

# A New Reliability Evaluation Methodology and its Application to Network-on-Chip Routers

Hamed S. Kia and Cristinel Ababei  
Department of Electrical and Computer Engineering  
North Dakota State University, Fargo ND, USA  
Email: {hamed.sajjadikia, cristinel.ababei}@ndsu.edu

**Abstract**—We propose a new circuit level reliability evaluation methodology. The proposed methodology is based on a divide and conquer approach, which enjoys the benefits of device level accuracy and of block level efficiency. At the core of the reliability estimation engine lies a Monte Carlo algorithm which works with failure times modeled as Weibull and lognormal distributions for major wearout mechanisms: time dependent dielectric break down, negative bias temperature instability, electromigration, thermal cycling, and stress migration. As a case study, we demonstrate how the proposed reliability evaluation technique can be applied to a Network-on-Chip router to identify the most vulnerable subblocks, which represent the reliability bottlenecks of the router.

## I. INTRODUCTION

Reliability has become a fundamental challenge in the design of integrated circuits [1]. One of the main difficulties in reliability aware design is the estimation of reliability. Evaluation of reliability is a challenging task because reliability is affected by numerous factors including aging mechanisms (e.g., time-dependent dielectric breakdown (TDDB) [2], negative bias temperature instability (NBTI) [3], electromigration (EM) [4], thermal cycling (TC), and stress migration (SM) [5]), process variations, dynamic power and thermal management, workload, and system architecture and configuration. While there has been significant work carried out to estimate reliability [6]–[18], we discuss next two approaches that are closest to our work. An extensive review of previous reliability simulation tools can be found in [6].

The RAMP approach [8], [9] models the mean time to failure (MTTF) of a processor microarchitecture as a function of temperature related failure rates of individual structures on chip. It divides the processor into several discrete structures (e.g., ALU, register files, etc) and applies analytical models to each structure. Then, it combines the structure level MTTFs to compute the overall MTTF of the entire processor assumed as a series failure system. Because the lifetime distributions of failure mechanisms are assumed to be exponential, the reliability is calculated by applying the sum-of-failure-rates (SOFR) model. This approach is not realistic because failure rates of units increase with time due to aging. To address this limitation of the SOFR model, RAMP 2.0 [10], [11] uses lognormal distributions, which are harder to deal with analytically. One of the main limitations of the RAMP approach as an architecture level approach is its accuracy. In addition, it may estimate equal MTTFs for blocks of different sizes but with activity

factors that cancel out the area proportionality factor.

Another, more recent, class of reliability evaluation approaches are based on Spice simulations. Failure rate based Spice (FaRBS) [12] and Maryland circuit reliability oriented (MaCRO) [13] are circuit level reliability simulation methods. The main advantage of this class of simulation methods is the device level granularity that enables reliability analysis at transistor level to identify the most vulnerable transistors. There are some issues related to the Spice based reliability simulation. These approaches do not consider the layout of the system and simulations are done under worst case temperature scenarios, which is not realistic. Besides, Spice circuit simulations tend to take long time especially when done for large circuits. In addition, both methods (FaRBS and MaCRO) are developed under the assumption that failure rate is constant. As discussed above this assumption is inaccurate.

## II. LIFETIME FAILURE MODELS

Many proposed lifetime reliability models assume a uniform device density on the chip and an identical vulnerability of devices to failure mechanisms. The lifetime distributions of failure mechanisms are usually assumed to be exponential [8]–[10], [15], [16]. As discussed in the previous section, this allows system-level reliability to be calculated by applying the sum-of-failure-rates (SOFR) model. However, this approach is not realistic because failure rates of units increase with time due to aging. To address this issue and to develop an accurate reliability model, more general lifetime distributions (e.g., Weibull and lognormal) must be utilized. On the other hand, when Weibull or lognormal distributions are utilized the prediction of reliability becomes more difficult and therefore Monte Carlo simulations must be employed. In this paper, we adopt Weibull distribution modeling for TDDB, NBTI, TC, and SM and lognormal distribution modeling for EM because these distributions have been found to best fit the corresponding wearout mechanisms [5].

Due to space limitations, below we only list the expressions of the mean time to failure that characterize the models of the aging mechanisms considered in this paper. For elaborate discussions of these failure mechanisms and their models the reader is referred to [5]–[7], [13], [14].

$$MTTF_{TDDB} \propto \left(\frac{1}{A}\right)^{\frac{1}{\beta}} \left(\frac{1}{F}\right)^{\frac{1}{\beta}} V_{gs}^{a+bT} e^{\left(\frac{c}{T} + \frac{d}{T^2}\right)} \quad (1)$$

$$MTTF_{NBTI} \propto V_{gs}^{-\frac{1}{\beta}} \left[ \frac{1}{1 + 2e^{-\frac{E_1}{kT}}} + \frac{1}{1 + 2e^{-\frac{E_2}{kT}}} \right]^{-\frac{1}{\beta}} \quad (2)$$

$$MTTF_{EM} \propto (J - J_{crit})^{-n} e^{\frac{E_a EM}{kT}} \quad (3)$$

$$MTTF_{TC} \propto \left( \frac{1}{T - T_{ambient}} \right)^q \quad (4)$$

$$MTTF_{SM} \propto |T_0 - T|^{-n} e^{\frac{E_a EM}{kT}} \quad (5)$$

The definitions of the variables and parameters involved in these expressions are well known and described in previous literature. In this paper, we utilize typical values of these parameters from [5], [14].

### III. NEW RELIABILITY EVALUATION METHODOLOGY

The proposed reliability evaluation methodology is an attempt to bridge the benefits of both RAMP like and Spice simulation based reliability evaluation approaches. The block diagram that illustrates the main steps of the proposed reliability evaluation methodology is shown in Fig.1.

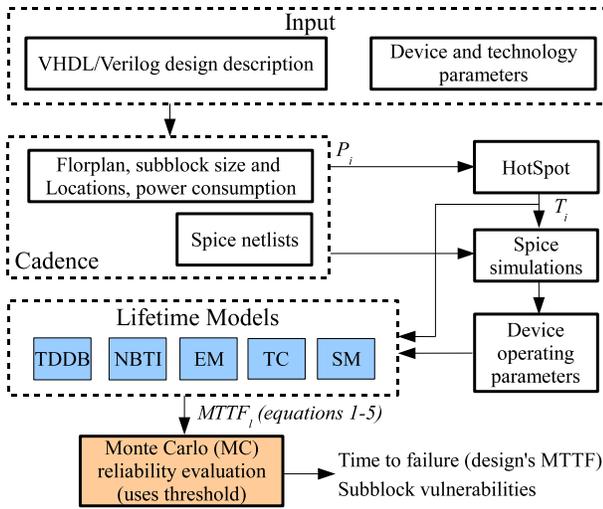


Fig. 1. Flow chart of the proposed reliability evaluation methodology.

The salient features of our methodology are as follows. First, in order to deal with complexity due to circuit size we adopt a divide and conquer approach. The hierarchy of the structure of a design is exploited by *zooming-in* to lower levels where the analysis is tractable within reasonable computational time. Second, similar to MaCRO method [13], we employ subblock level Spice simulations to derive transistor operating parameters. However, we conduct Spice simulations at realistic temperatures (different subblocks have different temperatures) rather than worst-case temperatures as it is done pessimistically in [13]. Third, we model failure times using Weibull and lognormal distributions that have been found to better fit experimental data. Fourth, the block level reliability (as MTTF) is estimated by aggregating the failure times of the individual subblocks. This process is implemented such that

the design hierarchy is *zoomed-out* to upper levels. Finally, the proposed method has the ability to identify the most vulnerable subblocks from a reliability point of view.

Following the flow chart from Fig.1, the main steps of the proposed reliability evaluation methodology are as follows:

*Step 1:* We start from a given hierarchical description of the design under consideration. This description can be in any hardware description language such as VHDL or Verilog. In addition, transistor and technology parameters are assumed to be given based on the technology node in which the design is to be fabricated.

*Step 2:* The design is synthesized, placed, and routed using Cadence tools [19], but any other CAD tool can be utilized. The resulting layout represents the block level floorplan, which is divided into individual structures or subblocks based on the initial structural description of the design. In this way, we basically obtain for each subblock its layout, location, and aspect ratio. In addition, power consumption estimates are also generated using Cadence tools.

*Step 3:* The floorplan and power estimates are then fed into HotSpot [20]. HotSpot is an accurate and fast thermal model based on an equivalent circuit of thermal resistances and capacitances that correspond to microarchitecture blocks. The output of the HotSpot simulation is a list with temperatures of each subblock. Our approach addresses one of the limitations of MaCRO like methods [13]. As mentioned earlier, instead of doing worst-case temperature simulations we work with actual operating temperature for each subblock. Therefore, reliability of each subblock can be evaluated more accurately.

*Step 4:* These temperatures are utilized together with circuit netlists generated from within Cadence tools to perform subblock level Spice simulations. These simulations provide us with the transistor operating parameters necessary to be plugged into the equations modeling the wearout mechanisms described in Section II.

*Step 5:* At this stage we have everything that is needed by the lifetime failure models described by equations (1)-(5). At the core of the proposed methodology we employ a Monte Carlo (MC) simulation algorithm (see Fig.1) implemented and run in Matlab [21]. Our technique is inspired from the RAMP method but executed at the subblock level where the elementary unit is the transistor.

The MC algorithm proceeds with the following main steps (1) for each failure mechanism run  $N = 10^7$  simulations: (a) for each transistor, generate failure time samples from the corresponding distribution and (b) use MIN analysis of these times by assuming the subblock as a series system to calculate the time to failure  $t_{f_{min}}^j$  of simulation  $j = 1, \dots, N$ . (2) calculate the overall subblock MTTF of time to failure for the current failure mechanism as  $t_{fi} = (\sum_{j=1}^N t_{f_{min}}^j) / N$ . (3) calculate the value of the overall subblock's time to failure as the minimum among the failure times due to each failure mechanism.

In our experiments, we found that in order to better differentiate between subblocks one only needs to focus on the most vulnerable transistors in a given subblock. Hence,

Algorithm: Reliability Evaluation	
1:	<b>In:</b> VHDL/Verilog description of design hierarchy. Device and technology parameters
2:	<b>Out:</b> Subblocks' vulnerabilities and times to failure, design's time to failure
3:	<b>Start</b>
4:	Synthesize design to generate its layout and floorplan
5:	Retrieve dimensions and location of each subblock
6:	Estimate power consumption of each subblock $P_i$
7:	Estimate operating temperature of each subblock $T_i$
8:	Generate Spice netlist for each subblock
9:	Simulate each subblock at estimated $T_i$ to derive operating parameters $V_k$
10:	Estimate times to failure $tf_i$ and percentage of vulnerable transistors:
11:	<b>for</b> $i \leftarrow 1$ to $S$ <b>do</b> // $S$ : number of subblocks
12:	<b>for</b> $l \leftarrow 1$ to $F$ <b>do</b> // $F$ : number of failure types
13:	Calculate $MTTF_l$ using equations 1-5 from Section II
14:	<b>for</b> $j \leftarrow 1$ to $N$ <b>do</b> // $N = 10^7$ Monte Carlo iterations
15:	$tf_{min}^j \leftarrow INF$ // Initialize
16:	$Counter_j \leftarrow 0$ // Initialize
17:	<b>for</b> $k \leftarrow 1$ to $D$ <b>do</b> // $D$ : number of devices
18:	$tf_k \leftarrow generate\_sample(MTTF_l)$
19:	<b>if</b> $tf_k < tf_{min}^j$ <b>then</b>
20:	$tf_{min}^j = tf_k$
21:	<b>end if</b>
22:	<b>if</b> $tf_k < threshold$ <b>then</b>
23:	$Counter_j = Counter_j + 1$
24:	<b>end if</b>
25:	<b>end for</b>
26:	$Frac\_below\_threshold_j = 100 \cdot \frac{Counter_j}{D}$
27:	<b>end for</b>
28:	$tf_i = \frac{\sum_j^N tf_{min}^j}{N}$
29:	$Frac\_below\_threshold = \frac{\sum_j^N Frac\_below\_threshold_j}{N}$
30:	<b>end for</b>
31:	$tf_i = MIN\{tf_i\}$
32:	<b>end for</b>
33:	$tf = MIN\_MAX\{tf_i\}$
34:	<b>End</b>

Fig. 2. Pseudocode of the proposed reliability evaluation methodology.

we introduce a *threshold* that helps to identify transistors whose lifetime samples are smaller than this threshold. As an indicator of how vulnerable a subblock is, we calculate the percentage of transistors whose lifetime sample is smaller than the selected threshold. This is illustrated in the pseudocode description of the algorithm presented in Fig.2. The threshold value is selected during the reliability qualification process as a function of the desired expected lifetime.

It is important to note that the pseudocode on lines 14-29 in Fig.2 is applicable to the TDDB and NBTI cases only. In this case, the proposed reliability methodology performs MC analysis at the transistor level. For the EM, TC, and SM wearout mechanisms the MC analysis is done at the subblock level similar to RAMP 2.0 [9]. The pseudocode description in these cases is simpler and we omit it here due to space limitations.

The output of the proposed reliability evaluation methodology consists of the actual estimate of the time to failure or MTTF of the design (line number 33 in Fig.2) and vulnerabilities of each individual subblock as percentage of transistors with average failure time shorter than the threshold (line number 29 in Fig.2). The overall MTTF of the design is estimated

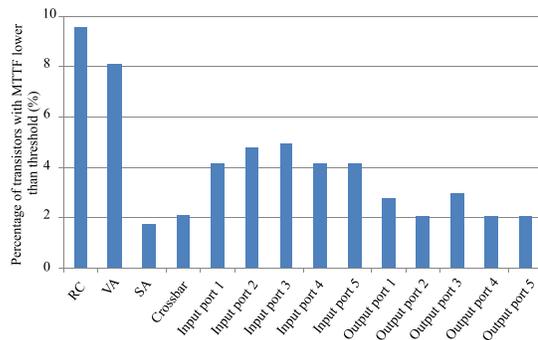


Fig. 3. Percentage of transistors with MTTF lower than threshold for TDDB case.

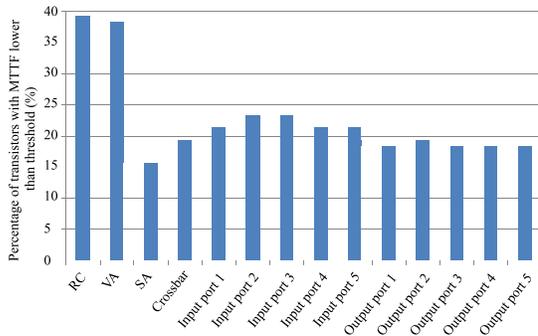


Fig. 4. Percentage of transistors with MTTF lower than threshold for NBTI case.

using a *MIN\_MAX* analysis similar to [9] in order to be able to handle redundant subblocks that might be introduced for improving reliability. This information can be useful to circuit and system designers to develop fault tolerant or robust circuits and systems. Armed with information about what are the reliability critical subblocks and transistors, designers can concentrate their design efforts [22] with wearout mechanism specific techniques only on those, thereby saving area and power.

#### IV. CASE STUDY AND SIMULATION RESULTS

While Network-on-Chip (NoC) has become the dominant communication architecture to cope with the ever-increasing complexity of multicore systems, its reliability has been studied significantly less compared to that of cores. Therefore, in this section, we demonstrate the proposed reliability evaluation methodology on an NoC router. Our goal is to analyze the microarchitecture of a typical router to identify its most vulnerable components. We focus our attention on one of the most popular pipelined router architectures described in [23]. The main components of this architecture are: routing computation (RC), virtual channel allocation (VA), switch allocation (SA), input ports, output ports, and crossbar switch.

We first code the router's specification in Verilog. Specifics of this description include: 5 input and 5 output ports, 2 virtual channels per port, 4 sets of registers for each virtual channel of each port, and 16 bites wide links. Then, we take the router design through the reliability evaluation methodology described in Fig.1. We utilize Nangate 45nm Open Cell

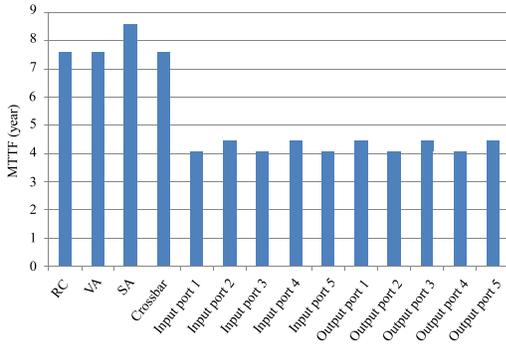


Fig. 5. MTTF of individual subblocks for EM case.

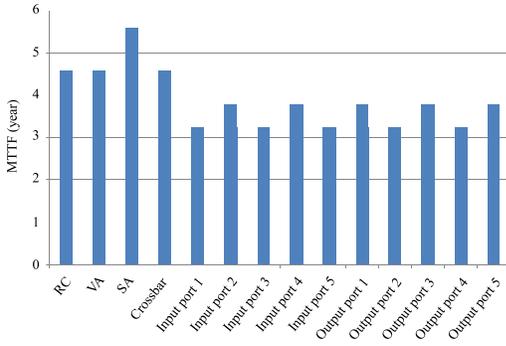


Fig. 6. MTTF of individual subblocks for TC case.

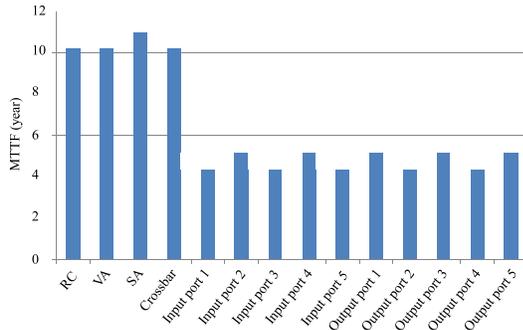


Fig. 7. MTTF of individual subblocks for SM case.

Library [24] within Cadence tools to synthesize and generate the floorplan of the router.

Fig.3 and Fig.4 show the results for TDDDB and NBTI cases as subblock vulnerabilities. We observe that RC and VA components contain the highest percentages of transistors with lifetime shorter than the selected threshold despite the fact that their area is smaller compared to for example the area of input registers. This can be explained by the fact that RC and VA components experience higher switching activities compared to the other router components, which in turn leads to higher temperatures. Note that this information could not be obtained with RAMP like reliability evaluation approaches.

Because for EM, TC, and SM the Monte Carlo algorithm is applied at subblock level (see discussion in Section III) we report the individual subblock MTTF in Fig.5, Fig.6, and Fig.7. We observe that input and output ports are the most vulnerable components to these three failure mechanisms.

## V. CONCLUSION

We proposed a new circuit level divide and conquer reliability evaluation methodology based on an MC algorithm. The proposed methodology can be very useful to circuit and system designers to predict circuit and system MTTF and to focus on cost effective design techniques targeted at specific parts of the design to improve its lifetime.

## ACKNOWLEDGMENT

This work was supported by the National Science Foundation (NSF) under grant CCF-1116022. Any opinions, findings, and conclusions or recommendations expressed herein are those of the authors and do not necessarily reflect the views of the NSF.

## REFERENCES

- [1] S. Borkar, "Designing reliable systems from unreliable components: the challenges of transistor variability and degradation," *IEEE Micro*, vol. 25, Nov. 2005.
- [2] J.H. Stathis, "Reliability limits for the gate insulator in CMOS technology," *IBM Journal of Research and Development*, vol. 46, 2002.
- [3] D.K. Schroder and J.A. Babcock, "Negative bias temperature instability: road to cross in deep submicron silicon semiconductor manufacturing," *Journal of Applied Physics*, vol. 94, July 2003.
- [4] S.M. Alam, C.L. Gan, D.E. Troxel, and C.V. Thompson, "Circuit-level reliability analysis of Cu interconnects," *ACM/IEEE ISQED*, 2004.
- [5] "Failure mechanisms and models for semiconductor devices," *JEDEC Publication JEP122E*, 2009.
- [6] J.B. Bernstein, M. Gurfinkel, X. Li, J. Walters, Y. Shapira, and M. Talmor, "Electronic circuit reliability modeling," *Microelectronics Reliability*, vol. 46, 2006.
- [7] M. White and J.B. Bernstein, "Microelectronics reliability: physics-of-failure based modeling and lifetime evaluation," *Jet Propulsion Laboratory*, NASA, 2008.
- [8] J. Srinivasan, S.V. Adve, P. Bose, J.A. Rivers, and C.K. Hu, "RAMP: a model for reliability aware microprocessor design," *IBM Research Report*, 2003.
- [9] J. Srinivasan, S.V. Adve, P. Bose, and J.A. Rivers, "Lifetime reliability: toward an architectural solution," *IEEE Micro*, vol. 25, 2005.
- [10] P. Ramachandran, S.V. Adve, P. Bose, J.A. Rivers, and J. Srinivasan, "Metrics for architecture-level lifetime reliability analysis," *IEEE ISPASS*, 2008.
- [11] J. Srinivasan, S.V. Adve, P. Bose, and J.A. Rivers, "Exploiting structural duplication for lifetime reliability enhancement," *IEEE ISCA*, 2005.
- [12] X. Li, B. Huang, J. Qin, X. Zhang, M. Talmor, Z. Gur, and J.B. Bernstein, "Deep submicron CMOS integrated circuit reliability simulation with SPICE," *ACM/IEEE ISQED*, 2005.
- [13] X. Li, J. Qin, B. Huang, X. Zhang, and J.B. Bernstein, "A new SPICE reliability simulation method for deep submicrometer CMOS VLSI circuits," *IEEE Trans. on Device and Materials Reliability*, June 2006.
- [14] X. Li, J. Qin, and J.B. Bernstein, "Compact modeling of MOSFET wearout mechanisms for circuit-reliability simulation," *IEEE Trans. on Device and Materials Reliability*, March 2008.
- [15] A.K. Coskun, T.S. Rosing, K. Mihic, G.D. Micheli, and Y. Leblebici, "Analysis and optimization of MPSoC reliability," *Journal of Low Power Electronics*, vol. 2, Apr. 2006.
- [16] J. Shin, V. Zyuban, Z. Hu, J. Rivers, and P. Bose, "A framework for architecture-level lifetime reliability modeling," *IEEE/IFIP DSN*, 2007.
- [17] Z. Gu, C. Zhu, L. Shang, and R.P. Dick, "Application-specific MPSoC reliability optimization," *IEEE TVLSI*, 2008.
- [18] J. Fang and S. Sapatnekar, "Scalable methods for analyzing the circuit failure probability due to gate oxide breakdown," *IEEE TVLSI*, 2011.
- [19] [www.cadence.com](http://www.cadence.com)
- [20] <http://lava.cs.virginia.edu/HotSpot>
- [21] <http://www.mathworks.com/products/matlab>
- [22] D. Fick, A. DeOrion, J. Hu, V. Bertacco, D. Blaauw, and D. Sylvester, "Vicis: a reliable network for unreliable silicon," *ACM/IEEE DAC*, 2009.
- [23] W.J. Dally and B. Towles, "Principles and practices of interconnection networks," *Morgan Kaufmann*, 2004.
- [24] <http://www.si2.org>