The Future of Embedded Systems at ESA: towards Adaptability and Reconfigurability

Luca Fossati, Jorgen Ilstad
{Luca.Fossati,Jorgen.Ilstad}@esa.int
European Space Agency - Keplerlaan 1, Noordwijk, 2200AG NL

Abstract

Embedded devices used in the space industry have to face stringent requirements imposed by their deployment environment, namely high reliability, limited resources, and challenges in heat dissipation. In addition, to make matters worse, the demand for advanced processing capability is growing steadily to meet future exploration and deep space mission requirements. At the European Space Agency, ongoing studies are focusing the use of reconfigurable and adaptable systems to help finding solutions to the aforementioned issues as well as to how reconfigurable systems can aid in mission cost reductions. After giving a general overview of the efforts spent in this direction, this paper presents two concrete examples of on going research activities funded by the European Space Agency.

1 Introduction

Embedded-systems are computer systems designed to perform one or a few dedicated functions, often with real-time computing constraints. They are embedded as part of a complete device often including also mechanical parts [1]. Such systems are more and more used in our everyday life [2] as well as in industrial applications and in the consumer market. An average person gets in touch with a sheer number of embedded systems every day, ranging from cellular phones, pagers, PDAs etc., to domotic devices and consumer electronic goods like digital video cameras, DVDs, game consoles. Most of today's embedded devices are heterogeneous systems composed of dedicated ICs, memories, A/D and D/A converters, sensors and microprocessor cores.

The design of such embedded systems is a difficult problem, requiring designers with skills and experience to identify optimal solutions. Moreover, embedded systems are special-purpose devices, dedicated to a restricted class of applications and they face stringent constraints including high reliability, low power, hard real-time, and low cost. Besides presenting a complex functionality, these systems have to comply with ever more severe non-functional constraints in various dimensions (power consumption, silicon area, cost, etc). To make the situation worse, architectural complexity (number of primitives) and heterogeneity (variety of primitives) increase with time, putting more pressure on the design process.

The landscape outlined so far is even more complicated when considering special application areas of embedded systems, such as the *space domain*. On the one hand space embedded systems have to withstand the harsh conditions of the space environment (radiation effects, limited resources, etc., see Section 1.1 for a more in depth description), on the other hand they have to provide high performance in terms of reliability, life-time, processing power, and adaptability to unforeseen situations and events (see Section 1.2).

The rest of this Section introduces the main challenges in space missions that motivate and drive research towards adaptivity and reconfigurability and Section 2 show how they can be used to solve such challenges. Finally, Section 3 presents some of the technologies that are being studied at the European Space Agency to enable adaptivity and reconfigurability.

1.1 Space Environment and Electronic Systems: Challenges

With respect the environment found on Earth, space is considerably different: there is little (or no) atmosphere, relatively high density of energetic particles, wide gradients of temperature, etc. [3]. Such differences necessarily determine some of the features that electronic components must have in order to survive in that environment. The main issues that must be taken into account are the *radiation environment*, the *limited availability of energy*, difficulties in *dissipating heat*, and mitigation towards *device failure*. If no corrective actions are put in place the reliability of the component and, as a consequence, of some instruments or even of the whole spacecraft could be severely affected.

Radiation Environment There is no space system in which radiation effects can be neglected. Reliability and longevity of embedded systems in the inhospitable space environment require careful consideration of the types of malfunctions that radiation can cause. Some radiation effects can be mission limiting when they lead to subsystem or system failure, or catastrophic system anomalies. Examples are damage of electronic components due to total ionising dose (TID), or high energy particles changing logic state of transistors i.e. so called single event upset (SEU). Others effects can be a source of interference, degrading the efficiency of the mission: examples are radiation ”background” in sensors or corruption of electronic memories [4].

Special measures have to be considered to harden the electronic devices against the space radiation environment, somewhat limiting the maximum obtainable performance and/or increasing the overall cost and power consumption. In most of the situations, COTS components cannot be used because of their sensitivity to radiation effects. Note that not only processing hardware can be affected, but also sensors' and instruments' behavior can be degraded. As Section 2 will explain in detail, including adaptability and reconfigurability properties in the hardware system can increase the overall reliability and minimize the effects of faults on the mission.

Power Consumption Electronic circuits consume power, which is, partly, translated into heat being created in the circuit; the fact can be easily experienced by touching a computer running a computationally intensive workload: the temperature on the CPU can easily reach 90 degrees Celsius [5].

In space it is not possible to have devices which consume such a great amount of power for two main reasons: (a) limited availability of power sources, and (b) limited heat dissipation capabilities. While the former is obvious, as most of the systems rely only on solar panels and batteries, the latter is mainly due to the absence of an atmosphere. That would ease the heat dissipation by natural air convection, but also due to the fact that active cooling introduces more system complexity and additional mass.

Lifetime Most of space missions have a duration of many years, typically 5 to 15 years depending on mission type, thus requiring the electronic devices within the spacecraft to have extreme reliability with only marginal performance degradation; this comes in addition to the harsh environmental conditions outlined before. This issue is complicated by the extreme difficulties in replacing/ repairing such systems once a component has failed due to wear-out or other permanent problems.

In general, single point of failure free systems are adopted, resulting in, at least, the replication of many of the on-board system's devices and functionalities, consuming precious resources and adding to mass and power budgets. Adaptable architectures, as presented in detail in Section 2, are of a great help in this respect enabling to dynamically reconfigure the system according to its health level, the performance needs, the mission status, etc.

1.2 Performance Requirement of Space Systems

Besides the hazards caused by the peculiarity of the space environment, other challenges exists as a result of the stringent performance requirements characterizing some of the current, and most of the future, missions. As of today, intensive computational power is required mainly to process payload data, as, for example, happens in the upcoming GAIA mission [6]. Future missions will feature growing demands in terms of autonomy [7, 8], which will be possibly satisfied through the use of powerful and adaptive execution units.

1.2.1 Autonomy in Future Avionic Systems

As briefly hinted above, future space undertaking will require more intelligence and on-board processing capabilities than what is available nowadays. Large, complex, and technologically challenging missions (like planetary exploration ones), involving robotics devices, rovers, and longer term missions (like Darwin [9], Mars Sample Return [10]) or human exploration ones, are driven by technology maturity, critical performance, and high reliability/availability requirements during certain phases of the mission. Autonomy is a key technology for satisfying such requirements, as it will allow to make up for i) the severe communication constraints (communication delays, limited visibility windows, limited bandwidth), ii) the high costs and demanding availability requirements for continuous ground operations, iii) the tight interaction with the environment (e.g. for planetary surface missions) and harsh environmental conditions (e.g. for Solar Orbiter), and iv) the high degree of operational uncertainty which characterize these missions.

1.2.2 Area, Mass, and Power Constraints

While, on the one hand, future space missions will require more and more advanced, possibly real-time, processing capabilities (to enable, among others, autonomy features), on the other hand they do and will impose severe constraints on the amount of area, mass, and power that can be used by the control and payload processing elements. For example, a major contributor to a space mission costs is the launcher costs driven by mass which, then, is again driven by the high reliability requirements leading to redundant systems.

At a first glance, when considering traditional computing paradigms, these two requirements of high performance and reduced mass/power consumption seem incompatible with each other. Fortunately, adaptive and reconfigurable systems enable overcoming some of these issues based on the observation that, often, different mission phases have different processing needs and that such needs are exclusive with each other: resources might, as such, be shared, saving energy (by temporarily deactivating the un-necessary functions), silicon area and mass (through time-multiplexing of the hardware resources). Adaptable onboard systems can also alleviate the need for 1 to 1 redundancy while, at the same time, offering high reliability and performance.

2 Reconfigurability and Adaptability: an Answer to the Needs of Future Spacecraft Systems

The previous Section introduced the main problems and challenges that have to be coped with by most of the electronic devices embedded in satellites, launch vehicles, planetary landers and, in general, spacecrafts. As we have seen, such challenges are going to become more difficult in the future as mission requirements increase. This Section starts from that to show how the introduction of adaptability and/or reconfigurability, besides bringing a drastic change with respect to traditional computing paradigms, helps in facing the mentioned issues.

2.1 Reconfigurable and Adaptive Hardware

Briefly stated, *adaptability is the ability of a system to accommodate to changes in its environment*. Adaptable (often also called self-aware) computer systems are capable of adapting their behaviour and resources thousands of times a second and automatically find the best way to accomplish a given goal despite changing environmental conditions and demands.

The first, more immediate and, probably, more intuitive solution to realize an adaptive system consists of the use of software running on top of a generic processing unit; software is flexible, allows programmability, may be designed with modularity in mind and can be re-used. However, software basically results in sequential execution of instructions, making it difficult to meet the already mentioned high performance requirements [11].

Increasing the processing frequency only partly solves the problem as the maximum increase is limited and the power consumption, which is described in Section 1.2 to be a strict design constraint, scales-up with higher frequencies. However, with the advent of dynamically reconfigurable hardware similar flexibility as in software can be applied to hardware devices. Field-Programmable Gate Arrays (FPGAs) is the typical device family used to implement adaptive hardware systems. Most of such devices, indeed, feature run-time and dynamic reconfiguration capabilities meaning that they can rapidly alter (on the fly) the functionalities of its components and of the interconnections between them.

2.1.1 Field Programmable Gate Arrays

Recent Field-Programmable Gate Arrays (FPGAs) [12] are parallel, distributed, and extremely advanced computing fabrics with a regular architecture of computational elements and memories. Each computational element (or logic block) consists of configurable combinatorial logic together with a few flip-flops. The configuration of the function of each logic block and its connections to other blocks are given by the so called "configuration bitstream" loaded from out-side the device and stored into the FPGA's memory (normally either SRAM or FLASH, depending on the device).

An important feature of most FPGAs consists of the capability to change their functionality over time, through modification of the configuration bitstream [13]. Some of these devices also allow to perform that during runtime, which means that they are able to change part of their functionality while they are operating. Hence the availability of reprogrammable FPGAs provides new perspectives for reconfigurable-evolvable on board systems, as compared to general-purpose processors. FPGAs often offer a better performance and greater flexibility than software, while they also outrival ASICs in terms of adaptability.

2.2 Dynamically Tuning Processing Resources

The hardware structures just introduced are the basis for building effective adaptive systems; in the context of spacecrafts, such systems help solving problems and issues on various fronts: (a) reliability and fault tolerance, (b) resources (area, mass, power) savings, and (c) autonomy.

These factors are in general not clearly separated, enabling design of systems with the ability to adapt and dynamically trade-off between performance, desired reliability level, and energy consumption depending on the environmental conditions, application requirements, etc. This means, for example, that in presence of faults, the system could gracefully degrade providing lower performance while still producing correct results. As we will see in Section 3, some systems might also be able to modify their hardware structure to correct the fault without any performance drawback and/or waste of resources. Unfortunately, the identified solutions cannot always be fully transparent with respect to the rest of the system. In particular, besides all the aspects cited so far and all the advantages that reconfigurable and adaptive systems might bring to spacecraft's electronics, safety, availability, and adherence to real-time constraints of the processing system must always be obeyed.

As a summary, *one of the the main aims of recent research at the European Space Agency consists of providing solutions to the ever increasing need for processing power, resources and autonomy, while guaranteeing high reliability and availability using reconfigurable (data processing) hardware based on intrinsically radiation sensitive components, such as reprogrammable FPGAs.*

2.2.1 Reliability

The reliability and robustness of on-board embedded systems is key for the success of space missions. Reliability is generally measured in terms of probability that the integrity of the satellite and of the mission is preserved even in the occurrence of failures. This must be ensured over the whole lifetime of the mission, which means even for long interplanetary and telecommunication missions with a lifetime sometimes exceeding 15 years. Unfortunately, it has become almost impossible to design totally fault-free chips [14] that are fully resistant to the defects hidden in the silicon or induced by aging or the harsh environmental conditions characterizing space missions, as described in Section 1.1. Because of the increasingly demanding mission requirements and, consequently, increased complexity of the electronic devices, coupled with the advance in semiconductor technology, systems-on-chip are now making use of smaller technology feature sizes. In this context, a system that adapts and reconfigures itself is able, if properly designed, to tolerate and recover from transient and permanent faults through self-healing.

The most commonly used devices in the implementation of adaptive and reconfigurable systems are SRAM-based FPGAs that store the configuration bitstream in on-chip SRAM, normally extremely prone to radiation induced faults and single event effects in particular. Appropriate precautions and techniques have to be developed and employed in order to mitigate single event effects [15]. Despite this strong drawback of SRAM-based FPGAs, i.e. being more sensitive to faults with respect to traditional solutions (e.g., ASICs or anti-fuse based FPGAs), they have higher capabilities of recovering from them, thanks to their reconfigurability and adaptability properties; when a fault is detected, the affected functionality is re-created on the FPGA device, thus restoring the device overall processing capabilities.

Adaptable hardware helps not only recovering from faults happening to the hardware component itself, but also to the ones affecting the rest of the system; for example, instruments' behaviour can change over time possibly as a result of the total radiation absorbed. The use of reconfigurable hardware also enables elegant and efficient graceful degradation of the overall system in case of defective/faulty sub-components and when the damage is such that it is not possible to restore the system to the original fully operational conditions. Dynamic partial reconfiguration enables hardware updates that reflect this change and, thus, allows more specific hardware parameters to be uploaded to compensate the system degradation.

Having multiple options for correcting a defect is expected to increase defect tolerance and to reduce the overall energy cost. Section 3 will analyze more in depth the concepts expressed in this paragraph, providing examples in the context of on-going projects.

2.2.2 Area, Mass, and Power Consumption

As well as increasing the overall reliability, adaptive and reconfigurable hardware can help in saving area, mass, energy, and cost. Indeed, a reconfigurable generic module can ensure maximum reuse of development across different systems by maintaining a generic common infrastructure and by adding ad-hoc reconfigurable modules for each specific design. Moreover, if the devices can be reconfigured dynamically, this adaptation can be done in an efficient manner during run-time, where parts of the device can be operative, while others, which are not been used, are changed. This allows for implementing time-sharing of the reconfigurable resources between different applications, thus increasing the area efficiency.

This concept can be extended even further by exploiting the fact that, often, planetary science and deep space missions have different mission phases; not all the phases will always require the same processing functionalities, each phase often having its dedicated set of functions. By adapting and reconfiguring the hardware with respect to the requirements of each phase, it would be possible to use the same exact hardware resources for each mission phase, with a consistent reduction in the area and mass used by the electronic sub-system.

2.3 Enabling Autonomy through Dynamic Reconfiguration

Autonomy is a key technology for the implementation of new and more advanced functionality in future avionics systems. Coming missions, like planetary exploration (in unknown environments) and formation flying, will have to rely on robust and fault tolerant implementation of autonomous functions. More in detail, the following onboard functions are needed to implement autonomy:

1. Autonomous on-board planning and scheduling of the mission operation will allow for timely response to unexpected conditions and evolution of the system operational state, providing the necessary flexibility in achieving the mission objectives;
2. a system- and environment-aware Fault Detection Isolation and Recovery (FDIR) with reasoning capabilities would provide for timely anomaly resolution in the operational context, greatly reducing the need for the Ground intervention and resulting in self-sustained systems;
3. the knowledge-based and prognostic approaches can allow for pro-active mission re-planning and operations adaptation, as well as preventive FDIR measures.

Reconfigurable hardware allows for changing or adapting payload processing during the flight mission. The systems thus can be freely adapted to several possible scenarios, even if it was not foreseen at system design time. In these cases, it is not uncommon that the data parameters to be measured differ from the final target of the instruments. A reconfiguration of the hardware which is more in-tune with the final target data sets can also widen the scope of a particular mission.

3 Examples

Technology assessment and proof of concept prototypes are a necessary first step before any new concept and idea can be deployed as a fully proved on-board system; this is especially true for adaptive and reconfigurable systems, whose structure and functionality is not fixed forever at design time, but which instead can and are meant to , *dynamically* change at run-time, during system operation. The Data Systems Division (TEC-ED) at ESA/ESTEC is actively investigating the use of adaptive systems into spacecrafts, mainly through the use of dynamically reconfigurable FPGAs. In particular, such efforts are centered around two activities, both focused on the use of multi-FPGAs systems, combining high performance processing, high reliability, redundancy, and run-time adaptability. The rest of this Section will briefly present these two activities.

3.1 Building Reliable Systems out of Not-Hardened Components

The aim of this work, developed in cooperation with Politecnico di Milano, consists of the definition of a methodology for the design and implementation of reliable embedded systems on multi-FPGA platforms [16, 17]. The target system is distributed onto the platform applying fault detection/mitigation techniques, so that each system portion is able to detect and signal when it is affected by an error. Fault mitigation is, then, achieved by exploiting the devices' dynamic reconfiguration capabilities, recreating the affected system portion. The overall goal consists of defining a reliable system, exploring different trade-offs between the achieved reliability properties (e.g., detection vs. tolerance, reconfiguration time) and the introduced overheads (e.g., area increase, latency).

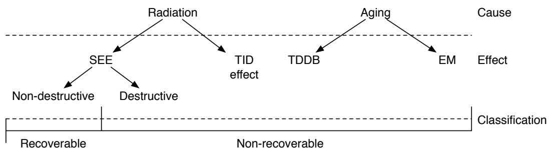

Figure 1: Overview of the potential fault types classification

Two classes of faults are taken into consideration, as shown in Figure 1: *recoverable* and *non-recoverable*; recoverable faults are the ones caused by radiations without a destructive effect, whereas non-recoverable faults are those caused by radiations with destructive or permanent degradation effects, radiations accumulation and device aging.

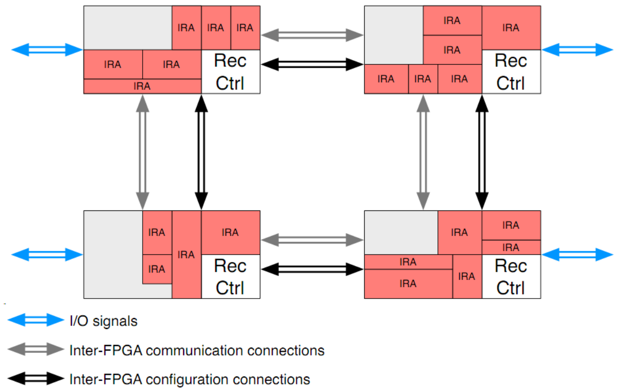

Figure 2: Overview of the proposed multi-FPGA reliable system. `IRAs` are areas independently reconfigurable from each other and containing the application logic; `Rec Ctrl` represents the block in charge of managing the reliability layer

Figure 2 shows a high-level overview of the proposed system; tasks (red blocks in the Figure) are initially distributed over the FPGAs and, during the life of the system, should a transient fault be detected, a recovery phase would take place, or should a permanent fault be detected, tasks shall be relocated to an available fault-free area of the FPGA, in order to avoid the faulty portion of the device. The *Reconfiguration Controller* (white block in Figure 2) takes care of monitoring the detected error signals and of performing the appropriate recovery actions. By exploiting the unique properties of partial dynamic reconfigurability of recent SRAM-based FPGA devices, the devised system is not only able to mitigate non-permannet fault effects in these devices, but it also provides an overall reliability potentially higher than other traditional solutions (e.g. ASICs), as it is an intrinsically redundant system. In addition, the use of reconfigurable hardware brings all the advantages widely discussed in Section 2, such as power and mass savings, along to being an enabling technology for building autonomous and adaptable

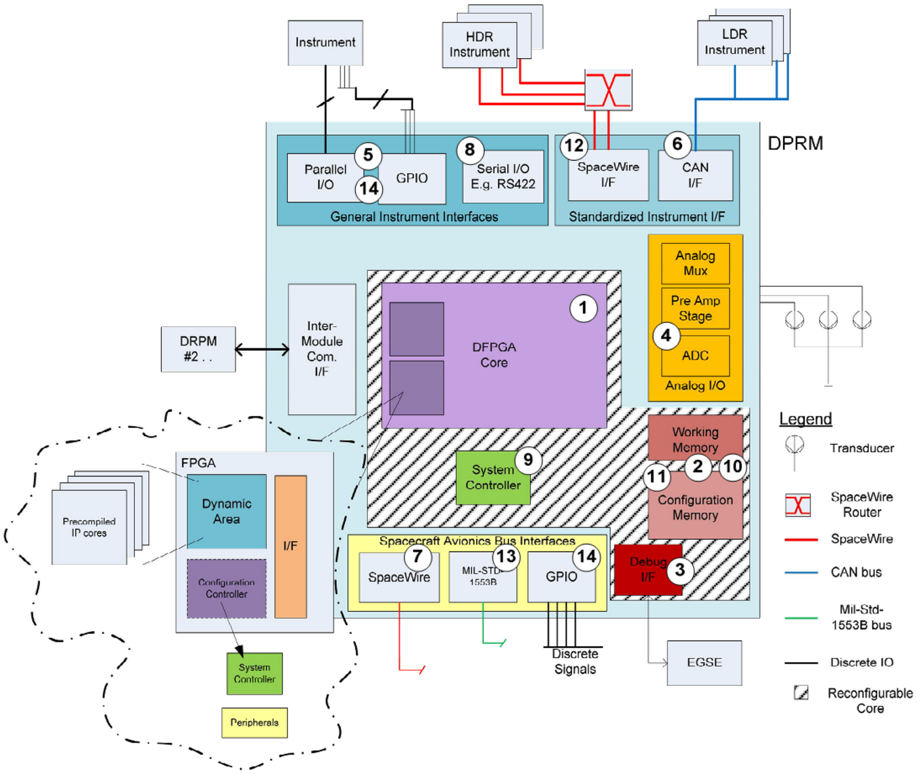

Figure 3: High level diagram of the Dynamically Reconfigurable Processing Module architecture

systems.

3.2 Paving the Way to Autonomy Through Multi-FPGA Adaptive Systems

The activity described in Section 3.1 is mainly centered on how to exploit the reconfigurable capabilities of a system to enhance its reliability, fault tolerance, and lifetime. While this is of fundamental importance, it is also necessary to understand how to deploy such system in the context of a space mission and when they are preferable to more traditional solutions like software running on top of general purpose microprocessors or custom ASICs.

To this aim ESA is studying and prototyping an architecture like the one depicted in Figure 3. It is composed of multiple modules (DFPGAs), and each one may be used as a stand-alone unit or multiple modules may be combined to form a system with more processing capacity and/or redundancy. In a flight design, each module would have a high degree of internal integration (system controller, reconfigurable FPGAs, memory etc). The modules are reconfigurable in the sense that their functionality can be changed, either completely or partially, by changing the software and/or firmware without changing the physical hardware of the unit. This can even be done dynamically, in other words while the module is operational (powered on and actively processing data inputs). Such architecture may be dedicated to a single application, for example a science instrument. Alternatively, it may support a mission with multiple applications, by providing a centralised processing resource for a multi-instrument payload. The centralization of all the payload processing brings advantages in terms of mass and energy savings, by enabling resource sharing among the various instruments and payloads. Sharing of data among the instruments themselves is also simplified, enabling a homogeneous and coherent management of the overall mission. To enable further flexibility and adaptability of the system, partial reconfiguration is considered, enabling different portions of the same device to be (re)configured independently and at different times, thus speeding up the reconfiguration process and eliminating the need to halt the computation of all the functionalities running on the device.

With the flexibility of processing tasks, this system lends itself well to missions that have extremely high data rates, in particular those employing state of the art imaging sensors as well as instruments such as Synthetic Aperture Radar (SAR) or hyperspectral image processing. In synthesis, the goal of this activity consists of developing a system that allows for the insertion of new hardware processing modules into a given architecture at run-time. This is enabled by the implementation of partially reconfigurable cores, which are embedded into a system hosting a central reconfiguration controller and a system controller providing suitable peripherals (interfaces, etc.) for space applications.

4 Conclusion

Space missions are becoming increasingly complex, and resource demanding. In order to execute and manage such missions the *brain* of the satellite has to be adequately dimensioned. In addition, it has to cope with the harsh space environment, limited possibilities for heat dissipation and reduced mass. Reconfigurable and adaptable embedded systems are currently being investigated at the European Space Agency as a mean to satisfy the requirements in terms of high processing power of future missions while maintaining a high degree of reliability and reduced mass and power budget. Two of such systems have been briefly presented, showing how the use of reconfigurable multi-FPGAs systems enabling to fulfill demanding requirements for next generation space mission.

References

[1] "Embedded systems glossary | netrino." [Online]. Available: http://www.netrino.com/Embedded-Systems/Glossary

[2] D. Sciuto, "Guest editor's introduction: design tools for embedded systems," *IEEE Design & Test of Computers*, vol. 17, no. 2, pp. 11–13, 2000.

[3] S. N. Lehr and V. J. Tronolone, "The Space Environment and Its Effects on Materials and Component Parts," *IRE Transactions on Reliability and Quality Control*, vol. RQC-10, no. 2, pp. 24 –37, 1961.

[4] *ECSS-E-ST-10-12C - Methods for the calculation of radiation received and its effects, and a policy for design margins*. European Cooperation for Space Standardization, 2008.

[5] D. J. Frank, "Power-constrained CMOS scaling limits," *IBM Journal of Research and Development*, vol. 46, no. 2.3, pp. 235 –244, 2002.

[6] J. Portell, X. Luri, and E. Garcia-Berro, "High-performance payload data handling system for GAIA," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, no. 2, pp. 421 – 435, 2006.

[7] W. Truszkowski, M. Hinchey, J. Rash, and C. Rouff, "Autonomous and autonomic systems: a paradigm for future space exploration missions," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 36, no. 3, pp. 279 –291, May 2006.

[8] N. Steiner and P. Athanas, "Hardware autonomy and space systems," in *IEEE Aerospace conference*, 2009, pp. 1 –13.

[9] A. Leger et al., "The DARWIN mission concept. Proposal to the ESA Horizon 2000 Plus Planning Process," 1993.

[10] B. Sherwood, D. Smith, R. Greeley, W. Whittaker, G. Woodcock, G. Barton, D. Pearson, and W. Siegfried, "Mars sample return: architecture and mission design," in *Aerospace Conference Proceedings, 2002. IEEE*, 2002.

[11] P. Yiannacouras, J. Steffan, and J. Rose, "Data parallel FPGA workloads: Software versus hardware," in *International Conference on Field Programmable Logic and Applications (FPL)*, 31 2009.

[12] J. Rose, A. El Gamal, and A. Sangiovanni-Vincentelli, "Architecture of field-programmable gate arrays," *Proceedings of the IEEE*, vol. 81, no. 7, pp. 1013 –1029, Jul. 1993.

[13] T. Todman, G. Constantinides, S. Wilton, O. Mencer, W. Luk, and P. Cheung, "Reconfigurable computing: architectures and design methods," *Computers and Digital Techniques, IEE Proceedings -*, vol. 152, no. 2, pp. 193 – 207, Mar. 2005.

[14] R. Baumann, "Radiation-induced soft errors in advanced semiconductor technologies," *IEEE Transactions on Device and Materials Reliability*, vol. 5, no. 3, pp. 305 – 316, 2005.

[15] L. Sterpone, M. Aguirre, J. Tombs, and H. Guzmán-Miranda, "On the design of tunable fault tolerant circuits on sram-based fpgas for safety critical applications," in *Proceedings of the conference on Design, automation and test in Europe (DATE)*, 2008, pp. 336–341.

[16] C. Bolchini, L. Fossati, D. Codinachs, A. Miele, and C. Sandionigi, "A Reliable Reconfiguration Controller for Fault-Tolerant Embedded Systems on Multi-FPGA Platforms," in *Defect and Fault Tolerance in VLSI Systems (DFT)*, 2010, pp. 191 –199.

[17] C. Bolchini and C. Sandionigi, "Fault Classification for SRAM-Based FPGAs in the Space Environment for Fault Mitigation," *Embedded Systems Letters, IEEE*, vol. 2, no. 4, pp. 107 –110, 2010.