Reliability Modeling and Management in Dynamic Microprocessor-Based Systems

Eric Karl, David Blaauw, Dennis Sylvester, Trevor Mudge
University of Michigan
1301 Beal Ave.
Ann Arbor, MI 48105
{ekarl,blaauw,dennis,tnm}@eecs.umich.edu

ABSTRACT

Reliability failure mechanisms, such as time dependent dielectric breakdown, electromigration, and thermal cycling have become a key concern in processor design. The traditional approach to reliability qualification assumes that the processor will operate at maximum performance continuously under worst case voltage and temperature conditions. However, the typical processor spends a very small fraction of its operational time at maximum voltage and temperature. In this paper, we show how this results in a reliability "slack" that can be leveraged to provide increased performance during periods of peak processor demand. We develop a novel, real time reliability model based on workload driven conditions. We then propose a new dynamic reliability management (DRM) scheme that results in 20-35% performance improvement during periods of peak computational demand while ensuring the required reliability lifetime.

Categories and Subject Descriptors

B.8.0 [**Hardware**]: Performance and Reliability – *General*

General Terms: Management, Reliability, Design, Performance

Keywords: Dynamic Reliability Management, Modeling, Oxide Breakdown, Electromigration, Thermal Cycling

1. Introduction

Traditional reliability qualification techniques project lifetime based upon failure mechanisms, such as dielectric breakdown or electromigration, assessed at constant worst-case operating conditions. While this approach is an accepted method of ensuring reliability, the limits it places on supply voltage leaves a significant and increasing reliability margin between processor performance at worst-case conditions and at typical conditions. Varying environmental conditions linked to portable products combine with power reduction techniques to exacerbate the limitation of conventional worst-case qualification methodology.

* This work was supported in part by the MARCO/DARPA Gigascale Systems Research Center. Their support is gratefully acknowledged

DAC 2006, July 24–28, 2006, San Francisco, California, USA.

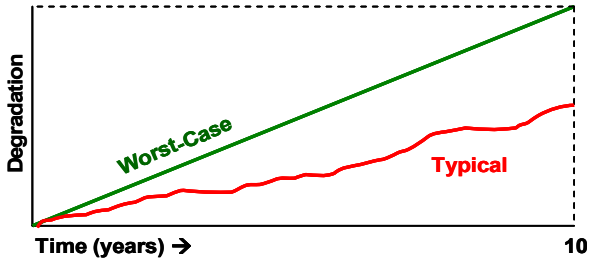

Figure 1. Reliability degradation over time.

In this paper, we propose the use of so-called *dynamic reliability management* (DRM), where real-time workloads and thermal information provides accurate inputs to reliability models for projecting the degradation caused by various failure mechanisms. The projected failure probability is used to control the maximum voltage in the dynamic voltage scaling (DVS) algorithm.

The concept of DRM is motivated in Figure 1. The line labeled *worst-case* shows accumulated damage due to a failure mechanism, such as oxide breakdown, over a 10-year time span under worst-case conditions. The accumulated damage from typical processor usage traces is shown by the line labeled *typical* and results in much lower final damage after 10 years. Hence, the failure probability for this typical usage is well below that of the specification, and the maximum allowed operating voltage is unnecessarily constrained, resulting in a loss of potential performance.

DRM was first introduced in [1] using a sum of failure rates method while considering multiple reliability mechanisms. Lu and co-authors [2] analyzed electromigration effects and suggested dynamic thermal management. However, both approaches focus on short time scales, and more critically do not explore the control system required to obtain performance gains.

In this paper, we implement a DRM system using physics-based failure models, expressed as incremental damage mechanisms using a linear cumulative damage model referred to as Miner's rule [3]. The performance impact of DRM in systems with DVS control techniques is analyzed with a focus on macro-level user-collected processor usage profiles rather than benchmark applications. The DRM system sets a maximum supply voltage based on the degradation characteristics modeled during operation, and exceeds the nominal supply voltage when possible while meeting the required reliability constraints. With the proposed implementation of a specific control DRM algorithm, this work demonstrates and quantifies the potential performance improvements of DRM for the first time.

2. Reliability Modeling

In order to implement a real-time dynamic reliability management scheme, we require models that can comprehend dynamic stress

behavior with minimal computational expense. High level compact models for oxide breakdown, electromigration and thermal cycling are addressed in the following sections. In our proposed approach, we cast all reliability models such that they express wear-out in terms of accumulated damage, or fraction of lifetime consumed. This allows simple projections of the failure rate at the desired lifetime and also allows the use of degradation dependent models for each reliability mechanism, a capability that is lost when dealing directly with probabilities.

2.1 Oxide Breakdown

Oxide breakdown, or dielectric breakdown, is a degradation mechanism that results in a low-impedance path through an insulating or dielectric barrier. Each tunneling charge has a probability of creating a defect when passing through the oxide. When a "critical" defect density is reached, there is a high probability that a low-impedance defect path exists in the oxide and a runaway current path through the insulating film will develop. The relationship between charge tunneling through the oxide and the defect density is expressed below in Equation 1, where N_{BD} is the defect density, P_{DG} is the probability of defect generation, and I_{tunnel} is the tunneling current [4].

$$N_{BD} \approx \int^{t} P_{DG}(V,T) I_{tunnel}(V,T) dt \tag{1}$$

A simple simulation methodology for estimating the critical defect density required for a low-impedance defect path was originally developed by Degraeve [5]. The PDF generated from Degraeve's percolation model is fit to a Weibull distribution and used to calculate the probability of entering the onset of defect-induced oxide breakdown for an individual device.

The tunneling current through a gate oxide is calculated using BSIM4 model equations. In this work, published defect generation relationships from an IBM technology node are used in the simulations [4]. This oxide breakdown model allows an incremental summation of defect density at variable supply voltage and temperature stress conditions. This closed-form, high-level oxide breakdown model is therefore ideal for a real-time DRM system considering dynamic stress conditions.

2.2 Electromigration

Black's formula [6] is a well-known relationship between the mean time to failure of an interconnect and the current density, temperature and physical dimensions of the wire as shown in Equation 2:

$$MTF = AJ^{-n} e^{(E_a / kT)} \tag{2}$$

The term A is a function of the materials and geometric structure of the wire and generally increases with both width and thickness of the structure. J is the current density, E_a is an activation energy for atom transport, k is the Boltzmann constant, and T is temperature.

Miner's rule [3] (linear cumulative damage) is used to estimate the EM lifetime of a conductor by adding the percentage of lifetime consumed during each period of varying stress.

$$\sigma_{life} = \sum_{t} \frac{MTF_{ref}}{MTF(J,T)} \Delta t \tag{3}$$

Equation 3 summarizes the adaptation to Black's formula that allows variable stress conditions to be expressed as a percentage of lifetime, σ_{life}. MTF_{ref} is a reference value that would be characterized at worst-case conditions for the design, and $MTF(J,T)$ is an MTF calculation that is performed with varying current density and temperature averaged over a time window, Δt.

A Weibull distribution is used to convert the percentage of lifetime figure (σ_{life}) to a probability of failure. A self-consistent temperature is calculated for wires in each layer considering the thermal effects of wire resistance and the current density at a given supply voltage [7]. The EM modeling in this work is limited to unidirectional currents (power/ground network) due to the greatly reduced experimental observation of failures in wires with bi-directional current [8].

2.3 Thermal Cycling

Thermal cycling is a mechanical stress mechanism that is manifested in many locations on an integrated circuit including solder connections and thin-film interfaces. Blish [9] related the number of cycles of thermal fatigue of various materials on a silicon die to the thermal swing via the well-known Coffin-Manson Equation [10]:

$$N_{cyc} \approx (\Delta T)^{-m} \tag{6}$$

$$N_{cyc} = \sum_{t} \frac{(\Delta T_{ref})^{-m}}{(\Delta T)^{-m}} \tag{7}$$

The number of cycles, N_{cyc}, before breakdown is related to the thermal swing ΔT and a coefficient depending upon the materials involved. In this work, we consider thin-film cracking damage with a Coffin-Manson exponent of 8.4 [9] and again use Miner's rule to express Equation 6 as a percentage of lifetime, seen in Equation 7. ΔT_{ref} is a reference thermal swing that the system is designed to withstand, and ΔT represents a measured thermal swing that may differ from the reference. Equation 7 is used to sum the damage caused by thermal swings to express their contribution in terms of equivalent cycles of a larger swing ΔT_{ref}.

3. System-Level Modeling

In order to derive the total projected system failure probability at the end of lifetime, t_{life}, we perform two tasks: 1) Based on the existing stress history and accumulated damage for a particular failure mechanism and device at the current time t_1, we project the probability of failure for that failure mechanism and device at t_{life}. 2) We combine the failure probabilities for all considered failure mechanisms and devices. We discuss each step below.

In the system, the damage (D_1) at time t_1 is extrapolated to the damage (D_{life}) at time t_{life} based on history information about the rate of damage up to time t_1 using the following simple linear extrapolation:

$$D_{life} = D_1 \frac{t_{life}}{t_1} \tag{8}$$

Equation 8 accounts for environmental conditions and workload history intrinsically, providing a lifetime projection that is tailored to the exact stress conditions experienced on the chip. Given the projected accumulated damage at t_{life}, the probability of failure of an individual structure is then calculated using a cumulative distribution function for the relevant reliability mechanism characterized for the given process technology.

Individual device reliability projections are used to compute a chip-level reliability projection across all devices and all failure mechanisms using the follow expression:

$$(1-P_{ox}) = \prod_{b}\prod_{d}(1-P'_{ox-b})$$
$$(1-P_{EM}) = \prod_{b}\prod_{l}\prod_{n}(1-P'_{EM-b-l}) \quad (9)$$
$$(1-P_{cyc}) = \prod_{b}(1-P'_{cyc-b})$$

P_{ox} is the probability of oxide failure, P_{EM} is the probability of an electromigration failure, and P_{cyc} is the probability of a thermal cycling failure. Oxide breakdown failure probability is calculated based on the number of devices per functional unit (d_b) with a specific failure rate for a device from each individual functional unit ($P'_{ox\text{-}b}$). Electromigration failure rate is projected from the individual failure rate ($P'_{EM\text{-}b\text{-}l}$) across the number of wires (n), in each layer (l) and block (b). Thermal cycling failure is calculated as a component from each block (b). Since the primary parameters that cause correlation between these failure mechanisms (voltage, temperature) are directly used in the calculation of accumulated stress over the simulated timespan, the correlation between these failure mechanisms is naturally considered. This observation allows the simple calculation of a total chip failure rate using probabilistic independence, which is used to drive the control algorithm in Section 4.

4. DRM System

Dynamic Voltage Scaling (DVS) is an ideal actuator for managing reliability concerns, since oxide breakdown and electromigration are both strongly voltage dependent and reductions in supply voltage greatly reduce the effect of these wear-out mechanisms. Thermal cycling damage can be indirectly limited by capping the maximum supply voltage, which limits the maximum temperature. Chip failure probability information generated from the models in Section 2-3 are used in a control algorithm to set a maximum allowable voltage for the DVS voltage assignment step.

A PID-based control algorithm is proposed as the key mechanism to provide maximum improvement in peak circuit performance when necessary, without affecting steady-state performance or compromising reliability. Equation 11 describes the behavior of a PID control system, where *e(t)* is an error signal, and *v(t)* is the output being controlled:

$$v(t) = v(t-1) + P\left[e(t) + R\int e(t)dt + D\frac{\partial e(t)}{\partial t}\right] \quad (11)$$

P is the proportional gain, *R* is the reset or integral gain, and *D* is the derivative gain. The proportional gain governs the response time of the controller, integral gain corrects for offset, and the derivative gain limits overshoot in the allowable voltage. In the proposed DRM system, *e(t)* is the probability of system failure projected to the lifetime, t_{life}, and *v(t)* is the maximum voltage available to the DVS algorithm.

Tuning the PID algorithm is dependent upon the desired response time and the length of time to correct offset issues. Overshoot, or selecting a maximum voltage that is too large, leads to a number of negative side-effects in a DVS system. To compensate for the excessive amount of wear-out damage, the algorithm will reduce the clock frequency below nominal which could limit performance in subsequent time periods. With poor proportional gain tuning, feedback instability leads to oscillation between high and low voltages. Setting a relatively high derivative gain (on the order of proportional gain) delivered good control performance in the proposed DRM system.

5. Results and Discussion

Workload data from several desktop computers was collected over several months to provide realistic processor utilization information with a wide-range of system behavior. A processor layout similar to the Alpha 21264 (16mm x 16mm) is used with process parameters based on 130nm industrial models. Initial power estimates are generated by Wattch and used with the processor utilization data collected to generate workload-based voltage/frequency/power traces. The PID controller assigns voltage/frequency pairs based upon the requested performance and the reliability state of the system. HotSpot 2.0 is used to calculate temperatures for each functional unit in the design using power numbers adjusted according to the selected supply voltage.

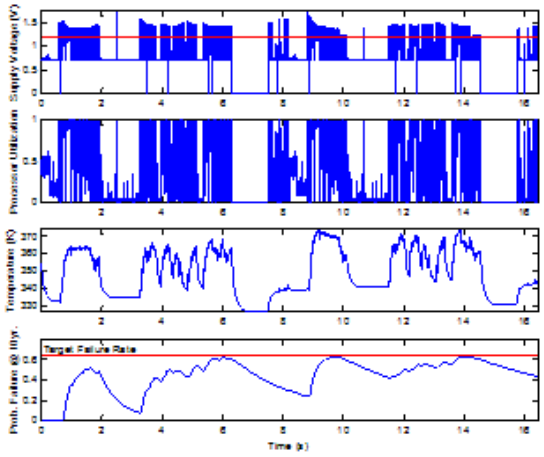

Figure 3. DRM Operation for Workload C630 (16 seconds)

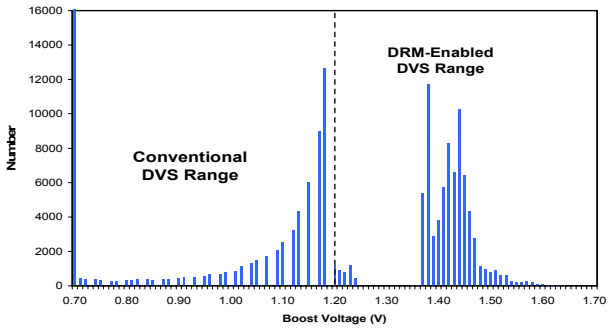

Figure 4. Detailed DRM-DVS voltage distribution.

The PID controller updates the maximum voltage every 50µs in the short time limit simulations in Figures 3-4 and every hour in lifetime simulations in Figure 5. A figure of merit to quantify the performance gains available with DRM that will be used throughout the results discussion is "peak performance improvement". This figure is a measure of the improvement in attainable frequency (%) during periods of *peak CPU demand*. This is a measure of how well DRM provides additional performance when it is needed most by the user.

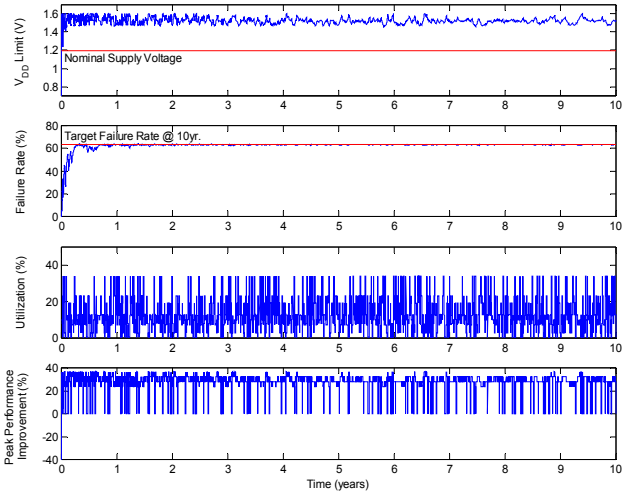

Figure 5. 10 year reliability simulation.

The traces in Figure 3 are of a DRM simulation spanning 16 seconds, allowing a detailed look at the interplay between the voltage, workload, temperature and projected failure rate at the 10 year lifetime. The histogram in Figure 4 shows the corresponding frequency of voltage assignments. The plot is bimodal since all tasks that require peak performance are executed at the maximum supply voltage allowed by the PID controller. Although there is no voltage limit upon the system, Figure 4 shows that an upper bound on voltage could be placed at 1.5V to accommodate power distribution or voltage regulator limitations with minimal impact.

Figure 5 displays the 10-year performance of the DRM control algorithm over a randomized selection of 10 representative 1-hour workloads collected from an actual desktop machine. The V_{DD} limit graph in each section of the plot represents the upper limit placed upon the DVS algorithm by the DRM mechanism, not the actual voltage during the entire trace. Peak performance gains over the 10-year workloads ranged from 20-35% compared to a nominal DVS controlled system. Specifically, the profile in Figure 5 shows a 26.7% peak performance improvement. In Figure 6, the peak performance improvement is plotted vs. the workload activity factor. The workload profile for this plot is constructed with oscillations between 0% and 100% utilization at a period of 5 min. with a variable duty cycle equaling the activity factor. For inactive systems, the voltage may be boosted dramatically above the nominal voltage, delivering a maximum of 34% improvement during periods of peak CPU demand. As the activity factor of the workload approaches 1.0, the peak performance gains approach 12.5%. At this point, the benefits of DRM are derived not from periods of low voltage stress in the workload, but from a lower operating temperature than specified in the worst-case reliability corner. The simulated chip is designed to operate at nominal voltage (1.2V) and a maximum on-die temperature of 125°C for 10 years. With an ambient temperature of 60°C, a 12.5% performance improvement is possible even with an extremely pessimistic workload. This plot demonstrates 15-20% peak performance improvement for workload profiles below 100% activity. Collected usage profiles indicate typical activity between 10-15% for which peak performance improvement is around 25%.

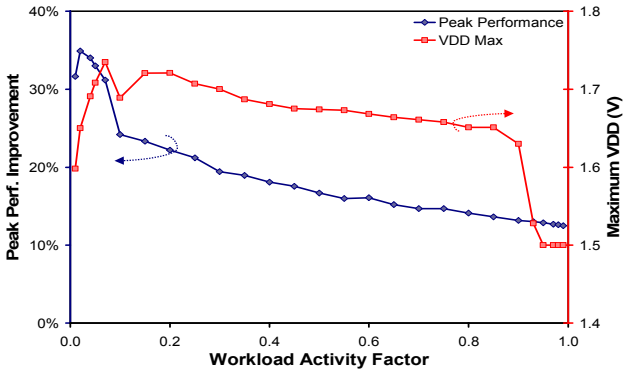

Figure 6. Peak Performance vs. Workload Activity

6. Conclusions

A detailed implementation of dynamic reliability management is presented for the first time, including a model for failure rate prediction under three common failure mechanisms and a PID-based control system that balances increased throughput in peak-demand periods with the remaining reliability lifetime. Workload and processor utilization information collected for typical users is used to quantify achievable gains in peak processor performance. On-chip real-time reliability monitoring allows supply voltages to be boosted beyond nominal values set during worst-case qualification, enhancing responsiveness during periods of critical computational demand. We observe typical peak performance gains of 20-35% over a variety of real-world workloads and usage profiles, without exceeding the specified lifetime budget.

7. References

[1] J. Srinivasan, *et al*, "The case for lifetime reliability-aware microprocessors." *Int. Symp. Computer Architecture,* 2004.

[2] Z. Lu, *et. al*, "Interconnect lifetime prediction under dynamic stress for reliability-aware design." *Int. Conf. on Computer-Aided Design,* 2004.

[3] M. A. Miner, "Cumulative damage in fatigue," *Journal of App. Mech.* **67**,1945.

[4] J. H. Stathis, "Physical and Predictive Models of Ultra Thin Oxide Reliability in CMOS Devices and Circuits," *IEEE Trans. Dev. & Mater. Reliabil.* **1,** 2001.

[5] R. Degraeve, *et al*, "A consistent model for intrinsic breakdown in ultra-thin oxides," *Int. Electron Devices Meeting*, p. 866, 1995.

[6] J. R. Black, "Electromigration failure modes for aluminum metallization in semiconductor devices," *Proc.of IEEE* **57**, 1969.

[7] M. Casu, *et al*, "An electromigration and thermal model of power wires for *a priori* high-level reliability prediction," *IEEE Tran. On VLSI Systems* **12,** 2004.

[8] A. Chandrakasan, ed., *Design of High-Performance Microprocessor Circuits*., IEEE Press, 2001.

[9] R. C. Blish, II, "Thermal cycling and thermal shock failure rate modeling," *Proc. Int. Reliability Physics Symp.,* 1997.

[10] Nelson, Wayne, *Accelerated Testing: Statistical Models, Test Plans and Data Analyses*, John Wiley & Sons, Inc., New York, 1990.

[11] K. Skadron, *et al*, "HotSpot: Techniques for modeling thermal effects at the processor-architecture level," *Int. Workshop on THERMal Investigations of ICs and Sys.,* 2002