

# Reliability, Thermal, and Power Modeling and Optimization

Robert P. Dick  
EECS Department  
University of Michigan  
Ann Arbor, Michigan  
dickrp@eecs.umich.edu

**Abstract**—This tutorial provides an overview of challenges to designing and implementing reliable integrated circuits and systems, and suggests areas for future study. It illustrates some concepts in detail, explaining the challenges of appropriately considering the impact of temperature on reliability in fault-tolerant systems. Finally, it points out considerations that may influence adoption of reliability modeling and optimization techniques and stresses the importance of considering the most relevant fault processes during reliability modeling and optimization.

## I. SOURCES OF NEW RELIABILITY PROBLEMS

Unexpected system-level reliability problems are rooted in specification and design flaws, including errors caused by variations in the fabrication process or runtime environment, for it is ultimately the responsibility of the system designer to understand and work in the context of these variations. In order to simplify the design process and permit greater specialization, it is common for system designers to assume that real-world operating environments are well approximated by simplified or static models. When these assumptions are incorrect, unreliable systems result. When changes to applications or fabrication technologies invalidate old simplifications or support new simplifications, there are opportunities for new system-level reliability modeling and optimization techniques.

What influences on reliability have been changing? There has been a trend toward increasing power density (and variation in power density), device count, and number of mostly-independent processor cores. There has recently been a significant amount of work on system-level and architectural reliability enhancement techniques, much of which has considered multi-core systems [1], [2], [3], [4], [5]. There is also increasing interest in non-CMOS device technologies due to anticipated eventual physical and economic limitations for CMOS [6]. Applications are changing, as well; complex portable computers equipped with numerous sensors and wireless communication interfaces (i.e., smartphones) are already in common use [7]. These environment- and behavior-aware platforms are enabling new applications. Applications and operating systems are now reaching the level of sophistication required for near-continuous on-line adaptation of resource use to the requirements of users and environments. This adaptation has the potential to improve efficiency and reliability but it complicates reliability modeling and optimization. Other applications (e.g., wireless sensor networks) require deployment of inexpensive computers into harsh environments [8], [9].

Thanks to Lan Bai, David Bild, Pai Chou, Peter Dinda, Paul Ferno, Russ Joseph, Brett Meyer, Seth Prejean, Li Shang, Yun Xiang, Lide Zhang, and Lin Zhang for prior discussions of system-level reliability modeling and optimization that influenced this summary.

## II. SUMMARY OF RELIABILITY PROBLEM TYPES

We now briefly summarize classes of reliability problems to provide context for subsequent discussion of reliability modeling and optimization techniques. A system or component *fails* when it does not meet its specified requirements [10]. A system experiences a *fault* when it enters an incorrect state, e.g., as the result of the failure of a component within the system. A *fault-tolerant* system can continue to honor its specifications even in the presence of a fault, e.g., by using redundant components when a component fails.

**Specification and design:** Errors in specifying the requirements of a system, and in refining these requirements until they have a form, efficiency, and level of detail consistent with the chosen implementation technology, are presently responsible for a very substantial portion (most likely the vast majority) of system faults [11]. The relevant problems are hard and receiving attention from researchers working on language design [12], formal verification [13], software engineering [14], operating system and middleware design [15], and hardware synthesis [3].

**Permanent faults** most commonly result when physical structures within an integrated circuit wear out. Example wear processes follow: electromigration [16], [17], time-dependent dielectric breakdown [16], [18], stress migration [16], thermal cycling [19], and negative-bias temperature instability [20]. The rates of many dominant integrated circuit wear processes depend strongly on temperature. As a result, temperature-aware reliability models are required, especially for systems experiencing substantial run-time temperature variation. Imprecision in characterization of integrated circuit fabrication and wear state and limitations in current fault modeling techniques result in the use of stochastic models. Although there exist some fundamental limits on knowledge of precise wear state, the primary limitation is economic. It is likely that stochastic models operating with incomplete state information will see continued use in the future. However, their distributions can be narrowed by using improved wear models as well as test-time [21] and run-time characterization [22]. Spatial redundancy can be used to compensate for permanent faults, but temporal redundancy is insufficient.

**Transient and intermittent faults** are the result of short-term or occasional changes in system state that do not cause permanent damage to the physical structures comprising the system. Single-event upsets in CMOS-based systems are increasing in importance due to decreasing operating voltages and charge storage node capacitances. In particular, the failure rate due to impact by high-energy neutrons produced by cosmic rays interacting with the atmosphere are rapidly increasing in rate as a result of process scaling [23]. Many non-CMOS device technologies are susceptible to transient and intermittent faults,

e.g., random background offset charge effects in single-electron tunneling transistors [24], [25].

Transient faults may also be caused by noise that was not predicted due to incomplete models of interaction among circuit components as a result of capacitive or inductive coupling [26] or voltage fluctuation in the power delivery network [27]. Difficult transient fault problems have a few themes. Their timing may be difficult to predict due to dependence on physical phenomena that cannot be controlled and are not practical to measure until after a fault has occurred, e.g., the impact of subatomic particles. In some cases, the faults are theoretically possible to predict, but the required modeling complexity is so high that the problem is computationally intractable, e.g., some types of crosstalk. Both temporal and spatial redundancy can be used to compensate for transient faults. Whether temporal redundancy is sufficient to compensate for intermittent faults depends on the durations between state changes.

**Influence of operating and fabrication parameter variation:**

Simplified (e.g., deterministic or static) models of operating parameters can be used to accelerate the design process. However, if parameters such as such as combinational logic delay and wear rate depend on parameters that vary (e.g., fabrication process variation or run-time temperature and current density variation), the cost of rapid design can be poor reliability or the need to use large guard bands, thereby degrading other metrics such as performance or energy consumption. Researchers have considered increasing the accuracy (and complexity) of permanent fault models used during system-level design. This requires either that stochastic models be used to capture the effects of uncertainty and/or that test-time or run-time measurements be used to (indirectly) capture variations of interest.

**III. TEMPERATURE-AWARE RELIABILITY MODELING AND OPTIMIZATION**

Modeling and optimization are tightly coupled in temperature-aware design of fault-tolerant systems. Adaptive changes in system state caused by or related to faults impact future fault probabilities. As a consequence, it is necessary for information to flow among fault modeling and optimization infrastructures. This can be easily illustrated by considering a fault tolerant architecture that adapts to permanent faults caused by temperature-dependent fault processes. Each adaptation influences performance, power consumption, temperature, and reliability in difficult to predict ways.

**III.A. Interactions Among Power Consumption, Temperature, and Reliability**

Figure 1 illustrates the relationships among power consumption, temperature, reliability, and process variation. Dashed boxes indicate the processes by which one attribute affects others. Dynamic and leakage power consumption are at the roots of many problems. Rapid changes in dynamic power consumption can result in transient voltage fluctuations as a result of the distributed inductance and resistance in off-chip and on-chip power delivery networks. These  $dI/dt$  effects, indicated by dashed boxes in Figure 1, can change logic combinational path delays, resulting in timing violations. This can cause transient faults or require reduction in processor frequency, i.e., guard-banding. High dynamic power consumption implies high current density, resulting in accelerated wear and permanent faults due to processes such as electromigration. Dynamic power consumption produces heat, which increases temperature. The effects of leakage

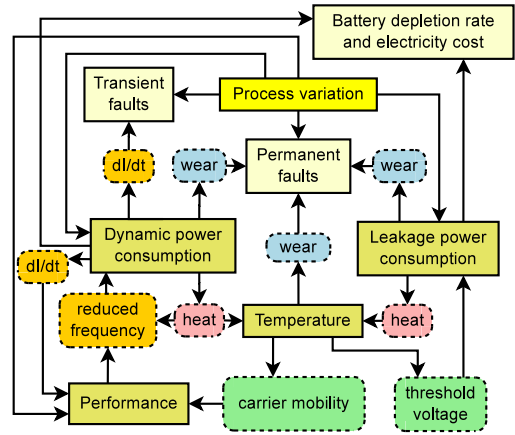


Figure 1. Overview of interactions among power consumption, temperature, and reliability (based on existing figure [28]).

power consumption are similar, with the exception of  $dI/dt$ -related problems. In general, leakage power variation is smaller than dynamic power variation. Both dynamic and leakage power consumption decrease battery lifespans in portable devices and increase server electric bills.

Increased power consumption increase temperature. Thermal profiles depend on the temporal and spatial distribution of power as well as the system cooling and packaging solutions. High temperature decreases threshold voltage, resulting in increased subthreshold leakage power consumption but improving transistor performance. In addition, it decreases charge-carrier mobility, decreasing transistor and interconnect performance. Finally, temperature has a large impact on many permanent fault processes.

Process variation influences transient fault rate via changes to critical timing paths; permanent fault rate via changes to numerous parameters such as wire and insulator dimensions; leakage power consumption via changes in dopant concentration; and dynamic power consumption.

As Figure 1 illustrates, the relationships among power, temperature, process variation, and reliability are complex. Controlling power, temperature, and reliability requires modeling across multiple design levels and consideration and control of runtime behavior. In summary, reliability models depend on performance [29], power consumption [30], and thermal models [31], [32], each of which is difficult to build and often computationally expensive to evaluate.

**III.B. Degrees of Modeling and Optimization Complexity**

We now describe the various degrees of reliability modeling complexity and accuracy that might be considered for use in the design and run-time management of a CMOS integrated circuit. We will start by contrasting the advantages and disadvantages of a number of potential levels of modeling sophistication.

- 1) **Ignore the possibility of faults:** This case is clearly the simplest and may be reasonable if the application being considered has weak reliability requirements.
- 2) **Static fault rate without redundancy:** In this case, it is assumed that the fault rates for components within the system do not change as a function of time. This assumption is reasonable if the effects of early lifetime faults can be neglected (perhaps due to burn-in) and the desired lifetime of the system is so short that wear will not have an opportunity to increase the fault rate.

- 3) **Wear modeling with static parameters:** In this case, the fault probabilities of devices in the system increase with time as a result of wear, but the influence of environmental parameters (e.g., temperature) on wear rate are ignored. This level of modeling detail may be appropriate for systems used in well-controlled environments. However, note that even if the average conditions of the environment during operation are known, the degree of variation from those conditions can significantly influence wear and therefore lifetime estimates. Consider a CMOS integrated circuit with a temperature that will vary around a known mean. A number of the dominant sources of wear in CMOS integrated circuits have associated wear rates ( $r_w$ ) that depend on temperature ( $T$ ) as follows:

$$r_w = k_1 e^{(-k_2/T)}, \quad (1)$$

where  $k_1$  and  $k_2$  are independent of temperature. Due to this non-linear dependence, accurate wear modeling requires knowledge of operating temperature distribution.

- 4) **Wear modeling with environment-dependent parameters:** At this level of modeling complexity, the specific values of parameters influencing fault rate and wear rate (e.g., temperature, current density, gate oxide thickness, or altitude) are considered during the design process. Moreover, these parameters may vary with time, from integrated circuit to integrated circuit, or even among devices within the same integrated circuit. It may first seem that knowing the distribution for each parameter would be sufficient even for wear rates with non-linear dependence on the parameters of interest (e.g., Equation 1). However, due to memory effects, some fault processes require time series for accurate wear modeling, e.g., thermal cycling [19] and negative-bias temperature instability [20]. As a result, accurate modeling requires prediction of parameter fluctuation. This leaves the designer in a quandary: the temporal variations in reliability model parameters have short timescales (e.g., seconds) while the timescales over which there are substantial changes in wear state are long (e.g., months). A naïve approach to this estimation problem would require extraordinary simulation time.

- 5) **Advanced wear modeling with fault detection and on-line adaptation:** It is at this point where we first consider non-trivial fault-tolerance techniques. Thus far, we have considered modeling the dependence of fault rate on environmental conditions and time. However, the system behavior was static. If, instead, the system adapts to faults dynamically, e.g., by changing the tasks assigned to components or transmissions associated with interconnects, the interaction between reliability models and optimization techniques becomes more complex. The chosen fault tolerance policy influences subsequent environmental conditions and therefore subsequent fault probabilities.

For example, consider a system containing four homogeneous processors, one idle and three active. If an active processor fails, and it is necessary for the system to preserve computational throughput, it would be possible to either increase the load on the remaining two active processors or leave their loads constant but activate the idle processor. Determining which decision would lead to more favorable reliability conditions would require understanding its implications for parameters impacting reliability, e.g., temperature. That would

require understanding the power management implications of the decision and evaluating power and thermal models. Note that the consequences of each decision must be evaluated, and that doing so may require other similar decisions leading up to eventual system failure. In short, temperature-aware reliability modeling of fault-tolerant systems is particularly challenging, especially in the presence of power management techniques.

- 6) **Advanced wear modeling with fault detection, on-line adaptation, and on-line uncontrolled state estimation:** The previously described analysis applies to systems in which all fault tolerance policies are pre-planned based on the predicted wear and fault patterns. However, additional run-time information may be available about fault locations and times (via fault detection [33]) and wear states (via on-line testing [22]). Fault tolerance policies might therefore be adaptively optimized on-line. Although the estimation necessary to support such a technique might at first seem computationally intractable, if the time spans over which substantial wear occurs are measured in months, not minutes, the overhead might be tolerable.

Note that all of these modeling and optimization problems are further complicated by significant variation in fault-relevant parameters among components.

This section has indicated some of the reasons why accurate reliability modeling can be challenging for systems with variable thermal profiles, particularly when the models are used to optimize fault tolerance policies.

#### IV. CONSIDERATIONS WHEN SELECTING RELIABILITY MODELING AND OPTIMIZATION TECHNIQUES FOR USE

The previous section might leave readers with the impression that developing increasingly detailed and accurate reliability models constitutes progress. However, the characterization and computational complexities of following this path to its logical conclusion would be unacceptable. It is therefore important to first determine which particular fault processes are relevant for the technology, environment, and system requirements under consideration. We now list a number of questions appropriate for reliability modeling or optimization techniques being considered for use.

- 1) Are the fault processes considered by the work dominant for the planned implementation technology?
- 2) Do the proposed reliability enhancement techniques operate at the most appropriate design levels, e.g., system through physical level for guaranteeing hard task deadlines and cell-level for tolerating fabrication errors in array structures.
- 3) Are the assumptions in the work consistent with the expected operating environment parameter values, e.g., temperature, current density, and altitude?
- 4) What is the net benefit of the proposed technique, considering increased design time, debugging complexity, and testing time, as well as changes to production processes and team organization?

The last question is critically important. Design and testing teams are under constant, extreme time pressure. Changes to the design and testing processes that increase complexity and time divert resources from other reliability enhancement efforts.

#### V. IDENTIFYING DOMINANT FAULT PROCESSES

Commonly, only the system designers have enough information to determine which fault processes will be dominant. In fact, it

is common for such information to be unknown even to system designers. This is especially true in the case of new ad hoc system designs that are not closely related to the designer's prior experience or case studies in the reliability literature. For example, consider the wireless sensor network deployed in the "Bird's Nest" Olympic Stadium in Beijing [34]. One might expect this system to fail as a result of battery energy depletion and component wear, focusing one's modeling efforts on these potential problems. In fact, a rain storm and leaky roof damaged a large portion of the system. We have heard many similar reliability related anecdotes from other system designers, many of which remain unpublished. It is important for system designers to carefully determine which fault processes to consider and model during the design process.

I would like to conclude this section with a plea to system designers. Whenever deploying a system in which reliability might be a concern, recording and publishing as much information about the faults and their causes as possible will enable reliability researchers to appropriately prioritize and focus their work. I would be happy to build a website for this purpose. Please contact the author if you are interested in having reliability case studies or anecdotes posted.

## VI. CONCLUSIONS

Research opportunities in reliable system design track technology and application changes that invalidate assumptions used to simplify modeling, or make new assumptions valid. We have pointed out a few ongoing trends, explained the challenges of temperature-dependent reliability modeling and optimization, indicated considerations influencing whether adopting a particular reliability modeling or enhancement technique is beneficial, and pointed out the importance of carefully selecting fault processes for consideration.

## REFERENCES

- [1] J. Srinivasan, S. V. Adve, P. Bose, and J. A. Rivers, "Lifetime reliability: toward an architectural solution," *IEEE Micro*, vol. 25, no. 3, pp. 70–80, May 2005.
- [2] A. K. Coskun, T. S. Rosing, K. Mihic, G. D. Micheli, and Y. Leblebici, "Analysis and optimization of MPSoC reliability," *J. Low Power Electronics*, vol. 2, no. 1, pp. 56–69, Apr. 2006.
- [3] M. Glaß, M. Lukasiewicz, T. Streichert, C. Haubelt, and J. Teich, "Reliability-aware system synthesis," in *Proc. Design, Automation & Test in Europe Conf.*, Apr. 2007.
- [4] B. H. Meyer and D. E. Thomas, "Reliability and cost in NoC-based MPSoCs," in *Proc. Semiconductor Research Corporation TECHCON*, Nov. 2008.
- [5] Y. Xiang, T. Chantem, R. P. Dick, S. X. Hu, and L. Shang, "System-level reliability modeling for MPSoCs," in *Proc. Int. Conf. Hardware/Software Codesign and System Synthesis*, Oct. 2010, to appear.
- [6] "International Technology Roadmap for Semiconductors," 2009, <http://www.itrs.net/links/2009ITRS/Home2009.htm>.
- [7] H. Falaki, R. Mahajan, S. Kandula, D. Lymberopoulos, R. Govindan, and D. Estrin, "Diversity in smartphone usage," in *Proc. Int. Conf. on Mobile Systems, Applications, and Services*, June 2010.
- [8] L. B. Ruiz, I. G. Siqueira, L. B. e Oliveira, H. C. Wong, J. M. S. Nogueira, A. A. F. Loureiro, and B. Horizonte, "Fault management in event-driven wireless sensor networks," in *Proc. Int. Symp. on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, Oct. 2004.
- [9] F. Koushanfar, M. Potkonjak, and A. Sangiovanni-Vincentelli, "On-line fault detection of sensor measurements," in *Proc. Int. Conf. on Sensors*, Oct. 2003, pp. 974–980.
- [10] D. P. Siewiorek and R. S. Swarz, *Reliable Computer Systems – Design and Evaluation*, 3rd ed. A. K. Peters, 1998.
- [11] H. A. Rahman, K. Beznosov, and J. R. Martí, "Identification of sources of failures and their propagation in critical infrastructures from 12 years of public failure reports," in *Proc. Int. Conf. on Critical Infrastructures*, Sept. 2006.
- [12] L. Lamport, *Specifying Systems: The TLA+ Language and Tools for Hardware and Software Engineers*. Addison-Wesley, MA, 2002.
- [13] G. Klein, et al., "seL4: formal verification of an OS kernel," in *Proc. Symp. on Operating Systems Principles*, Oct. 2009, pp. 207–220.
- [14] T. Ball, "The verified software challenge: A call for a holistic approach to reliability," *Verified Software: Theories, Tools, Experiments – Lecture Notes in Computer Science*, vol. 4171, pp. 42–48, 2008.
- [15] A. S. Tanenbaum, J. N. Herder, and H. Bos, "Can we make operating systems reliable and secure?" *IEEE Computer*, vol. 39, no. 5, pp. 44–51, May 2006.
- [16] "Failure mechanisms and models for semiconductor devices," Joint Electron Device Engineering Council, Tech. Rep., Aug. 2003, JEP 122-B.
- [17] J. R. Black, "Electromigration—a brief survey and some recent results," *IEEE Trans. Electron Devices*, vol. 16, no. 4, pp. 338–347, Apr. 1969.
- [18] J. Srinivasan, S. V. Adve, P. Bose, and J. A. Rivers, "Exploiting structural duplication for lifetime reliability enhancement," in *Proc. Int. Symp. Computer Architecture*, June 2005, pp. 520–531.
- [19] Y. Zhang, M. L. Dunn, K. Gall, J. W. Elam, and S. M. George, "Suppression of inelastic deformation of nanocoated thin film microstructures," *J. of Applied Physics*, vol. 95, no. 12, pp. 8216–8225, June 2004.
- [20] M. Alam and S. Mahapatra, "A comprehensive model of PMOS NBTI degradation," *Microelectronics Reliability*, vol. 45, no. 1, pp. 71–81, Jan. 2005.
- [21] K. S. Kim, S. Mitra, and P. G. Ryan, "Delay defect characteristics and testing strategies," *IEEE Design & Test of Computers*, vol. 20, no. 5, pp. 8–16, Sept. 2003.
- [22] K. Constantinides, O. Mutlu, T. Austin, and V. Bertacco, "A flexible software-based framework for online detection of hardware defects," *IEEE Trans. on Computers*, vol. 58, no. 8, pp. 1063–1079, Aug. 2009.
- [23] P. E. Dodd and L. W. Massengill, "Basic mechanisms and modeling of single-event upset in digital microelectronics," *IEEE Trans. on Nuclear Science*, vol. 50, no. 3, pp. 583–602, June 2003.
- [24] M. Furlan and S. V. Lotkhov, "Electrometry on charge traps with a single-electron transistor," *Physica Rev. B*, vol. 67, p. 205313, 2003.
- [25] H. Wolf, F. J. Ahlers, J. Niemeyer, H. Scherer, T. Weimann, A. B. Zorin, V. A. Krupenin, S. V. Lotkhov, and D. E. Presnov, "Investigation of the offset charge noise in single electron tunneling devices," *Trans. on Instrumentation and Measurement*, vol. 46, no. 2, pp. 303–306, Apr. 1997.
- [26] K. Agarwal, D. Sylvester, and D. Blaauw, "Modeling and analysis of crosstalk noise in coupled RLC interconnects," *IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems*, vol. 25, no. 5, pp. 892–901, May 2006.
- [27] R. Joseph, Z. Hu, and M. Martonosi, "Wavelet analysis for microprocessor design: Experiences with wavelet-based dl/dt characterization," in *Proc. Int. Symp. High-Performance Computer Architecture*, Feb. 2004.
- [28] D. Brooks, R. P. Dick, R. Joseph, and L. Shang, "Power, thermal, and reliability modeling in nanometer-scale microprocessors," *IEEE Micro*, pp. 49–62, May 2007.
- [29] T. L. Pillage, R. A. Rohrer, and C. Visweswariah, *Electronic Circuit System Simulation*. McGraw-Hill Book Company, NY, 1995.
- [30] D. Brooks, V. Tiwari, and M. Martonosi, "Wattch: A framework for architectural-level power analysis and optimizations," in *Proc. Int. Symp. Computer Architecture*, June 2000, pp. 83–94.
- [31] W. Huang, S. Ghosh, S. Velusamy, K. Sankaranarayanan, K. Skadron, and M. Stan, "HotSpot: a compact thermal modeling methodology for early-stage VLSI design," *IEEE Trans. VLSI Systems*, vol. 14, no. 5, pp. 501–524, May 2006.
- [32] S. V. J. Narumanchi, J. Y. Murthy, and C. H. Amon, "Boltzmann transport equation-based thermal modeling approaches for hotspots in microelectronics," *Heat Mass Transfer*, vol. 42, pp. 478–491, 2006.
- [33] M. A. Goma and T. N. Vijaykumar, "Opportunistic transient-fault detection," in *Proc. Int. Symp. Computer Architecture*, June 2005, pp. 172–183.
- [34] "Personal communication with Professor Lin Zhang (Tsinghua University, Beijing)," 2009.