

# Design Trade-off Between Performance and Fault-Tolerance of Space Onboard Computers

M S Gorbunov<sup>1,2</sup> and A A Antonov<sup>1</sup>

<sup>1</sup> Scientific Research Institute of System Analysis of Russian Academy of Sciences (SRISA), Nakhimovsky prospect 36-1, Moscow, 117218, Russia

<sup>2</sup> National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Kashirskoe highway 31, Moscow, 115409, Russia

E-mail: gorbunov@niisi.msk.ru, antonov@niisi.msk.ru

**Abstract.** It is well known that there is a trade-off between performance and power consumption in onboard computers. The fault-tolerance is another important factor affecting performance, chip area and power consumption. Involving special SRAM cells and error-correcting codes is often too expensive with relation to the performance needed. We discuss the possibility of finding the optimal solutions for modern onboard computer for scientific apparatus focusing on multi-level cache memory design.

## 1. Introduction

The complexity of space onboard computers grows constantly. Technology node scaling leads to the necessity of additional protection against so called Single Event Effects (SEE), which are caused by a single ion's strike to the sensitive volume of integrated circuit (IC) [1]. Table 1 shows the performance characteristics of some onboard microprocessors [2]-[5]. One can see that despite their using in recent 10-15 years, their characteristics are close to each other (except RAD5545) and correspond to the parameters of 1990 generation commercial microprocessors. They use external or single level of cache memory and their clock frequency is less than 100-200 MHz. The characteristics of space microprocessors are often optimized taking into account the fault tolerance and power consumption, not the performance. RAD5545 multi-core microprocessor from BAE Systems demonstrates the qualitative leap compared to previous generations: it is manufactured using special radhard 45 nm Silicon-on-Insulator (SOI) process, it has 3 levels of on-chip cache memory and uses multi-level approach of fault protection. It leads to rather good performance, which can be compared to commercial devices: up to 5.6 giga operations per second (GOPS)/3.7 GFLOPS.

All microprocessors mentioned above are extremely expensive and not always available on market due to the International Traffic in Arms Regulations (ITAR) restrictions. Special radhard processes are also unavailable for international design centres. Nevertheless, the well-known design concept called radiation hardening by design (RHBD) could help to overcome these restriction using architecture, schematics and layout level techniques and standard commercial CMOS technology. It is important to note that the price for such microprocessors could be reasonable only if their performance along with power consumption and fault tolerance would be attractive for users. So the performance must be involved in the optimization procedures during the design process.



An optimization procedure is based on a mathematical model of the microprocessor or its main building blocks. It is mentioned in [6] that one needs about 130 parameters to describe the two-level cache system.

**Table 1.** Performance characteristics of onboard microprocessors.

	RAD6000	RAD750	LEON3FT-RTAX	Mongoose V	RAD5545
Manufacturer	BAE Systems	BAE Systems	Aeroflex Gaisler /MicroSemi	Synova	BAE Systems
Architecture	32-bit IBM System/6000	32-bit IBM PowerPC 750	32-bit SPARC v8	32-bit MIPS R-3000	32/64 bit Power
Clock frequency	33	200	25 MHz	15 MHz	N/A
Performance	35 MIPS	>400 MIPS	20 MIPS	N/A	3 DMIPS / MHz
Cache memory	8 KB	32 KB+32 KB	8 KB+4 KB	Up to 256 KB + 256 KB	32 KB+32 KB L1 512 KB L2 1024 KB L3
RH methods	RHBP	RHBP	RHBP FPGA	RHBD, RHBP SOI	RHBD RH45™ circuit library, RHBP 45nm SOI
Apparatus	Spirit Opportunity Coriolis Deep Space-1 Gravity Probe B HESSI MARS 98 SIRTF SMEX-Lite Swift Triana	Curiosity Deep Impact Mars Reconnaissance Orbiter	SCOC3	New Horizons EO-1/WARP IceSat Glas MAP	N/A

In this paper we propose the simplified analytical model for two-level cache memory system that is exposed to SEE at the given orbit. The model takes into account the experimental data for SRAM Single Event Upset (SEU) tolerance for a given technology node, the time penalty coming from the used Error Correcting Codes (ECC) and can be further extended for more common cases.

## 2. Cache memory analytical model

To a first approximation, the access time to Li (i=1, 2...) cache  $t_{Li}$  can be expressed as follows:

$$t_{Li} = t_{logic} + t_{cell} + t_{ED} + t_{EC} , \quad (1)$$

where  $t_{logic}$ ,  $t_{cell}$ ,  $t_{ED}$ ,  $t_{EC}$  are delays at controlling logic, at cells switching, at error detection and correction correspondingly. More precisely, we can assume that  $t_{logic}$  is the time needed for address decode and access,  $t_{cell}$  is the delay from WL signal to the appearance of signal at the outputs of sense amplifier.

Note the following dependencies:

$$t_{logic} = t_{logic}(C, B, A, C_{type}), \quad (2)$$

$$t_{ED} = t_{ED}(C, CM), \quad (3)$$

$$t_{EC} = t_{EC}(CM, P_{error}(C, SM), SER), \quad (4)$$

where  $C$  is cache size (in words),  $B$  is block (fetch) size,  $A$  is associativity,  $C_{type}$  is cell type (6T, DICE [7], etc.),  $CM$  is error correction method,  $SM$  is scrubbing method (passive or active),  $P_{error}$  is the probability that an error occurs in fetched data,  $SER$  is Single Event Rate value depending on orbit and space weather.

The following main cases can occur:

1. Undetected error leads to Single Event Functional Interruption (SEFI). We do not consider such cases in this paper.
2. Detected and corrected error leads to the additional delay in total access time, but does not lead to the cache miss.
3. Detected but not corrected error leads to the additional delay in total access time and to the cache miss.

Let us consider the case of two-level cache (without explicit separation to instruction and data cache) with the external memory. The performance-tolerance factor, which is the total unhardened to hardened cache access time ratio  $K$ , is following:

$$K = \frac{N_{read} \cdot (t_{L1uh} + m_{L1uh} \cdot (t_{L2uh} + m_{L2uh} \cdot t_{MM})) + T_{store}}{N_{read} \cdot (t_{L1h} + m_{L1h} \cdot (t_{L2h} + m_{L2h} \cdot t_{MM})) + T_{store}}, \quad (5)$$

where  $N_{read}$  is the number of reading from cache,  $T_{store}$  is the number of stores,  $t_{MM}$  is the time needed to obtain data from the external memory. These parameters depend on the running program and cache and external memory features. SEE-dependent parameters are following:  $m_{L1uh}$  ( $m_{L2uh}$ ) and  $m_{L1h}$  ( $m_{L2h}$ ) are miss rates,  $t_{L1uh}$  ( $t_{L2uh}$ ) and  $t_{L1h}$  ( $t_{L2h}$ ) are L1 (L2) cache access time for unhardened and hardened cache correspondingly. When the number of cache access is higher than the number of stores, equation (5) can be rewritten in the following simplified form:

$$K' = \frac{t_{L1uh} + m_{L1uh} \cdot (t_{L2uh} + m_{L2uh} \cdot t_{MM})}{t_{L1h} + m_{L1h} \cdot (t_{L2h} + m_{L2h} \cdot t_{MM})}. \quad (6)$$

Miss-rate depends on cache size as follows:

$$m = m_0 \cdot C^{-\alpha}, \quad (7)$$

where  $m_0$  is microarchitecture dependent parameter,  $\alpha=0.3...0.7$  [8]-[9].

SEEs add an additional component to miss rates:

$$m_{SER} = m_0 \cdot C^{-\alpha} + SER \cdot P_{error} \cdot t_{CPU}, \quad (8)$$

where  $t_{CPU}$  is CPU cycle time.

Equation (8) assumes that all errors occurred are not corrected, but detected, which is suitable for special cache SRAM cells placement (e.g., physical separation of logically neighbouring cells from the same word) or using error detection techniques. Otherwise, a system error occurs. Equation (8) shows that SEE can be neglected in the following condition:

$$SER \ll m_0 \cdot C^{-\alpha} / (P_{error} \cdot t_{CPU}). \quad (9)$$

Equation 9 shows that large caches (L2, L3 and so on) and poor spatial locality lead to stronger vulnerability. The faster microprocessors (lower  $t_{CPU}$ ) have lower portion of cache misses due to SEE. The typical case for direct-mapped cache from [10] gives the following parameters values:  $m_0=3.5$  words<sup>a</sup> (assuming that 1 word equals 1 byte),  $\alpha=0.432$ . For simplicity, we consider hereafter the case when all errors are occur in fetched data and detected. For 8 KB and 512 KB cache and  $t_{CPU}=1$  ns, the critical SER values thus equal  $0.7 \cdot 10^8 \text{ s}^{-1}$  and  $0.1 \cdot 10^8 \text{ s}^{-1}$ . Note that critical SER values should be considered for different cache levels separately. Equation (9) assumes that the active scrubbing procedure is applied, so there is no error accumulation, which is important for large caches.

Single Event Rate can be calculated using CRÈME [11], SPENVIS [12] or other instruments. In our model, we use the figure of merit from [13] for geosynchronous orbit:

$$SER = 200 \cdot \frac{\sigma_L}{L_{0.25}^2} \left[ \frac{\text{upsets}}{\text{bit-day}} \frac{(\text{MeV} / \text{mg} / \text{cm}^{-2})^2}{\text{cm}^2} \right], \quad (10)$$

where  $\sigma_L$  is SEU saturation cross section,  $L_{0.25}$  is the linear energy transfer (LET) corresponding to the cross section that is 0.25 of  $\sigma_L$ .

The estimation for the rather typical values (corresponding to 65 nm technology node)  $\sigma_L \sim 10^{-8} \text{ cm}^2/\text{bit}$ ,  $L_{0.25} = 1 \text{ MeV} \cdot \text{cm}^2/\text{mg}$  [14] gives  $1.51 \cdot 10^{-6} \text{ s}^{-1}$  and  $9.7 \cdot 10^{-5} \text{ s}^{-1}$  for 8 KB and 512 KB cache correspondingly. The comparison with the results obtained from equation (9) shows that SEE can be neglected in the performance account if all errors occurred are detected, i.e. all needed error detecting and correcting solutions are applied.

Let us now estimate the performance penalties due to the using of ECC and special types of the cells. Table 2 shows the parameters for SRAM blocks designed using bulk 65 nm and SOI 250 nm CMOS process design kits. Special layout with 2 (N+ and P+) guard rings were used in 65 nm versions.

**Table 2.** Timing characteristics of SRAM cells.

	6T cell		DICE cell	
Technology node, nm	250	65	250	65
Area per 1 bit, $\mu\text{m}^2$	17.28	2.34	46.08	4.68
$t_{logic}$ , ns	2.83	1.50	2.87	1.55
$t_{cell}$ , ns	1.55	0.50	2.00	0.49

One can see from Table 2 that DICE cell is usually twice as large as 6T cell. Also, DICE cell is not necessarily slower than its 6T counterpart, but its  $t_{logic}$  value is 1-3% higher. We consider the additional delay by 10% for error detection and correction as the worst case in our simulations.

Using of DICE cells usually leads to additional power consumption, but the consideration of this factor is beyond the scope of this work.

### 3. Simulation results

We assume that the total chip area must be constant (e.g., for using the same package), so the sizes of DICE-based caches are two times lower than of 6T-based. We simulate four cases:

- 8 KB 6T L1 and 512 KB L2 caches without ECC;
- 8 KB 6T L1 and 512 KB L2 caches with ECC;
- 8 KB 6T L1 and 256 KB DICE L2 caches with ECC;
- 4 KB DICE L1 and 256 KB L2 caches with ECC.

The first case is not suitable for special applications, but we consider it here for the comparison. Note that for all cases except the first one the fault tolerance is the same (provided the assumptions from the previous section).

Table 3 shows the simulation results for the factor calculated using equation (6), where “unhardened” case is the first one mentioned above. We used  $t_{MM} = 20 \text{ ns}$  for main memory access for all cases. One can see that the introduction of ECC almost does not influence the performance, but the DICE-based cache reduction can drastically decrease the performance not due to the cell speed (it is almost the same), but due to the increased cache misses even when SER is neglected.

Note that besides the performance-tolerance factor, the total access time for the cache system can be calculated.

**Table 3.** Simulated performance-tolerance factors.

	65 nm	250 nm
Case 1	0.96	0.94
Case 2	0.79	0.83
Case 3	0.61	0.65

#### 4. Conclusions

We propose the simplified analytical model for two-level cache memory system that is exposed to SEE at the given orbit. The model takes into account the experimental data for SRAM Single Event Upset (SEU) tolerance for a given technology node, the time penalty coming from the used Error Correcting Codes (ECC). The fault-tolerance is important factor affecting performance, chip area and power consumption. Involving special SRAM cells and error-correcting codes is often too expensive with relation to the performance needed. The introduction of ECC almost does not influence the performance, but the DICE-based cache reduction can drastically decrease the performance not due to the cell speed, but due to the increased cache misses even when SER is neglected. The proposed model can be further extended for more common cases.

#### 5. Acknowledgments

This research was supported by the Russian Foundation for Basic Research grant to the Scientific Research Institute of System Analysis (14-29-09268).

Authors would like to thank P. Dolotov, Y. Katunin and P. Stepanov from SRISA, Moscow, Russia, for valuable information about SRAM characteristics.

#### References

- [1] Dodd P E *et al.* 2010 *IEEE Trans. Nucl. Sci.* **57** 1747-63
- [2] <http://www.baesystems.com/>
- [3] <http://www.cpushack.com/>
- [4] <http://www.esa.int/>
- [5] RAD5545 Multi-core System-on-chip Power Architecture Processor. Datasheet
- [6] Przybylski S A 1990 *Cache and Memory Hierarchy Design: A Performance-Directed Approach* (San Mateo: Morgan Kaufmann Publishers, Inc.)
- [7] Calin T *et al.*, 1996 *IEEE Trans. Nucl. Sci.* **43** 2874-8
- [8] Chow C K 1976 *IEEE Trans. Comput.* **25** 157-64
- [9] Hartstein A *et al.* 2008 *J. Instruction-Level Parallelism* **10** 1-22
- [10] Hennessy J L and Patterson D A 2011 *Computer Architecture, Fifth Edition: A Quantitative Approach* (The Morgan Kaufmann Series in Computer Architecture and Design)
- [11] <https://creme.isde.vanderbilt.edu/>
- [12] <https://www.spennis.oma.be/>
- [13] Petersen E 2011 *Single Event Effects in Aerospace* (John Wiley & Sons)
- [14] Gorbunov M S *et al.* 2014 *IEEE Trans. Nucl. Sci.* **61** 1575-82