardening is performance. The insertion of delays greatly impacts clock to Q and setup timing. While the master latch is transparent, upsets on the input are passed to the HOLD node to be filtered. In the extreme case, the upset begins right at the required setup time for the feedback loop. Since it will take at least one delay to clear the upset, and the setup time of the loop is one delay, plus the inversions around the loop; the minimum setup time for resilience to be achieved must be greater than 2x the delay. This additional overhead sets a ceiling on throughput. Also of concern is the fact that the delays themselves can quickly become

18
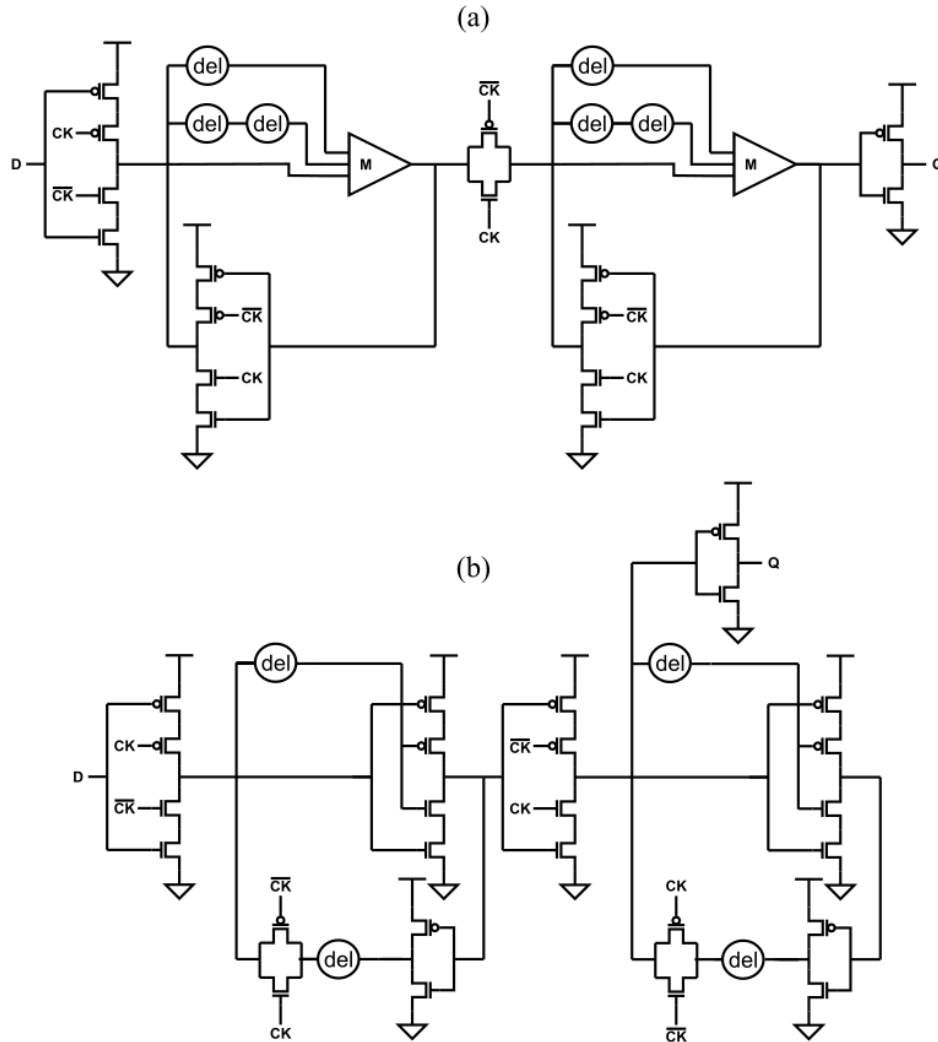
Figure 1.10 Temporally filtered flip-flops, (a) Mavis-Eaton design, (b) Knudsen-Clark design.

larger than the latch. As shown in figure 1.9 even implementing the minimum sized delay more than doubles the transistor count of the flip flop. Actual implementations will require much larger delays. This can drive the total power and area cost higher than some spatially redundant flip-flops. The primary advantage is there is no need for redundant logic, control, or error handling in order to achieve both SEU and SET resilience.

19

1.5.2.1　4CEFF

　　　　The four C-element flip flop (4CEFF) (Shambhulingaia, Lieb and Clark 2015) is

an extension of simpler temporal designs that capitalizes on the use of multiple C-

elements to reduce the number of delay elements needed as shown in Fig 1.11. This

dramatically reduces the area and power overhead of the design while maintaining a high

level of resilience. The single internal delay element protects both the hold and setup

nodes of the latch, the second C-element after the setup node is required to prevent an

upset from propagating around the loop and tri-stating the driving node. Failure to

account for this possibility will leave the latch vulnerable to SEU on the feedback path.
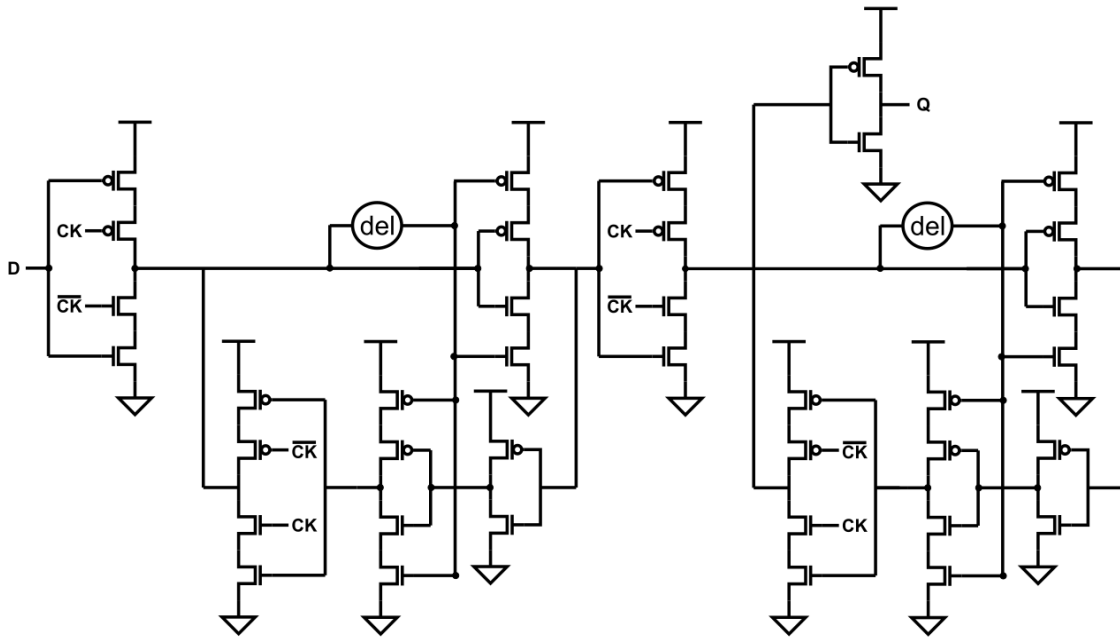
Figure 1.11 Four c-element flip-flop schematic proposed by Shambhulingaiah, Lieb and
Clark, where DEL components represent location of selected delay circuit. Note
symmetric master and slave half latches.

Upsets entering the hold node from an SET on the data or clock port are filtered by the first C-element. the hold node from an SET on the data or clock port are filtered by the first C-element. To prevent MNCC from disrupting tri-stated nodes the outputs of the c-elements must be spatially separated from their inputs as well as other c-elements in their feedback path. Similarly, the two inputs to any c-element should be spatially separated to prevent simultaneous upset. This high level in internal node separation is the primary design hurdle of successful implementation of the 4CEFF.

1.5.3   Hybrid Flip Flops

In addition to these two broad families of flip-flops there are many designs that combine these techniques. For instance, the DF-DICE latch adds delay filters to the clock and data inputs of a standard DICE latch. Other designs use a temporally filtered master latch and a DICE slave. TMR designs can have their data inputs offset with delays to create an overall delay filter. DICE methodology has also been extended to sets of four nodes in the Quatro latch. And also, there are several c-element based dual redundant flops with their internal nodes interlocked.

CHAPTER 2

14 NM SEE CHARACTERIZATION

For this work a version of the 4CEFF was implemented in a 14 nm finFET node. A methodological design approach was used to ensure radiation resilience in the target node. This process began with technology characterization of SEE response in 3D-TCAD simulation and was followed up with library SEE response characterization via SPICE upset simulation. The primary goals of this effort were to characterize the switching time and duration of upsets at various LET energy levels in representative worst-case circuits. These timing factors became the primary design metrics for subcomponent and device selection in the final flip-flop micro-architecture.

2.1 Technology Characterization

Sizing a delay element for temporal filtering first begins with understanding the expected duration of upsets at the target LET for the technology and process. The duration of a pulse is dependent on primarily the amount of charge collected, node voltage and capacitance and the restoring current. Any charge collected in excess of Q_{crit} will need to be cleared by the restoring current, and thus extend the duration of the upset pulse. To characterize the total charge collection, output pulse shape, and duration a pair of cascaded inverters was simulated in mixed mode using 3D-TCAD.

The TCAD structure described above was placed into a mixed-mode simulation consisting of two cascaded inverters as shown in Fig 2.1. The series of two fin inverters was chosen for simplicity of simulation and because it represents the configuration of
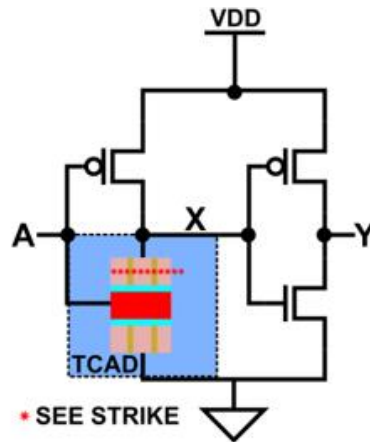
Figure 2.1 Mixed mode circuit model for nMOS strikes.

most nodes within the standard DFF layout. The TCAD transistor model was substituted

into the driving inverter to act as the struck node. An upset campaign was then conducted

at a range of LETs and the resultant voltage and current waveforms for the struck node

and the output were recorded. Integration of the restoring current for the duration of the

pulse was used to estimate Q_{coll}. All ion tracks were simulated at normal incidence to the

device with the location varied across its surface. The maximum charge collection was

observed when the strike passed through the centerline of the fin near at the gate spacer.

This is consistent with other published work (Calomarde, et al. 2020). This maximum

Q_{coll} for each LET was recorded and used to program the SPICE SEE macro-model

described above.

Figure 2.2 displays the voltage waveforms for an LET = 10 MeV·cm^2/mg ion

strike on an nMOS fin near the gate spacer. The black line shows the voltage upset

induced on node X by the impinging particle strike. The voltage is rapidly driven low and

remains a diode voltage below ground for approximately 20ps before the restoring device

23
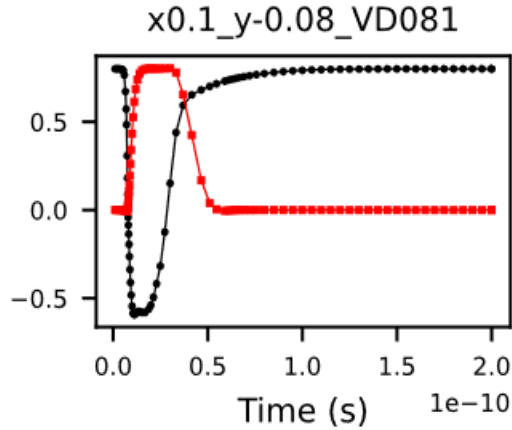
Figure 2.2 Voltage response of node X (black) and Y (red) in mixed mode simulation for an LET 10 MeV·cm^2/mg nMOS strike at normal incidence. Strike location was centerline of the fin at the fin / gate spacer junction.
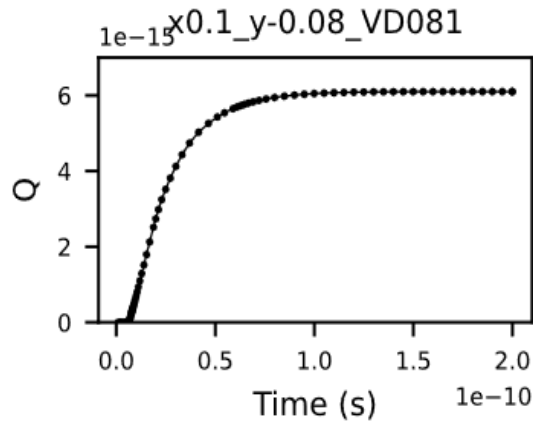
Figure 2.3 Charge collection on node X form the same strike as in Fig 2.2 calculated using the cumulative trapezoids method of numerical integration of the drain current of the restoring device.

can begin to clear the fault. The downstream inverter has a transition time on the order of

10 ps and thus reacts rapidly to the upset as shown in red. Figure 2.3 plots the charge

collection on node X over time. For this strike the device collected approximately 6 fC of

charge. These Q_{coll} and I/V waveform data for worst case ion strike locations were then

passed to SPICE for calibration of the SEE macro model.

24

2.2 Library Characterization

TCAD simulation is ill suited for SEE characterization in more complex circuits.
The total solve time for a single transistor can easily reach into the hundreds to thousands
of CPU hours. Therefore, transitioning to SPICE for circuit level simulation is a logical
and prudent next step in characterizing the SEE performance of a target cell library.
Having determined the expected durations of SEE in cascaded two fin inverters the next
step was to generalize the data to the worst case SEE within the standard DFF and the
standard cell logic.

From first principles we can determine the critical parameters driving pulse
duration. Given $I = \frac{dQ}{dt}$, and the total charge collected can be expressed as the sum of the
critical charge to upset plus the excess charge collected, $Q_{coll} = Q_{crit} + Q_{ex} = CV +
It + Q_{ex}$. We can solve for the upset duration directly, $t_{upset} = \frac{CV + Q_{ex}}{I_{restore}} + t_{transition}$.
Therefore, excluding pulse widening during propagation, the longest expected pulse
length for a given amount of charge collection will occur at the node with the weakest
drive current and most vulnerable node will be the one with the lowest capacitance.

Using this information, worst case pulse length estimates were made for SET
(SEE occurring in standard cell logic chains) and SEU (SEE occurring within the FF) in
the target library. For SET, the lowest restoring currents are provided by the largest
NAND and NOR gates available at the smallest gate width. The lowest load capacitance
is provided by placing the minimum sized inverter directly adjacent to the output and

wiring at M1. For SEU, the weakest dive node is the hold node. The lowest capacitance node is the setup node, but the additional drain capacitance of the hold node is sufficiently small compared to the gate capacitances of the feedback inverter that the reduced drive current of the hold node dominates the variation in pulse length.

Figure 2.3 depicts the two SET characterization circuits. A minimum sized inverter is driven by either a four input NAND or NOR gate. The NAND gate is driven by an input tied to VDD, sensitizing the PMOS drain areas. The restoring current is therefore supplied by the four series NMOS devices. The NOR gate is driven by an input tied to VSS, sensitizing the NMOS drain areas. The restoring current is therefore supplied by the four series PMOS devices. Figure 2.4 depicts the two SEU characterization circuits. The circuit consists of an unhardened D-latch. The clock inputs were tied such that the latch was transparent and thus there was no feedback path to latch an upset, and therefore mask the time it takes for the restoring current to clear an upset. Since the drive strength of the feedforward and feedback paths on MHOLD are identical, the duration of the SEU event while in retention can be approximated as the duration of an SET event while transparent. The setup was run with the data input set to both VDD and VSS to sensitize both the P and the N drain areas of MHOLD. In each case the SET macro-model is connected to the sensitized node and the total charge for collection is varied. The resultant voltage transient pulse lengths are recorded to estimate the delay length needed to filter a given amount of charge collection.
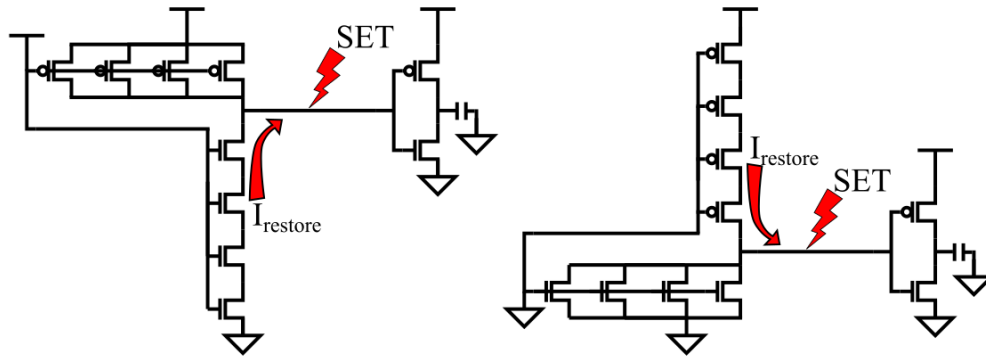
Figure 2.4 Schematic of the SPICE test bench for worst case SET pulse length characterization. On the left a NAND4 with its inputs tied high. On the right a NAND4 with its inputs tied low.
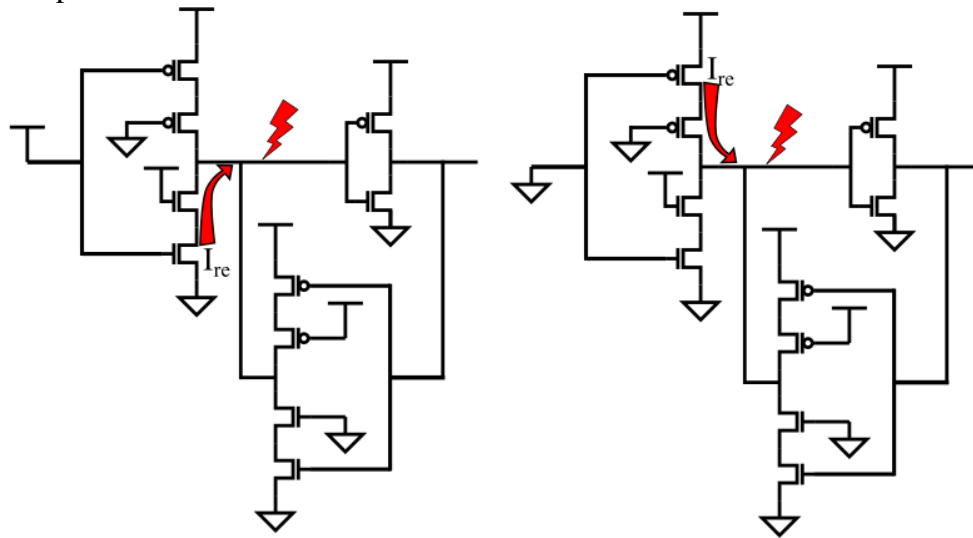
Figure 2.5 Schematic of the SPICE test bench for worst case SEU pulse length characterization. A transparent D-latch representative of the master latch of a DFF with data tied high in the left, and low and the right.

Figure 2.5 shows the results of these two experiments. An interesting finding is that the worst SET duration is longer than the worst SEU duration for a given amount of charge. This demonstrates the need to characterize both the target library that will be used to build the logic feeding the inputs and the circuit to be hardened. Hardening just to

27

expected SEU durations would have left the circuit venerable to SET at lower LET levels. As expected, in both cases pulse duration increases linearly with increased charge collection. Charge collection values from TCAD at specific LET energy levels are overlaid on the data to estimate the energy rejection capabilities of a filter with the given delay length. Based on the CREAM96 "worst day" model and the GCR at solar maximum models of particle flux in GEO orbit, there is a two order of magnitude per decade drop off above LET 1 MeV·cm^2/mg, and additional two orders of magnitude drop at LET > 30 MeV·cm^2/mg (Xapsos 2018, Mangaret 2018). With temporal filtering designed to reject upsets for energies below LET 10 MeV·cm^2/mg we expect less than 1.0 % of particle events to cause an error, filtering at LET 40 MeV·cm^2/mg we expect less than 1E-8 % of particle events to cause an error. Based on the results of the SPICE simulations, I targeted delays in the range of 20 to 50 ps for filtering these energy levels, but a designer may choose any delay length required to meet their specific application needs.
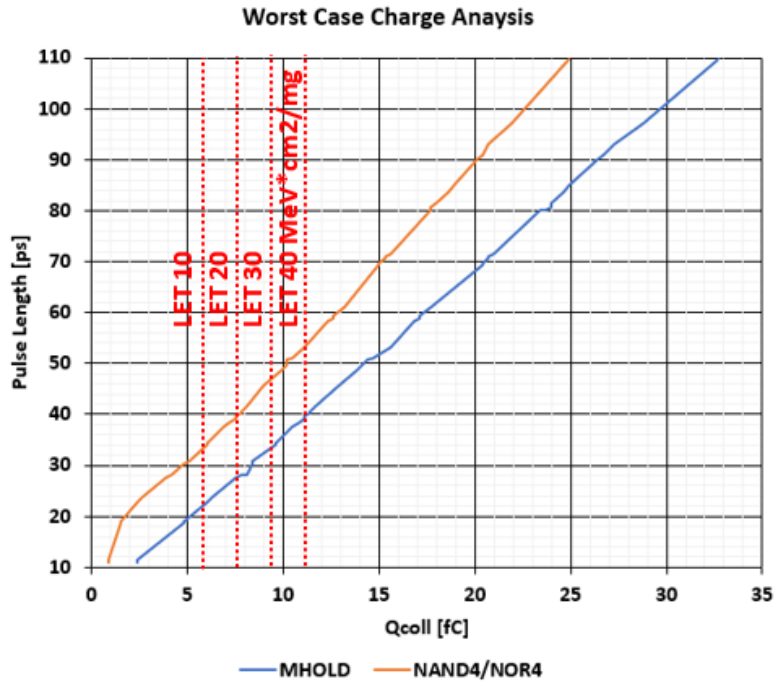
Figure 2.6 Charge collection Vs. upset duration for worst case nodes.

CHAPTER 3

DELAY DESIGN IN 14 NM

Having determined the range of delay length required to filter upsets in the 10 to

40 LET MeV·cm^2/mg range, the design of a delay for the filter can begin. All temporal

filters are highly dependent on the performance of their constituent delay cells. For good

SEE performance several parameters are of high concern. First the delay element should

introduce the fewest number of additional targets to the design as possible. Second any

new target areas should be at least as resilient as the weakest nodes in the unhardened

circuit. Failure to achieve these goals will result in reduced radiation performance instead

of enhanced.

Failure to minimize the number of additional target nodes will raise the saturation

cross section of the circuit and increase the error rate at high LET, once the delay filter is

overwhelmed there are more targets. Introducing target nodes that are less resilient will

lower the threshold LET of the overall circuit and increase the error rate at low LET,

introducing a new weakest node extends the vulnerability into the low LET range. In

practice one must tradeoff between these competing goals and thus temporal hardening

schemes typically aim to raise the threshold LET with the smallest increase in saturation

cross-section.

3.1 Delay Types

CMOS delay elements are circuits that produce a digital output equivalent to their

input after some specific amount of time. These types of circuits can be broadly

categorized as series inverters or current-starved inverters (also known as voltage-controlled inverters). As with all CMOS circuits total area and power-dissipation are critical parameters for evaluating design performance. Additionally, for SEE performance the drive strength of internal nodes, which is directly proportional to the charge that can be cleared in a given time, is of high concern (Knudsen and Clark 2006).

3.1.1 Series and Stacked Inverters

The simplest delay structure is a series of cascaded inverters. In this scenario the total delay of the series is the combined propagation delays of each stage, and thus scales down with technology. The propagation delay of an inverter is proportional the load capacitance and inversely proportional to its drive strength. The capacitance can be tuned by adding additional gate capacitances to the outputs of the constituent inverters. For these simple delay schemes the total delay is proportional to the number of delay stages, and for advanced nodes implies the need for deep chains to achieve long delays. The leakage of these delays is typically low and thus the total power dissipation is primarily due to switching power. Additionally, since back-to-back inverters are the basic building block of a latch, a simple inverter chain should be at least as resilient as the unhardened nodes of the latch.

This simple scheme is the least area efficient method of producing a delay however and requires the largest total number of stages, and thus new target nodes, to deploy. This limitation can be partially overcome by implementing stacked transistors in the inverter design. The single transistors of the pull up and pull-down networks are

replaced with several transistors in series. This increases the load capacitance on each node and reduces the drive strength of the constituent inverters. This reduced drive strength can lead to pulse expansion within the delay element. To counter this, every other inverter can be set back to the unit sized inverter. The tallest transistor stacks located within a standard DFF are two transistors high. Thus, stacks in excess of this risk becoming less resilient than the base DFF.

3.1.2 Current Starved Inverters

Current-starved or voltage-controlled inverters add a header and footer transistor to the pull-up and pull-down networks respectively. Typically, these header and footer transistors are shared across several stages of the delay. The gates of these transistors are biased such that they limit the drive strength of the inverters. This design is highly area and power efficient for creating a delay and has the added benefit of being programmable in theory. The obvious down sides are the reduced drive strength creating nodes more susceptible to upsets than the unhardened circuit and the need to design and implement a radiation tolerant biasing circuit. SEE in the biasing circuit can cutoff power to entire sections of a delay or alter the timing through the delay. TID effects can also cause the biasing point and thus delay timing to dramatically shift over the lifetime of the part. Due to these limitations this family of delays was excluded from further analysis.

3.1.1 Low-Swing Pass-Transistor

Shambhulingaiah et al presented a novel delay element that was demonstrated to be both low power and area efficient (2011). Originally taped out in 130 nm the delay

was programmed by adjusting the channel length of the pass transistors. The delay was implemented in parallel to overcome long upset pulses generated within the circuit. In finFET implementations channel lengths are set and so series transistors in the transmission gate are used to create long delays. This produces two effects, the charge removal current of any struck node remains high, and charge sharing across multiple nodes occurs in longer delays. These effects it will be shown have mitigated the long pulse generation issue.

3.2   Delay Performance

Any of the above delay types can be used to form a delay filter. So long as the delay is long enough to filter the expected upsets the filter should reject upsets on its inputs. The addition of these filters to a circuit entails a large area and power cost. It is therefore important to understand the various performance characteristics of each circuit to choose a circuit with the minimal overall performance impact on the DFF. In addition to these basic performance metrics, when considering SEE resilience adding elements to a circuit creates additional target area. Care must be taken to ensure that any circuit elements that are added to the DFF do not create new weaker nodes that will reduce the circuits resilience instead of increasing it.

Thirty-seven variations of delay from the series inverter family of delay elements and twelve variations of delay from the pass transistor family of delays were designed and simulated. The designs were constrained to have an area no greater than the standard library flip-flop, and transistor stacks no deeper than four, the maximum within the

standard cell library. Several simulations were run on each circuit. Rising and falling propagation delay was measured, along with switching power. Then SEE upsets were simulated on the weakest internal nodes to measure the SEE performance of the delay. Total susceptible area is approximated by fin count, and total area is estimated by the number of gate pitches needed for the layout. Table 1 summarizes the performance and physical design characteristics of all the delay designs. Figure 3.1 displays a selected number of delay schematics to demonstrate the naming conventions. Each of these metrics will be examined in detail in the following sections.

TABLE 1: Delay Performance Metrics

Delay Element	Delay 0 to 1	Delay 1 to 0	Ave Delay (ps)	Power 1 to 0	Power 0 to 1	Ave Energy (fJ)	Sensitive Fins (N)	Size (PP)
11	1.137E-11	1.146E-11	11.415	1.89E-06	3.88E-06	0.867	8	3
21	1.850E-11	1.904E-11	18.77	2.329E-06	4.381E-06	1.007	12	4
31	2.678E-11	2.816E-11	27.47	2.780E-06	4.872E-06	1.148	16	5
41	3.616E-11	3.881E-11	37.485	3.238E-06	5.358E-06	1.289	20	6
1111	2.269E-11	2.279E-11	22.74	3.388E-06	5.380E-06	1.315	16	6
2111	3.006E-11	3.059E-11	30.325	3.806E-06	5.865E-06	1.451	20	7
2121	3.888E-11	3.998E-11	39.43	4.261E-06	6.348E-06	1.591	24	8
3111	3.863E-11	3.999E-11	39.31	4.251E-06	6.351E-06	1.590	24	8
3121	4.771E-11	4.969E-11	48.7	4.694E-06	6.822E-06	1.727	28	9
3131	5.728E-11	6.027E-11	58.775	5.155E-06	7.304E-06	1.869	32	10
4111	4.831E-11	5.093E-11	49.62	4.705E-06	6.829E-06	1.730	28	9
4121	5.766E-11	6.098E-11	59.32	5.136E-06	7.291E-06	1.864	32	10
4131	6.737E-11	7.174E-11	69.555	5.591E-06	7.772E-06	2.004	34	11
4141	7.787E-11	8.368E-11	80.775	6.054E-06	8.248E-06	2.145	40	12
111111	3.402E-11	3.411E-11	34.065	4.881E-06	6.873E-06	1.763	24	9
211111	4.140E-11	4.191E-11	41.655	5.300E-06	7.356E-06	1.898	28	10
212111	5.046E-11	5.154E-11	51	5.742E-06	7.830E-06	2.036	32	11
212121	5.927E-11	6.093E-11	60.1	6.192E-06	8.309E-06	2.175	36	12
311111	4.996E-11	5.132E-11	50.64	5.745E-06	7.841E-06	2.038	32	11
312111	5.928E-11	6.126E-11	60.27	6.173E-06	8.304E-06	2.172	36	12
312121	6.810E-11	7.065E-11	69.375	6.623E-06	8.787E-06	2.312	40	13
313111	6.914E-11	7.212E-11	70.63	6.631E-06	8.780E-06	2.312	40	13
313121	7.823E-11	8.183E-11	80.03	7.068E-06	9.256E-06	2.449	44	14
313131	8.777E-11	9.241E-11	90.09	7.527E-06	9.731E-06	2.589	48	15
411111	5.965E-11	6.228E-11	60.965	6.197E-06	8.323E-06	2.178	36	12
412111	6.925E-11	7.255E-11	70.9	6.616E-06	8.775E-06	2.309	40	13
412121	7.806E-11	8.194E-11	80	7.064E-06	9.256E-06	2.448	44	14
413111	7.924E-11	8.359E-11	81.415	7.064E-06	9.241E-06	2.446	44	14
413121	8.833E-11	9.330E-11	90.815	7.503E-06	9.716E-06	2.583	48	15
414111	9.005E-11	9.584E-11	92.945	7.526E-06	9.718E-06	2.587	48	15
11p11	1.106E-11	1.100E-11	11.03	3.315E-06	5.353E-06	1.300	16	6
2111p2111	3.066E-11	3.097E-11	30.815	7.117E-06	9.258E-06	2.456	40	14
2121p2121	3.920E-11	4.007E-11	39.635	8.018E-06	1.022E-05	2.736	48	16
21p21	1.879E-11	1.909E-11	18.94	4.159E-06	6.296E-06	1.568	24	8
3111p3111	3.999E-11	4.109E-11	40.54	7.990E-06	1.020E-05	2.729	48	16
31p31	2.757E-11	2.863E-11	28.1	5.057E-06	7.267E-06	1.849	32	10
41p41	3.735E-11	3.963E-11	38.49	5.973E-06	8.218E-06	2.129	40	12
PG_1	2.421E-11	3.197E-11	28.09	2.209E-06	4.050E-06	0.939	12	5
PG_11	5.062E-11	6.609E-11	58.355	4.023E-06	5.695E-06	1.458	24	10
PG_2	3.475E-11	4.585E-11	40.3	2.628E-06	4.266E-06	1.034	16	7
PG_21	6.205E-11	8.070E-11	71.375	4.468E-06	5.923E-06	1.559	28	11
PG_22	7.245E-11	9.468E-11	83.565	4.897E-06	6.144E-06	1.656	32	14
PG_3	4.632E-11	6.020E-11	53.26	3.490E-06	4.475E-06	1.195	20	9
PG_31	7.436E-11	9.551E-11	84.935	4.918E-06	6.148E-06	1.660	32	14
PG_32	8.472E-11	1.095E-10	97.11	5.345E-06	6.364E-06	1.756	36	16
PG_33	9.593E-11	1.238E-10	109.865	5.766E-06	6.572E-06	1.851	40	18
PG_P1	2.093E-11	2.680E-11	23.865	3.979E-06	5.721E-06	1.455	16	9
PG_P2	3.004E-11	3.876E-11	34.4	4.876E-06	6.162E-06	1.656	24	13
PG_P3	4.020E-11	5.129E-11	45.745	5.681E-06	6.582E-06	1.839	32	17

35
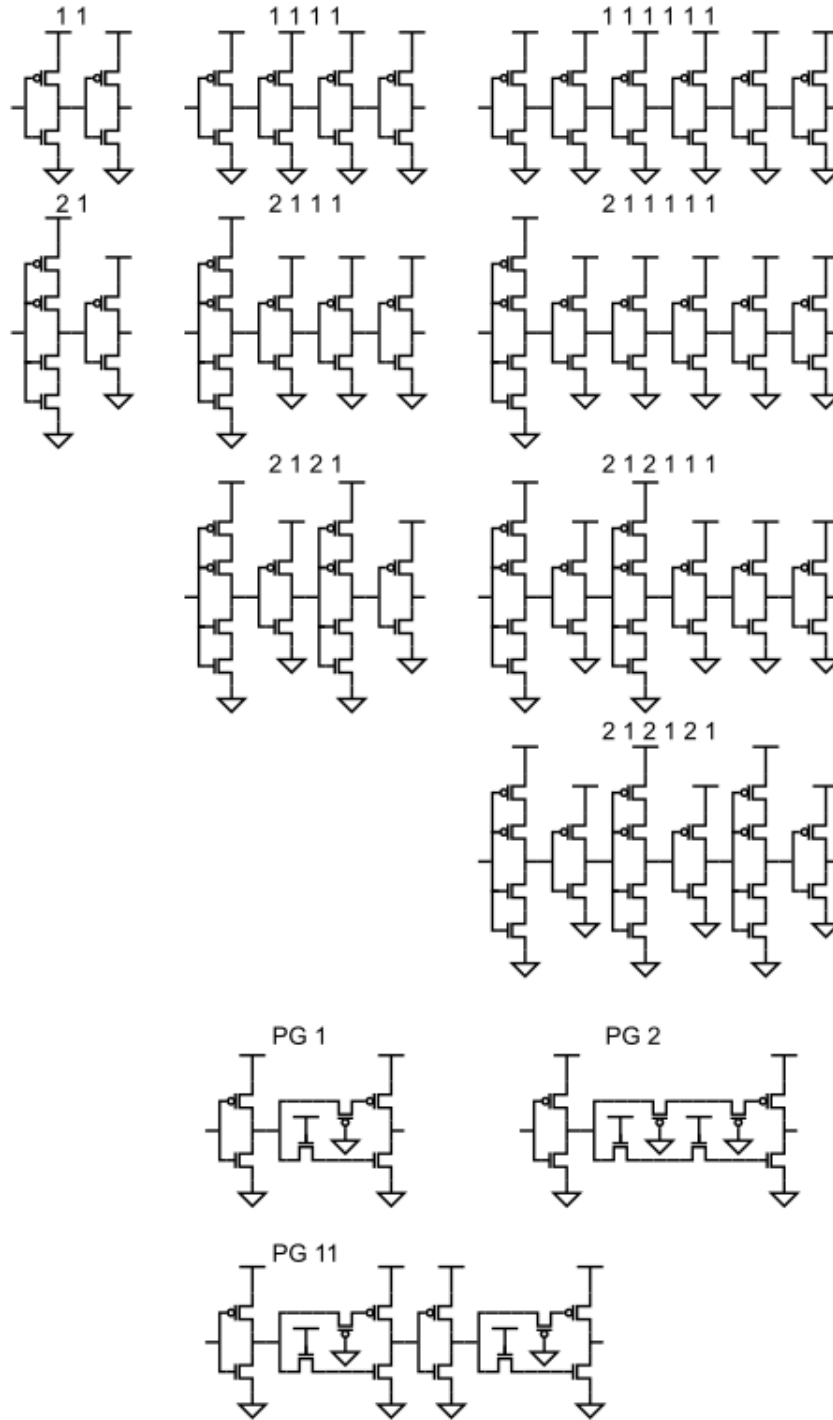
Figure 3.1 Selected delay schematics for reference and naming convention clarity. PG delays are Low-Swing Pass-Transistor delays.

3.2.1 SEE Immunity

      Temporal filters increase the threshold LET of the circuits they protect. This

increased SEE immunity can be quantified by testing how much charge is required to

produce upsets longer than the propagation delay of the delay element. Thus, one can

therefore compare the worst-case pulse length measurements from the library

characterization simulations to the delays of each delay element and determine the

maximum Q_{coll} that the delay can filter out. Figure 3.2 displays the SEU immunity of

SEU Immunity
SEU Charge (SEU_Qeq) = WC Charge: Q_MHOLD @(pulse length == delay)
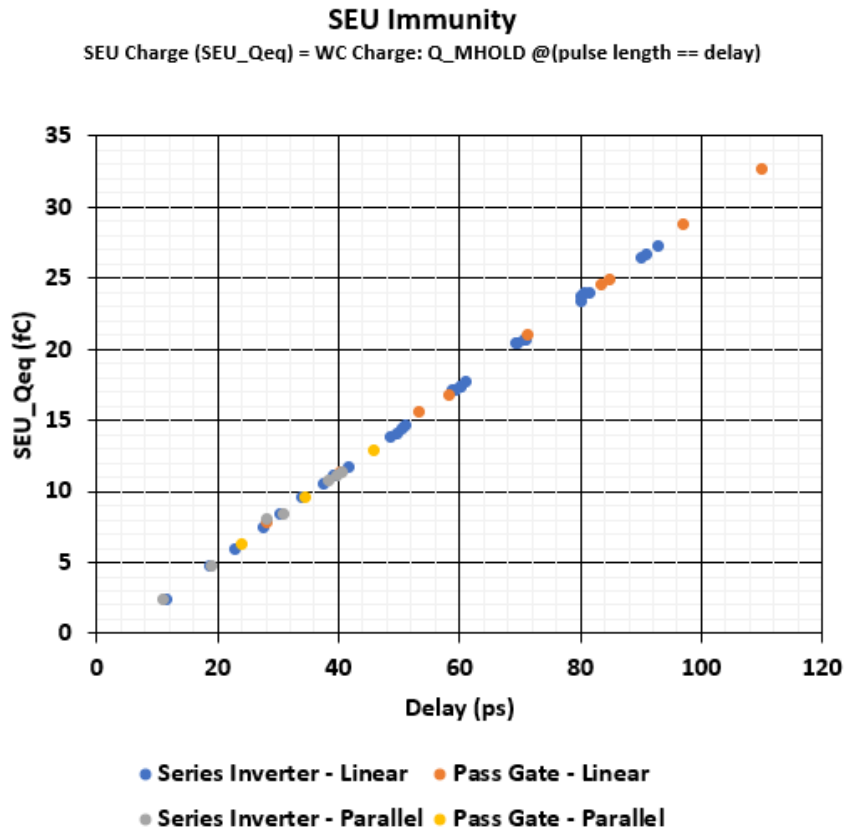
Figure 3.2 SEU immunity of the delay elements where the critical charge clearing
capability is estimated as the charge collection needed to generate a pulse longer than the
delay when upsetting MHOLD in the unhardened flip flop.

each of the delay elements. The minimum Q_{coll} on MHOLD of an unhardened flipflop to

generate a pulse longer than a given delay is shown. Figure 3.3 displays the SET

immunity of each delay element. The minimum Q_{coll} on the output of a four input NAND

or NOR driving a minimum sized inverter needed to generate a pulse longer than a given

delay is shown.

It can clearly be seen that either family of delay element can be sized to provide

the desired delay time and thus tune the threshold LET of the circuit. The relative costs of
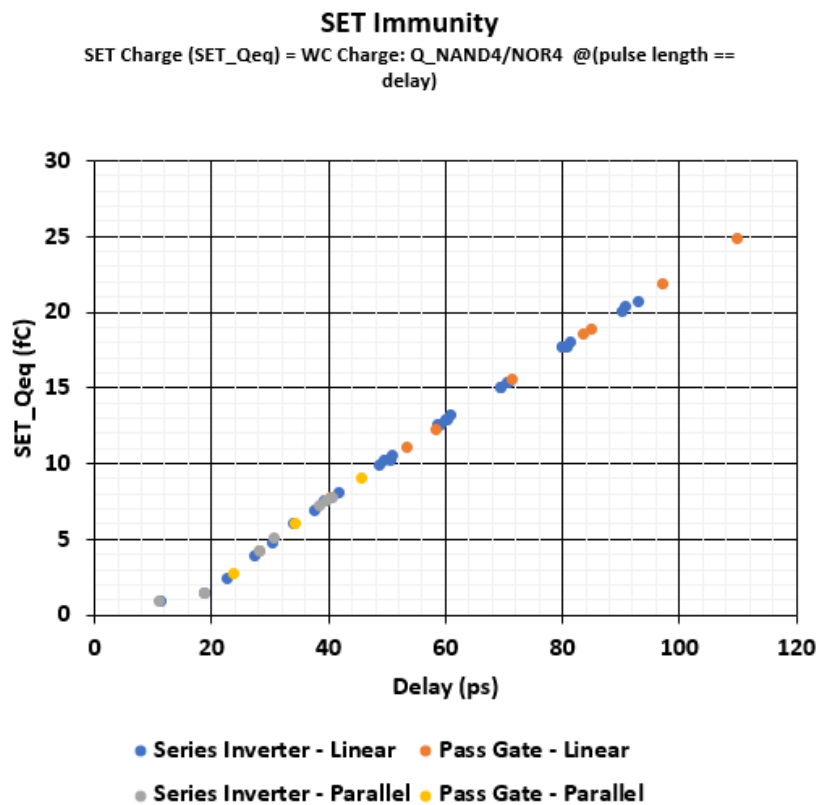
Figure 3.3 SET immunity of the delay elements where the critical charge clearing capability is estimated as the charge collection needed to generate a pulse longer than the delay when upsetting the output on the weakest drivers in the library.

38

this protection for each delay needs to be considered in order to down select to the optimal delay.

3.2.2.   Energy

The dynamic power consumption of a delay chain is directly proportional to the total gate capacitance that must be charged or discharged during a transition. Thus, deep chains or chain stages with high fanouts will drive up dynamic power consumption quickly. Fig. 3.4 displays the average switching energy vs delay for each design. The Pass-Gate family of delay's low number of driven gates dramatically reduces switching power for a given delay length. The Pass Gate family of delays had minimum 17% and.
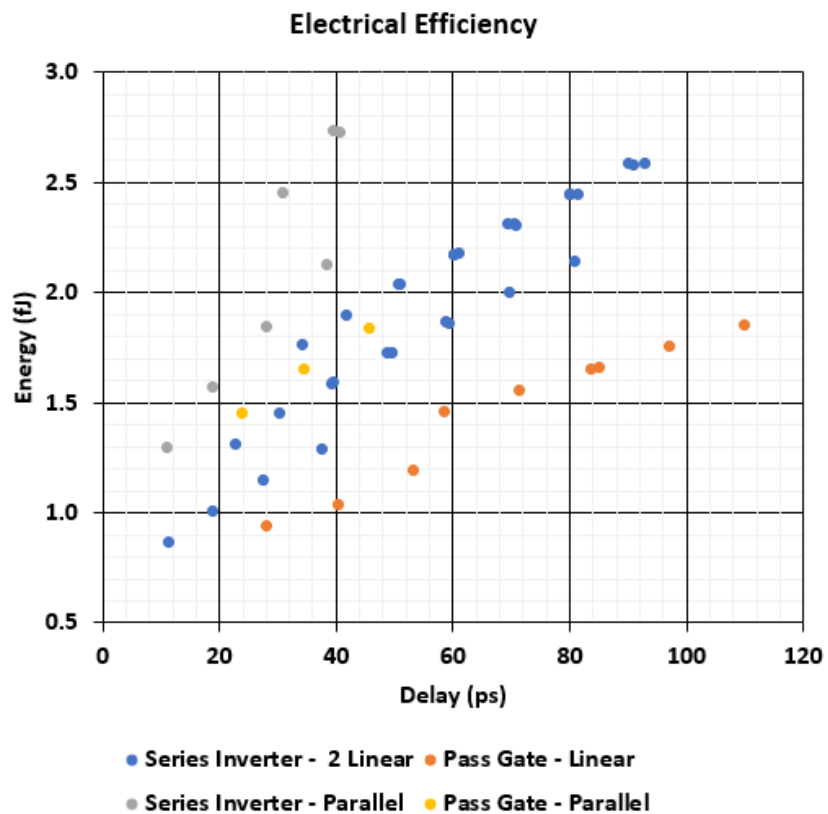
Figure 3.4 Switching energy for each delay.

39

maximum of 41% reduction in power for a given delay length Additionally, the number

of static leakage paths remains two regardless of the delay length

3.2.3   Area

The total layout area is proportional to the total number of transistors and the

number of required diffusion breaks for a given delay design. Once again long delay

chains or chains dependent on tall transistor stacks will rapidly drive up the area. In a

typical delay chain, a diffusion break will be required every two stages, the Pass-Gate
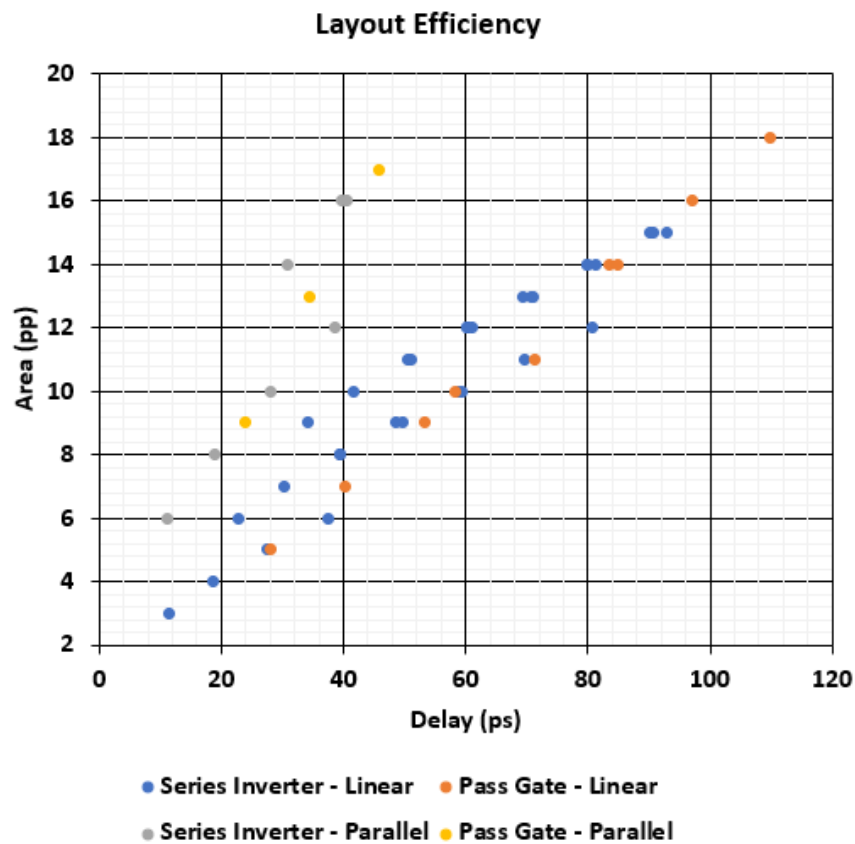
Figure 3.5 Layout area for each delay as expressed in number of standard cell height poly pitches.

family of delays never require a diffusion break, and only a single series transistor stack. Fig. 3.5 displays the area of the delay design as a function of its propagation delay. All the designs in the Pass-Gate family lie on the frontier of the design space.

3.2.4 Pulse Generation

One is tempted to believe that pulses originating within the delay are of little concern since they will immediately pass to the c-element and be filtered. This view is true given infinite time to clear an upset, however in clocked circuits this is not the case. During the period where the filters output is tri-state the latch isn't writable, extending the duration of these periods has direct impact on the setup timing and minimum input pulse length as described above. Any internally generated pulse longer than the input data pulse length can mask the data update. Since the setup and hold timing, and thus the minimum pulse width, are dependent on the delay length an optimal delay will not generate output pulses longer than its input to output delay. This provides a critical criterion for evaluating the SEE performance of a delay circuit.

As described above the estimates of SEU immunity were based on the worst-case pulse length on the weakest node in the DFF. To analyze the relative SEU vulnerability of the delay circuits the internal nodes were upset using the following procedure. First, define Q_{upset} to be the amount of charge needed to produce an upset duration at MHOLD (the weakest node in the DFF) equal to the length of the delay. Then inject Q_{upset} into the internal nodes of the delay circuit and measure the output pulse length. If the output pulse is longer than the circuit delay, then the internal nodes are more susceptible to SEU than

41

MHOLD. Implementing such a circuit will require extra setup and hold timing margin be added for soft errors originating within the delay.

Nearly every inverter chain produced an output pulse longer than its delay making it the new weakest node in the circuit. Only series inverter chains with a maximum fan out of one or parallel chains with a maximum fanout of two produced pulses shorter than their delay. Thus, long chains of inverters would be required for long delays. The Pass-Gate delays outperformed all other delays, none of them produced pulses longer than their delay. In fact, the parallel versions did not produce output pulses at all when Q_{upset}
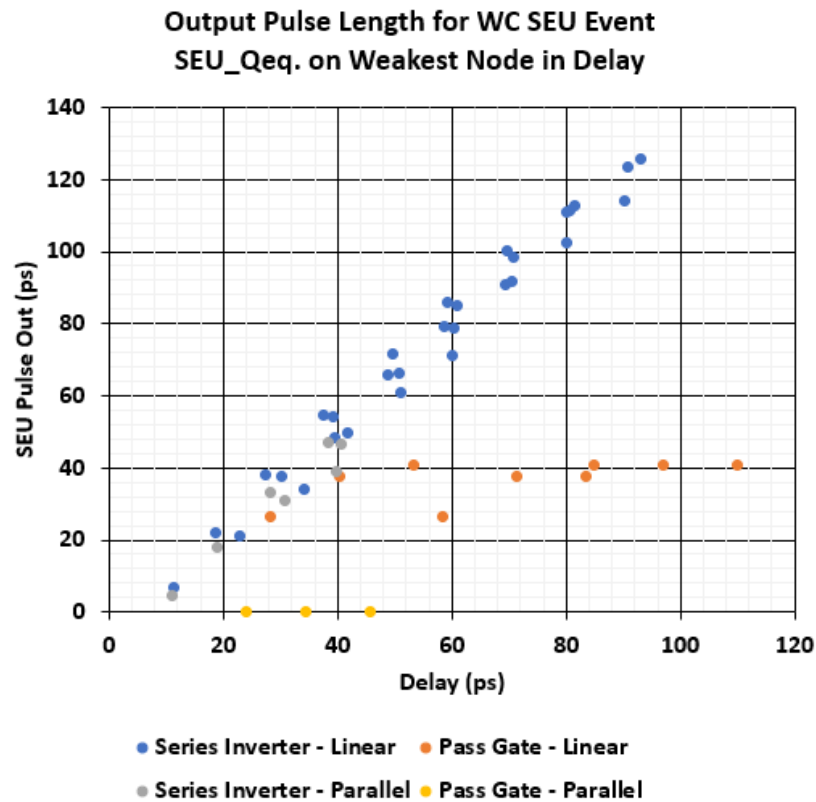
Figure 3.6 Output pulse length given worst case SEU charge collection.

was injected onto internal nodes. This does not mean that the pass gate circuits do not generate pulses from internal strikes. The pulse durations necessary to propagate simply require more charge collection than Q_{upset}, meaning the threshold LET of the delay element itself is much higher than the unhardened flip flop it is designed to protect. Fig 3.6 displays the output pulse length vs delay of each tested delay element.

To understand the performance advantage of the pass gate circuit let's analyze a strike within the PG1 delay shown below in Fig 3.7. For strikes on node X it is clear the drive strength of the restoring device is greater than the two stack in the clocked inverters driving the HOLD node and the four stack transistors of a NAND or NOR four. Similarly, the total load and timing arc from X to Y are greater than those provided by the INVx0p5 in both the SET and SEU models. So, it is expected that Q_{crit} of node X is be higher than Q_{upset}. It is not so clear for nodes C1 and C2.

Looking at node C1 let us investigate how a strike will propagate around the circuit. For C1 to be sensitized, then its initial value must be logic zero. Therefore, nodes A and Y are logic one, and node X is logic zero at the time of the strike. At this point transistors P2, N1, and N2 are all on. In the first moments of the strike, V(C1) will rise to logic one and tri-state node Y, not upset it. $I_{ds}(P2)$ will begin to drive current from node C1 to X and V(X) will begin to rise. Since N1 and P2 are both on, they will settle into a ratio circuit configuration. V(C2) will rise matching V(X) with $I_{ds}(N2)$ near cutoff throughout. Due to the ratio configuration on X and the good matching of the process, V(X) and V(C2) will settle just below VDD/2 if the upset persists long enough. Only once V(C2)
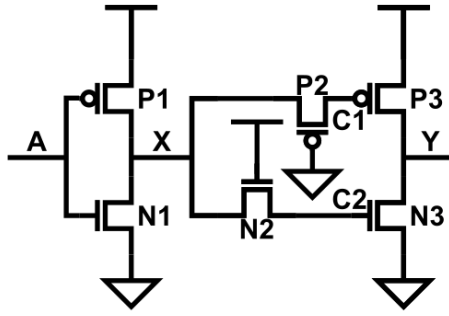
Figure 3.7 Delay element PG1with annotated device naming. Note the timing arc from an upset at C1 to Y is through P2, N2, and then the N3 switching time. Also note that in the event of an upset at C1, C1 and N1 form a ratio circuit and dive X to mid rail.

exceeds V_{tn} will N3 turn on and begin to pull down Y. Since V(C2) is below VDD/2 the

drive on node Y is minimal. The overdrive voltage (V_{sat}) is only about 100 mV. This

results in a very slow slew rate on Y. This added transit time from the upset location to

the switching device N3 coupled with voltage starving the gate of N3 mean much longer

upsets are required before an output pulse can propagate. In the parallel case the other

delay can always source enough current to overwhelm the voltage starved transistor and

keep the output near threshold and thus fail to flip the downstream node.

3.2.5   Down Select

The high levels of area and power efficiency coupled with the low sensitivity of the

pass-gate family of delays makes it an ideal candidate for use in temporal filtering in this

process node. Figure 3.8 compares all the delays' weighted performance against their

delay length. weighted performance was defined as, $A * E * G$. Where 'A' is the area 'E'

is the switching energy, and 'G' is the pulse generation ratio defined as $(\frac{\tau_{pulse\_gen}}{\tau_{delay}})$. The

44
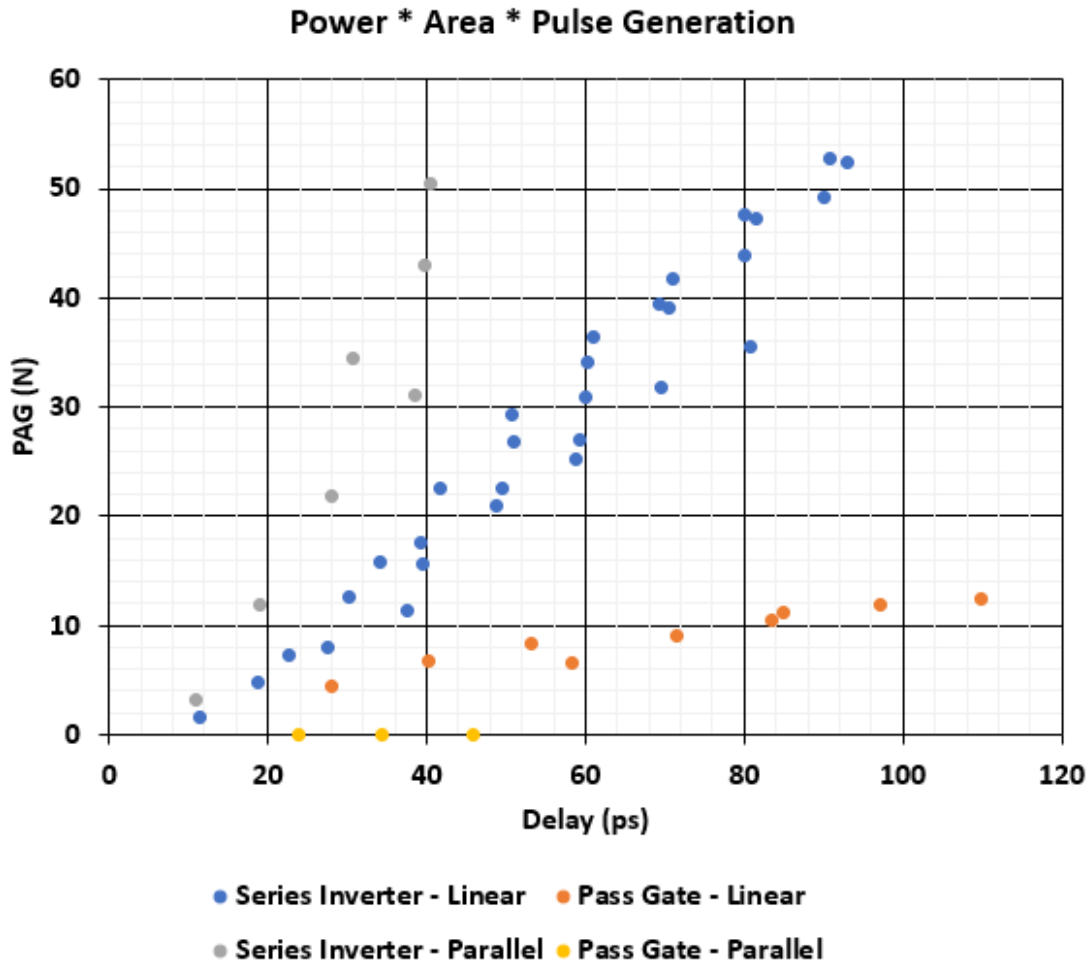
Figure 3.8 Overall performance comparison of all delay elements. With lower arcs representing the optimal curve of the design frontier.

three parallel pass gate circuits had zero output pulse and thus appear optimal on this plot, but their large area and power use make them less desirable than the non-parallel designs of the same delay length. Since the non-parallel designs do not introduce a weaker node to the circuit, fully eliminating pulse generation provides little benefit. Based on this data the three single stage pass-gate delay elements were selected for use in the final design.

CHAPTER 4

DESIGN OF THE 4CEFF IN 14 NM

A design for experiment approach was implemented for the 4CEFF, and three levels

of multi-bit interleaving, and three sizes of delay were built. The following chapter

outlines the design methodology and process for the physical design of the nine

configurations of the 4CEFF

Physical design of the 4CEFF was completed in the Cadence Virtuoso tool suite,

simulated in HSPICE, and verified with the Mentor Graphics Caliber DRC and LVS

tools. Chip level integration of the test structures, comprised of multiple shift registers,

was completed in Cadence Genus and Innovus. The design followed a standard circuit

development flow, beginning with schematic design and simulation, followed by layout

and post layout simulation. Recursive adjustments were made throughout the process.

Finally, the verified layout was integrated into a shift register block by exporting a LEF

and timing file to Genus for synthesis. The structural netlist and timing constraints files

generated were passed to Innovus for block-level placement and routing. The final blocks

were integrated into a 3mm square test chip and packaged for heavy ion testing. The

following section details the important considerations and design parameters of each of

these steps.

4.1   Schematic Design

A schematic design of the 4CEFF was developed in Virtuoso for sizing and

simulation. Fig. 4.1 depicts the final schematic design using a PG1 delay with transistor
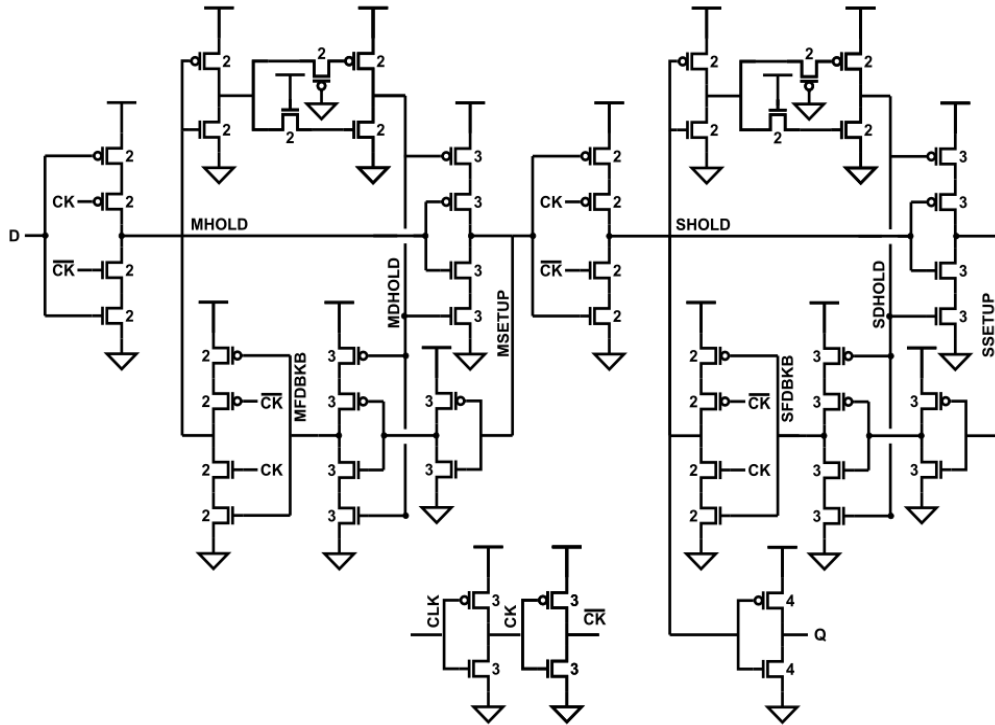
Figure 4.1 Schematic design of 4CEFF with clock circuitry showing transistor sizing. The configuration shown includes a PG1 delay element.

sizing shown. The component transistors of the C-elements and feedback inverter were

sized at 3 fins to increase drive strength and therefore SEE resilience. The crossover

MUX that comprises the drivers to the hold node had to be kept at 2 fins to accommodate

the tight routing constraints in the area. This sizing is typical to that of a unit sized

standard cell D-FF. The increased drive to all other nodes helps to ensure that no other

node presents a more susceptible target than the weakest node in the standard D-FF.

4.1.1   Pre-Layout Simulation

The completed schematic design was exported as an HSPICE netlist from Virtuoso

and simulated in HSPICE. Figure 4.2 displays a section of the functional testing where

data pulses are successfully latched. The latch initially stores a one while the input is held

47
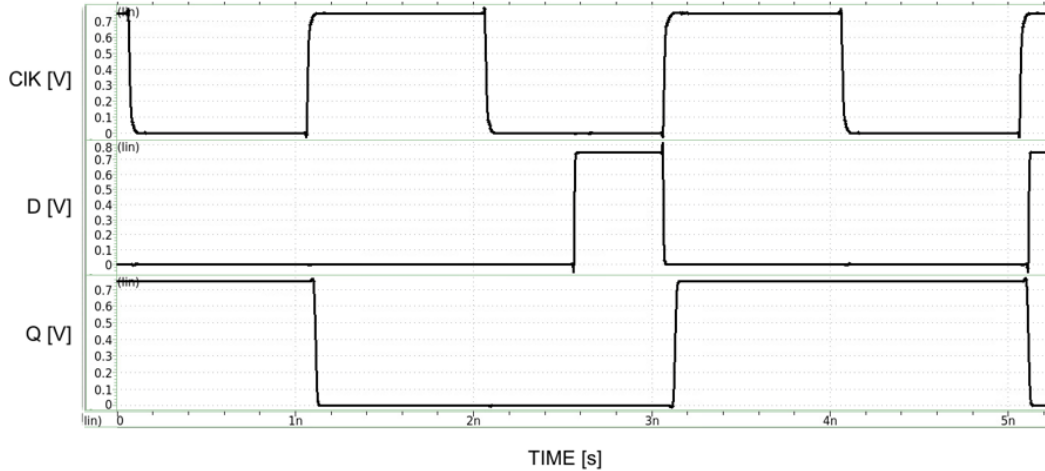
Figure 4.2 Functional simulation of the 4CEFF pre-layout netlist.

low in the first clock cycle. At the first rising edge the output updates to zero having

successfully latched the new data. In the second cycle the data is pulsed high and then

drops low again just before the next rising edge of the clock. The output nevertheless

successfully latches the high input. This is because the delay filter rejects the short

transition low just before the clock edge and retains the previous data. In the final clock

cycle the is left low until after the next rising clock edge and the latch successfully

updates to a low output. The 4CEFF using the PG1 delay had a setup time (including

padding of one extra delay time) of 81 ps, a hold time of 9ps and a Clock to Q of 22 ps.

4.1.2 Critical Charge Comparisons

In addition to functional testing upset testing was conducted on every node in the

4CEFF and standard DFF. Upsets were injected at clock high, clock low, rising and

falling clock edges, and just after setup time and just before hold time. Data was held in

the high and low state and transitioned from low to high and high to low. Both the nMOS

48

TABLE 2: Critical Charge Estimates

Node	Q_Crit fC		Node	Q_Crit fC	
	4CEFF	DFF		4CEFF	DFF
CLK	>20	8	CLKN	>20	10
CLKB	>20	8	D	18	6
MDHB	>20		SDHB	>20	
MDHBN	>20		SDHBN	>20	
MDHBP	>20		SDHBP	>20	
MDHOLD	>20		SDHOLD	>20	
MFDBKB	6		SFDBKB	6	
MFDBKBN	18		SFDBKBN	8	
MFDBKBP	16		SFDBKBP	18	
MFDBKN	10	2	SFDBKN	18	2
MFDBKP	10	2	SFDBKP	8	2
MHOLD	6	2	SHOLD	6	2
MHOLDN	12	2	SHOLDN	6	
MHOLDP	8	2	SHOLDP	12	
MSETUP	10	4	SSETUP	>20	2
MSETUPB	>20		SSETUPBN	>20	
MSETUPN	>20		SSETUPN	>20	
MSETUPP	>20		SSETUPP	>20	

and pMOS were struck with the appropriate pull up or pull-down strike simulation. This produced a matrix of 1872 test cases for each level of charge collection. 10 levels of charge collection were simulated from 2 to 20 fC for the 4CEFF for a total simulation campaign of 18,720 runs. Errors were recorded by simultaneously simulating two instances of the 4CEFF with identical inputs, but upsets only injected on the first. The output data was compared at the end of the simulation to determine if an SEU had occurred. A similar campaign was run on the standard DFF with a smaller range of charge collection for 1 to 10 fC. Table 2 summarizes the results. Of note is that the

weakest nodes in the 4CEFF using the PG1 delay have a critical charge three times greater than the weakest nodes in the standard DFF. All added nodes are also at least three times more resilient than the nodes of the standard flop according to this analysis.

4.2   Circuit Layout

Having validated the schematic design, the 4CEFF layout was developed in Virtuoso from the ground up. For compactness and efficiency of interconnect the component parts, crossover mux, c-elements, and inverters were laid out in one block to minimize the number of diffusion breaks and maximize metal one routing. Before circuit layout could begin, several design considerations needed to be addressed. Namely how to avoid MNCC within the 4CEFF. Multi-node charge collection is of great concern in advanced finFET technology. With node spacing significantly smaller than the effective charge column radius the chances of multiple nodes collecting charge from a single event is greatly increased. To properly form a mitigation plan two things must be considered, which nodes can be co-sensitized and how far should they be separated from one another.

4.2.1   Co-Sensitized Node Separation

In the 4CEFF several pairs of nodes can be co-sensitized. First, if the nodes on both the input and output of the delay are simultaneously struck then the pulses will arrive at the c-elements simultaneously and defeat the upset filtering. Second, if both nodes on the input of the feedback c-element are simultaneously struck, the filter is similarly defeated. Third, if the delayed input to the c-element is struck its output will tri-state, if that tri-stated node is then struck there will be no restoring current to clear the upset and the error

may propagate around the loop and latch in the upset given the delay is already passing

erroneous data. These error modes were confirmed in simulation by initiating a multi-

node upset campaign in HSPICE that systematically upset every possible pair of nodes.

Once these co-sensitized nodes were identified the circuit was broken down into six

zones of separation as shown in Fig. 4.3. The delays were placed with the opposite latch

and the feedback paths were swapped as well. The CLK and Q circuits were used as

spacers to separate the delays from the rest of the latches. This interleaving of the master

and slave latches resulted in considerable local interconnect routing and resulted in the

4CEFF becoming a two-row height cell with local interconnect up to M3 and being

completely wire bound on M2.
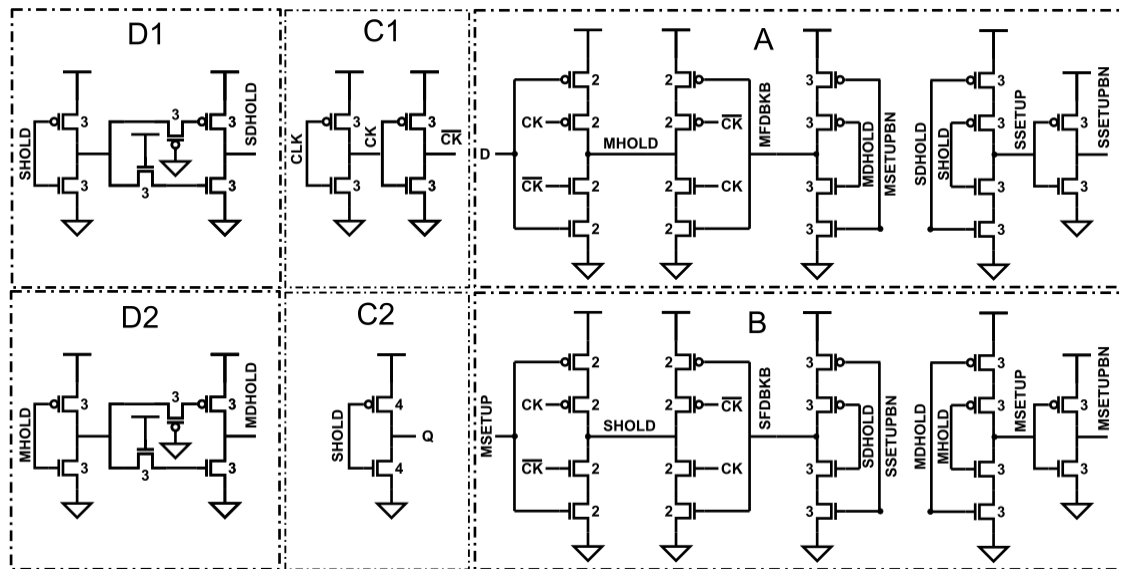
Figure 4.3 Co-sensitized node separation zones. Note the delays are on the row opposite
the hold node they are fed from as are the feedback and setup nodes. Clock and Q
circuitry was sued to increase spacing between delays and the hold node.

4.2.2 Multi-bit Interleaving

To further ensure that co-sensitized nodes were sufficiently spaced from one another, two levels of additional multi bit interleaving were laid out. The 4CEFF single bit flip flops were arranged into four-bit cells. One with four non-interleaved flops, one with two sets of two-bit interleaved flops, and one four-bit interleaved flop as diagrammatically shown in Fig 4.4. The four-bit interleaving left the area completely wire-bound on M3 as shown in Fig 4.5. These varying levels of interleaving were chosen to help experimentally determine spacing requirements for co-sensitized nodes due to the limitations of the TCAD simulation. The small size of the simulation block and the implementation of a single transistor pair meant that full extent of high LET charge deposition could not be quantified. In short, estimates of the minimum distance required to prevent MNCC could not be made beforehand utilizing this method.
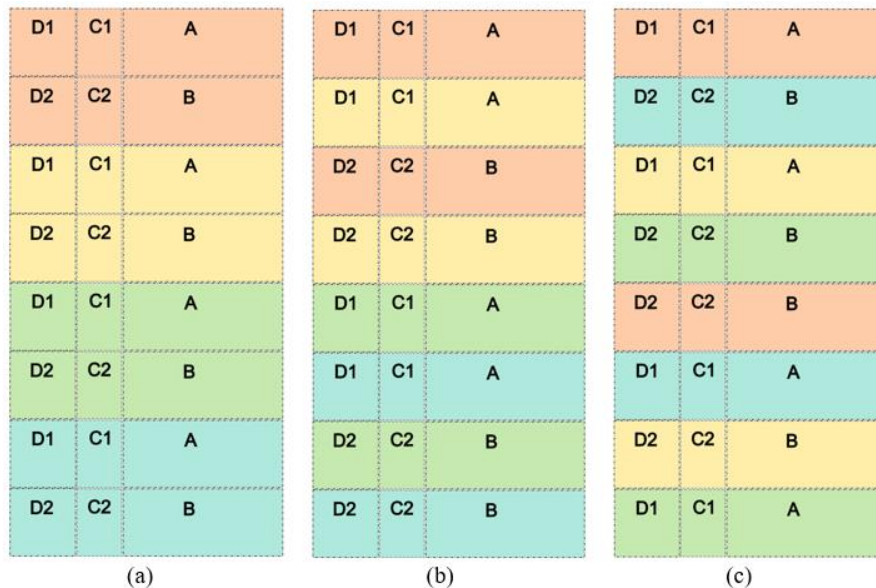
Figure 4.4 4-Bit floorplan: (a) not interleaved, (b) 2-bit interleaving, and (c) 4-bit interleaving.
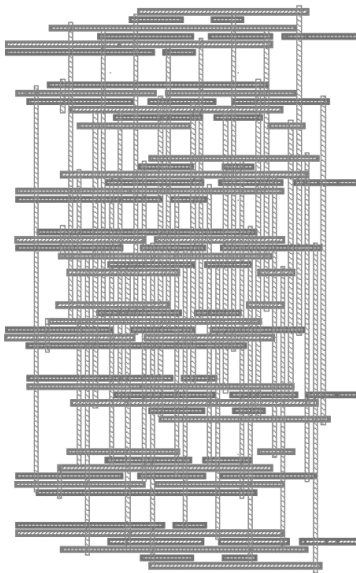
Figure 4.5 M2 and M3 routing for four-bit interleaving. Local routes for both layers are completely wire bound. M2 follow rails for power routing not shown.

4.3   Shift Register Block

For SEE testing, the nine 4CEFF layouts were implemented as shift registers. Three blocks of twenty-four-bit shift registers were designed, one for each delay length. Within each block were six four-bit sift registers organized as three sets of two, one set for each level of bit interleaving. The following sections outline the steps required to implement the 4CEFF in the design.

4.3.1   LEF Generation

Automatic Place and Route (APR) tools expect a rectangular footprint and simplified layout that only shows the pin locations and routing blockages. To facilitate this, the edges of the layouts were squared up with minimum sized fill cells and the data, clock, and Q pins were brought up to metal two and extended toward the edges for

maximum access. The new layouts were passed to cadence Abstract where detailed LEF files were generated for metal three and complete blockage was used for metal one and two. Failure to utilize detailed routing on M3 will result in spacing and forbidden pitch DRC violations on M3 in the post APR block after import back into virtuoso for physical verification.

4.3.2   Timing Constraints

In order for the synthesis and APR flow to function, a timing file had to be generated for the four-bit 4CEFF flop. Due to the complexities of the internal feedback structures, full evaluation in a software like Liberate was forgone and a custom *.lib file was created by hand. Naive use of standard timing software will produce erroneous timing data for the setup and hold time constraints. Specifically, the setup time calculated will equal the delay time plus the timing arc from output of the delay through to MFDBKB, and the hold time will be negative and equal to the timing arc from the output of the first C-element to MHOLD. While this timing information is correct for the extreme limits of operation, it will defeat the radiation robustness to clock upsets and data upsets during the data capture window. As stated previously an upset on the data input can happen right at the edge of setup time masking the data write and thus the 4CEFF requires at least two delay times for rejection of upsets in the data capture window. Similarly with the negative hold time, what is measured is the minimum pulse to drive the two C-elements and propagate around the loop, not the minimum hold time to guarantee accurate capture of data. The input data must be held until after the clock to

ensure that the correct data can be latched if an upset at the edge of setup time occurs. Thus, the hold time should be equal to the timing arc from MFDBKB to MHOLD during the clock transition. Clock upsets appear at the input to the delay filter as data inputs and thus the same constraints apply to filtering upsets on the clock nodes.

4.3.3   Synthesis

Each of the three shift register blocks were synthesized top-down in Cadence Genus. The top block consisted of a toggle flip flop to output a half clock "alive" signal, and 24 parallel shift registers grouped into three groups of eight bits. Each group of eight bits corresponded to one of the three levels of interleaving and consisted of two parallel 4-bit shift registers. The four-bit shift registers were built from 1,500 buffered 4CEFF blocks in series with a flat clock. Each buffered 4CEFF block instantiates one 4CEFF and two hold buffers in series on each of the Q ports to prevent race through from clock skew. Each chain was driven by an x1 inverter and used an x8 inverter as an output driver.

The design was synthesized at 500 MHz with clock uncertainties of 20 ps on setup and 40 ps on hold. A single worst case timing corner was used for setup and similarly for hold to simplify synthesis. The structural netlist was generated with the hierarchy preserved for ease of debugging. All sequential elements were preserved, and boundaries were preserved to prevent optimization of the shift chain or the delay buffers.

4.3.4   Automatic Place and Route

The structural netlist and SDC files produced by Genus were passed to Innovus along with the generated LEF and liberty files for APR. The assigned floor plan area was 500 µm square. Top and bottom rows were all fill. Well-Taps were inserted every 45 µm in column fashion. M2 follow rails were added but power via stapling was held until after placement. M3 power stripes were arranged such that they did not interfere with the placement of the multi row cells. The power grid was routed up to M7 before placement. Input pins entered from the south and output pins exited north. The clock pin was placed in the center. All Pins were placed on M7 for chip top integration. Color aware placement was used and a minimum placement gap of 2 poly pitches was enforced. Fig. 4.6 shows the completed block, where 67% density was achieved.
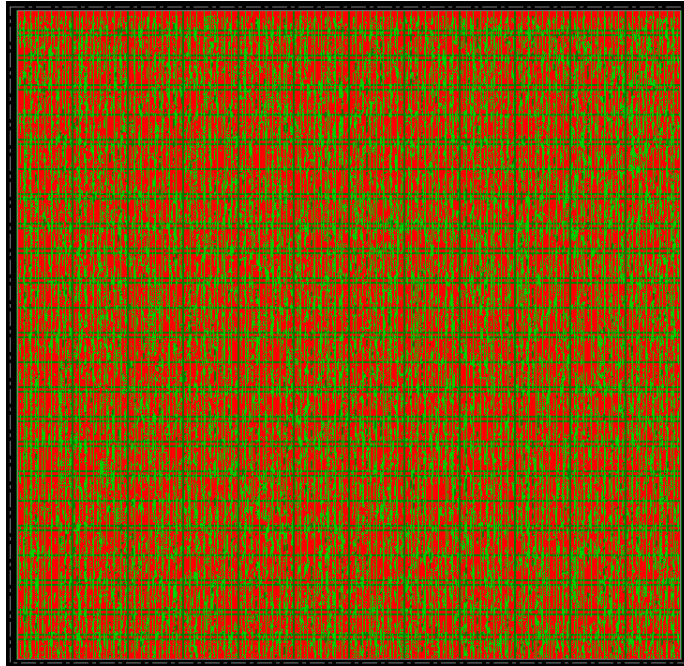
Figure 4.6 APR layout of a typical shift register block. Three blocks were generated, one for each delay size.

CHAPTER 5

SEE TESTING RESULTS

The finished design was fabricated in 14 nm FinFET technology and integrated on

a 3x3 mm die. The packaged chip ASU SNL TC4 was socketed into a custom daughter

card and the I/O was routed to a Xilinx FPGA for external clock generation and data

control. The completed test setup as shown in Fig 5.1 was taken to the Lawrence

Berkeley National Lab (LBNL) for heavy ion testing (Lawrence Berkeley National

Laboratory 88-Inch Cyclotron 2021).

Figure 5.1 Packaged part socketed into the daughter card and connected to the Xilinx
FPGA controller.

5.1   LBNL 88in Cyclotron Heavy Ion Testing

The 4CEFF shift registers were tested in dynamic, alternating data on the input and

an active clock during irradiation, and two static modes. The static modes consisted of

fully loading the shift registers with either ones or zeros, halting the clock before

irradiation, and then shifting out the data after the exposure run was complete. Dynamic

testing was done at 50 MHz The 88 in. cyclotron provided 16 AMeV beams. A baseline

DFF shift chain on a separate chip, ASU SNL TC3 containing a standard DFF shift

register, was irradiated during the same trip. Figure 5.2 shows the experimental setup

mounted in air and centered in front of the beam aperture. TC3 and TC4 shared identical

packages and I/O and allowed for identical test setup.

Forty-two individual test runs were conducted by setting a max fluence and then

reading out the data between runs. The Xilinx FPGA was cleared and reinitialized

between each run. Table 3 summarizes the settings and data for each run.

Figure 5.2 Test setup mounted and centered in the beam.

Table 3: Heavy Ion Test Data

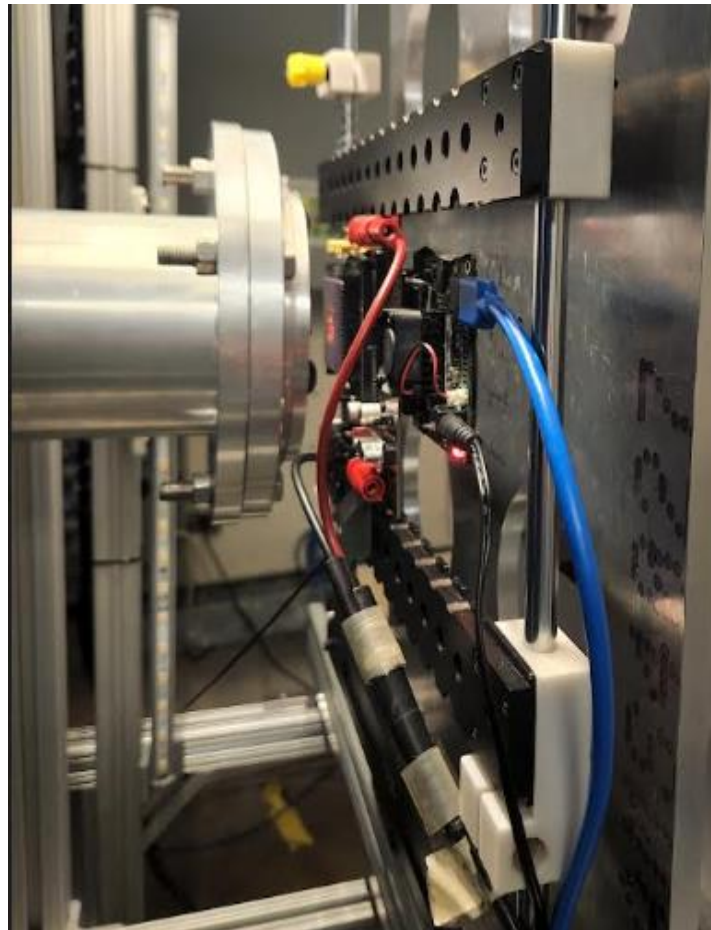| Run	Ion	LET	Fluence	Test Type	Target	Run	Ion	LET	Fluence	Test Type	Target
1001	Kr	25	1.01E+07	Dynamic	StdDFF	1034	V	10.9	1.00E+07	Dynamic	DUT3
1002	Kr	25	1.00E+07	Dynamic	StdDFF	1035	V	10.9	1.01E+07	Dynamic	DUT2
1003	Kr	25	1.01E+07	Static 0s	StdDFF	1036	V	10.9	1.01E+07	Dynamic	DUT1
1005	Kr	25	1.01E+07	Static 0s	StdDFF	1037	V	10.9	1.01E+07	Static 0s	DUT3
1006	Kr	25	1.01E+07	Dynamic	StdDFF	1038	V	10.9	1.01E+07	Static 0s	DUT2
1007	Ar	7.27	1.00E+07	Dynamic	StdDFF	1039	V	10.9	1.01E+07	Static 0s	DUT1
1008	Ar	7.27	1.00E+07	Static 0s	StdDFF	1040	V	10.9	1.01E+07	Static 1s	DUT3
1009	Ne	2.39	1.00E+08	Static 0s	StdDFF	1041	V	10.9	1.01E+07	Static 1s	DUT2
1010	Ne	2.39	3.01E+07	Dynamic	StdDFF	1042	V	10.9	1.01E+07	Static 1s	DUT1
1011	N	1.16	3.01E+07	Dynamic	StdDFF	301	Ar	7.27	3.00E+07	Static 1s	DUT1
1012	N	1.16	3.01E+07	Static 0s	StdDFF	302	Ar	7.27	3.00E+07	Static 1s	DUT2
1013	N	1.16	3.01E+07	Static 0s	StdDFF	303	Ar	7.27	3.00E+07	Static 1s	DUT3
1014	N	1.16	3.01E+07	Static 0s	StdDFF	304	Ar	7.27	3.00E+07	Static 0s	DUT1
1015	Xe	49.3	1.00E+07	Dynamic	DUT3	305	Ar	7.27	3.00E+07	Static 0s	DUT2
1016	Xe	49.3	1.00E+07	Dynamic	DUT2	306	Ar	7.27	3.00E+07	Static 0s	DUT3
1017	Xe	49.3	1.00E+07	Dynamic	DUT1	307	Ar	7.27	3.00E+07	Dynamic	DUT1
1018	Xe	49.3	1.00E+07	Static 0s	DUT3	308	Ar	7.27	3.00E+07	Dynamic	DUT2
1019	Xe	49.3	1.00E+07	Static 0s	DUT2	309	Ar	7.27	3.00E+07	Dynamic	DUT3
1020	Xe	49.3	1.00E+07	Static 0s	DUT1	311	Ne	2.39	3.00E+07	Static 1s	DUT2
1021	Xe	49.3	1.00E+07	Static 1s	DUT3	312	Ne	2.39	3.00E+07	Static 1s	DUT1
1022	Xe	49.3	1.00E+07	Static 1s	DUT2	313	Ne	2.39	3.00E+07	Static 1s	DUT3
1023	Xe	49.3	1.00E+07	Static 1s	DUT1	314	Ne	2.39	3.00E+07	Static 0s	DUT1
1025	Kr	25	1.00E+07	Dynamic	DUT3	316	Ne	2.39	3.00E+07	Static 0s	DUT2
1026	Kr	25	1.01E+07	Dynamic	DUT2	317	Ne	2.39	3.00E+07	Static 0s	DUT3
1027	Kr	25	1.00E+07	Dynamic	DUT1	318	Ne	2.39	3.00E+07	Dynamic	DUT1
1028	Kr	25	1.01E+07	Static 0s	DUT3	319	Ne	2.39	3.00E+07	Dynamic	DUT2
1029	Kr	25	1.00E+07	Static 0s	DUT2	320	Ne	2.39	3.00E+07	Dynamic	DUT3
1030	Kr	25	1.01E+07	Static 0s	DUT1	357	Xe	49.3	1.70E+06	Dynamic	stdDFF
1031	Kr	25	1.01E+07	Static 1s	DUT3	358	Xe	49.3	1.00E+07	Dynamic	stdDFF
1032	Kr	25	1.01E+07	Static 1s	DUT2	359	Xe	49.3	1.00E+07	Dynamic	stdDFF
1033	Kr	25	1.01E+07	Static 1s	DUT1	360	Xe	49.3	1.00E+07	Static 0s	stdDFF

5.1.1   Results

Error counts for the various runs were tabulated by shift register and totaled for the nine configurations. The cross sections were then calculated and normalized by the number of flip-flops of each configuration which was 12,000. Dynamic testing proved promising as demonstrated by Fig 5.3 shown at the end of this section. The threshold

59

LET shifted up around 6 MeV and the saturation cross section remained slightly lower than the baseline DFF at low LET. The largest delay was the only one that appeared to have a higher saturation cross section at high LET and only for the single bit non-interleaved 4CEFF. Table 4 summarizes the collected cross section data for the dynamic runs. Values in red indicate they are limiting cross sections. No errors were detected, values were generated assuming the next particle would cause an upset.

Table 4: Dynamic Cross Sections in cm^2

LET	Cross Section cm^2				
MeV·cm^2/mg	49.3	25	10.9	7.3	2.4
Dly3 P4	6E-10	1.52E-10	9.5E-12	3.2E-12	3.2E-12
Dly3 P2	7.67E-10	1.58E-10	8.3E-12	2.8E-12	2.8E-12
Dly3 P1	1.15E-09	3.58E-10	3.3E-11	2.8E-12	2.8E-12
Dly2 P4	5.24E-10	3.3E-10	9.5E-12	1.6E-11	3.2E-12
Dly2 P2	6.92E-10	1.32E-10	1.7E-11	8.3E-12	2.8E-12
Dly2 P1	7.73E-10	4.21E-10	1.7E-11	1.4E-11	2.8E-12
Dly1 P4	6.14E-10	1.81E-10	8.5E-11	6.3E-12	3.2E-12
Dly1 P2	7.33E-10	2.58E-10	8.3E-11	2.8E-12	2.8E-12
Dly1 P1	8.42E-10	4.42E-10	1.6E-10	1.7E-11	2.8E-12

Static testing revealed significant differences depending on the data state of the 4CEFF. Table 5 summarizes the cross-section data for the static zero testing runs. When storing zeros, longer delays and higher levels of interleaving dramatically reduced cross section compared to the baseline DFF. The four-bit interleaved flip flops didn't record a single upset even at LET 50. The two-bit interleaved flip flops only recorded four upsets

total over all ten exposures. Only the non-interleaved flops failed to outperform as shown in Fig 5.4. Error counts for even the non-interleaved flops dropped into the single digits for all LET values below 25 MeV·cm^2/mg. Detected errors using Argon at LET 7.3 MeV·cm^2/mg after seeing zero errors using 10.9 MeV·cm^2/mg Vanadium can be attributed to tripling the flux for the low LET runs.

Table 5: Static Cross Sections Storing Zeros

LET MeV·cm^2/mg	Cross Section cm^2				
	49.3	25	10.9	7.3	2.4
Dly3 P4	4.76E-12	4.71E-12	4.71E-12	3.17E-12	3.17E-12
Dly3 P2	8.33E-12	4.13E-12	4.13E-12	2.78E-12	2.78E-12
Dly3 P1	5.33E-10	3.30E-10	4.13E-12	2.78E-12	2.78E-12
Dly2 P4	4.76E-12	4.76E-12	4.71E-12	3.17E-12	3.17E-12
Dly2 P2	4.17E-12	4.17E-12	4.13E-12	2.78E-12	2.78E-12
Dly2 P1	4.89E-10	2.25E-10	4.13E-12	2.78E-12	2.78E-12
Dly1 P4	4.76E-12	4.71E-12	4.71E-12	3.17E-12	3.17E-12
Dly1 P2	4.17E-12	8.25E-12	4.13E-12	2.78E-12	2.78E-12
Dly1 P1	6.82E-10	3.21E-10	8.25E-12	1.11E-11	2.78E-12

The limiting cross sections for the interleaved flip flops set about an order of magnitude below the baseline DFF at LET 7.3 MeV·cm^2/mg and two orders of magnitude below at LET 25 MeV·cm^2/mg. The baseline DFF's cross sections in identical conditions were 7.94E-11 and 2.53E-10 cm^2 respectively.

Unfortunately, the same cannot be said for the stored ones condition. All nine configurations failed to outperform the baseline DFF at LET energies above 25 MeV·cm^2/mg. High error counts persisted down well into the low LET range regardless

of levels of interleaving. Longer delays produced higher saturation cross sections and the two-bit interleaving performed the worst for all three delays at LET 49.3 MeV·cm^2/mg. Low LET performance was comparable to the dynamic case with subthreshold LET shifting up moderately, increasing gains as with longer delays as expected. Table 6 summarizes the cross-section data for the static ones testing runs.

Table 6: Static Cross Sections Storing Ones

LET	Cross Section cm^2				
MeV·cm^2/mg	49.3	25	10.9	7.3	2.4
Dly3 P4	1.43E-09	3.19E-10	9.43E-12	3.17E-12	3.17E-12
Dly3 P2	2.1E-09	2.72E-10	4.1E-12	2.78E-12	2.78E-12
Dly3 P1	1.51E-09	3.65E-10	2.48E-11	2.78E-12	2.78E-12
Dly2 P4	1.14E-09	4.04E-10	6.6E-11	2.22E-11	3.17E-12
Dly2 P2	1.19E-09	4.36E-10	1.73E-10	1.39E-11	2.78E-12
Dly2 P1	1.16E-09	5.09E-10	1.07E-10	1.94E-11	2.78E-12
Dly1 P4	1.06E-09	5.27E-10	2.07E-10	1.07E-10	3.17E-12
Dly1 P2	1.24E-09	4.79E-10	2.15E-10	1.16E-10	2.78E-12
Dly1 P1	1.13E-09	4.95E-10	2.88E-10	9.69E-11	2.78E-12

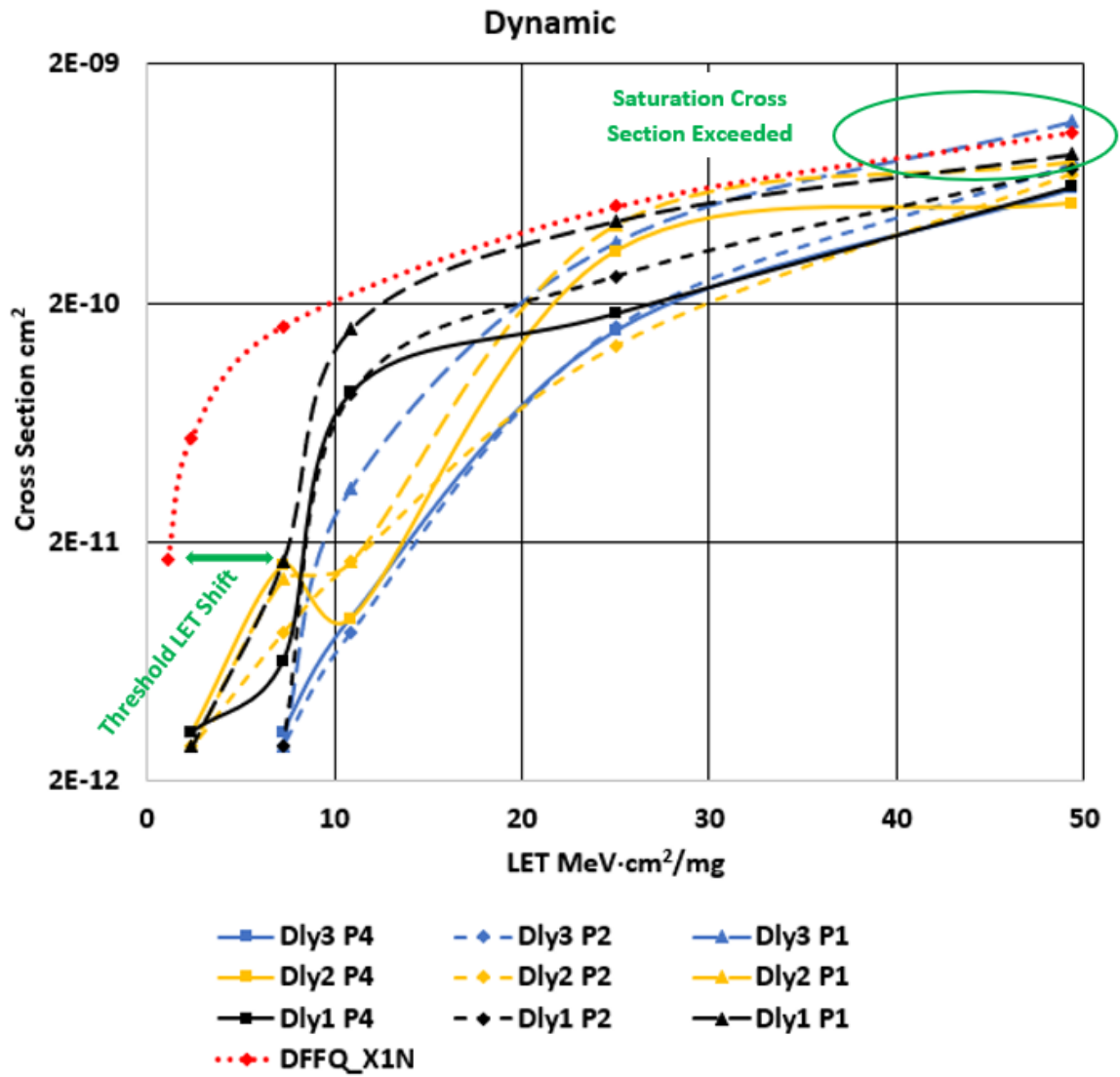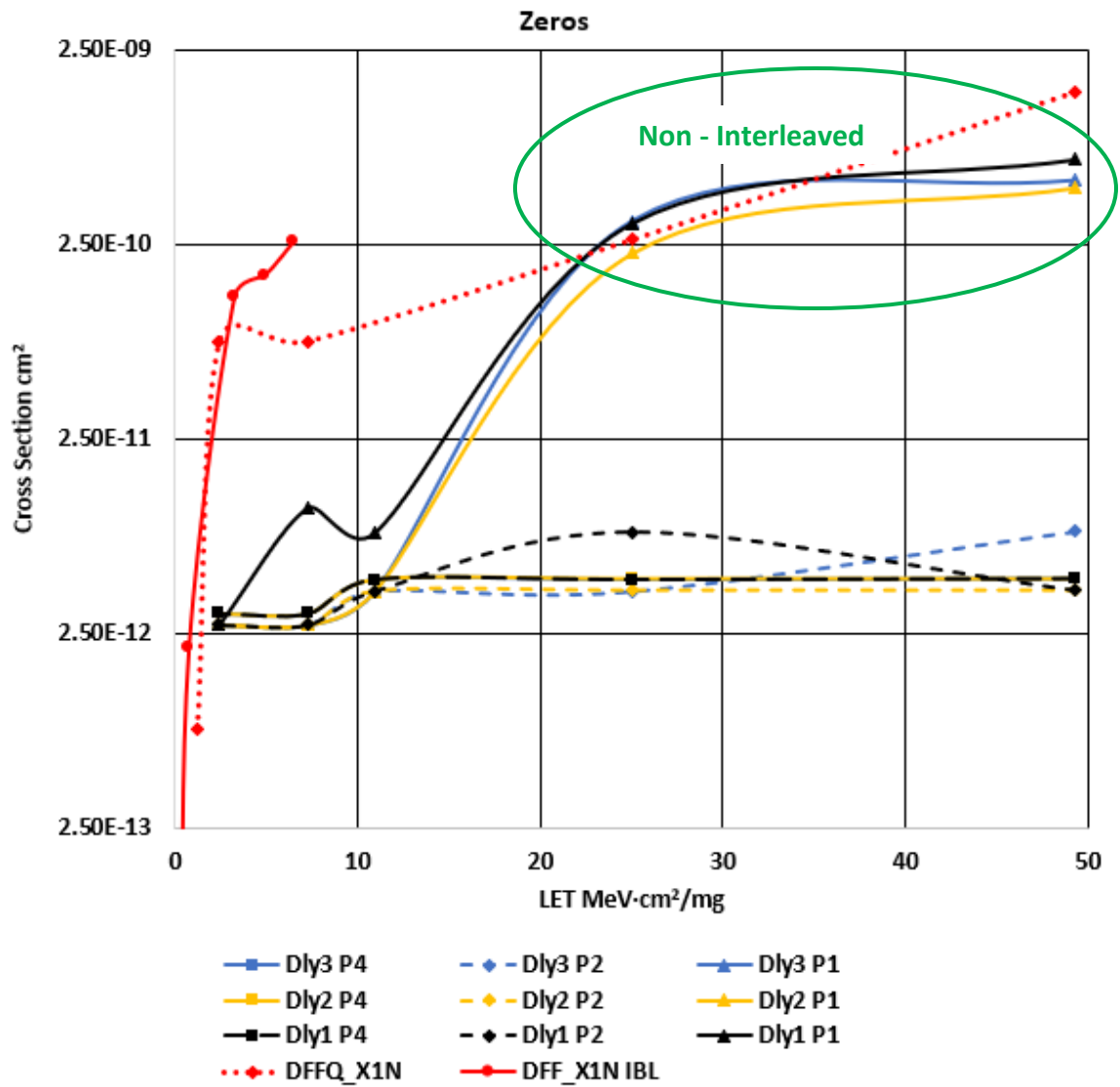Figure 5.3 Dynamic cross sections.

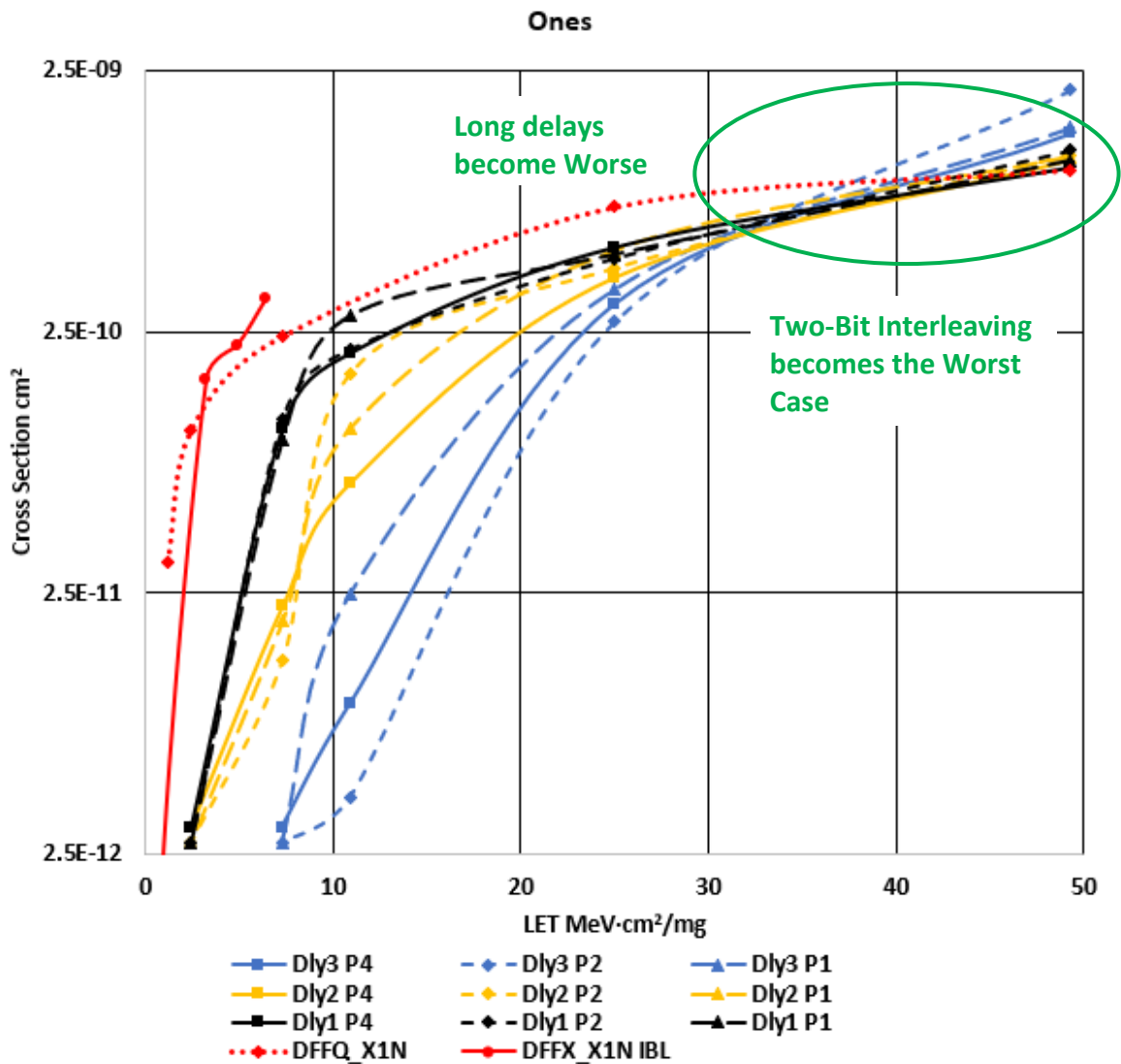Figure 5.4 Static cross sections storing zeros.

Figure 5.5 Static cross sections storing ones

5.1.2   Analysis

While the 4CEFF did successfully raise the threshold LET as intended and

increases to the saturation cross section were moderate in the target LET range, the upset

dependency on storage state presents and unforeseen error mode. Inspection of the data

reveals several trends immediately that help enlighten the cause. First, based on the error counts for the different levels of interleaving the best option for performance is the four-bit interleaved circuit. In all three test cases this level of interleaving produced the lowest error counts. While the two-bit interleaving performed well with the dynamic and static zeros case, it had the worst performance in the static ones case. This implies that the upset mechanism is area dependent and increasing layout area exacerbates it for a low level of interleaving before a critical safe node distance is reached. Second, the discrepancy between the static zero and ones cases. Third, there appears to be an inverse relationship between delay length and SEE performance in the static one case. In the 4CEFF increased delay time is directly proportional to the expected length of tri-state periods on the output of the C-elements, a particularly vulnerable period for that node. These trends taken together, spreading the flip flop a small amount making it worse, storage state dependency, and the period that the circuit is expected to be tri-stated begin to inform the nature of the error mechanism, and a likely source has been previously presented in the literature.

Substrate charge collection in highly scaled finFET technology is widely documented (Fang and Oates 2011, El-Mamouni, Zhang and Pate, et al. 2011, El-Mamouni, Zhang and Schrimpf, et al. 2011). At LET values below 10 MeV·cm^2/mg cross sections for finFETs at normal incidence have been measured that are smaller than the actual geometry of the sensitized area. It has been shown that up to 70% of the charge deposited by a strike is lost to substrate diffusion (Nsengiyumva, et al. 2017). As LET

raises, this diffused charge is collected by sensitized drain areas and can result in upsets from near misses. At high LET the SEE response of finFET circuits begins to match that of planar devices and location dependency of the strike ceases to matter as charge diffusion within the substrate transports charge to the edge of the space charge region of the sensitized nodes for collection over a large area. This secondary collection mechanism is the primary driver of MNCC for nodes beyond the charge deposition radius of the initial ion track.

This substrate charge collection mechanism is likely the culprit for the failure of the 4CEFF in the static ones condition. The 4CEFF is made from two symmetric half latches. Fig 5.6 demonstrates the biasing condition of the master latch with data equal to logic one. In this configuration the outputs of both c-elements, MSETUP and MFDBKB, are at logic one. If there is a strike on any node in the latch, one or both c-elements will tri-state. This cuts off any restoring current to MSETUP and/or MFDBKB. If there is charge in the substrate from the same or another strike, diffusion will distribute the charge over a large area. The excess holes in the drain implant will attract the excess electrons diffusing across the substrate and collect them at the edge of the space charge region. This substrate collection is small, on the order of leakage currents that the driving gate could normally sink or source without issue. Diffusion of deposited charge within the substrate has a long duration though and with the node tri-stated the collection will result in the node steadily leaking down to zero. This helps explain why the phenomena is exacerbated by the longer delays and small increases in area. The extended delays result

67

in longer periods in tri-state and thus extend the vulnerable window of time for charge

collection. Small increases in area expose the entire circuit to a larger area of substrate

without moving them far enough to be beyond the rage of the charge. In the case of the

static zeros, there appears to be insufficient charge collection within the n-well to

facilitate a similar action. This is likely the result of the shallow n-well and small sub-fin

area not having sufficient volume to collect enough charge to impact the tri-stated nodes.

If the performance in the ones case could be made to match the zeros case the 4CEFF

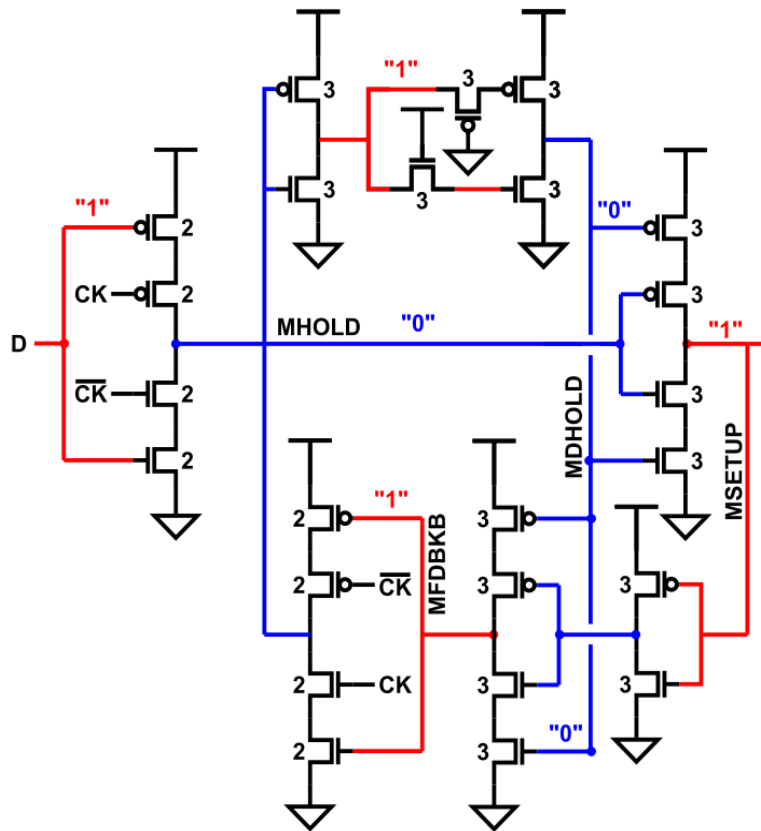would meet or exceed all of our goals.

Figure 5.6 4CEFF half latch biasing in static ones operation.

CHAPTER 6

MITIGATION PLAN

The vulnerability of tri-stated outputs of C-elements or guard gates to SEE is documented in the literature as mentioned above and several means to alleviate the problem have been proposed. I believed nodal separation and redundant c-element usage would mitigate this issue but failed to account for collection from substrate diffusion. Planar devices in previous nodes employed guard rings and annular gate designs but these layouts are no longer possible in the nanoscale finFET technology. Another possibility would be to utilize a second delay and implement a majority voter. This would nearly double the size and power requirements of the Flip Flop and result in significant local routing congestion in an already wire bound design. An initial routing investigation suggests that in order to maintain co-sensitized node separation, up to 25% of the layout would be dark silicon solely for the purpose of routing, a problem already seen with DICE based flops in advanced nodes. The BISER DFF relies on a jam-latch keeper to stabilize and maintain its C-element output. This simple idea requires only local wiring and minimal additional logic.

6.1   P-Keeper Circuit

Using a keeper circuit to provide constant restoring current to the c-element should mitigate the failure mode described above. Similar circuits have utilized jam latches to accomplish this but since the flip flop only fails in the one to zero direction a simple p-keeper circuit will suffice to hold the node at logic one. This keeper circuit must

be sized such that it can be overwhelmed by the nMOS during transition to keep the half-latch writable and must be driven by a complementary node. The driver must also not be a node that will tri-state the same output if it is upset and simultaneously turn off the keeper defeating its purpose. Lastly, if possible, the driving node should not be on the same row as any node it can tristate.

Fortunately, the first downstream inversion from the c-elements satisfies all these conditions as shown in Fig 6.1. Let us analyze the master latch. The feed forward c-element drives node MSETUP and can be tri-stated by MHOLD or MDHOLD. The first down stream inverter drives MSETUPBN. If either MHOLD or MDHOLD are upset, then MSETUPBN will drive the keeper and provide restoring current. If MSETUPBN is upset the keeper will turn off but the pulse will be caught by the second c-element and tri-state MFDBKB. MFDBKB is on a separate row and should not be drawn down by substrate charge collection and is also protected by MHOLD which has not been upset. Consequently, MSETUP does not tri-state. Similar analysis can be done for the second c-element with one additional consideration. A strike to MHOLD will turn off the keeper for MFDBKB but the node will not tri-state until after the pulse propagates through the delay and upsets MDHOLD. For pulses shorter than the delay, MHOLD will have restored the keeper before MDHOLD tri-states MFDBKB. Thus, for all upsets the circuit is designed to filter, the keeper is maintained. To ensure MNCC resilience four-bit interleaving should be employed to maintain maximum separation of the tri-state node drivers from their keeper circuit. One may also want to shift the delay interleaving by two
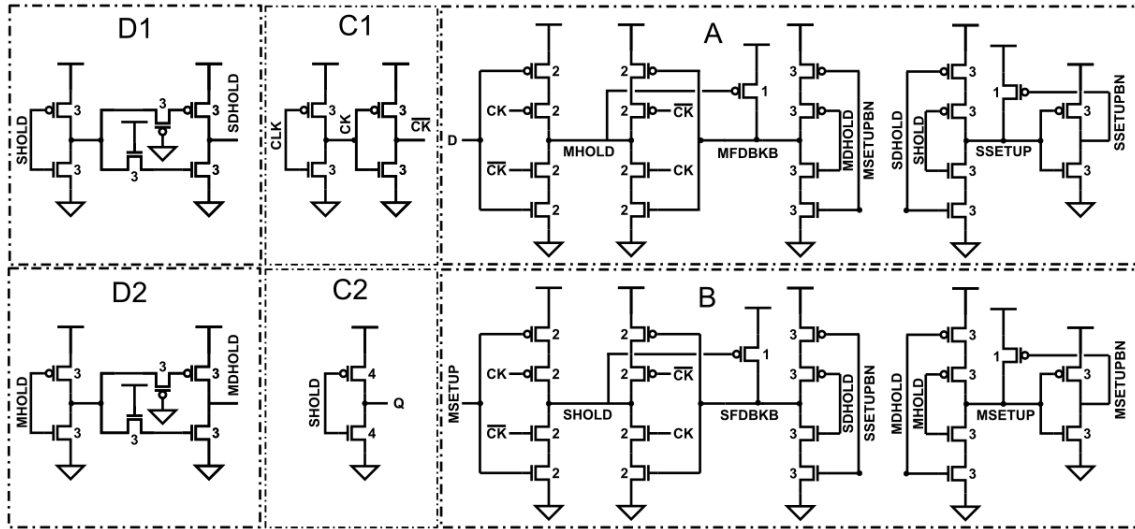
70

Figure 6.1 4CEFF schematic with keepers added.

rows to make sure no delay shares a row with a tri-state node it drives. This would require additional M3 routing over the delays, one of the few areas where such routes are still available.

6.2   Simulation

To ensure functionality and minimal degradation to the pull-down timing the c-element was simulated to determine the impact of the keeper. A single fin keeper transistor was simulated in the test configuration shown in Fig 6.2. The inputs were toggled to generate input high, low, and both tristate conditions. The results of the simulation are shown in 6.3. The half VDD point was reached 5 ps slower and the 80-20 pull-down timing increased by 10 ps, but this had little impact on overall functionality. Fig 6.4 shows the functionality of the 4CEFF V2.0 using the same test bench as the original design. Functionality remains intact with almost no timing impacts. Measured

71

differences in clock to q timing showed an average 1.5 ps (6.8%) increase. Setup times

increased by an average of 6 ps (7.4%) and hold times by an average of 3 ps (33%). In

addition, only four fins of SEE target area was added to the layout or 6.6E-11 cm^2 in the

worst case.
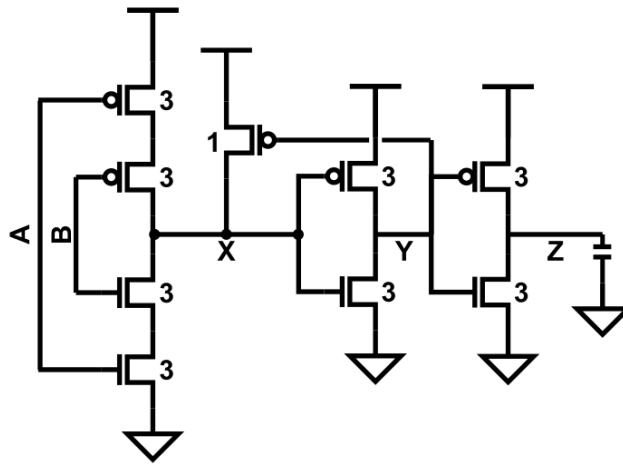
Figure 6.2 C-element test schematic with keepers and load added.

TIME [s]

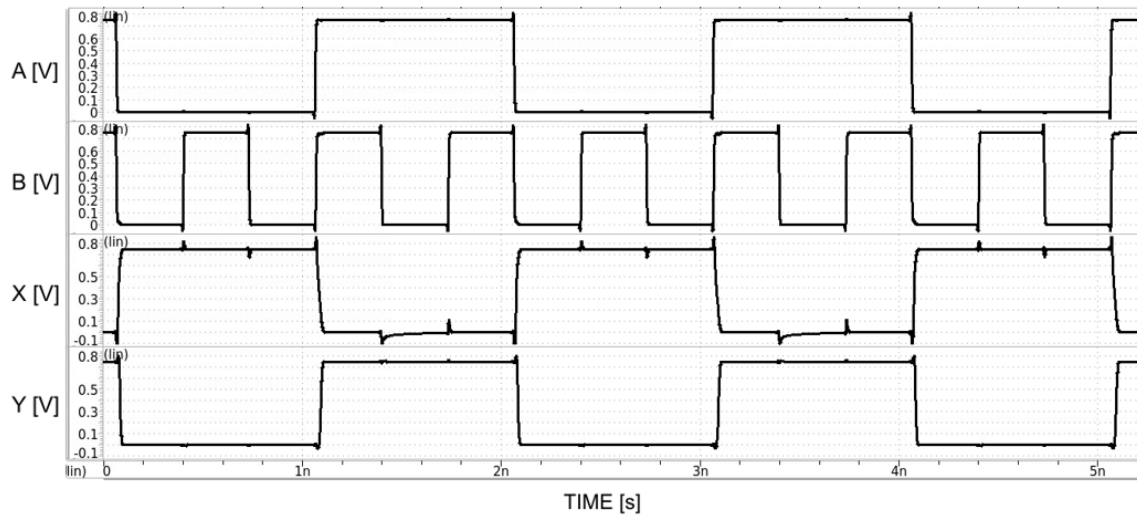Figure 6.3 C-element test waveforms demonstrating functionality and minimal impact on
pulldown slop. Inputs A and B were toggled to generate input high, low, and both tristate

conditions. Node X successfully transitions and drives node Y which provides the feedback to drive the keeper.
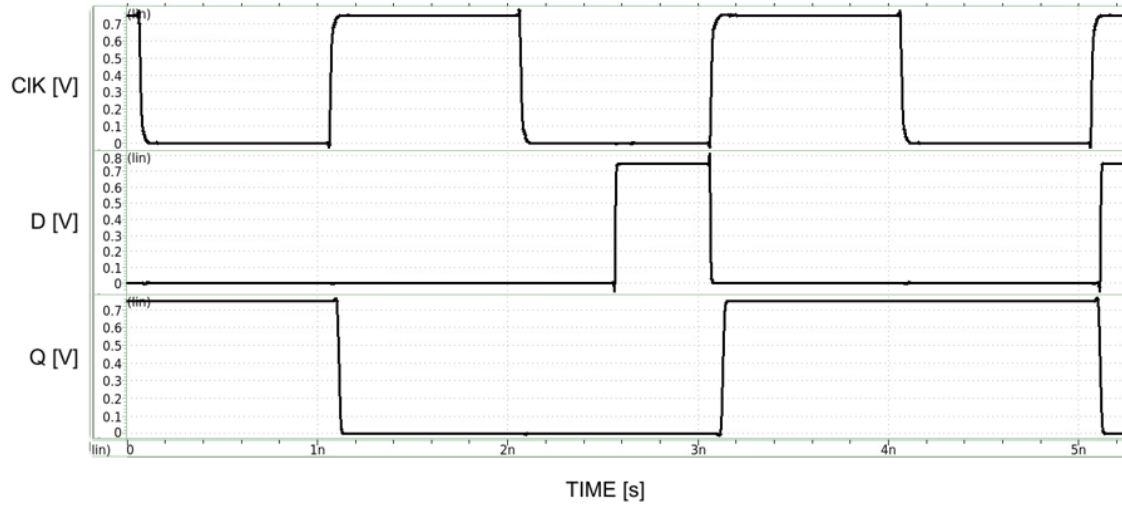

Figure 6.4 4CEFF V2 functional simulation. Setup and Clock to Q timing impacts were on the order of 5 ps.

73

CHAPTER 7

CONCLUSION

This work has demonstrated a methodological approach to RHBD design in advanced nodes starting with characterization of the target technology and library and continuing through the design and implementation and test of an RHBD test circuit. A summary of the state of the field and the various flip flop hardening design techniques was presented with discussion of the advantages and disadvantages of the various methods. The details of a full implementation including heavy ion testing were presented with results and analysis. Finally, a failure mode was described and analyzed and a circuit-based remedy was designed and simulated.

For a proper basis of design to be established, characterization of the target technology and cell library must be performed beginning at the device level. To facilitate this SEE have been characterized in TCAD for the target 14nm technology and expected charge collection and pulse durations were experimentally determined. This data was used to tune a SPICE level macro model to then analyze the worst case SEE scenarios in the cell library. These simulations were used to properly analyze the performance of various delay cell layouts but can be generalized to understand the SEE performance of the cell library itself. Any representative group of cells can be used to estimate the SEE response of an arbitrary complex circuit using the same method.

With this understanding, over forty delay designs were simulated and analyzed for power area and resilience to SEE. Care was taken to ensure that the insertion of the

74

delays would not introduce a new weakest node to the circuit and that the power and area impact was minimized. Analysis of the delays revealed that the pass gate family of single stage delays offered the highest performance. Of primary benefit was significant reductions in power and increased resilience to internal upsets. Use of the macro models allowed for the characterization the new circuits pre-silicon for qualitative comparison, greatly reducing the risk and overhead required to explore new design tradeoffs. In the final analysis, three different sizes of delay were selected for use based on the characterization effort. Specifically, those whose durations would filter the upsets below our target threshold LET range of 10 to 40 MeV·cm^2/mg.

Upset campaigns were then undertaken for the 4CEFF and standard DFF. Estimates showed a 3x increase in minimum energy required to upset the 4CEFF when using the smallest selected delay. Additional systematic MNCC upset simulation identified pairs of co-sensitized nodes that were then separated in the layout. Increased nodal separation was then added via three levels of multi bit interleaving. This information was presented through pre-layout nodal separation schematics and final layout routing the interleaved designs. This level of intra and inter flip flop node isolation resulted in a completely wire bound design on M2 and M3 that required access pins be extended to the edges of the cell layout. This demonstrates the difficulty of maintaining cell density while achieving RHBD mitigation of MNCC. This problem is typical of any redundant design in advanced nodes and is a limiting factor for DICE based designs where subcomponents are not easily isolated and separated from one another. Early

75

identification of co-sensitized nodes is therefore critical to layout planning and successful local routing, something made possible by this systematic multi-node upset scheme.

The nine 4CEFF variations were then successfully integrated into three sets of shift registers for heavy ion testing. The results revealed that the threshold LET had raised 6 MeV·cm^2/mg while only moderately increasing the saturation cross section. There was however an unforeseen error mode when storing ones. Subthreshold charge collection caused the tri-stated nodes of the delay filters to leak down while storing a logic one, allowing errors to propagate through downstream filters. This process was exacerbated by long delays that increased the period the nodes were tri-stated and thus increased subthreshold collection time. This caused the cross sections in the ones case to be between 1.5 and 2x higher than the zeros case and rendered their performance worse than the standard DFF. Our systematic design of experiment approach, utilizing multiple levels of interleaving and delay length, provided the information necessary to trace the error to a likely cause. Finally, based on the analysis of our test data a new design integrating a p-keeper feedback circuit was presented along with functional simulation to demonstrate its viability. Impacts on performance were limited with setup time increasing 6 ps, hold increased 3ps, and clock to q increased 1.5 ps.

The main contributions of this work demonstrate of a systematic approach to RHBD design. This work selected the 4CEFF circuit for implementation in 14nm finFET technology, but the same methods and procedures can be applied to any circuit design process. The work presents a unified approach from base technology through to complex

circuit design that elucidates the design decision space at every stage. The importance of a design of experiment approach is also made clear via the failure analysis only made possible by our muti-variable multiple test vector experimentation. Integration of this data allowed for the rapid application of our process to facilitate a subsequent improved design. As nodes continue to scale and we push computing resources farther into space and other radiation environments, the need to reliably design resilient systems will only increase. This work has presented one path to addressing that need.

Summary of Contributions:

- TCAD characterization of SEE in 14nm finFETs.

- SPICE simulation of SEE in target cell library.

- Pre-Silicon qualitative comparisons of resilience between baseline and proposed design.

- MNCC aware layout designed via automated multi-node upset simulation.

- Physical design and implementation of 9 RHBD multi-bit flip-flop variations.

- Heavy ion test data and analysis of RHBD and baseline circuits.

- An Improved design for future implementation.

REFERENCES

14nm Lithography Process. Accessed January 23, 2020.
    https://en.wikichip.org/wiki/14_nm_lithography_process.

Andjelkovic, M, A Ilic, Z Stamenkovic, M Krstic, and R Kraemer. 2017. "An Overview
    of the Modeling and Simulation of the Single Event Transients at the Circuit
    Level." *Proc. 2017 IEEE 30th International Conference on Microelectronics.* Nis:
    IEEE. 35-44.

Artola, L, M Gaillardin, G Hubert, M Raine, and P Paillet. 2015. "Modeling Single Event
    Transients in Advanced Devices and ICs." *IEEE Transactions on Nuclear Science*
    62 (4): 1528-1539.

Barnaby, H J. 2006. "Total-Ionizing-Dose Effects in Modern CMOS Technologies."
    IEEE Transactions on Nuclear Science 53 (6): 3103-3121.

Baumann, R C. 2013. "Landmarks in Terrestrial Single-Event Effects." *NSREC Short
    Course.* San Francisco, CA: IEEE.

Baumann, R. 2001. "Single Event Transients in Fast Electronic Circuits." *NSREC Short
    Course.* Vancouver: IEEE.

Bazizi, E M, A Zaka, T Herrmann, F Benistant, J H M Tin, J P Goh, L Jiang, M Joshi, H
    van Meer, and K Korablev. 2014. "Advanced TCAD for Predictive FinFETs Vth
    Mismatch Using Full 3D Process/Device Simulation." *Proc. 44th European Solid
    State Device Research Conferance.* Venice. 341-344.

Black, J D, D R BALL II, W H Robinson, D M Fleetwood, R D Schrimpf, R A Reed, D
    A Black, et al. 2008. "Characterizing SRAM single event upset in terms of single
    and multiple node charge collection." *IEEE Transactions on Nuclear Science* 55
    (6): 2943-2947.

Buturla, E M, P E Cottrell, B M Grossman, and K A Salsburg. 1981. "Finite-Element
    Analysis of Semiconductor Devices: The FIELDAY Program." *IBM Journal of
    Research and Development* 25 (4): 218-231.

Calomarde, A, A Rubio, F Moll, and F Gamiz. 2020. "Active Radiation-Hardening
    Strategy in Bulk FinFETs." *IEEE Access* 8: 201441-201449.

Clark, L T. 2010. "Microprocessors and SRAMs for Space: Basics, Radiation Effects and Design." *NSREC Short Course.* Denver, CO.: IEEE.

Diggins, Z J, N J Gaspard, N N Mahatme, S Jagannathan, T D Loveless, T R Reece, B L Bhuva, et al. 2013. "Scalability of Capacitive Hardening for Flip-Flops in Advanced Technology Nodes." *IEEE Transactions on Nuclear Science* 60 (6): 4394-4398.

Dodd, P E. 1999. "Basic Mechanisms for Single Event Effects." *NSREC Short Course.* Albuquerque, NM: IEEE.

El-Mamouni, F, E X Zhang, N D Pate, N Hooten, R D Schrimpf, R A Reed, K F Galloway, et al. 2011. "Laser- and Heavy Ion-Induced Charge Collection in Bulk FinFETs." *IEEE Transactions on Nuclear Science* 58 (6): 2563-2569.

El-Mamouni, F, E X Zhang, R D Schrimpf, R A Reed, K F Galloway, D McMorrow, E Simoen, C Claeys, S Cristoloveanu, and W Xiong. 2011. "Pulsed laser-induced transient currents in bulk and silicon-on-insulator FinFETs." *Proc. 2011 International Reliability Physics Symposium.* Monterey: IEEE. SE.4.1-SE.4.4.

Esposito, M G, J E Manuel, A Privat, T P Xiao, D Garland, E Bielejec, G Vizkelethy, et al. 2021. "Investigating Heavy-Ion Effects on 14-nm Process FinFETs: Displacement Damage Versus Total Ionizing Dose." *IEEE Transactions on Nuclear Science* 68 (5): 724-732.

Fageeha, O, J Howard, and R C Block. 1994. "Distribution of Radial Energy Deposition Around the Track of Energetic Particles in Silicon." *Journal of Applied Physics* 75 (5): 2317-2321.

Fang, Y P, and A S Oates. 2011. "Neutron-induced charge collection simulation of bulk FinFET SRAMs compared with conventional planar SRAMs." *IEEE Transactions on Device and Materials Reliability* 11 (4): 551-554.

Gwyn, C W, D L Scharfetter, and J L Wirth. 1967. "The Analysis of Radiation Effects in Semiconductor Junction Devices." *IEEE Transactions on Nuclear Science* 14 (6): 153-169.

Hamed, E A, and I Lee. 2021. "Categorization and SEU Fault Simulations of Radiation-Hardened-by-Design Flip-Flops." *Electronics* 10 (13): 1572-1591.

Hindman, N D, L T Clark, D W Patterson, and K E Holbert. 2011. "Fully Automated, Testable Design of Fine-Grained Triple Mode Redundant Logic." *IEEE Transactions on Nuclear Science* 58 (6): 3046-3052.

IEEE. 1967. "1967 Radiation Effects Conference." *IEEE Transactions on Nuclear Science* 14 (6): 6-7.

James, D. 2016. "Moore's Law Continues into the 1 x-nm Era." *Proc. 27th Annual SEMI Advanced Semiconductor Manufactoring Conferance.* Saratoga Springs. 1-10.

Knudsen, J, and L T Clark. 2006. "An Area and Power Efficient Radiation Hardened by Design Flip-Flop." *IEEE Transactions on Nuclear Science* 53 (6): 3392-3399.

Koga, R, K B Crawford, P B Grant, W A Kolasinski, D L Leung, T J Lie, D C Mayer, S D Pinkerton, and T K Tsubota. 1993. "Single ion induced multiple-bit upset in IDT 256 K SRAMs." *Proc. Radiation Effects on Components and Systems.* Saint-Malo: IEEE. 485-489.

2021. *Lawrence Berkeley National Laboratory 88-Inch Cyclotron.* October 7. Accessed April 29, 2022. http://cyclotron.lbl.gov/.

Mangeret, R. 2018. "Radiation Hardness Assurance: How Well Assured Do We Need to Be?" *NSREC Short Course.* Kona, HI: IEEE.

Matush, B, T Mozdzen, L T Clark, and J Knudsen. 2010. "Area-Efficient Temporally Hardened by Design Flip-Flop Circuits." *IEEE Transactions on Nuclear Science* 57 (6): 35883595.

Mavis, D, and P Eaton. 2002. "Soft Error Rate Mitigation Techniques for Modern Microcircuits." *Proc. IEEE International Reliability Physics Symposium.* Lake Tahoe: IEEE. 216-225.

Munteanu, D, and J L Autran. 2008. "Modeling and Simulation of Single Event Effects in Digital Devices and ICs." *IEEE Transactions on Nuclear Science* 55 (4): 1854-1878.

Nagel, L W, and D O Pederson. 1973. "SPICE (Simulation Program with Integrated Circuit Emphasis)." *Memorandum no. ERL-M382.* University of California, Berkeley, April.

Nagel, L, and R Rohrer. 1971. "Computer Analysis of Nonlinear Circuits, Excluding Radiation (CANCER)." *IEEE Journal of Solid-State Circuits* 6 (4): 166-182.

Naseer, R, and J Draper. 2006. "DF-DICE: A Scalable solution for Soft Error Tolerant Circuit Design." *Proc. IEEE International Symposium on Circuits and Systems.* Island of Kos: IEEE. 3890-3893.

Nsengiyumva, P, L W Massengill, M L Alles, B L Bhuva, D R Ball, J S Kauppila, T D Haeffner, W T Holman, and R A Reed. 2017. "Analysis of Bulk FinFET Structural Effects on Single-Event Cross Sections." *IEEE Transactions on Nuclear Science* 64 (1): 441-448.

Petrovic, V, and M Krstic. 2015. "Design Flow for Radhard TMR Flip-Flops." *Proc. 18th International Symposium on Design and Diagnostics of Electronic Circuits & Systems.* Belgrade: IEEE. 203-208.

Privat, A, and L T Clark. 2015. "Simple and Accurate Single Event Charge Collection Macro Modeling for Circuit Simulations." *Proc. IEEE International Symposium on Circuits and Systems.* Lisbon: IEEE. 1858-1851.

Raine, M, G Hubert, P Paillet, M Gaillardin, and A Bournel. 2011. "Implementing Realistic Heavy Ion Tracks in a SEE Prediction Tool: Comparison Between Different Approaches." *2011 12th European Conference on Radiation and Its Effects on Components and Systems* 59 (4): 950-957.

Raine, M, H Guillaume, M Gaillardin, L Artola, P Paillet, S Girard, J Sauvestre, and A Bournel. 2011. "Impact of the Radial Ionization Profile on SEE Prediction for SOI Transistors and SRAMs Beyond the 32-nm Technological Node." *IEEE Transactions on Nuclear Science* 58 (3): 840-847.

Sexton, F W. 2003. "Destructive Single Event Effects in Semiconductor Devices and ICs." *IEEE Transactions on Nuclear Science* 50: 603-621.

Shambhulingaiah, S, C Lieb, and L T Clark. 2015. "Circuit Simulation Based Validation of Flip-Flop Robustness to Multiple Node Charge Collection." *IEEE Transactions on Nuclear Science* 62 (4): 1577-1588.

Shambhulingaiah, S, L T Clark, T J Mozdzen, N D Hindman, S Chella, and K E Holbert. 2011. "Temporal sequential logic hardening by design with a low power delay element." *Proc. 12th European Conference on Radiation and Its Effects on Components and Systems.* Sevilla: IEEE. 144-149.
81

Synopsys. 2013. "Sentaurus Process User Guide." Mountain View, CA.

2001. "Triple Module Redundancy Design Techniques for Virtex FPGAs." *Xilinx App. Note 197*. November.

Van Lint, V A J, J H Alexander, D K Nichols, and P R Ward. 1967. "Computerized Model for Responce of Transistors to a Pulse of Ionizing Radiation." *IEEE Transactions on Nuclear Science* 14 (6): 170-178.

Warren, K M, A L Sternberg, J D Black, R A Weller, R A Reed, M H Mendenhall, R D Schrimpf, and L W Massengill. 2009. "Heavy Ion Testing and Single Event Upset Rate Prediction Considerations for a DICE Flip-Flop." *IEEE Transactions on Nuclear Science* 56 (6): 3130-3137.

Xapsos, M. 2018. "A Brief History if Space Climatology: From the Big Bang to the Present" *NSREC Short Course.* Kona, HI: IEEE.

Zhang, M, S Mitra, T M Mak, N Seifert, N J Wang, Q Shi, K S Kim, N R Shanbhag, and S J Patel. 2006. "Sequential Element Design With Built-In Soft Error Resilience." *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* 14 (12): 1368-1378.