

Enhancement of System-Lifetime by Alternating Module Activation

Frank Torres

► **To cite this version:**

Frank Torres. Enhancement of System-Lifetime by Alternating Module Activation. Gunar Schirner; Marcelo Götz; Achim Rettberg; Mauro C. Zanella; Franz J. Rammig. 4th International Embedded Systems Symposium (IESS), Jun 2013, Paderborn, Germany. Springer, IFIP Advances in Information and Communication Technology, AICT-403, pp.82-91, 2013, Embedded Systems: Design, Analysis and Verification. <10.1007/978-3-642-38853-8_8>. <hal-01466695>

HAL Id: hal-01466695

<https://hal.inria.fr/hal-01466695>

Submitted on 13 Feb 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Enhancement of System-Lifetime by Alternating Module Activation

Frank Sill Torres

Dept. of Electronic Engineering, Federal University of Minas Gerais, Brazil
franksill@ufmg.br

Abstract. Reliability and robustness have been always important parameters of integrated systems. However, with the emergence of nanotechnologies reliability concerns are arising with an alarming pace. The consequence is an increasing demand of techniques that improve yield as well as lifetime reliability of today's complex integrated systems. It is requested though, that the solutions result in only minimum penalties on power dissipation and system performance. The approach Alternating Module Activation (AMA) offers both extension of system lifetime and low increase of power and delay. The essential contribution of this work is an analysis to which extent this technique can be improved even more. Thereby, components that enable partial concurrent error detection as well as Built-in self-test functionality are included. Further, a flow for comparison of system's lifetime on cell-level is presented. Final results indicate an improvement of the system's lifetime of up to 58 % for designs in which the expected instance lifetime differs by factor 2.

Keywords: Robustness; Redundancy; Sleep Transistor, Modeling, BIST

1 INTRODUCTION

CMOS is still the predominating technology for digital designs with no identifiable concurrence in the near future. Driving forces of this leadership are the high miniaturization capability and the robustness of CMOS. The latter, though, is decreasing with an alarming pace against the background of technologies with sizes at the nanoscale. Such technologies, with device dimensions in the range of a few nanometers, suffer from an increased susceptibility to different kinds of failures during operation [1]. In contrast to previous technology generations, solutions within the manufacturing process are not sufficient anymore to deal with these kinds of issues. Accordingly, reliability concerns are not only an issue of manufacturing anymore, but also have to be considered in all abstraction layers of the design process. Thereby, three main strategies can be identified: (I) design techniques that detect errors [2], (II) techniques that detect and correct errors [3] and (III) those techniques that try to avoid or at least prolongate errors [4][5]. As those techniques of strategy (I) require another mechanism to cope with the detected error, they do not increase the expected lifetime of the designs as aimed at in this work.

We proposed in previous works [6][7] a design technique that relates to strategy (III) and combines Sleep Transistors with the idea of modular redundancy to extend lifetime reliability of integrated circuits. In this work, we propose how this approach can be combined with techniques of strategy (II) to cope with errors as they can finally still occur.

The remainder of this contribution is organized as follows: section 2 summarizes the initial approach while section 3 presents the proposed extensions of the design technique. The subsequent section 4 introduces an extended flow on cell-level in order to compare lifetime reliability of integrated circuits. The following section 5 presents and discusses simulation results before section 6 concludes this work.

2 Alternating Module Activation

This section describes the fundamentals of the previously developed approach Alternating Module Activation as well as requirements for the necessary control logic.

2.1 Basic Idea

An essential characteristic of power gating with Sleep Transistors [8] is its ability to dynamically disconnect the power supply during the runtime of integrated systems. Hence, during the disconnected state the gated logic is ideally without any inherent currents and voltages, and thus electromagnetic fields. Furthermore, local temperatures are reduced as there is no switching activity present. It should be noted that these are key parameters for several lifetime decreasing effects, like electromigration [9], gate-oxide breakdown [10], and negative bias temperature instability [11], of integrated circuits. Thus, during an idle phase of a gated logic these effects are eliminated or at least strongly reduced. As a consequence, the mean time to failure (MTTF), which is the average time that a system operates until it fails, is prolonged approximately by the time that the design is in the idle phase. This relation is applied by the proposed approach Alternating Module Activation (AMA). Hereby, each gated module (i.e. each logic block) is implemented at least two times (see also Fig. 1). During the runtime though, only one of these instances is active while the others are disconnected from the power supply. Consequently, the resulting ideal mean time to failure $MTTF'_{AMA}$ of a module realized with the proposed approach can be expressed by:

$$MTTF'_{AMA} = N \cdot MTTF_{min} \quad (1)$$

where N is the number of redundant instances and $MTTF_{min}$ is the minimum MTTF over all module instances. Equation (3) refers to the ideal case where any additional logic is neglected and the gated modules are completely disconnected from the power supply.

It could be shown in previous works [7] that there is a moderate increase in dynamic power dissipation of ca. 6 %, while the leakage and area are roughly doubled.

2.2 Control circuitry

In order to properly work, additional logic is required to multiplex the results from the currently active instance to the subsequent module. This is implemented by multiplexers that are placed behind the redundant instances as depicted in Fig. 1. Here, a simple 2:1 multiplexer is shown to forward the correct signals from the redundant instances of the module A to the subsequent module B.

Commonly, power gated logic requires additional clock cycles before the logic can be fully operated again (i.e. wake-up time [8]). Hence, it is not feasible to connect the signals controlling the Sleep Transistors (here /Sleep1 and /Sleep2) also directly to the multiplexers. Instead, a control signal scheme as shown in Fig. 1 should be applied to ensure data consistency. Thereby, it has to be assured that before a transition of the multiplexed outputs both instances are active (/Sleep1 and /Sleep2 are logically '1').

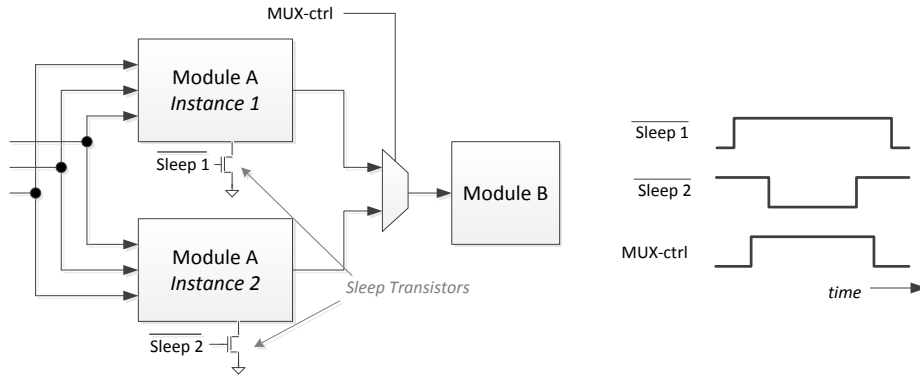


Fig. 1. The initial AMA approach with two redundant instances, whereas the results of the active instance are forwarded by the subsequent multiplexer, and related control signal scheme

Considering the transition time the mean time to failure $MTTF''_{AMA}$ results to:

$$MTTF''_{AMA} = N \cdot (1 - p_{trans,i}) \cdot MTTF_i \quad \text{with: } MTTF_i = MTTF_{min} \quad (2)$$

with p_{trans} is the probability that the instance i is in the transition phase but its output is still not forwarded by the multiplexer.

For a comprehensive investigation, it has to be considered that the lifetime of the system also depends on the MTTF of the multiplexers. The multiplexers though are realized as transmission gates [6], whereas only one path is active at a time. Thus, the impact of failure mechanisms, like gate-oxide breakdown or electromigration [12], is also correspondingly smaller. Nevertheless, it is reasonable to apply special design strategies for the multiplexers as well, like transistors with thicker gate oxide and wider wires.

3 Enhanced Alternating Module Activation

This section proposes extensions of the AMA approach that increase the lifetime in case of faultiness of one of the instances. Beside this, error detection capability and Built-in self-test (BIST) functionality are added.

3.1 Partial Concurrent Error Detection

A missing function of the initial version of the Alternating Module Activation approach is error detection capability. Hence, it is proposed to add comparators to each multiplexer. Its function is the verification that all multiplexer's inputs have the same value, and thus, whether all instances of a module produce equal results (comparator C-M in Fig. 2). However, only during the transition phase, i.e. when one instance is disconnected from supply while another is connected (see Fig. 1), more than one instance is active at the same time. This presents a limitation as only during this phase concurrent error detection (CED) is possible. The intention of this partial CED though, is not the identification of transient faults [13]. In contrast, its purpose is the detection of permanent faults.

It is recommended to modify the transition phase in the way that all instances are connected for a limited time. Thus, the probability of detection of an error can be increased. Considering this change, the mean time to failure $MTTF'_{EAMA}$ of a module results to:

$$MTTF'_{EAMA} = N \cdot (1 - p_{trans}) \cdot MTTF_{min} \quad (3)$$

where p_{trans} denotes the probability that the instances are in transition phase and its outputs are not forwarded by the multiplexer. Consequently, the increase of p_{trans} reduces the time between the occurrence of a permanent failure and its detections.

3.2 Selective Complete Deactivation Of Instances

One major drawback of the initial version of the AMA approach is the complete function loss if one of the instances fails. Thus, it is proposed to utilize the existence of at least two instances of each module. The basic idea is the complete deactivation of an instance in case of failure, i.e. the control algorithm stops to consider the defective instance. This deactivation can be pushed so far that the single instance configuration is reached. Thus, the expected life time of the circuit can be increased by the difference of the mean time to failure of the instances of each module. Considering exclusively the MTTF of the instances of the module the resulting $MTTF_{mod_EAMA}$ can be estimated with:

$$\begin{aligned}
\text{MTTF}_{\text{EAMA}} &= \left[N \cdot \text{MTTF}_1 + (N-1) \cdot (\text{MTTF}_2 - \text{MTTF}_1) + \dots + (\text{MTTF}_N - \text{MTTF}_{(N-1)}) \right] \cdot (1 - p_{\text{trans}}) \\
&= (1 - p_{\text{trans}}) \cdot \sum_{i=1}^N \text{MTTF}_i \quad \text{with: } \text{MTTF}_i > \text{MTTF}_{(i-1)}
\end{aligned} \tag{4}$$

Hence, in case of different mean time to failures of the instances the increase $\Delta\text{MTTF}_{\text{mod_EAMA}}$ results to:

$$\Delta\text{MTTF}_{\text{EAMA}} = (1 - p_{\text{trans}}) \cdot \sum_{i=2}^N \text{MTTF}_i \quad \text{with: } \text{MTTF}_i > \text{MTTF}_{(i-1)} . \tag{5}$$

These differences of the MTTF result from variations of process parameters, aberrations of layout parameters, on-die temperature distribution, and effects through neighboring blocks.

3.3 Built-In Self-Test for Faulty Instance identification

Another missing function of the initial approach is the identification of a faulty instance. Hence, it is proposed to add a memory based Built-In Self-Test (BIST) mode. Therefore, test input and output vectors for each module have to be generated and stored in a memory block whose inputs are multiplexed to the module inputs. Further, the outputs of the memory and the module are connected to comparators (see Fig. 2). Thus, in case of detection of an error by the partial CED the proposed BIST structure can be applied for successive tests of each instance for identification of the faulty one.

3.4 Final Architecture and Control Scheme

Fig. 2 shows the final architecture of the extended approach whereas the initial blocks are greyed out. For reasons of lifetime extension both kinds of comparators as well as the memory can be switched off by Sleep Transistors when it is not needed. Fig. 3 depicts the new control structure which is extended by two phases of error detection. As described in subsection 3.1 the partial CED is only active during the transition phase. Further, the design changes to the BIST mode only in case of the detection of an error. During that mode the system has to be halted as no correct functionality can be guaranteed. After detection of a faulty instance it is removed from the list of possible active instances and the system returns to normal operation.

4 Technique for MTTF Comparison on Cell-Level

This section proposes a new technique on cell-level for the comparison of mean time to failure of integrated designs.

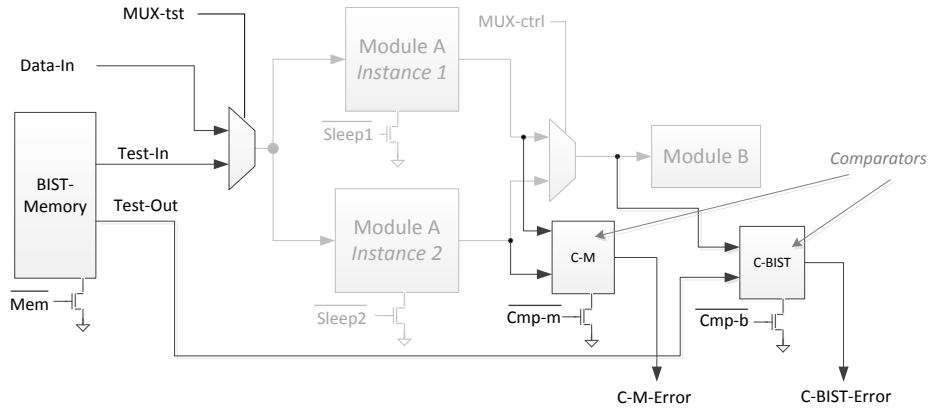


Fig. 2. Structure of Enhanced AMA (blocks of the initial AMA are greyed out)

4.1 Types of modeling of failure mechanisms

Several models for individual failure mechanisms within integrated circuits can be found in the literature [8][9][12], whereas SPICE simulations are reported as the most accurate approach used by circuit designers. However, the accuracy comes together with major computational efforts and simulation times, which limits the maximum number of elements within an investigation. In contrast, approaches on higher levels decrease drastically the effort in computation, allowing the analysis of considerably more complex design [14][15]. However, this gain comes with the price of reduced accuracy.

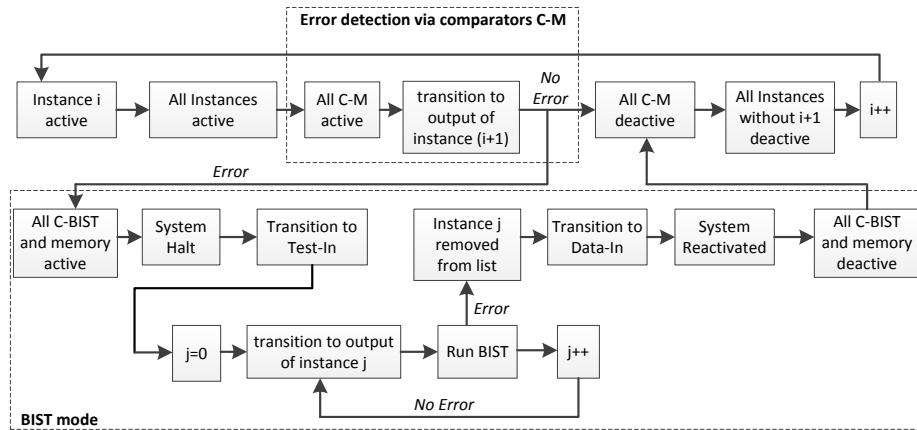


Fig. 3. Control flow for extended AMA approach, enhanced by an error detection phase during module transition and a BIST mode

4.2 Cell models for MTTF Comparison

The proposed approach for modeling of the meant time to failure of integrated designs is an extension of the in [6] presented mixed-signal method. In contrast to the solution on SPICE level the new approach applies models on cell-level in which the function of the logic cells deteriorates over time. Thereby, the level of degradation is parameterized for each cell and bases on results from studies on SPICE level [6]. Further, all cells receive an additional input that defines whether the cell is active or deactivated. This allows different level of degradation depending on the state of connected Sleep Transistors.

5 RESULTS OF THE SIMULATIONS

In this section, the setup of the test environment is presented before the obtained simulation results are discussed.

5.1 Setup of the test environment

The presented simulation results are based on designs from the ISCAS benchmark suite (c1355 and c3450) [15], the ITC99 benchmark suite (b05, b15 and b21) [17], and two proprietary designs, i.e. a 32-bit multiplier (mult) and a simple 8-bit MIPS-like processor (MIPS). The applied library consists of 8 standard cells described as VERILOG modules with a hard degradation limit that takes the cell active time into consideration. The levels of cell degradation are based on simulation results obtained from the test environment presented in [7]. Thereby, all cells were realized in a predictive 16 nm technology [18] and simulated with the same error models as in [7]. Next, all cells were simulated with different parameters for the error models, and with connected Sleep Transistors in on- and off-mode. Thereby, the values of the parameters were chosen in a manner that for each cell five different MTTF could be defined. The multiplexer are implemented as transmission gates with increased transistor dimensions that elevate the MTTF of these components.

In the current implementation the control circuitry is included in the test environment and not part of the analyzed designs. In future works several robust design strategies shall be analyzed for these block. The probability that an instance is in a transition phase without having its outputs forwarded, defined by frequency and length of a transition, was set to 1 %. Due to random values for degradation all simulations were executed 100 times. Further, the automated duplication of module instances as well as the insertion of Sleep Transistors, multiplexer and comparators is done by a tool specifically written for these tasks.

The number of redundant instances is limited to two as we consider solutions with higher numbers of instances as too costly in terms of area.

5.2 Results and discussion

In a first step it was verified whether the results of the initial AMA approach (see section 2) can be reproduced in the proposed test environment. Therefore, for each design the MTTF of the raw version without any redundant blocks was estimated. Subsequently, those designs were modified according to the AMA approach presented in section 2. Thus, each design was duplicated and complemented with the multiplexer, while the test environment was extended by the related control logic. Further, for both experiments the cell degradation models that lead to the longest MTTF were chosen. At these simulations, a design was considered as defective with the appearance of the first wrong result at the design outputs. The results presented in Fig. 4 show that the improvements of the initial approach could be reproduced whereas the MTTF could be increased by an average of factor 1.98.

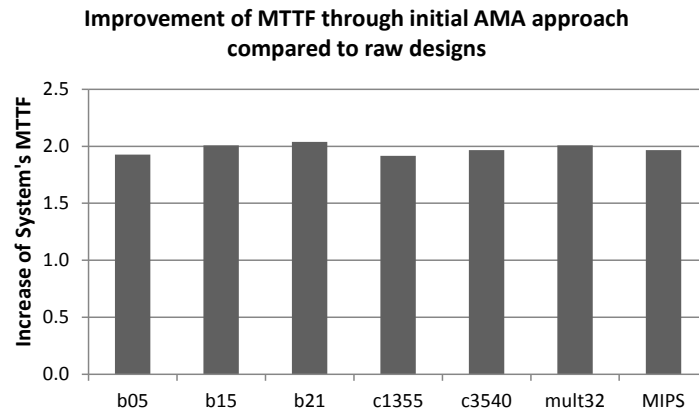


Fig. 4. Increase of Mean Time To Failure (MTTF) of designs realized with the initial AMA approach (each module with two instances) compared to the raw versions

In the next step the proposed extension of the AMA approach (see section 3) was analyzed. Initially, the designs were simulated with each module realized as single instance and for all five degradation classes of the cells. Next, comparators and a memory based BIST were added while the control logic was extended by error verification. Then, each modified design was simulated five times whereas the cell degradation models of one instance of each module were varied.

The results of this analysis are depicted in Fig. 5. Here, the increase of the mean time to failure of the extended approach compared to the initial one is shown for varying relation between the MTTFs of the instances. The depicted curves show the minimum, average, and maximum improvement of the system's MTTF compared to the initial approach. It follows that for equal distributed MTTFs of the instances the proposed extension leads only to a negligible increase of expected system lifetime (average: 1 %). In contrast, already with one group of instances having a 25 % lower mean time to failure the system's MTTF can be increased in average by 8 %. If one group of the instances has a MTTF this is by factor 2 shorter the improvement increases up

to 51 % (average: 50 %). Hence, it could be shown that the presented approach can considerably increase the system lifetime.

It should be noted that these simulations cannot classify the increase of robustness against errors that are not based on temporary degradation but abrupt failures, e.g. based on high temperature peaks, extreme overvoltage due to electro-static discharge, or infant mortality effects.

For this analysis we consider a comparison of the MTTF of raw designs with the extended version of the approach as not conclusive. This is due to fact that the choice for the cell degradation model for the raw versions would be only random.

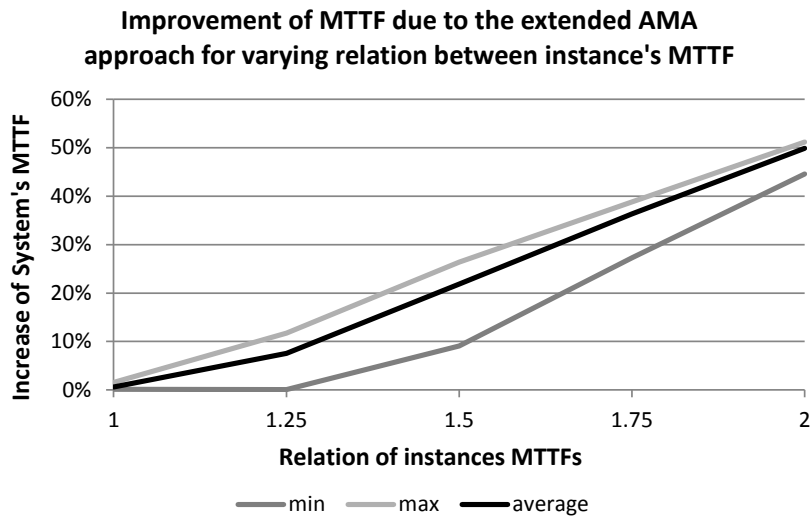


Fig. 5. Improvement of Mean Time To Failure (MTTF) of extended approach compared to the initial approach under variation of relation of instance's MTTFs (each module with 2 instances)

6 CONCLUSION

Integrated circuits realized in nanometer technology are continuously more susceptible to severe failure mechanisms. This alarming development necessitates design techniques to improve the lifetime reliability. Hence, the presented work proposes an extension of an approach that combines the ideas of Sleep Transistors and modular redundancy in a beneficial way. Thereby, the approach aims at increased lifetime reliability while the impact on delay and power dissipation is kept to a minimum. Due to proposed extensions of this work it is possible to detect permanent errors. Furthermore, the modifications lead to extension of the expected system's lifetime as faulty instances can be identified and disconnected. In order to compare system lifetime, we also proposed a modeling technique on cell-level. Finally, simulation results show that the proposed improvements of the design approach can extend the system's Mean Time To Failure (MTTF) in average by 52 % if the instance's MTTF differ by factor 2.

7 REFERENCES

1. J. Srinivasan, S. Adve, P. Bose, and J. Rivers, "The impact of technology scaling on lifetime reliability", Proceedings of IEEE International Conference on Dependable Systems and Networks, (2004).
2. P. Bernardi, L. M. V. Bolzani, M. Rebaudengo, M. S. Reorda, F. L. Vargas, and M. Violante, "A new hybrid fault detection technique for Systems-on-a-Chip", IEEE Transaction on Computers, (2006). 55, 2, pp. 185-198.
3. S. Mitra, N. Seifert, M. Zhang, Q. Shi, and K. S. Kim, "Robust System Design with Built-In Soft-Error Resilience," Computer, (2005), Vol. 38, Issue 2, pp. 43-52.
4. T. Inukai, T. Hiramoto, and T. Sakurai, "Variable threshold CMOS (VTCMOS) in series connected circuits" Proceedings of the International Symposium on Low Power Electronics and Design, (2001), pp. 201-206.
5. J. Tschanz et al., "Adaptive body bias for reducing impacts of die-to-die and within-die parameter variations on microprocessor frequency and leakage", IEEE Journal of Solid-State Circuits, (2002), vol. 37, pp. 1396-1402.
6. F. Sill Torres, C. Cornelius, D. Timmermann, "Reliability Enhancement via Sleep Transistors", Proceedings of 12th IEEE Latin-American Test Workshop (2011), p. 1-6.
7. C. Cornelius, F. Sill Torres, D. Timmermann, "Power-Efficient Application of Sleep Transistors to Enhance the Reliability of Integrated Circuits", Journal of Low Power Electronics, (2011), v. 7, n° 4, p. 552-561.
8. M. Powell, S.-H. Yang, B. Falsafi, K. Roy, and T. N. Vijaykumar, "Gated-Vdd: A circuit technique to reduce leakage in deep-submicron cache memories", Proceedings of International Symposium on Low Power Electronics and Design, (2000), pp. 90-95.
9. J. Srinivasan, S. V. Adve, P. Bose, J. Rivers, and C.-K. Hu, "RAMP: A Model for Reliability Aware Microprocessor Design", IBM Research Report, RC23048 (2003).
10. J. Stathis, "Reliability limits for the gate insulator in cmos technology", IBM Journal of Research & Develop, (2002), Vol. 46, N° 2/3, pp. 265-286.
11. E. Maricaud and G. Gielen, "NBTI model for analogue IC reliability simulation", Electronics Letters, (2010), Vol. 46, N° 18.
12. "Failure Mechanisms and Models for Semiconductor Devices", JEDEC Publication JEP122-A, Jeduc Solid State Technology Association (2002).
13. R. Possamai Bastos, F. Sill Torres, G. Di Natale, M. Flottes, B. Rouzeyre, "Novel transient-fault detection circuit featuring enhanced bulk built-in current sensor with low-power sleep-mode", Microelectronics Reliability, 52, 9-10, (2012), 1781-1786.
14. D. Lorenz, M. Barke, U. Schlichtmann, "Aging analysis at gate and macro cell level," Proceedings of Computer-Aided Design, (2010), pp.77-84.
15. J. Xiao; J. Jiang; X. Zhu; C. Ouyang, "A Method of Gate-Level Circuit Reliability Estimation Based on Iterative PTM Model", Proceed. Dependable Computing, (2011), pp.276-277.
16. M. Hansen, H. Yalcin, and J. P. Hayes, "Unveiling the ISCAS-85 Benchmarks: A Case Study in Reverse Engineering", IEEE Design & Test, (1999), Vol. 16, N° 3, pp. 72-80.
17. L. Basto, "First results of ITC'99 benchmark circuits," IEEE Design & Test of Computers, vol.17, no.3, (2000), pp. 54-59.
18. W. Zhao, and Y. Cao, "New generation of Predictive Technology Model for sub-45nm early design exploration," IEEE Transactions on Electron Devices, (2006), Vol. 53, N°. 11, pp. 2816-2823.

Acknowledgements. This work was supported by grants from CNPq, CNPq/DISSE, CAPES, FAPEMIG, and UFMG/PRPq.