Reliability Analysis of Field-Programmable Gate-Array-Based Space Computer Architectures

Justin A. Hogan†
S2 Corporation, Bozeman, Montana 59715
Raymond J. Weber‡
Flat Earth, Inc., Bozeman, Montana 59718

and

Brock J. LaMeres‡
Montana State University, Bozeman, Montana 59717

DOI: 10.2514/1.I010481

This paper presents an analysis of the radiation tolerance of field-programmable gate-array-based space computers. The primary failure mechanism studied in this paper is single-event effects due to high-energy ionizing radiation. The analysis is performed on the most common architectures deployed on field-programmable gate-array-based systems including simplex, triple modular redundant, inclusion of spares, and configuration memory scrubbing. The reliability of each system is modeled using a Markov chain to predict its mean time to failure. Orbital dependencies are discussed in addition to a comparison of reliability across different process nodes.

Nomenclature

\( L \) = system lifetime, \( s \)
\( p \) = probability of failure
\( R \) = reliability
\( T \) = transition matrix
\( \alpha \) = rate of context switch, \( s^{-1} \)
\( \lambda \) = radiation-induced fault rate, \( s^{-1} \)
\( \mu \) = repair rate, \( s^{-1} \)

I. Introduction

SPACE science and exploration missions drive the need for increasingly powerful computer systems capable of operating in harsh radiation environments. Advances in sensor technology and overall space system complexity necessitate ever-increasing processing power in order to acquire, store, process, and transmit scientific data from the far reaches of the solar system to researchers back on Earth. Existing radio downlinks do not have sufficient bandwidth to handle the data rates that are proposed for future missions. As such, it is desired that space computers perform onboard processing of real-time data and perform data compression procedures before transmitting information back to Earth. This further motivates the need for increased computation within space systems.

To achieve the reliability required by space-based electronic systems, adverse effects caused by ionizing radiation must be mitigated. These effects can be divided broadly into two categories: cumulative effects, and transient effects. As the name suggests, cumulative effects accumulate within a device over time. The rate at which these effects accumulate is a function of the severity of the radiation environment and the duration of exposure. The total ionizing dose (TID) is a measure of cumulative effects in space electronics; it represents the amount of energy deposited in a material per unit of mass and carries the unit of rad [1]. The dose survivable by an electronic device is often in the range of kilorad to megarad. Transient effects are the immediate effects of radiation interactions with electronics. Termed single-event effects (SEEs), these interactions are brief in duration, but their effects may linger in an affected circuit as the state changes to memory elements. SEEs result from interaction with a single particle of ionizing radiation. SEEs are of primary concern in space electronics. They are induced by the transient charge generated by a high-energy particle passing through a semiconductor material. These interactions create erroneous voltages within semiconductor substrates, which may be latched into digital memory elements. The severity of SEEs, in terms of its effect on the operation of the affected system, is highly dependent upon where it occurs within the system. In synchronous digital systems, the time of the strike relative to the system clock edge is also a critical factor in determining the end result of the interaction.

Many modern digital integrated circuits are designed using a metal-oxide-semiconductor field-effect transistor (MOSFET) as the fundamental design building block. Complementary metal-oxide-semiconductor (CMOS) circuits have gained wide acceptance in integrated circuit technology as a direct result of their low power consumption, reliability, and ability to efficiently implement complex logic functions [2]. The performance of CMOS devices continually increases as a result of improvements in manufacturing processes: primarily, the ability to manufacture transistor features on continually smaller scales in accordance with Moore’s law [3]. The MOSFET is a voltage-controlled current device for which the proper operation depends on well-known charge distributions within the device. These charge distributions include the doping concentrations of the source and drain, as well as the base semiconductor material, which constitutes the bulk substrate within which a conduction channel is induced. These electron–hole carrier concentrations, along with the feature dimensions (namely, channel length, width, and gate oxide thickness) define the voltage-current relationship of the device. Changes to the charge distribution due to ionizing radiation adversely affect device performance.

Received 13 June 2016; revision received 24 January 2017; accepted for publication 15 March 2017; published online 6 April 2017. Copyright © 2017 by the American Institute of Aeronautics and Astronautics, Inc. All rights reserved. All requests for copying and permission to reprint should be submitted to CCC at www.copyright.com; employ the ISSN 2327-3097 (online) to initiate your request. See also AIAA Rights and Permissions www.aiaa.org/randp.

*Hardware Design Engineer, 2310 University Way No. 4.
†Research Engineer, 985 Technology Blvd.
‡Associate Professor, Electrical and Computer Engineering Department, 533 Cobleigh. Member AIAA (Corresponding Author).
Due to their popularity, and the fact that they essentially define the state of the art in modern electronic systems, designers have sought to use MOSFETs and other semiconductor devices in space applications. Unfortunately, MOSFETs are particularly susceptible to the harsh radiation environment of space. On Earth, electronic devices are shielded from the vast majority of naturally occurring ionizing radiation thanks to the atmosphere and the magnetosphere [4]. On Earth, there are few natural phenomena capable of randomly injecting charge into MOSFET devices aside from exceptionally high-energy highly penetrating cosmic rays. In space, the reliability of the MOSFET is diminished as a direct consequence of interactions with high-energy ionizing radiation wherein extraneous charge is generated and sometimes trapped within the semiconductor materials. As the reliability of the basic building block decreases, so too does that of the electronic system as a whole.

In this paper, we present a reliability analysis of a variety of radiation fault mitigation approaches deployed in space computer systems. We focus our study on programmable, CMOS commercial-off-the-shelf (COTS) devices due to their increased popularity of late. We first present the motivation for moving toward a programmable COTS solution and then provide an analysis of reliability on the most common architectures used today. We also present a new architecture that includes spare circuitry and dynamic, partial reconfiguration of a device to improve upon the state of the art in reliability.

II. Motivation

A. Radiation Effects on Electronics

In electronic system design for space applications, there is a tradeoff between performance and reliability: particularly, reliability considerations surrounding faults due to the harsh space radiation environment. Electronics in space are not afforded the protection of the Earth’s atmosphere and magnetosphere, and as such are under constant bombardment by high-energy particles, including an assortment of heavy ions, electrons, and protons. Particles possessing sufficient energy are capable of inducing bit errors in computer systems through atomic-level ionization mechanisms. As ionizing radiation transits a semiconductor substrate, it leaves in its wake a charge track consisting of electron–hole pairs. Generated electrons and holes dissipate quickly through the process of recombination. The creation and subsequent dissipation of these electron–hole pairs in the semiconductor are the root cause of radiation effects in modern electronic devices.

Over time, some of the charge generated by the ionization process becomes trapped in gate and field oxides. This accumulation of charge (also known as TID) contributes to performance degradation because it interferes with proper operation of transistors and increases power consumption due to increased leakage currents. Defects in the form of radiation-induced broken bonds give rise to persistent holes located in the oxide layers [5]. These holes tend to accumulate in material defects in the oxide near the silicon/oxide boundary, and similarly in the field oxide regions [1]. Should sufficient charge accumulate in the gate oxide, a conduction channel will be induced in the same way as it is when a control voltage is applied to the gate terminal of the device. Removal of the gate control voltage no longer has the ability to eliminate the conduction channel, and the MOSFET is effectively stuck in an active state. In a negative-channel metal-oxide-semiconductor (NMOS) transistor, this equates to an inability to turn the device off. In a positive-channel metal-oxide-semiconductor (PMOS) transistor, this equates to an inability to turn the device on.

During the process of charge recombination, electron–hole pairs flow within the semiconductor material. This flow of charge constitutes a short-lived current or voltage within the circuit. It is this transient flow of charge that gives rise to SEEs in complimentary metal-oxide-semiconductor circuits. A SEE that is transient in nature and only temporarily affects combinational logic circuits is known as a single-event transient (SET). An SET that occurs in a storage device such as a latch, flip-flop, or other bistable circuit element and results in a stored bit error is known as a single-event upset (SEU). In computer systems, SEUs in semiconductor-based memory become most probably due to the expansive area consumed for data and program storage on the die. When SEUs occur within a portion of the circuit, such as the memory, which results in improper operation of the computer that a reset alone cannot resolve, it is referred to a single-event functional interrupt (SEFI). A common example of this would be a SEU in the program memory of a computer that alters an operation code (hereafter referred to as opcode). In this example, resetting the computer will not put the computer back into an operational state because the opcode itself is corrupted and opcodes are not restored during a reset. In modern programmable logic devices, banks of static random access memory (SRAM) are used to hold the hardware configuration of the fabric. These devices are uniquely susceptible to SEFIs because any SEU that occurs in the configuration memory alters the hardware configuration of the system and ultimately causes overall failures. Figure 1 shows a graphical depiction of how ionizing radiation causes failures in a CMOS device.

B. Traditional Radiation Mitigation Approaches

The most intuitive approach to radiation mitigation is shielding. Shielding from radiation involves the use of a physical barrier between the electronic device and the external radiation environment. The use of shielding is common, and required, in many terrestrial nuclear applications for the purpose of protecting the health of personnel as well as preventing radioactive material from polluting the environment at nuclear sites [1]. In terrestrial applications, the shielding can be of any size and arbitrarily thick. Any type of material, regardless of weight, can be used to achieve shielding specifications. In space applications, the amount of shielding used can become cost-prohibitive due to the cost of transporting the increased mass. In addition, although shielding is effective for lower-energy radiation (less than 30 MeV/amu, atomic mass unit), it can have diminishing returns for higher-energy particles [6]. In fact, shielding may worsen the radiation environment for the electronics as higher-energy cosmic rays (greater than 100 MeV/amu) generate a stream of secondary particles as a result of its interaction with the shielding material.

TID mitigation is performed through various manufacturing techniques designed to minimize imperfections within the transistor gate oxide, at the silicon/oxide interface, and within the crystal lattice of the substrate. As a semiconductor device is exposed to ionizing radiation, small portions
of the generated charge become trapped in oxide layers, at semiconductor/oxide boundaries, and within crystal lattice imperfections throughout the device. Because the gate oxide is particularly susceptible to charge accumulation, reducing its capacity to trap charge in the first place is paramount to eliminating TID. The ability of an oxide to trap charge is a function of its thickness, and it is diminished in modern manufacturing process nodes [7,8]. The hardening of integrated circuit devices against TID currently uses local oxidation of silicon [9] or shallow trench isolation [10] to isolate transistors and reduce leakage effects. These modifications in design layout fall under a category of radiation hardening known as radiation hardened by design (RHBH). Another type of radiation hardening, known as radiation hardened by process (RHBP), seeks to reduce charge trapping in the semiconductor materials by minimizing defects in the substrate and oxide layers during the manufacturing process. Polishing of wafer surfaces and careful control of oxidizing conditions helps minimize oxide interface defects, thus reducing the number of possible charge traps in the device. RHBP and RHBH processes have successfully achieved TID immunity up to 1 Mrad [11].

As gate oxide thicknesses have decreased with advances in manufacturing technology, TID has become less of a consideration in general because both the amount of charge buildup in an oxide as well was the resultant threshold voltage shift are directly proportional to oxide thickness. Manufacturing techniques have significantly reduced the total ionizing dose as a consideration when using parts designed at 45 nm or less [7]. For example, the Virtex-6 implemented in 40 nm technology has achieved 380 krad and 1 Mrad with reduced timing [12]. As a result, off-the-shelf components are becoming more inherently TID immune. Modern gate oxide thicknesses effectively eliminate TID as a design consideration when using cutting-edge off-the-shelf components.

The primary effect to be mitigated in modern CMOS devices are SEEs. These upsets vary in severity, depending upon where they occur, but must be mitigated globally in order to achieve the reliability required by space systems. A handful of approaches exist for mitigating SEEs at the fabrication level. These approaches typically focus on minimizing the amount of charge generated within the semiconductor substrate. For example, if the transistor is built on top of an insulating material, it becomes less susceptible to SEEs due to the reduced amount of charge generated by radiation because the semiconductor volume passed through is reduced. This technology is known as silicon on insulator. Electrons in insulators require much more energy to be excited into the conduction band than in a semiconductor. This insulator property results in less charge generation under radiation exposure [13].

In addition to process mitigation, there are many techniques for eliminating SEEs at the architectural level [14,15]. These techniques can be implemented on the custom integrated circuit (IC) materials mentioned previously or in traditional COTS fabrics. A common system-level technique for SEU mitigation is triple modular redundant (TMR) [16]. TMR systems run three identical components in parallel. The outputs of these concurrent operations are voted upon by a majority-rules voting circuit to determine if any of the system components are faulty and to prevent erroneous outputs from propagating through the system. TMR has been applied with varying granularity ranging from bit-level triplication of circuits to system-level triplication of major electronic components [17]. This approach to mitigating radiation effects through redundancy is rooted in early theoretical work concerning how perfectly reliable computing machines could be created from inherently unreliable components, e.g., mechanical relays [18]. Though the technology has changed, this technique for mitigating component-level faults has remained relevant.

Another technique for imparting radiation hardness on stored program computer systems is to ensure the integrity of memory contents through the use of memory scrubbing. In this process, the contents of memory locations are periodically rewritten with known good data. This prevents errors in memory from accumulating, and it reduces the likelihood of using corrupted data values in computations. Scrubbing can either be blind or use readback technology. Blind scrubbing simply overwrites the contents of the memory, regardless of its validity. A readback scrubbing process reads the contents of a memory location, compares it to the desired value, and only performs a write operation if there is a discrepancy. In field-programmable gate-array (FPGA) systems, memory scrubbing is used as a way to maintain the configuration memory contents. The presence of a scrubber significantly increases the overall reliability of a system [19].

The use of memory systems that are inherently more radiation tolerant is also an approach to mitigating SEEs. Systems based on antifuse memory are inherently immune to SEEs due to the removal of a transistor in the circuit being used to store the logic level. Systems based on electrically erasable programmable read only memory (EEPROM) or FLASH memory also provide a higher level of immunity to SEEs due to the higher amount of energy needed to cause a SET. FPGAs with FLASH-based configuration memory are commercially available from companies such as Microsemi (i.e., the RTG4).

C. Drawbacks of Existing Mitigation Techniques

TID mitigation techniques lead to reduced performance and increased cost. Shielding adds mass to a system, and mass is the driver for launch services cost. Adding excessive shielding can quickly cause a space system to be impractically expensive. Because realistic thicknesses of shielding are only effective for lower-energy particles, it also only aids in TID mitigation and leaves SEEs to be mitigated using different techniques.

The TID hardening techniques discussed in the previous section rely on design-specific features or process-level modifications to achieve adequate hardness. Though effective, these techniques add substantial costs to the manufacturing process. Once designed and manufactured, new devices must undergo extensive testing to demonstrate their radiation hardness. These tests are very expensive, and yet another cost that is passed on to the customer. Because there is simply not a large market for such devices, radiation-hardened parts do not benefit from cost reductions associated with volume manufacturing. The end result is radiation-hardened devices that are significantly more expensive than their off-the-shelf counterparts. Radiation-hardened computer systems can cost as much as 40 times more than an equivalent COTS system [20].

In addition to being more expensive, radiation-hardened components, in general, exhibit lower performance than commercial devices. The performance lag is a result of the manufacturing techniques used to protect the devices from radiation. The techniques use older process nodes resulting in larger minimum feature sizes, and consequently slower switching times, greater power consumption, and lower performance. Also, these design and layout techniques add area to the circuitry, thus decreasing the performance further. Radiation-hardened microprocessors generally lag commercial devices in performance by 10 or more years [21].

SEE mitigation approaches implemented at the architectural level (i.e., TMR and scrubbing) add additional circuitry to the system. This additional circuit area increases the chances of active circuitry being struck by ionizing radiation. In fact, if implemented improperly, they can actually decrease the SEE reliability of the system [22]. As such, care must be taken when deploying such techniques. Systems based on antifuse memory lack the flexibility to be reconfigured during a mission and are becoming less common in modern designs. The primary drawback of systems based on FLASH or EEPROM memory is that they have slower memory access cycles. Additionally, they are still susceptible to SEEs and require traditional mitigation approaches such as scrubbing.

D. Commercial-Off-the-Shelf FPGA-Based Computing as a Solution

A move to COTS devices in space applications is desirable due to their substantially lower cost and performance superiority to radiation-hardened devices. However, if a COTS solution is used, it must be resilient to both TID and SEE failure mechanisms. Historically, the feature sizes of integrated circuits used to implement space computers were such that TID was the primary concern with respect to radiation. Larger devices had thick oxide insulators that were highly susceptible to charge trapping because of their relatively large volumes within the device.
Simultaneously, the diffusion regions of the older devices were large enough that a high-energy particle strike did not cause sufficient energy in order to change the state of a logic gate. This was because the radiation particle sizes were relatively small compared to the diffusion region volume and a strike could not create a sufficient amount of charge to switch the device. In modern integrated circuits (e.g., below the 65 μm process node), the feature sizes have been reduced to the point where TID is no longer the dominating failure mechanism. This is because the oxide thicknesses of the devices are so thin that the statistical probability of charge getting trapped is minimal. Thus, modern ICs are becoming inherently tolerant to TID. For example, modern FPGAs are achieving TID tolerance levels greater than 300 krad when implemented in the 65 nm process node and as much as 600 krad when implemented in a 22 nm node [23]. As TID immunity is increased with each subsequent process node, so is the susceptibility to SEEs. In modern devices, faults caused by SEEs are now the greatest concern [24]. This is because the diffusion regions of modern devices have been reduced in size to the point where the charge caused by a radiation strike is sufficient to cause a state change.

FPGAs, particularly SRAM-based devices, provide a unique flexibility to aerospace systems. A single FPGA can be used to implement multiple system functions by loading different configuration bitstreams based on current system needs or operating mode. This allows hardware sharing by nonconcurrent processes that would otherwise require independent hardware systems, resulting in an overall reduction in component count and system complexity. One of the most limited resources on space systems is electrical power. Reducing the amount of power used in computation allows increased system runtime. More advanced configuration features, such as active partial reconfiguration, allow specific portions of an FPGA to be reprogrammed without affecting the operation of the rest of the FPGA. This allows hardware peripherals to be instantiated on an as-needed basis, resulting in power savings through an overall reduction in device resource utilization.

The remaining element of an FPGA-based space computer system is implementing an SEE mitigation strategy. As such, much work of late has been in developing and testing FPGA-based flight hardware [25, 26]. The most widely adopted technique for fault mitigation is to use a combination of TMR (which detects faults and prevents errors from propagating through the system) and configuration memory scrubbing (which prevents faults from accumulating in the TMR system by maintaining the integrity of the device configuration SRAM). The combination of these techniques is commonly referred to as “TMR+scrubbing.” The benefits of implementing TMR in terms of device reliability have been demonstrated [27, 28]. More recently, support for error detection and correction codes to protect the block random-access memory on Xilinx FPGAs has been included in the device architecture [29]. Additionally, configuration memory error detection and correction are implemented in configuration primitives available to system designers. These primitives enable detection of configuration faults and correction by a user design.

The focus of this paper is to present a reliability analysis of the most common FPGA-based computer architectures used in aerospace applications. This includes simplex, TMR, and TMR + scrubbing. We also analyze a TMR + scrubbing + spares, which is a novel contribution of our own research team [30]. We focus our study on the Xilinx Virtex family because they were the first family of SRAM-based FPGAs to support partial configuration and have a higher level of adoption in aerospace missions.

### III. Reliability Analysis

#### A. Reliability

In space applications, we are concerned with the probability that a system will operate as designed for a specified period of time. This period of time is highly dependent on the nature of the mission or the function of the system under study. For example, a system controlling the thrusters of a rocket-propelled probe must function properly over the duration of the mission. If this system fails, it does so at the expense of the entire mission. High-reliability communications links (e.g., combat satellite communications systems) cannot afford to be down during a battle on the ground. High-reliability systems must be sufficiently reliable to perform their specified task for a known period of time. Reliability is defined as the probability of (a system) working for a specified period of time [31]. To estimate system reliability, statistical tools are put to use. In space systems, the effects due to ionizing radiation contribute to device performance interruptions and degradation. As a result, these external environmental factors must also be considered in addition to the normal device reliability statistics.

To calculate the reliability of a system, one starts with a probability density function representing the operational lifetime of a device. The exponential distribution is commonly used to represent the lifetime of complex systems [31]. Taking this distribution as an example, the steps for calculating system reliability are shown. The mathematical equation for the exponential density function is shown in Eq. (1). This equation provides the relative frequency of lifetimes of a given system:

\[
\begin{align*}
  f(t) &= \begin{cases} 
    \lambda e^{-\lambda t}, & t \geq 0 \\
    0, & t < 0 
  \end{cases} 
\end{align*}
\]

(1)

In Eq. (1), the parameter \( \lambda \) represents the failure rate of the system. In the reliability analyses performed here, \( \lambda \) represents the radiation-induced fault rate. This failure rate describes a Poisson distribution, but it is estimated as a constant value using available modeling tools. In fact, \( \lambda \) varies significantly with time and location within an orbit. To determine the likelihood of a system operating for a specific duration, the probability distribution function is used. This estimate is calculated by integrating the density function up to the specified value, as shown in Eq. (2). This distribution provides the probability of failure versus time. For a system with a lifetime represented by random variable \( L \), the value of the distribution function \( F(t) \) is the probability of failure at time \( t \):

\[
F(t) = P(L < t) = \int_{-\infty}^{t} f(t) \, dt
\]

(2)

\[
F(t) = P(L < t) = \int_{-\infty}^{t} \lambda e^{-\lambda t} \, dt
\]

(3)

\[
F(t) = 1 - e^{-\lambda t}
\]

(4)

Previously, the reliability was defined as the probability of a system functioning as specified for a given time interval. Because the distribution function gives the probability of failure, it follows that the reliability is represented by Eq. (5):

\[
R(t) = 1 - F(t)
\]

(5)
For the exponential distribution, Eq. (5) becomes the expression shown in Eq. (6) and, ultimately, Eq. (7):

\[ R(t) = 1 - F(t) = 1 - (1 - e^{-\lambda t}) \]

(6)

\[ R(t) = e^{-\lambda t} \]

(7)

To analyze the entire system reliability consisting of individual subsystems, the reliability of each component is first found individually. The total system reliability is then the product of the individual reliabilities. As an example, consider a TMR architecture. There will be a reliability associated with the triplicated subsystem and another reliability associated with the voter. Equations (8–10) illustrate this:

\[ R_{TMR} = e^{-\lambda_1 t} \]

(8)

\[ R_{voter} = e^{-\lambda_2 t} \]

(9)

\[ R_{sys} = R_{TMR} \cdot R_{voter} \]

(10)

This shows that, as time increases, the reliability decreases exponentially. Intuitively, this makes sense because the likelihood of failure increases as time goes on. This is observed on a daily basis with computer systems. Older devices tend to fail more frequently than newer devices.

B. Markov Chains

This technique for arriving at the expression for system reliability in the prior section is useful for single, simple systems. As the complexity of a system increases and multiple sources of failure are accounted for independently, this approach to calculating the reliability becomes cumbersome. A related but more appropriate technique is to use Markov modeling to calculate the total system reliability. A Markov chain models a system as a state diagram and a set of probabilities describing the likelihood of transitioning among states. McMurtrey et al. at Brigham Young University presented a comprehensive explanation in [32] for a procedure at arriving at an equivalent expression for reliability as Eq. (7) using Markov chains of complex systems. This approach serves as the model for how our analysis evaluates various FPGA-based computer architectures. We consider the worst-case condition in our analysis, being that a radiation strike will cause a fault. This is due to the critical nature of aerospace systems that must operate remotely throughout their mission duration. Designers of such systems are often most concerned about the worst-case analysis because its impact can result in mission failure.

C. Simplex System Model

In a Markov chain, the system is modeled as a sequence of states. The likelihood of transitioning from the current state to an adjacent state is determined by a probability of occurrence. These probabilities are related to the failure rate of the system, and they are conveniently represented in matrix form. As an example, consider a simple system consisting of two states: healthy and failed. S1 denotes the healthy state, where the system is functioning properly; and S2 denotes the failed state, where the system has failed. The state transition diagram for this system is shown in Fig. 2.

A transition matrix \( T \) is defined that contains the probabilities of transitioning between each state. The entry in position \( m, n \), where \( m \) is the row of \( T \) and \( n \) is the column of \( T \), provides the probability of transitioning from state \( m \) to state \( n \). This matrix is defined in Eq. (11):

\[
T\{m, n\} = \begin{bmatrix}
t_{1,1} & \cdots & t_{1,n} \\
\vdots & \ddots & \vdots \\
t_{m,1} & \cdots & t_{m,n}
\end{bmatrix}
\]

(11)

For the system in Fig. 2, the transition matrix will be Eq. (12):

\[
T = \begin{bmatrix}
(1 - \lambda \cdot \Delta t) & (\lambda \cdot \Delta t) \\
0 & 1
\end{bmatrix}
\]

(12)

Equations (13–15) show how to compute the probability of residing in state \( S (S = 1) \) after time step \( k \):

\[
p_s(t = k\Delta t) = [p_1(0) \ldots p_n(0)] \left[ \begin{bmatrix} t_{1,1} & \cdots & t_{1,n} \\
\vdots & \ddots & \vdots \\
t_{m,1} & \cdots & t_{m,n} \end{bmatrix} \right]^k
\]

(13)

\[
p_s(t = k\Delta t) = [p_1(0) \ p_2(0)] \left[ \begin{bmatrix} t_{1,1} & t_{1,2} \\
t_{2,1} & t_{2,2} \end{bmatrix} \right]^k
\]

(14)

Fig. 2  Markov chain for simplex system consisting of two states: S1 = healthy and S2 = failed.
The reliability is then computed using the definition given in Eq. (5):

\[ R(t) = 1 - p_n(t = k\Delta t) \]  \hspace{1cm} (17)

In an absorbing Markov chain, the final state in the vector \( p_n \) is a failure state that requires system reset for recovery. In our analysis, the \( n \)th state in a system with \( n \) states is absorbing. The probability of being in this state \( [p_n(t = k\Delta t)] \) represents the probability of system failure. Thus, the reliability is \( 1 - p_n(t = k\Delta t) \). Figure 5 shows a plot of the reliability for the simple, nonredundant system shown in Fig. 2. The numerical technique presented in Eqs. (11–17) for estimating system reliability is used here. In this example, the fault rate \( \lambda \) was chosen somewhat arbitrarily at the upper range of expected values for a modern, 90 nm device in low-Earth orbit \( (\lambda = 1 \times 10^{-3} \text{ SEU per millisecond}) \). The fault rate describes the frequency with which SEUs are expected to occur in a device. We can use an arbitrary \( \lambda \) because we are comparing the reliability of different architectures. This allows a fair comparison.

The reliability can be used to find the primary metric of interest for system, which is the mean time to failure (MTTF). The MTTF is defined as the point \( t \) at which the reliability \( R(t) = 0.5 \).

### D. TMR System Model

Redundancy implies the use of multiple independent systems to perform an identical task. In aerospace applications, this is often realized as completely separate, sometimes identical hardware components being tasked with the same work. The basis for using redundant hardware is found in basic statistical analysis techniques. The basic structure of a TMR system is shown in Fig. 4, and the Markov chain is shown in Fig. 5.

In a TMR system, it is required that at least two of the redundant modules operate properly. Mathematically, any time two systems must operate concurrently, they are considered to be in series, and their total reliability is the product of their individual reliabilities. The system is considered to be operating properly if all three modules are healthy \( (S_1) \) or if any pair of modules is working \( (S_2) \). The probability of these conditions being met are shown in Eq. (18) [33]:

\[ R(t) = R^1(t) + 3R^2(t) \cdot (1 - R(t)) = 3R^2(t) - 3R^3(t) \]  \hspace{1cm} (18)

Figure 6 shows the reliability of a TMR system compared to a simplex system with the same fault rate of \( (\lambda = 1 \times 10^{-3} \text{ SEU per millisecond}) \). Of note is the fact that the reliability is higher for TMR systems before the MTTF \( (R = 0.5) \), but the reliability is actually lower than a simplex
system after the MTTF. This is attributable to the fact that, after one unit fails in a TMR system, it is required that both remaining systems remain operational. The added area effectively doubles the fault rate over a simplex system, hence the reduced reliability [34]. The MTTF for TMR and simplex systems is the same.

E. TMR with Repair Capability

One of the attractive aspects of using FPGAs in space applications is their ability to be repaired through configuration memory scrubbing. Scrubbing can restore faulted hardware without the need of a system reset. Furthermore, the use of partial reconfiguration of only the affected region can lead to improved reliability due to lower repair times. Repair becomes an essential part of TMR systems because the penalty for adding the extra area to the circuit is minimized. Faults are prevented from accumulating in the system, which essentially allows the system to move backward through the Markov chain to a healthier state. Figure 7 shows a Markov model for a TMR system with repair capability.

In FPGA systems, the time it takes to fully scrub a device is known as the scrub rate, or repair rate. The repair rate, represented here by $\mu$, is the number of repairs per second. The maximum repair rate is dependent on a variety of factors, including device size (in terms of logic resources), the bandwidth of the configuration interface, the bandwidth of the storage device containing the golden configuration data, and the latency of fault detection. These parameters vary on a per-design basis. The repair rate for our work is taken from a Xilinx ML605 Virtex-6 evaluation board. The repair was performed using partial reconfiguration of only the affected circuit. The empirical repair rate was found to be $\mu = 100 \text{ ms/repair}$. The results of a Markov model are demonstrative of the qualitative advantage of using a repair process in a TMR system. Figure 8 shows that the MTTF and reliability in general are much greater for a TMR with a repair system compared to TMR and simplex systems.

F. TMR Plus Spare System Model

To build on the standard TMR architecture, spare resources can added to the system. In this approach, the spares are not active in the circuit and are typically held in reset to save power. This is different from simply adding more redundant circuits, as in an N-modular redundancy (NMR) approach. The rationale for adding spare processors is that more faults are required for the system to reach a failure state, thus increasing the

![Fig. 7 Markov chain for a TMR system with repair capability.](image)

![Fig. 8 Exponential reliability curve for TMR system using a scrubber repair circuit ($\lambda = 1 \times 10^{-3} \text{ SEU ms}^{-1}, \mu = 100 \text{ ms}$).](image)
MTTF. In a high-fault-rate environment, as processors are faulted, the system gradually degrades until there are no longer a sufficient number of tiles for the system to operate properly. In a TMR system, that number is one. When there is only one tile remaining, there is no longer any way for the system to check the output, and thus no way for the system to determine whether or not it has experienced a design-level fault. TMR with repair provides a substantial increase in reliability over a TMR system. The premise underlying the TMR+spares with repair architecture is that it is faster to change from a faulted resource to a healthy resource (i.e., bring a spare out of reset and synchronize it with the two undamaged circuits within the TMR triad) than it is to repair a faulted resource using scrubbing. This leads to a further increase in reliability compared to the TMR with repair system. Figure 9 shows the Markov chain for a TMR + spare system with repair capability.

In this architecture, $S_1$ represents the healthy state wherein no resources are faulted. There is one spare and three healthy processing tiles. There are two failure paths out of $S_1$: a fault in the spare tile transitions the system to $S_3$, and a fault in one of the active tiles transitions the system to state $S_2$. In state $S_2$, the system must undergo a context switch to replace the faulted active tile with a healthy spare. This is performed at a rate of $\alpha$ swaps per millisecond (measured at 20 ms on a representative, redundant core system). If a fault is experienced in either of the two other active tiles, then the system cannot be restored to a functional state. Thus, the system transitions to state $S_5$, which is an absorbing state. A system reset is required for recovery. From state $S_3$, where the spare tile is faulted, the system can transition back to state $S_1$ through the repair process. Recovery occurs at repair rate $\mu$. If an active tile is faulted before the spare tile is repaired, the system transitions to state $S_4$. In $S_4$, the system is technically still functional because the faulted active tile outputs are mitigated by the voter circuit. From $S_4$, the faulted active tile can be repaired to transition back to state $S_3$. Should a fault occur in either of the remaining two active tiles in $S_4$, the system transitions to the failure state $S_5$ and a reset is required. Figure 10 shows the reliability of the TMR + spare system.

**IV. Orbital Fault Rates**

The fault rate is perhaps the most critical parameter in the Markov chain models used to estimate the reliability of the system. It is also a very difficult parameter to estimate because it is dependent on a range of factors. The factors fall into two categories: device factors and environment factors. Environment factors include the severity of the radiation environment and the breakdown of the radiation spectrum. The radiation environment begins with the particle flux. This describes the number of particles passing through a particular area as a function of time. The particle flux represents a spectrum of particles of varying types, energies, and linear energy transfer in silicon. The particles of greatest concern are heavy ions and high-energy protons trapped in the van Allen belts. Particles below a certain energy are unlikely to induce SEUs in a device. The particle flux is highly dependent on orbit. In this work, we limit our analysis to electronics residing in low Earth orbit (LEO), which has the properties listed in Table 1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum altitude</td>
<td>160 km</td>
</tr>
<tr>
<td>Maximum altitude</td>
<td>2000 km</td>
</tr>
<tr>
<td>Minimum orbit period</td>
<td>90 min</td>
</tr>
<tr>
<td>Maximum orbit period</td>
<td>127 min</td>
</tr>
</tbody>
</table>
The SEU cross section is a parameter used to determine the threshold particle energy required for inducing an upset. As the linear energy transfer increases for a given particle, the probability of inducing an upset increases until it reaches a saturation value. This saturation value is often taken as the bit cross section, and is given in squared centimeters per bit. Thus, the number of bits used in a design, often referred to as the area of a design, is the cross-sectional area of importance in determining the number of upsets to expect given a particle flux spectrum. Device factors include the manufacturing process node, the transistor architectures, isolation techniques, doping levels of the silicon substrate, critical charge values, etc. Many of these parameters are not immediately available from manufacturers because they are majorly irrelevant to designers in most applications. However, estimation of the fault rate using modeling tools depends on some of these parameters. Therefore, the accuracy with which the fault rate is estimated for a given orbit location under specified solar conditions is limited by the accuracy with which the device parameters can be estimated. Fault rates in this work were studied across the Virtex FPGA family from Xilinx, Inc. Prior work done in [35] had measured/simulated the fault rates for the Virtex, Virtex-2, and Virtex-4 FPGAs. Our team used the CREME96 tool [36] to predict the SEU fault rates for the Virtex-5 and Virtex-6 FPGAs using published data about the devices. For the Virtex 5/6 simulations in CREME96, the critical charge, sensitive volume, and bit cross-sectional area were used. The cross-sectional area was taken to be $5.73 \times 10^{-4} \text{μm}^2/\text{bit}$ [37] for the Virtex-5 and $2.75 \times 10^{-4} \text{μm}^2/\text{bit}$ for the Virtex-6. The sensitive volume, which is the region of each transistor within which single-event effects originate, was taken as the cube of the process node (65 and 40 nm, respectively). The critical charge was calculated using Eq. (19), which is from [38]:

$$Q_{\text{crit}} = 0.023 \frac{pC}{\mu m^2} \cdot L^2 \quad (19)$$

Table 2 gives the fault rates for the FPGAs evaluated in this work from the Xilinx Virtex family. These values are the estimates for $\lambda$ used in the Markov models when evaluating the MTTF for specific devices. For each device, the fault rates are presented for five solar conditions: solar minimum, solar maximum, worst week, worst day, and peak 5 min. Solar minimum conditions model the base radiation environment without the presence of solar flares. This environment varies with the solar cycle, so it is represented differently during solar maximum than at solar minimum. To account for this, a fault rate for solar maximum is also included. For more harsh conditions, the peak rates over a week, day, and 5 min are included. Historical measurements form the basis for these conditions, the details of which are provided in [36].

V. Mean Time-to-Failure Analysis

The next analysis that was performed was determining the MTTF of each architecture versus the incoming fault rate $\lambda$. The device used was again the Virtex-6 FPGA with $\mu = 100$ ms and $\alpha = 20$ ms. Figure 11 shows this analysis. As expected, the simplex and TMR without repair have the worst MTTF performance. Adding scrubbing significantly increases the MTTF, and adding a spare increases it further. Also shown is how the impact of adding more than one spare to the system impacts the MTTF. This analysis shows that the largest gain over TMR + scrubbing is including one spare but that additional spares do not significantly improve the MTTF.

![Fig. 11 MTTF for different architectures versus fault rate.](image-url)
The impact of additional spares can be observed more clearly by plotting the MTTF versus the number of spares for a variety of fault rates. Figure 12 shows this analysis. Note that adding one spare provides the most gain in the MTTF; however, there is an inflection point between one and two spares, and additional spares provide no more improvement. This is due to the increased area of the spares outweighing the gains of the fast context switch.

Another analysis of interest is how the process node impacts the MTTF. Each subsequent process node is more susceptible to SEEs, as is evident in Table 2. At the same time, the smaller process node allows a particular circuit to be implemented in a smaller area on the die. As the area is reduced, the likelihood that it will be struck by radiation is also reduced. Figure 13 shows the MTTF for the same circuit (in this case, a 64 bit voter) when implemented on each of the Virtex devices from Table 2.

This figure highlights that the MTTF for a particular circuit is nearly flat when implemented for different modern process nodes. It should be noted that, when moving toward smaller process nodes, the increase in available area provided by the smaller transistor size is typically used to increase the sophistication of system. This increases the area of the circuit of interest and results in a decrease in the MTTF. Note that Fig. 13 shows the MTTF for the same circuit implemented in different process nodes.

VI. Conclusions

The analysis presented in this paper revealed some intuitive and nonintuitive aspects of the reliability of field-programmable gate-array (FPGA)-based space computers. First, triple modular redundancy (TMR) by itself did not provide a significant gain in reliability over a simplex system. It was only when a repair mechanism was incorporated that TMR provided notable improvement in reliability. Next, the impact of adding
spare circuitry was analyzed, and it was found that this had another significant gain in reliability as long as repair capability was included. A spare circuit was advantageous as long as the time to bring on the spare was less than the repair time (t < μ). Adding more than one spare resulted in diminishing returns in reliability due to the increased area of the additional circuitry, outweighing any gains by introducing a shorter recovery time through activating a healthy circuit. The impact of orbit was also analyzed across multiple FPGAs within the Xilinx Virtex family. As expected, more modern process nodes were more susceptible to single-event effects (SEEs). Also of note was that, even in low Earth orbit, under worst-case solar conditions, the single-event upset rate could be significant with peak faults approaching 2000 per day in a 40 nm process node. Finally, an analysis of the mean time to failure (MTTF) was performed when implementing the same circuit (a 64 bit voter) in different process nodes. It was discovered that, even though the SEE susceptibility increased when moving to a modern process node, the decrease in area of the circuit due to TMR/spare being the most advantageous architecture.

Acknowledgments

The authors would like to thank NASA for financial support of this work under Cooperative Agreement Notice (CAN) numbers NNX10AN32A, NNX10AN91A, NNX12AM50G, NNX13AR03A, and NNX14AL03A. The authors would also like to thank the NASA Goddard Space Flight Center, Science Data Processing Branch for collaborating on this work.

References


doi:10.1109/TNS.2009.2033381


doi:10.2514/1.I010106


doi:10.1147/rd.62.0200


doi:10.2514/3.26418

G. P. Brat
Associate Editor