Interconnect Lifetime Prediction with Temporal and Spatial Temperature Gradients for Reliability-Aware Design and Runtime Management: Modeling and Applications

UNIV. OF VIRGINIA DEPT. OF COMPUTER SCIENCE TECH. REPORT CS–2006–23
OCTOBER 2006

Zhijian Lu[†], Wei Huang[†], Mircea Stan[†], Kevin Skadron[‡], John Lach[†]
Departments of [†]Electrical and Computer Engineering and [‡]Computer Science, University of Virginia
Charlottesville, VA 22904

Abstract

Thermal effects are becoming a limiting factor in high performance circuit design due to the strong temperature dependence of leakage power, circuit performance, IC package cost and reliability. While many interconnect reliability models assume a constant temperature, this paper analyzes the effects of temporal and spatial thermal gradients on interconnect lifetime in terms of electromigration. For temporal thermal variations, we present a physics-based dynamic model for estimating interconnect lifetime for any time-varying temperature/current profile, and this model returns reliability equivalent temperature and current density that can be used in traditional reliability analysis tools. For spatial temperature gradients, we give close bounds in terms of uniformly distributed temperatures to estimate the lifetime of interconnects subject to non-uniform temperature distribution. Our results are verified with numerical simulations and reveal that blindly using the maximum temperature leads to very inaccurate or too pessimistic lifetime estimation. In fact, our dynamic model reveals that when the temporal temperature variation is small, average temperature (instead of worst-case temperature) can be used to accurately predict interconnect lifetime. Therefore, our results not only increase the accuracy of reliability estimates, but they also enable designers to reclaim design margin in reliability-aware design. In addition, our dynamic reliability model is useful for improving the performance of temperature-aware dynamic runtime management by modeling reliability as a resource to be consumed at a stress-dependent rate. *This report supersedes TR CS-2005-10.*

Index Terms

lectromigration, reliability-aware design, dynamic stress, temperature gradients, dynamic thermal/reliability management.lectromigration, reliability-aware design, dynamic stress, temperature gradients, dynamic thermal/reliability management.E

I. Introduction

Due to increasing complexity and clock frequency, temperature has become a major concern in integrated circuit design. Higher temperatures not only degrade system performance, raise packaging costs, and increase leakage power, but they also reduce system reliability via temperature enhanced failure mechanisms such as gate oxide breakdown, interconnect fast thermal cycling, stress-migration and electromigration (EM). The introduction of low-k dielectrics in the future technology nodes will further exacerbate the thermal threats [1]. In this paper, we focus on temperature-related EM failure on interconnects. Other failure mechanisms will be investigated in the future.

The field of temperature-aware design has recently emerged to maximize system performance under lifetime constraints. Considering system lifetime as a resource that is consumed over time as a function of temperature, dynamic thermal management (DTM) techniques [2], [3] are being developed to best manage this consumption. While the dynamic temperature profile of a system is workload-dependent [3], [4], several efficient and accurate techniques have been proposed to simulate transient chip-wide temperature distribution [4], [5], [6], providing design-time knowledge of the thermal behavior of different design alternatives. Currently, DTM studies assume a fixed maximum temperature, which is unnecessarily conservative. To better evaluate these techniques and explore the design space, designers need better information about the lifetime impact of temperature.

Failure probability in VLSI interconnects due to electromigration is commonly modeled with lognormal reliability functions. The variability of lifetime is strongly dependent on the interconnect structure geometries and weakly dependent on environmental stresses such as current and temperature [7], while median time to failure (MTF) is determined by current and temperature in the interconnect. In this paper, we use MTF as the reliability metric and investigate how it is affected by temporal and spatial thermal gradients. Historically, Black [8] proposed a semi-empirical temperature-dependent equation for EM failures:

$$T_f = \frac{A}{j^n} exp\left(\frac{Q}{kT}\right) \tag{1}$$

where T_f is the time to failure, A is a constant based on the interconnect geometry and material, j is the current density, Q is the activation energy (e.g., $0.6eV$ for aluminum), and kT is the thermal energy. The current exponent, n, has different values according to the actual failure mechanism. It is assumed that $n = 2$ for void nucleation limited failure and $n = 1$ for void growth limited failure [9]. Black's equation is widely used in thermal reliability analysis and design.

However, Black's equation assumes a constant temperature. For interconnects subject to temporal and/or spatial thermal gradients, two questions need to be answered: 1. Is Black's equation still valid for reliability analysis in these cases? 2. If Black's equation is valid, what temperature should be used? Though in absence of clear answers in the literature, in practice, Black's equation is still widely assumed, and the worst-case temperature profile is usually used to provide safeguard, resulting in pessimistic estimations and unnecessarily restricted design spaces. As an example, we use the *Hotspot* toolset [4], an accurate architecture-level compact thermal model, to simulate a processor running the Spec2000 benchmarks. The temperature and the power of the hottest block (i.e., the integer unit) for one benchmark are plotted in Figure 1. In this case, the substrate temperature varies between 110^oC and 114^oC, and the maximum power is more than 1.5 times the minimum power. We can see that for only a small portion of time is the program running at the worst-case temperature.

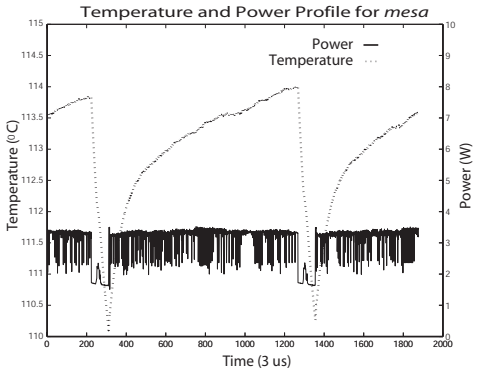


Fig. 1. A simulated temperature/power profile for an integer unit running the *mesa* Spec2000 benchmark. [10]

In the first part of this paper, we answer the above two questions. We find that, for EM subject to time-varying stresses, Black's equation is still valid, but only with the reliability equivalent temperature/current density that returns from a dynamic reliability model presented in this paper [10]. For EM subject to spatial thermal distribution, Black's equation cannot be applied directly. Instead, we give the bounding temperatures which can be used in Black's equation to bound the actual lifetime subject to non-uniform temperature distribution. Therefore, our results can be seamlessly integrated into current reliability analysis tools based on Black's equation [11]. In addition, while designers are currently constrained by constant, worst-case temperature assumptions, the analysis presented in this paper provides more accurate, less pessimistic interconnect lifetime predictions. This results in fewer unnecessary reliability design rule violations, enabling designers to more aggressively explore a larger design space. One limitation in the application of our results is that our analysis is currently based on two-terminal interconnects, such as those seen in global signal interconnects and power/ground distribution networks. Recently Alam *et al.* [11] proposed lifetime predictions for multi-terminal interconnects. Our future work will include extending our current findings to multi-terminal interconnects.

Worst-case power dissipation and environmental conditions are rare for general-purpose microprocessors. Designing the cooling solution for the worst case is wasteful. Instead, the cooling solution should be designed for the worst "expected" case. In the event that environmental or workload conditions exceed the cooling solution's capabilities and temperature rises to a dangerous level, on-chip temperature sensors can engage some form of "dynamic thermal management" (DTM) [4], [12], [13], which sacrifices a certain amount of performance to maintain reliability by reducing circuit speed whenever necessary. Existing DTM techniques do not consider the effects of temperature fluctuations on lifetime and may unnecessarily impose performance penalties.

In the second part of this paper, we propose runtime dynamic reliability management (DRM) techniques based on our dynamic reliability model [14]. By leveraging this model, one can dynamically track the "consumption" of chip lifetime during operation. In general, when temperature increases, lifetime is being consumed more rapidly, and vice versa. Therefore, if temperature is below the traditional DTM engagement threshold for an extended period, it may be acceptable to let the threshold be exceeded for a time while still maintaining the required expected lifetime. In effect, lifetime is modeled as a resource that is being "banked" during periods of low temperature, allowing for future withdrawals to maintain performance during times of higher operating temperatures. Using electromigration as an example, we show the benefits of lifetime banking by avoiding unnecessary DTM engagements while meeting expected lifetime requirements.

The concept of dynamic reliability management is first introduced by Srinivasan *et al.* [3]. In their work, they proposed a chip level reliability model and showed the potential benefits by trading off reliability with performance for individual applications. They assumed an oracular algorithm for runtime management in their study, and they did not consider the effects of inter-application thermal behaviors on reliability. Later work from the same authors [15] refined their reliability model and showed the improvement in reliability using redundant components. In this paper, we focus on practical runtime management techniques for the worst-case on-chip component (i.e. hottest interconnect) to exploit both intra- and inter-application temperature variations. The combination of their model and our techniques is expected to bring more advantages and is open for future investigation. Ramakrishnan and Pecht [16] proposed to monitor the life consumption of an electronic system and project the system lifetime based on the monitoring results. We take a similar approach to monitoring the stresses on the circuit continuously, and we also intelligently adapt the circuit operation to maximize the circuit performance without reducing reliability.

The rest of the paper is organized as follows. Section II introduces a stress-based analytic model for EM, which serves as the base model in this paper. In Section III, we extend this model to cope with time-varying stresses (i.e., temperature and current) and derive a formula to estimate interconnect lifetime, which we analyze in Section IV. In Section V, we analyze the impact of non-uniform temperature distribution on lifetime prediction due to EM. We illustrate how designers can use our analysis to reclaim some design margin by considering runtime variations in Section VI. In Section VII, we exploit our proposed dynamic reliability model in runtime thermal management and propose a banking-based dynamic reliability management technique to improve system performance while maintaining lifetime constraints. Finally, we summarize the paper in Section VIII.

II. An analytic model for EM

In this section, we describe the basic EM model used in the paper. In the following sections, we will extend this basic EM model to predict interconnect lifetime under dynamic thermal and current stresses.

Clement [17] provides a review of 1-D analytic EM models. Several more sophisticated EM models are also available [9], [18]. In this paper, we only discuss the EM-induced stress build-up model of Clement and Korhonen [19], [20], which has been widely used in EM analysis and agrees well with simulation results using a more advanced model by Ye *et al.* [21].

EM is the process of self-diffusion due to the momentum exchange between electrons and atoms. The dislocation of atoms causes stress build-up according to the following equation [19], [20]:

$$\frac{\partial \sigma}{\partial t} - \frac{\partial}{\partial x}\left(\left[D_a\left(\frac{B\Omega}{kTl^2\varepsilon}\right)\right](\frac{\partial \sigma}{\partial x} - \frac{qlE}{\Omega})\right) = 0 \tag{2}$$

where $\sigma(x,t)$ is the stress function, and an interconnect failure is considered to happen when $\sigma(x,t)$ reaches a threshold (critical) value σ_{th}. D_a is the diffusivity of atoms, a function of temperature. B is the appropriate elastic modulus, depending on the properties of the metal and the surrounding material and the line aspect ratio. Ω is the atom volume. ε is the ratio of the line cross-sectional area to the area of the diffusion path. l is the characteristic length of the metal line (i.e., the length of the effective diffusion path of atoms). q is the effective charge. E is the applied electric field, which is equal to ρj, the product of resistivity and current density. The term $\frac{qlE}{\Omega}$ corresponds to the atom flux due to the electric field, while $\frac{\partial \sigma}{\partial x}$ corresponds to a back-flow flux created by the stress gradient to counter-balance the EM flux. And the total atomic flux at a specific location in the interconnect is proportional to the sum of these two components:

$$J = \left[D_a\left(\frac{B\Omega}{kTl^2\varepsilon}\right)\right](\frac{\partial \sigma}{\partial x} - \frac{qlE}{\Omega}) \tag{3}$$

Equation (2) states that the mechanical stress build-up at any location is caused by the divergence of atomic flux at that point, or $\frac{\partial \sigma}{\partial t} = \nabla J$. If we assume a uniform temperature across the interconnect characteristic length and let $\beta(T) = D_a\left(\frac{B\Omega}{kTl^2\varepsilon}\right)$ (which we refer to as the temperature factor throughout the paper) and $\alpha(j) = \frac{qlE}{\Omega}$, we obtain the following simplified version, the solution of which depends on both temperature and current density:

$$\frac{\partial \sigma}{\partial t} - \beta(T)\frac{\partial}{\partial x}\left(\frac{\partial \sigma}{\partial x} - \alpha(j)\right) = 0 \tag{4}$$

Clement [19] investigated the effect of current density on stress build-up using Equation (4), assuming that temperature is unchanged (i.e., $\beta(T) = constant$), for several different boundary conditions. He found that the time to failure derived from this analytic model had exactly the same form as Black's equation (1). The exponential component in Black's equation is due to the atom diffusivity's (D_a's) dependency on temperature by the well-known Arrhenius equation: $D_a = D_{ao} exp\left(\frac{-Q}{kT}\right)$.

Applying the parabolic maximum principles [22] to Equation (4), we know that at any time t, the maximum stress along a metal line can be found at the boundaries of the interconnect line. Figure 2 shows the numerical

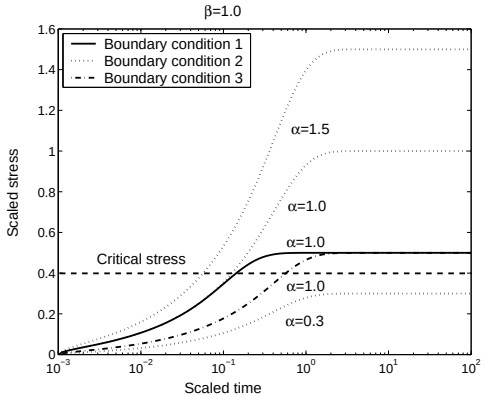


Fig. 2. EM stress build-up for different boundary conditions and α values. All processes have $\beta = 1$ (α and β are defined in Equation (4)). [10]

solutions for Equation (4) at one end of the line (i.e., $x = 0$) for different boundary conditions and α values, all with $\beta = 1$. The three boundary conditions shown here are similar to those discussed in [19] for finite length interconnect lines. It indicates that both boundary conditions and current density (α) affect the stress build-up rate (i.e., the larger the current, the faster the stress builds up.). Also seen from the figure is that the stress build-up saturates at a certain point. This is because, in saturation, the atom flux caused by EM is completely counterbalanced by the stress gradient along the metal line. It is believed that the interconnect EM failure occurs whenever the stress build-up reaches a critical value, σ_{th} (as shown in Figure 2). If the saturating stress is below the critical stress, no failure happens. In the following discussion, we assume that the saturating stress in an EM process is always above the critical stress.

III. EM UNDER DYNAMIC STRESS

In this section, we first show that the "average current" model can be used to estimate EM lifetime under dynamic current stress while the temperature is constant. Then we derive a formula to reveal the effect of time-dependent temperature on EM. Finally, based on these two results, we generalize an EM lifetime prediction model accounting for the combined dynamic interplay of temperature and current stresses.

A. Time-dependent current stress

Clement [19] used a concentration build-up model similar to the one discussed here to verify that in the case in which temperature is kept constant, the average current density can be used in Black's equation for pulsed DC current. As for AC current, an EM effective current is used by the Average Current Recovery (ACR) model [23], [24]. In this paper, we do not distinguish between these two cases. We only consider the change of EM effective current due to various causes (e.g., phased behaviors in many workloads). This is because the time scale of the current variation studied in this paper is usually much longer than that of the actual DC/AC current changes in the interconnects.

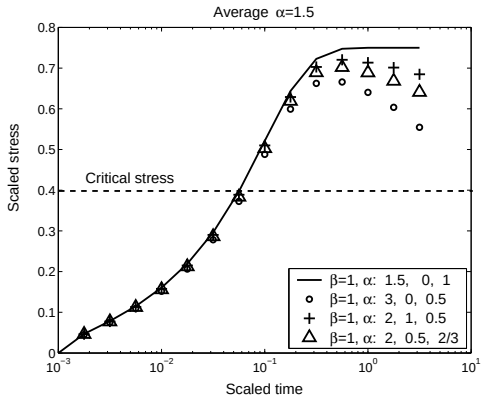


Fig. 3. EM stress build-up under time-dependent current stress. In each EM process, α (defined in Equation (4)) oscillates between two values with different duty cycles. The time dependence of α is given in the legend.[2]All curves have the same average value of α. The solid line is the stress build-up with a constant value of α. [10]

[2]For example, the numbers after the circle represent the case in which α is a square-wave function and varies between 3 and 0 with a duty cycle of 0.5. This representation of the time-dependent square-wave function is used in other figures throughout the paper.

We numerically solve Equation (4) with different time-dependent α functions, and the results are plotted in Figure 3. The stress build-ups for all EM processes in Figure 3 overlap before saturations (or before reaching the critical stress), since they have the same average current. Thus, the EM process under time-varying current stress can be well approximated by average current. Note that the curves in Figure 3 diverge after they reach their maximum stress. This is because the time-varying current could not create a stable counterbalancing stress gradient for EM. However, we are only interested in the EM process before reaching the critical stress when EM failure occurs.

B. Time-dependent thermal stress

If the temperature (β) of the interconnect is time-dependent, we can derive the EM stress build-up expression indirectly based on the following theorem.

Theorem 1: Consider stress build-up Equation (4) with constant values for β and α. Let $\sigma_1(x,t)$ be the solution for the equation with $\beta = \beta_1$ under certain initial and boundary conditions and $\sigma_2(x,t)$ be the solution with $\beta = \beta_2$ for the same initial and boundary conditions. If the solutions for Equation (4) are unique for those initial and boundary conditions, we have

$$\sigma_2(x,t) = \sigma_1(x, \left(\frac{\beta_2}{\beta_1}\right) t)$$

Proof: Since $\sigma_1(x,t)$ is the solution for the equation, we have $\frac{\partial \sigma_1}{\partial t}(x, \left(\frac{\beta_2}{\beta_1}\right) t) - \beta_1 \frac{\partial}{\partial x}\left(\frac{\partial \sigma_1}{\partial x}(x, \left(\frac{\beta_2}{\beta_1}\right) t) - \alpha(j)\right) = 0$. On the other hand, let $\sigma_2(x,t) = \sigma_1(x, \left(\frac{\beta_2}{\beta_1}\right) t)$, we have $\frac{\partial \sigma_2}{\partial t}(x,t) = \left(\frac{\beta_2}{\beta_1}\right) \frac{\partial \sigma_1}{\partial t}(x, \left(\frac{\beta_2}{\beta_1}\right) t)$ and $\frac{\partial \sigma_2}{\partial x}(x,t) = \frac{\partial \sigma_1}{\partial x}(x, \left(\frac{\beta_2}{\beta_1}\right) t)$. This leads to $\frac{\partial \sigma_2}{\partial t}(x,t) = \beta_2 \frac{\partial}{\partial x}\left(\frac{\partial \sigma_2}{\partial x}(x,t) - \alpha(j)\right)$, which demonstrates that $\sigma_1(x, \left(\frac{\beta_2}{\beta_1}\right) t)$ is the solution for the stress build-up equation with $\beta = \beta_2$, under the same initial and boundary conditions. ■

Theorem 1 tells us that the stress build-up processes in the interconnect are independent of the value of β in Equation (4). The value of β only determines the build-up speed of the process. For example, at time $\left(\frac{\beta_2}{\beta_1}\right) t$, the stress build-up of an EM process with $\beta = \beta_1$ *sees* the stress build-up of an EM process with $\beta = \beta_2$ at time t. In other words, *it is possible to use the expressions for stress build-up under constant temperature to describe the EM process under time-varying thermal conditions*.

Consider that temperature varies over time, and EM effective current doesn't change. We can divide time into segments, such that temperature is constant within each time segment. In other words, β in Equation (4) is a segment-wise function, described as:

$$\beta(t) = \begin{cases} \beta_1, & t \in [0, \Delta t_1] \\ \beta_2, & t \in (\Delta t_1, \Delta t_1 + \Delta t_2] \\ \cdots & \\ \beta_i, & t \in \left(\sum_{k=1}^{i-1} \Delta t_k, \sum_{k=1}^{i} \Delta t_k\right] \\ \cdots & \end{cases}$$

We denote $M0$ as the metal line of interest. Imagine that there is another metal line, denoted by $M1$, having the same geometry and EM effective current as $M0$. $M1$ has a constant value of β equal to $\beta 1$, while $M0$ will experience a time-dependent function of $\beta(t)$. Let $\sigma_0(t)$ and $\sigma_1(t)$ be the stress evolution on metal line $M0$ and $M1$ respectively. During the first time segment, the stress build-ups on both metal lines are the same. Thus, at the end of this time segment, we have $\sigma_0(\Delta t_1) = \sigma_1(\Delta t_1)$. $M0$ will continue to build up stress with β_2 during the second time segment. According to Theorem 1, the stress evolution of $M0$ during Δt_2 will be the same as that of $M1$, except that it will take $M1$ a time period of $\frac{\beta_2}{\beta_1}\Delta t_2$ to achieve the same stress. Similar analysis can be applied to other time segments. As a result, at the end of the ith time segment, the stress build-up in $M0$ will be equal to that in $M1$ after a total time of $\sum_{k=1}^{i}(\frac{\beta_k}{\beta_1})\Delta t_k$. In other words, we can convert the stress evolution under time-varying thermal stress into EM stress evolution with constant temperature.

It follows that at the end of the ith time segment, the stress in $M0$ is specified as: $\sigma_0(\sum_{k=1}^{i} \Delta t_k) = \sigma_1\left(\sum_{k=1}^{i}(\frac{\beta_k}{\beta_1})\Delta t_k\right)$.

As $\Delta t_i \to dt$, $\beta_i \to \beta(T(t))$, we obtain the integral version for the stress build-up function:

$$\sigma_0(t) = \sigma_1\left((\frac{1}{\beta_1})\int_0^t \beta(T(t))dt\right) \tag{5}$$

If we assume that the stress build-up reaches a certain threshold (σ_{th}) at which an EM failure occurs, we have:

$$\int_0^{t_{failure}} \beta(T(t))dt = \varphi_{th} \tag{6}$$

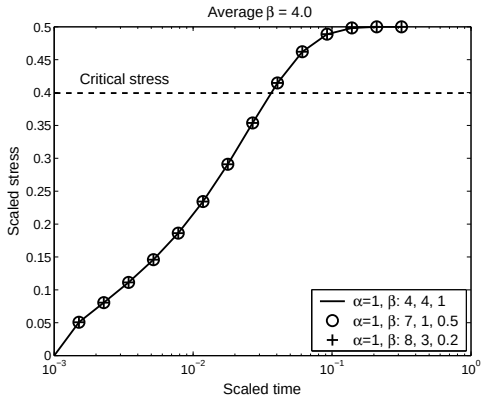


Fig. 4. EM stress build-up at one end of the interconnect with different time-dependent β functions (square waveform). The solid line is the case with a constant value of β equal to the average value of β in other curve. [10]

where φ_{th} is a constant determined by the critical stress (i.e. $\varphi_{th} = \sigma_1^{-1}(\sigma_{th})\,\beta_1$). If an average value of $\beta(t)$ exists, we obtain a closed form for the time to failure:

$$t_{failure} = \frac{\varphi_{th}}{E(\beta(T(t)))} \tag{7}$$

where $E(\beta(t))$ is the expected value for $\beta(t)$, and $\beta(t)$ is the temperature factor, as defined in Equation (4), having the form $\beta(T(t)) = A'\left(\frac{exp(-\frac{Q}{kT(t)})}{kT(t)}\right)$ where A' is a constant. In comparison with Black's equation, Equation (7) indicates that the average of temperature factor β should be used.

One way to interpret Equation (6) is to consider interconnect time to failure (i.e., interconnect lifetime) as an available resource, which is consumed by the system over time. Then the $\beta(t)$ function can be regarded as the consumption rate.

Let $MTF(T)$ be the time to failure with a constant temperature T. We have $\beta(T) = \frac{\varphi_{th}}{MTF(T)}$ by Equation (7). Substitute this relation in Equation (7) again and consider the time-varying temperature, and we obtain an alternative form for Equation (7):

$$t_{failure} = \frac{1}{E(1/MTF(T))} \tag{8}$$

Equation (8) can be used to derive the absolute time to failure provided that we know the time to failure for different constant temperatures (e.g., data from experiments).

By calculating the second derivative of $\beta(T)$ as a function of temperature, it can be verified that $\beta(T)$ is a convex function within the operational temperatures. By applying Jensen's inequality, we have $E(\beta(T)) \geq \beta(E(T))$, which, according to Equation (7), leads to an interesting observation: constant temperature is always better in terms of EM reliability than oscillating around that temperature (with the average temperature the same as the constant temperature).

Similar to the methods for verifying the "average current model", we obtain numerical solutions for the stress build-up equation using different square waveforms for β. Figure 4 compares these results and shows that the time to failure will be the same as long as the EM processes exhibit the same *average* value of β.

C. Combined dynamic stress

In reality, both temperature and current change simultaneously. In most cases, the variation of temperature on the chip reflects changes in power consumption, thus directly relating to current flow in the interconnects. In order to describe the EM process in this general case, we can, again, divide time into multiple small segments, and in each time segment, assume that both current and temperature are constant. The temperature and current stresses on the interconnect within time segment Δt_i is denoted by a pair of values (α_i, β_i). Following the same technique as for the time-varying thermal stress, we compare the EM processes in two metal lines ($M0$ and $M1$), and one ($M0$) of which is under time-varying thermal and current stresses. We construct an EM process in the second metal line ($M1$) such that $M1$ is subject to a constant thermal stress ($\beta_{M1} = \beta_1$). Applying Theorem 1 reveals that the stress evolution of $M0$ within Δt_i, under (α_i, β_i), is the same as that of $M1$ under stress (α_i, β_1) for a time period of $\frac{\beta_i}{\beta_1}\Delta t_i$. Thus, at the end of the ith time segment, the stress build-up of $M0$ is equal to the stress evolution of $M1$ at the time $\sum_{k=1}^{i}(\frac{\beta_k}{\beta_1})\Delta t_k$. Notice that the current stress on $M1$ is time-dependent (i.e, $\alpha_{M1} = \alpha_i$ for a time period of $\frac{\beta_i}{\beta_1}\Delta t_i$). In order to find the stress of $M1$ at $\sum_{k=1}^{i}(\frac{\beta_k}{\beta_1})\Delta t_k$, the current profile (i.e., α as a function of time) for

$M1$ should be considered:

$$\alpha_{M1}(t) = \begin{cases} \alpha_1, & t \in \left[0, \frac{\beta_1}{\beta_1}\Delta t_1\right] \\ \alpha_2, & t \in \left(\frac{\beta_1}{\beta_1}\Delta t_1, \frac{\beta_1}{\beta_1}\Delta t_1 + \frac{\beta_2}{\beta_1}\Delta t_2\right] \\ \cdots & \\ \alpha_i, & t \in \left(\sum_{k=1}^{i-1}\frac{\beta_k}{\beta_1}\Delta t_k, \sum_{k=1}^{i}\frac{\beta_k}{\beta_1}\Delta t_k\right] \end{cases}$$

Since the stress evolution in $M1$ is under constant thermal stress, we may apply the "average current model". As $\Delta t_i \to dt$, $\beta_i \to \beta(T(t))$ and $\alpha_i \to \alpha(t)$, we derive the EM reliability equivalent current for $M0$ (or the average current for $M1$) as:

$$j_{equivalent} = \frac{\int_0^T j(t)\beta(t)dt}{\int_0^T \beta(t)dt} = \frac{E\left[j(t)\beta(t)\right]}{E\left[\beta(t)\right]} \tag{9}$$

where T is a relatively large time window, and $j(t)$ is the corresponding current density for $\alpha(t)$. Thus, the EM process in $M0$ can be approximated by an EM process with constant stresses (i.e., $j = j_{equivalent}$ and $\beta = \beta_1$). Using a similar derivation as for Equations (5), (6), and (7), combined with Black's equation, we obtain the time to failure for $M0$:

$$t_{failure} = \frac{C}{j_{equivalent}^n E(\beta(T(t)))} \tag{10}$$

where $j_{equivalent}$ is defined by Equation (9), and C is a constant.

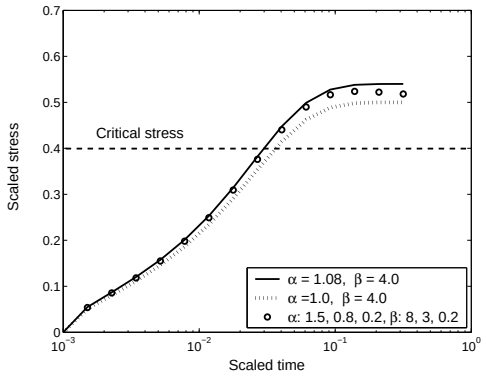


Fig. 5. EM stress build-up at one end of the interconnect with time-varying α (current) and β (temperature) functions (i.e., square waveforms). The circles represent the numerical solution for time-varying α and β. The solid line is with a constant value of α calculated according to Equation (9) and a constant value of β equal to the average value of that in the time-varying case. As a comparison, the EM process (dotted line) simply using the average current of the time-varying case is also shown. These results show that EM process under dynamic stresses (circles) can be well approximated by a process with constant stresses (solid line). [10]

Figure 5 compares the stress build-ups for different dynamic current and temperature combinations. These results illustrate that the EM process under dynamic stresses can be well approximated by an EM process with a constant temperature (i.e., $E(\beta)$) and a constant current (i.e., $I_{equivalent}$ as defined in Equation (9)). Therefore, for an interconnect with concurrent time-dependent temperature and current stresses, time to failure has the same form as Black's equation, except that the reliability-equivalent current (the actual current modulated by the temperature factor β (i.e., weighted averaging by β)) and the mean value of the temperature factor are used.

In fact, if the current and the temperature are statistically independent, we have $\frac{E[j(t)\beta(t)]}{E[\beta(t)]} = E\left[j(t)\right]$ in Equation (9). In this case, the reliability equivalent current will be reduced to the average current and we get back to the "average current model". On the other hand, if the current is constant, Equations (9) and (10) will lead us to Equation (7). Finally, if both temperature and current are time invariant, Black's equation (Equation (1)) is obtained.

IV. Analysis of the proposed model

Equations (9) and (10) form the basis of our proposed EM model under concurrent time-varying temperature and current stress. In this section, we use these equations to evaluate EM reliability. Specifically, we compare the reliability of constant temperature with that of fluctuating temperature, and we show the difference of lifetime projection between our model and the traditional worst-case model.

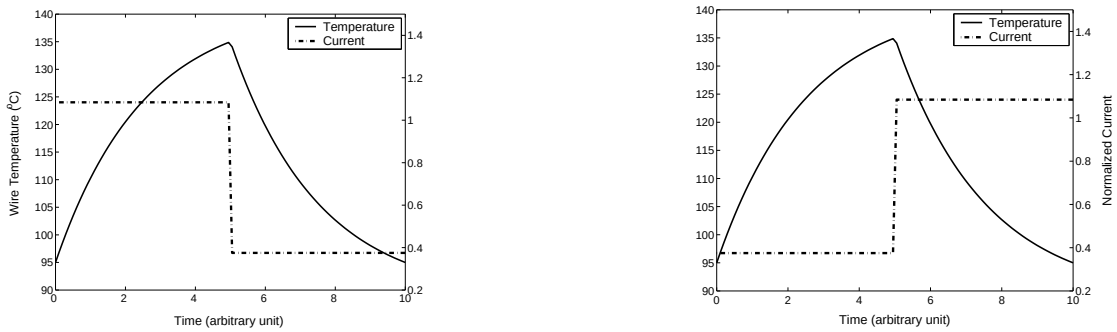


Fig. 6. Temperature and current waveforms analyzed in the paper: (a) in phase current/temperature, (b) out of phase current/temperature. [10]

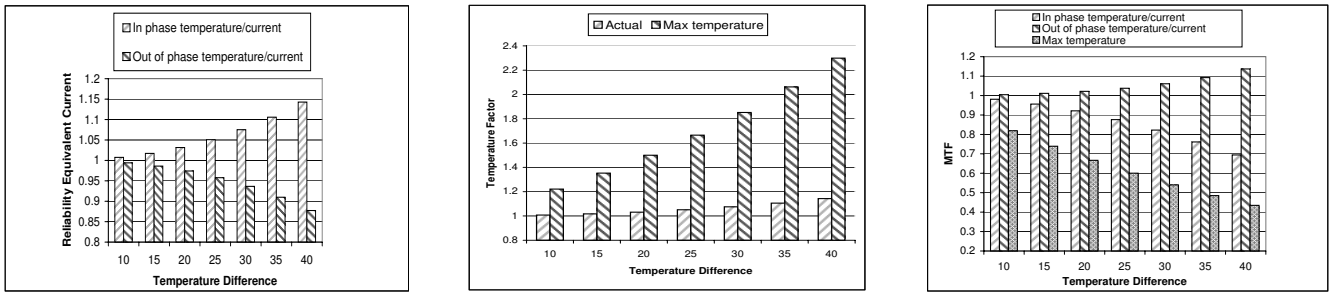


Fig. 7. Comparison of electric current, temperature factor (β) and MTF for different peak to peak temperature cycles. All results are normalized to the average current and/or temperature case. (a) Ratio of reliability equivalent current (our model) to average current. Both cases of current variation (in and out of phase with temperature) are included. (b) Ratios of temperature factor (β) using average temperature, max temperature, and our model. (c) Comparison of MTF for four different calculations: average temperature/average current, maximum temperature/average current, our model for current in phase with temperature, and our model for current out of phase with temperature. [10]

For any two temporal temperature and current profiles we can easily compare the EM reliability, using our model, by:

$$\frac{MTF_1}{MTF_2} = \frac{j^2_{equivalent2} E(\beta(T_2(t)))}{j^2_{equivalent1} E(\beta(T_1(t)))}$$

where MTF_1 is the time to failure under time-varying temperature profile $T_1(t)$ and electric current profile $j_1(t)$.

As shown in Figure 1, in real workload execution, temperature changes along with the changes in power consumption (i.e. current). It is interesting to see how the interactions between temperature and current profiles affect the interconnect lifetime. The possible interactions between temperature and current form a spectrum, and the plots in Figure 6 show the two extremes of this spectrum. In this figure, a simple assumption is made that the current is proportional to the difference between the steady substrate temperature and the ambient temperature (i.e., 40^oC). The temperature difference between the substrate and the interconnects is fixed to be 21^oC, which is a reasonable assumption for high-layer interconnects [25]. Using the data from Figure 1, the maximum temperature of the substrate is assumed to be 114^oC (i.e., 135^oC at the interconnects), and we change the minimum temperature to obtain different temperature/current profiles. Using these profiles, we can compare the reliability equivalent current with the average current, compare the temperature factor using our model with those of average and maximum temperatures, and finally compare the MTFs in these cases (i.e., average current/average temperature, reliability equivalent current/average temperature factor (β), and average current/maximum temperature).

Our results are reported in Figure 7, and we summarize our observations as follows:

- As the peak to peak temperature difference is small, both the reliability equivalent current and the temperature factor predicted by our dynamic stress model are very close to those calculated from using average current and average temperature. That is because the temperature factor function (β), although an exponential function of temperature, can be well approximated by a linear function of temperature within a small temperature

range. Thus, the MTF predicted by using average temperature/current provides a simple method for reliability evaluation with high accuracy.

- As the temperature difference increases, we can no longer simply use average temperature/current for MTF prediction. Both the reliability equivalent current and the temperature factor increase (degrading reliability) quickly as the temperature difference increases.
- On the other hand, using maximum temperature always underestimates the lifetime, resulting in excessive design margins.
- One interesting phenomenon arises in the case in which the current is out of phase with temperature variation. Recall that the reliability equivalent current is actually a temperature factor weighted average current, and high temperature increases the weights for the accompanied current. Thus, the reliability equivalent current is reduced compared to the case in which temperature/current are synchronized. This brings a non-intuitive effect on the reliability projection—MTF even slightly increases as the temperature cycling magnitude increases.

In the above discussion, the duty cycle of the current waveform is fixed (i.e., 0.5). We also investigated the effects of different duty cycles, but the data is not shown here due to space limitations. In general, when the temperature change is small (e.g., within 10^oC), using the average temperature to predict lifetime is still a good approximation (less than 5% error) regardless of the duty cycle. While the temperature variation increases, the difference between our model and using average temperature is largest at a duty cycle of about 0.4. On the other hand, the smaller the duty cycle, the larger the difference between our model and using maximum temperature. Thus, using maximum temperature is reasonable only when the duty cycle is large (i.e., higher temperature dominates almost the entire cycle).

V. Electromigration under spatial temperature gradients

In addition to temporal temperature variations, large temperature differences across the chip are commonly seen in modern VLSI design. Ajami *et al.* [26] showed that non-uniform temperature has great impacts on interconnect performance. In this section, we will illustrate the importance of considering spatial temperature gradients for interconnect reliability.

A. EM model with spatial thermal gradients

Due to the exponential dependence of diffusivity on temperature, EM in interconnects with spatial temperature gradients has quite different characteristics than those with constant temperature. Guo *et al.* [27] reported that EM in aluminum interconnect is strongly affected by the relative direction of electron wind and thermal gradients, while Nguyen *et al.* [28] found that temperature gradients greatly enhance EM in aluminum interconnect. Following the stress build-up model introduced in Section II, the atomic flux due to EM can be modeled by $J = \beta(T)\left(\frac{\partial \sigma}{\partial x} - \alpha(j)\right)$, and the stress build-up at a specific location is caused by the divergence of atomic flux at that location, i.e. $\frac{\partial \sigma}{\partial t} = \nabla J$. When the temperature is uniform across the interconnect, i.e. $\beta(T)$ is independent of location, Equation (4) is obtained. When the temperature is not uniform, the following equation is derived to describe the stress build-up under thermal gradients:

$$\frac{\partial \sigma}{\partial t} - \beta(T(x))\frac{\partial}{\partial x}\left(\frac{\partial \sigma}{\partial x} - \alpha(j)\right) - \frac{\partial \beta(T(x))}{\partial x}\left[\frac{\partial \sigma}{\partial x} - \alpha(j)\right] = 0 \tag{11}$$

where σ, β and α are defined in Section II. When compared with Equation (4), Equation (11) introduces a third term $\frac{\partial \beta(T(x))}{\partial x}\left[\frac{\partial \sigma}{\partial x} - \alpha(j)\right]$, which captures the atomic flux divergence induced by spatial thermal gradients along the interconnect. Though temperature gradient itself will cause migrations of atoms from high temperature to low temperature, a phenomenon called thermomigration (TM), the atomic flux due to TM is generally believed to be much smaller than that due to EM [28]. Therefore TM is not explicitly modeled in Equation (11). Jonggook *et al.* [29] investigated EM in aluminum (Al) interconnects subject to spatial thermal gradients. They modeled EM from a different approach but yielded an equation with a form similar to ours. Since dual-damascene Cu interconnects have become the mainstream technology in modern VLSI design and have quite different EM characteristics from Al [30], in the following, we focus on EM failure in copper interconnects.

Various experiments [7], [31] showed that, in copper interconnect, voids tend to nucleate at the cathode end (near the via), and void growth is the dominant failure process because the critical mechanical stress for void nucleation in copper is much smaller than that for aluminum. With spatial thermal gradients in the interconnect, it is possible that the location of void nucleation is no longer at the cathode end. However, in this case, void growth tends to be slower than that at the cathode, because there are atomic fluxes both going into and coming from the void in the middle of the interconnect [31]. Bearing these observations in mind and assuming a void-growth dominated failure, we choose a boundary condition for Equation (11) to model void growth at the cathode such that the mechanical

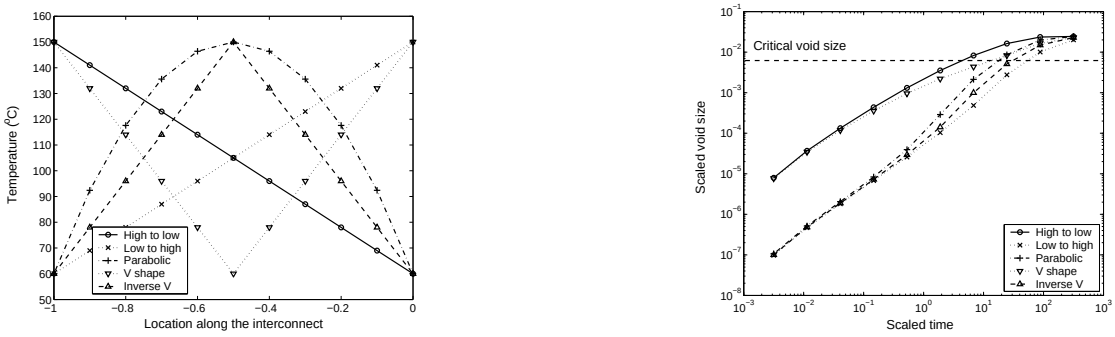
stress at the cathode end is zero (free stress at void) and the atomic flux at the other end is zero (complete blockage for atomic flux), or:

$$\sigma(x=-l,t)=0,\ J(x=0,t)\Rightarrow\left[\frac{\partial\sigma}{\partial x}-\alpha(j)\right]\Big|_{x=0}=0$$

where $x=-l$ is the cathode end. This boundary condition is consistent with that used by Clement [17] to model void growth due to EM. The void size at time t can be approximated by [17]:

$$\Delta l\approx\int_{-l}^{0}\frac{-\sigma(x,t)}{B}dx$$

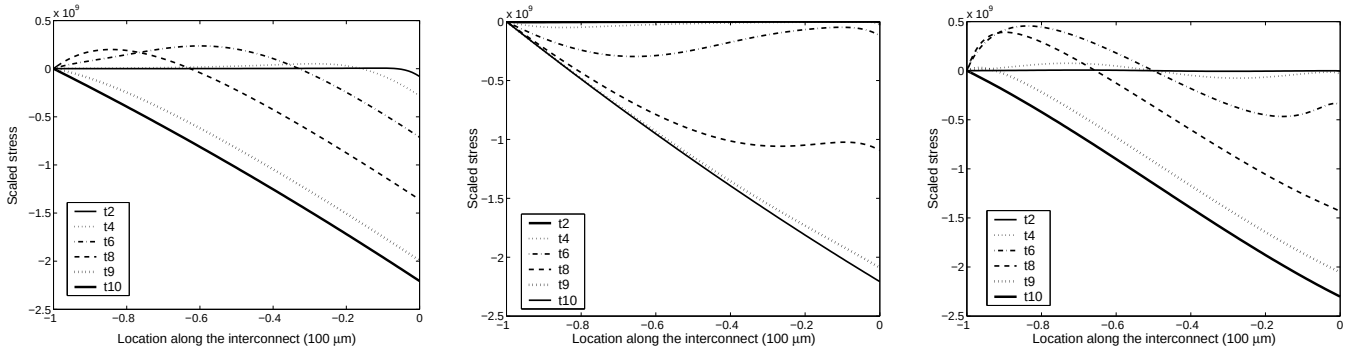
where $\sigma(x,t)$ is the mechanical stress (tensile stress) developed along the interconnect at time t and B is the elastic modulus. Because we are unaware of any closed form solution for Equation (11) with the above boundary condition, we use numerical solutions to analyze the impact of thermal gradients on electromigration.



(a) (b)

Fig. 8. Effects of non-uniform spatial temperature distribution on EM induced void growth. (a) Various temperature profiles along a $100\mu m$ copper interconnect (left end is the cathode). (b) Void growth with different spatial temperature profiles.

The temperature spatial profile along an interconnect is the combined effects of joule heating and substrate temperature distributions. Figure 8 (a) plots several temperature profiles used in our study and their effects on EM induced void growth. The length of the interconnect is $100\mu m$, and electrons are assumed to flow from the left end (cathode) to the right end of the interconnect. Though all temperature profiles have the same maximum and minimum temperatures, their void growth differs greatly due to the different thermal gradients along the interconnect (Figure 8 (b)), resulting quite different failure time. In order to investigate how thermal gradients affect EM induced void growth, we also plot, in Figure 9, the mechanical stress build-up along the interconnect at different times, with different thermal profiles. In spite of different temperature profiles on the interconnect, in the final EM process stage ("t10" in Figure 9), a steady stress gradient is built up to counter-balance the driving force of electron wind, i.e. $\frac{\partial\sigma}{\partial x}-\alpha(j)=0$, resulting in voids with comparable saturation sizes (Figure 8 (b)).



(a) (b) (c)

Fig. 9. Stress build-ups at different time points along the interconnect under spatial thermal gradients. (a) Low to high temperature profile. (b) High to low temperature profile. (c) Parabolic temperature profile. Electrons flow from the left to the right, causing compressive stress (negative in the figures) on the right side of the interconnect. The left end (cathode) is stress free to model the growth of a void. The time points ("t2" through "t10") are corresponding to the time points in the plots with the same temperature profile in Figure 10.

However, as shown in Figure 9, the kinetic aspects of stress build-up for different temperature profiles are quite different, especially when the relative direction of electron flow and temperature gradients changes. For a "low to high" temperature profile, the temperature increases linearly from the cathode end to the anode end (shown in Figure 8 (a)). At the early EM stage, as indicated by "t2" and "t4" in Figure 9 (a), the stress gradient near the cathode is negligible. Therefore the atomic flux at the cathode is only determined by the electron wind at the temperature of that location, because the atomic flux is the sum of the fluxes induced by the stress gradient and the electron wind (Equation (3)). Therefore, in this case, the void growth at the cathode is subject to almost the same kinetics as those with a uniform temperature across the interconnect. Later on in the EM process, the effect of thermal gradients begins to play its role. As shown by "t6" and "t8" in Figure 9 (a), tensile (positive) stress is built up from the cathode end towards the other end, due to the increasing temperature from the cathode end. The stress gradient created by this tensile stress distribution forms an atomic flux in the same direction as the electron wind. Thus, void growth in the cathode end is enhanced later by the increasing temperature. On the contrary, Figure 9 (b) shows quite different kinetics for EM process with "high to low" temperature profile, where the temperature is decreasing linearly from the cathode end (shown in Figure 8 (a)). In the early stage, as in the case for a "low to high"profile, void growth is similarly dominated by the temperature at the cathode end, as illustrated by the stress distributions at "t2" and "t4" in Figure 9 (b). Subsequently, compressive (negative) stress is built up from the cathode towards the anode, because of the decreasing temperature from the cathode, as shown by the stress distributions at "t6" and "t8" in Figure 9 (b). The stress gradient due to the compressive stress distribution in this case creates an atomic flux in the opposite direction of the electron wind, retarding the void growth at cathode. The stress distributions induced by EM with a "parabolic" temperature spatial profile at different time points are drawn in Figure 9 (c). The kinetics of stress build-up in this case are similar to those in a "low to high" temperature profile, because both temperature profiles have similar temperature gradients near the cathode. On the other hand, in the late stage of the EM process (as indicated by "t9" and "t10" in Figure 9), regardless of the temperature distribution across the interconnect, significant stress gradient is formed in the opposite direction of atomic flux and slows down the void growth, and finally the steady state of the EM process is reached (or void growth saturates). In summary, in the early stage of the EM process, the void growth is largely dependent on the temperature at the cathode, while later on, the void growth is enhanced or retarded depending on the temperature gradient near the cathode. Finally, void growth is suppressed by the back-flow stress gradient just like in the case of the EM process with a uniform temperature distribution.

B. Empirical bounds for void growth with non-uniform temperature distribution

In Section III, our analysis reveals that the EM process with time-varying temperature variations can be approximated by an EM process using a constant reliability equivalent temperature T_{eq}, as long as $\beta(T_{eq}) = E[\beta(T(t))]$. However, in the case where there is a non-uniform temperature across the interconnect, we cannot find a similar reliability equivalent temperature, due to the difference in the EM kinetics in the different stress build-up stages as shown in Figure 9. Instead, we try to find two constant temperatures, such that the void growth due to EM with non-uniform temperature can be bounded by the void growth with uniformly distributed temperature equal to these two bounding temperatures respectively. The reason for our approach is as follows. Because Black's equation is only valid for a uniform temperature distribution, and many existing reliability analysis tools are based on Black's equation, by providing the bounding temperatures for interconnects subject to spatial thermal gradients, one can still use these tools to evaluate the effects of non-uniform temperature distributions.

Following our previous discussion on Figure 9, one can expect that the cathode temperature can serve as the lower/upper bound temperature for void growth with increasing/decreasing temperature towards the anode end. On the other hand, the void size is proportional to the amount of atoms moved from the cathode end and the void growth rate is determined by the atomic flux at the cathode. We would like to find the other bounding temperature by bounding the atomic flux at the cathode. Consider an interconnect of length l subject to a certain spatial temperature profile $T(x)$, with both ends at zero stress $\sigma(0,t) = \sigma(l,t) = 0$. In the steady state, there is a uniform atomic flux flowing through the interconnect, expressed as (see Appendix):

$$J_{steady} = -\frac{\int_0^l \alpha(T(x))dx}{\int_0^l \frac{1}{\beta(T(x))}dx}$$

By examining the steady-state stress distribution along the interconnect in the above case, it can be further verified that when the temperature is increasing from the cathode, tensile stress is built up from the cathode, and the atomic flux at the cathode is enhanced by the stress gradients, a situation similar to "t6" and "t8" in Figure 9 (a). When

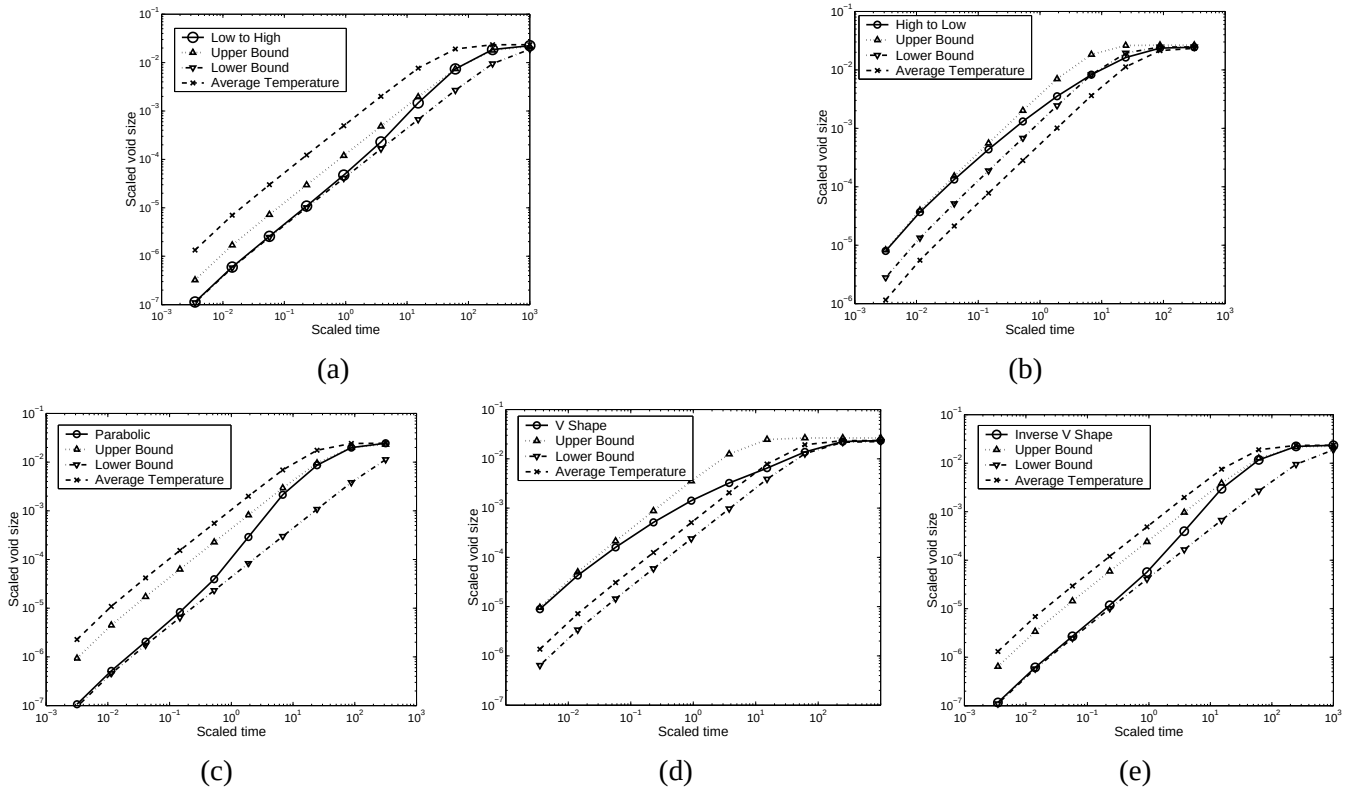


Fig. 10. Void growth with non-uniform temperature distribution is bounded by those with a uniformly distributed temperature. (a) Low to high temperature profile. (b) High to low temperature profile. (c) Parabolic temperature profile. (d) V shape temperature profile. (e) Inverse V shape temperature profile.

the temperature is decreasing from the cathode end, atomic flux at the cathode is retarded by the stress gradient, similar to "t6" and "t8" in Figure 9 (b). Therefore we propose to use J_{steady} to bound the atomic flux at the cathode of the interconnect subject to a non-uniform temperature distribution. However, the stress at the anode is not free (Figure 9), which does not satisfy the boundary condition for J_{steady}. We instead propose to use half length (the half from the cathode) of the interconnect to calculate J_{steady}, as Figure 9 shows that the stress at the middle of the interconnect is close to zero most of the time during the EM process. Finally the bounding temperature is determined in such a way that the atomic flux due to electron wind at this temperature is equal to the calculated J_{steady}.

Temperature gradient at cathode	Lower Bound	Upper Bound
Increasing temperature in the current direction	$T_{lb} = T_{cathode}$	$\beta\left(T_{ub}\right)\left(\alpha(T_{ub})\right) = \frac{\int_{-l}^{-\frac{l}{2}} \alpha(T(x))dx}{\int_{-l}^{-\frac{l}{2}} \frac{1}{\beta(T(x))}dx}$
Decreasing temperature in the current direction	$\beta\left(T_{lb}\right)\left(\alpha(T_{lb})\right) = \frac{\int_{-l}^{-\frac{l}{2}} \alpha(T(x))dx}{\int_{-l}^{-\frac{l}{2}} \frac{1}{\beta(T(x))}dx}$	$T_{ub} = T_{cathode}$

TABLE I

PROPOSED BOUNDING TEMPERATURES FOR VOID GROWTH IN AN INTERCONNECT WITH LENGTH l SUBJECT TO A NON-UNIFORM TEMPERATURE DISTRIBUTION. T_{lb} IS THE LOWER BOUND. T_{ub} IS THE UPPER BOUND. $T(x)$ IS THE TEMPERATURE PROFILE. $x = -l$ AND $x = 0$ ARE THE LOCATIONS OF THE CATHODE AND THE ANODE, RESPECTIVELY (AS SHOWN IN FIGURE 8).

We would like to point out that since void growth at the cathode is only dependent on the atomic flux nearby, the temperature gradient near the cathode plays the major role in determining (enhancing or retarding) the void growth, while the temperature distribution far away from the cathode is not as important. This observation is verified by testing with various temperature profiles. (Due to space limitations, we cannot show them all here.)

Therefore, in the above discussion, we focus on the temperature gradient near the cathode without assuming any specific temperature distribution along the second half of the interconnect. Table I numerates the proposed formulas to calculate bounding temperatures for void growth with non-uniform temperature distributions, and only the temperature gradient near the cathode is used to choose the appropriate bounding formula. The void growth with different temperature profiles as well as those with uniform temperatures are compared in Figure 10. In these plots, the void growth with interconnect thermal gradients is closely bounded by the void growths with the proposed uniform bounding temperatures. Blindly using the average temperature to evaluate the EM lifetime will either overestimate (e.g. Figure 10 (a)) or underestimate (e.g. Figure 10 (b)) the void growth, let alone using the maximum temperature. Wachnik *et al.* [32] demonstrated that it is possible to construct an electromigration resistant power grid by using shorter interconnect segments because of the Blech effect [30]. This finding seems to imply that, under normal operating conditions, the critical void size causing EM failure should be at a similar order of magnitude as that of the saturation void size (e.g., as the case shown in Figure 8). Therefore, for increasing/decreasing temperature at the cathode, the upper/lower bound temperature could serve as a good estimation of the interconnect lifetime with a non-uniform temperature distribution.

Joule heating in an interconnect usually results in a symmetric temperature distribution with the maximum temperature in the middle, due to the much lower thermal resistance of the vias on both ends. Therefore, the symmetric temperature distributions along the interconnect are of more practical interest. The"parabolic" and "inverse V shape" temperature profiles shown in Figure 8 (a) are used to approximate this kind of temperature distributions. Interestingly, for these temperature distributions, as indicated by Figure 10 (c) and (e), even the upper bound temperature for void growth is lower than the average temperature. In the EM measurements of copper interconnects performed by Meyer *et al* [33], they considered the non-uniform temperature distribution due to self-heating, and tried to fit their measurements with Black's equation by using different temperatures (e.g. maximum, average, weighted average, via (minimum) temperature). They reported that the best fit temperature is strongly weighted to the via temperature. Their findings agree with our analysis presented here.

C. Effects of combined temporal and spatial temperature gradients

So far we have discussed the interconnect lifetime prediction under temporal and spatial temperature distributions separately. In practice, due to circuit activity variations, one might expect the spatial temperature distribution over an interconnect would change over time. The electromigration diffusion equation (Equation (2) or Equation (11)) can be extended to capture this situation by assuming that temperature T is a function of both time and interconnect location. However, we cannot obtain a closed form analytic solution in this highly complex scenario. Instead, we propose to combine the results we have found so far in the cases of temporal gradients only and spatial gradients only to estimate the interconnect lifetime subject to both temporal and spatial temperature gradients.

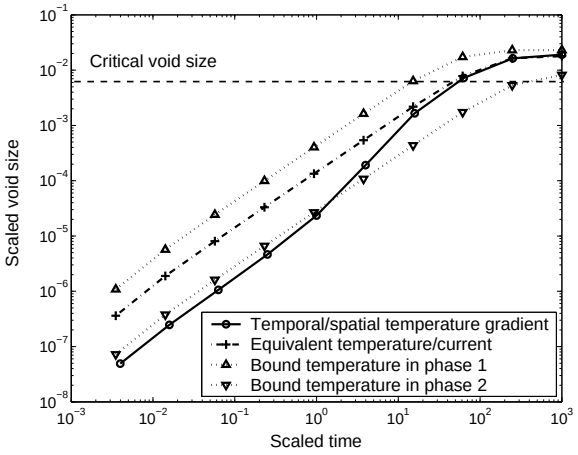


Fig. 11. Void growth subject to both temporal and spatial thermal gradients can be bounded by that using uniform temperature and current.

At time t, the temperature profile across an interconnect is denoted by $T(t, x)$, and the current density is $j(t)$. Using the formulas in Table I, we can find the bounding temperature at t, denoted by $T_b(t)$. Applying Equations (9) and (10) for the case of temporal temperature gradients to $T_b(t)$ and $j(t)$, we could find an equivalent uniform temperature and current to approximate the void growth subject to both temporal and spatial temperature gradients. Figure 11 shows one example. In this example, the interconnect Cu line is subject to two "parabolic" temperature profiles, with each one for half the time (i.e., 50% duty cycle), denoted by "phase 1" and "phase 2" in the figure.

We solve Equation (11) numerically with temperature being a function of both time and space, and we plot the void growth. This figure indicates that the void growth subject to the time varying temperature profile can be bounded by that using time invariant uniform temperature and current, as calculated according to the procedures proposed here. As a comparison, we also plot the void growth at the bound temperature of each temperature profile alone. If the critical void size is close to the saturation void size, as shown in the figure, one can use the calculated equivalent temperature and current to estimate the interconnect lifetime subject to both temporal and spatial thermal gradients, using Black's equation (Equation (1)). We have also testes this for other temperature profiles and duty factors, and the results are similar but are not presented here due to space limitations.

VI. Design time optimization considering runtime stress variations

In the traditional IC design flow, static and dynamic analyses are performed for the initial design to determine current loading information. Then this information is combined with the worst-case temperature to find those design points violating the reliability specification [34]. However, as we have shown above, using worst-case temperature is too conservative and could result in excessive design margins. Here we propose a design flow incorporating runtime stress information as shown Figure 12. In this design flow, the actual or projected current and temperature loads are fed into an accurate reliability model, such as the one proposed in this paper. We expect that the reliability projection from these models will generally enable more relaxed design constraints and provide a wider design space.

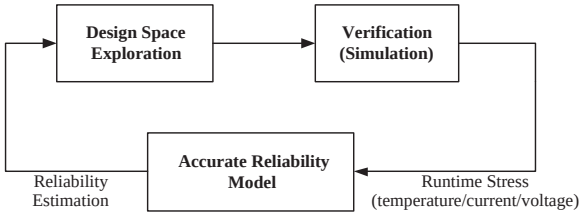


Fig. 12. A proposed design flow incorporating runtime stress information. [10]

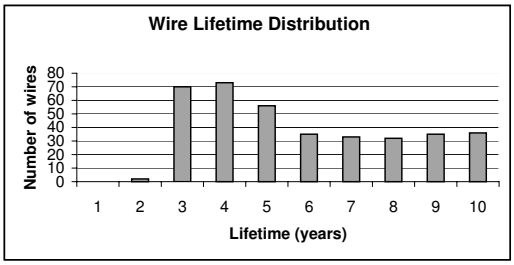
For instance, when temperature fluctuates within a relatively small range (e.g., 10^oC), our model predicts that using average temperature is good enough for reliability evaluation. Therefore, we could potentially reduce the number of design points falsely flagged for design rule violations when using the worst-case temperature. One example is illustrated in Figure 13 using data from a power grid design [35]. In this example, the worst-case temperature of a design is 135^oC, and Wang *et al.* [35] showed that there were a total of 372 wires violating the reliability requirement by using that worst-case temperature. However, if runtime stress information is available at design time, we can move some wires that are outside the specified reliability threshold (10 years of MTF at 135^oC in this example) into the reliable bins by re-calculating the lifetime distribution using our dynamic reliability model. Equivalently, we can shift the reliability threshold towards fewer years on the original wire lifetime distribution diagram. Using the results in Figure 7(c), we can estimate the benefits obtained, in terms of design margin reclamation, by considering runtime temperature fluctuations. These results are shown in Figure 13(b).

This example only illustrates some potential advantages in design optimization offered by our dynamic reliability model. As part of future work, we will integrate our model into existing reliability-aware design flows, such as the power grid optimization method proposed by Wang *et al.* [35].

VII. Runtime reliability-aware thermal management

Recently, many DTM techniques [4], [12], [13] have been proposed to ensure that a chip will never operate above some temperature threshold. However, these techniques do not explicitly study the effects of transient behaviors on system reliability, and instead implement a temperature upper-bound at the expense of degraded performance. By modeling lifetime as a resource to be consumed over time, we can manipulate chip lifetime directly at runtime.

High temperature limits the circuit performance directly by increasing interconnect resistance and reducing carrier mobility. However, it has been shown that ([4]) using DTM to compensate the temperature dependency of clock frequency induces very mild performance penalty. On the other hand, Banerjee *et al.* [1] showed that temperature induced reliability issue tends to limit the circuit performance in future technology generations. In the following discussion, we assume that the temperature threshold is set solely for reliability specification, and circuits can operate correctly above this threshold whenever allowed. Although extreme high temperature may cause immediate thermal damage for IC circuits, we study a range of operating temperatures only with long-term reliability impacts (i.e. temperature induced aging). High temperatures causing immediate or unrecoverable damage are assumed to be far above the range of normal operating temperatures studied here (e.g. the temperature used in accelerated EM test



(a)

Temperature variation (^oC)	New reliability threshold (years)	Number of wires reduced	Percentage reduction
5	9.06	33	8.8
10	8.34	59	15.9
15	7.73	79	21.2
20	7.24	95	25.5
25	6.85	107	29.8

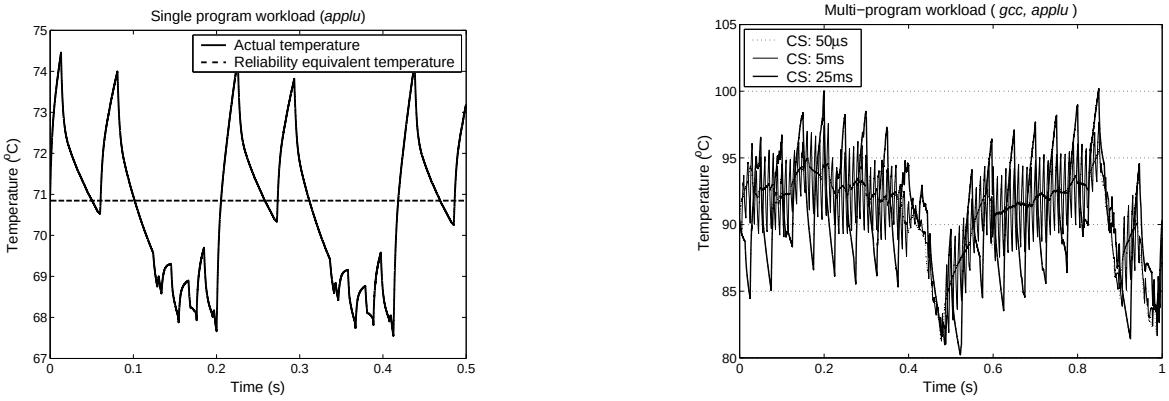
(b)

Fig. 13. (a) Distribution of wires violating the MTF specification using maximum temperature (data extracted from [35]) with a total of 372 wires. (b) Reduction of the number of wires violating the MTF specification under different temperature variations (maximum temperature: 135^oC). [10]

is usually around 200^0C [32]). A monitoring and feedback mechanism is implemented at runtime to ensure that circuits operate well below such temperatures.

A. Lifetime banking opportunities

Due to activity variations, the power consumptions of on-chip components (i.e. caches, FP/INT units, branch predictor, etc.) are not constant. Therefore, there exists not only chip-wise spatial temperature gradients but also temporal temperature gradients for each component.



(a) (b)

Fig. 14. Temporal temperature variation. (a) Single program workload. (b) Two-program workload with context switching. [14]

Figure 14 depicts the temperature profiles for two different workloads that are commonly seen in general purpose computing. Figure 14(a) represents a single program workload and Figure 14(b) represents a multi-program workload with context switching. In the single program workload, temperature changes over time due to the phased behavior in the executed program. In the multi-program workload, besides the execution variations within each program, inter-program thermal differences also affect the overall thermal behavior of the workload. For example, in Figure 14(b), the workload is composed of one cold program (*applu*) and one hot program (*gcc*). Thus the temperature fluctuation in Figure 14(b) is quite different with various context switching intervals. Though there are different thermal behaviors

for different workloads, one can still find some common characteristics as compared with server workloads, which we will discuss in Section VII-D. In Figure 14, temperature variations occur in a manner with small granularity in both magnitude and time interval. More formally, the temperature profile can be decomposed into a constant temperature component (average temperature) and a high frequency component. The analysis in Section IV reveals that the constant temperature component in the temperature profile is approximately equal to the reliability equivalent temperature, as shown in Figure 14(a). It is the high frequency component that provides opportunities for lifetime banking. When the actual temperature is under the reliability equivalent temperature, the lifetime is consumed with a slower speed, which allows subsequent execution above the reliability equivalent temperature.

B. Dynamic Reliability Management Based on Lifetime Banking

In Section III, we derived the lifetime model for electromigration subject to dynamic stresses (Equation (9) and (10)). Considering void growth limited failures such as those in dual-damascene Cu interconnects [30], let current exponent $n = 1$, we can rewrite MTF by combining Equation (9) and (10) as:

$$T_f \propto \frac{1}{E\left[j(t)\left(\frac{exp(\frac{-Q}{kT(t)})}{kT(t)}\right)\right]}$$

Or equivalently, by eliminating the expected-value function, one can express the MTF in an integral form:

$$\int_0^{T_f} j(t)\left(\frac{exp(\frac{-Q}{kT(t)})}{kT(t)}\right)dt = D \tag{12}$$

where D is a constant determined by the structure of the interconnect. Equation (12) models interconnect time to failure (i.e., interconnect lifetime) as a resource consumed by the system over time. Function $r(t) = \left[j(t)\left(\frac{exp(\frac{-Q}{kT(t)})}{kT(t)}\right)\right]$ can be regarded as the consumption rate. In DSM copper technology, void growth failure (e.g. at vias) is the major EM induced failure mechanism [36], and $r(t)$ can be regarded as the void growth rate (i.e. the atom drift rate at the cathode) in this case. Equation (12) provides a model to capture the effect of transient behaviors on system lifetime. One interesting case is when $j(t) = 0$, which occurs when the system is inactive as commonly seen in systems with non-server, user-driven workloads. When this happens, the atomic flux becomes zero while the effect of the back-flow diffusion near the cathode created by the EM atomic flux in active periods is worth careful examination. If the inactivity happens in the early stage of the void growth, the back-flow diffusion is negligible and the void simply stops growth during the power-down periods. If the back-flow diffusion is comparable to the normal EM atomic flux, which happens at a very long time after EM begins (e.g. at the order of several years). This back-flow diffusion tends to reduce the void size at the inactivity periods by refilling the void with atoms. However, in order to have significant impact on the void size already formed, this healing process has to last for a duration comparable to the time it took to grow to the current void size, e.g. several years. The inactivity period in normal usage is usually much less than this time scale. Therefore, the void size is essentially unaffected in inactivity. Our simulations confirm this observation and more detailed discussion on this aspect is out of the scope in this paper. In summary, the void size remains unchanged during the inactive period if the inactive period is much less than the total active time. Equation (12) accurately models this phenomenon by specifying $r(t) = 0$ during the inactive periods.

Ideally, we would like to monitor the temperature and current for each individual interconnect to build an exact full chip reliability model. In practice, only a limited number of temperature sensors are available on die, and a detailed and complex full chip reliability model is not suitable for runtime management due to the computation overhead. In this study, we use the maximum temperature measured across the chip at runtime, together with the worst-case current density specified at design time, to calculate the dynamic consumption rate. This is a conservative but safe approach. Thus, the results obtained in this study provide a lower bound for the potential benefits delivered by the proposed DRM method. Further refinement of the full chip reliability model will be part of future work. When DVS is applied, the worst-case current density in the IC interconnects should be scaled according to the voltage/frequency setting used. The relationship between current density, supply voltage and clock frequency can be modeled by transferred charges per clock cycle [37]: $j \propto \frac{CV}{T} = CVf$, where C is the effective capacitance.

When a chip is designed, usually an expected lifetime (e.g., 10 years) is specified under some operating conditions (e.g., temperature, current density, etc.). We use $r_{nominal}$ to denote the lifetime consumption rate under the nominal conditions (e.g. reliability constrained temperature threshold). During runtime, we monitor the actual operating conditions regularly, calculate the actual lifetime consumption rate $r(t)$ at that time instance, and compare the actual rate with the nominal rate $r_{nominal}$ by calculating $\int(r_{nominal} - r(t))dt$, which we call the "lifetime banking deposit". When $r(t) < r_{nominal}$, the chip is consuming its lifetime slower than the nominal rate. Thus, the chip's lifetime deposit is increased. When $r(t) > r_{nominal}$, the chip is consuming its lifetime faster than the nominal, and

the lifetime banking deposit will be reduced. According to Equation (12), as long as the lifetime deposit is positive, the expected lifetime will not be shorter than that under the nominal consumption rate $r_{nominal}$. Figure 15 illustrates this S_DRM technique. For example, in the interval $[t0, t1]$, the reliability of the chip is banked, while in $[t1, t2]$, the banking deposit is consumed. At time instance $t2$, the banking deposit becomes less than some threshold, and a cooling mechanism has to be engaged to quickly pull down the lifetime consumption rate to the nominal rate, just as is done in conventional DTM techniques. In other words, our S_DRM technique adopts DTM as a bottom-line guarding mechanism.

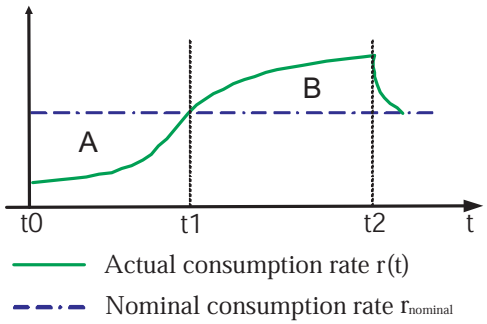


Fig. 15. Simple dynamic reliability management (S_DRM). [14]

Therefore, the difference between conventional DTM and our S_DRM lies in the case where the chip's instantaneous consumption rate is larger than its nominal rate. In DTM, the lifetime consumption rate is never allowed to be larger than the nominal. In S_DRM, before we engage thermal management mechanisms we first check to see if the chip currently has a positive lifetime balance. If enough lifetime has been banked, the system can afford to run with a lifetime consumption rate larger than the nominal rate. Otherwise, we apply some DTM mechanism to lower the consumption rate, thus preventing a negative lifetime balance. In this study, we use dynamic voltage/frequency scaling as the major DTM mechanism. Since S_DRM only needs to monitor the actual lifetime consumption rate and to update the lifetime banking deposit, the computation overhead is negligible compared to that of DTM.

C. Experiments and analysis for general-purpose computing workloads

1) Experimental set-up: We run a set of programs from the Spec2000 benchmark suite on a processor simulator (SimpleScalar [38]) with the characteristics similar to a 0.13μm Alpha 21364. We simulate each program for a length of 5 billion instructions, and obtain both dynamic and static (leakage) power traces, which are fed as inputs to a chip-level compact thermal model *Hotspot* [4] for trace-driven simulation. In our trace-driven simulations, we include the idle penalty due to frequency/voltage switching, which is about 10us in many real systems [4]. Furthermore, since leakage power is strongly dependent on temperature, we scale the leakage power trace input dynamically according to the actual temperature obtained during runtime, using a voltage/temperature-aware leakage model [39]. Since the *Hotspot* model is highly parameterized, one can easily run experiments on a simulated processor with different thermal package settings. In order to obtain meaningful results, one should carefully choose the initial temperature setting for the *Hotspot* model. For each new thermal package setting, we obtain its initial temperatures by repeating the trace-driven simulations until the steady temperatures of the chip are converged, as suggested in [4].

We implement both DTM and S_DRM in the *Hotspot* model and set 110°C as the temperature threshold for both runtime management techniques. Both schemes use a feedback controlled dynamic voltage/frequency scaling mechanism to guard the program execution. For example, in DTM, when the actual temperature is above a certain temperature threshold, a controller is used to scale down the frequency/voltage, ensuring the program will never run at a temperature higher than 110°C. Our S_DRM scheme uses 110°C as the nominal temperature for the lifetime consumption rate. If the program never runs at a temperature less than that of the nominal (i.e., without banking opportunity), our S_DRM scheme will perform the same as thermal threshold-based DTM as the DTM policy is always engaged. On the other hand, if the program never exceeds the nominal temperature with full CPU speed, neither mechanism is engaged. Finally, we record the simulated execution times for fixed length power traces as the system performances under the two runtime management techniques, and use "performance slow-down" , defined as

$$\frac{(\text{simulated time w/ runtime management} - \text{simulated time w/o runtime management})}{\text{simulated time w/o runtime management}}$$

as the metric to compare both techniques.

2) Single-program workload: Figure 16 plots the dynamic process for both conventional DTM and the proposed S_DRM techniques for benchmark *gcc*. The feedback controller in DTM effectively clamps the temperature within the target temperature (110^oC) by oscillating the clock frequency between 1.0 and 0.9, resulting in a reliability equivalent temperature less than that, and causing unnecessarily frequent clock throttling (Figure 16 (a)). On the other hand, our S_DRM technique can exploit reliability banking opportunities during the cool phase, and delay the engagement of throttling, while maintaining the specified reliability budget, as proved by the reliability equivalent temperature shown in Figure 16 (b).

Figure 17 shows the performance penalty for both DTM and S_DRM with the same thermal configuration. Only those benchmarks subject to performance penalties due to runtime management are shown here. As clearly indicated in the figure, performance penalty with the S_DRM scheme is always less than that with DTM scheme, when the thermal configuration is the same. On average, the S_DRM technique reduces the performance penalty by about 40% of that due to DTM (from 7% to 4%). Also shown in the figure is the performance of DTM with a more expensive thermal package whose convection thermal resistance is only one third of the other's. As one can expect, a more expensive thermal package can reduce the performance penalty. Figure 17 shows that, on average, S_DRM with a higher thermal resistance can achieve a performance very close to that of DTM with a lower thermal resistance. These results imply that, if the tolerable performance lost is fixed, the application of S_DRM allows the usage of a much cheaper thermal package than that required by the conventional DTM technique.

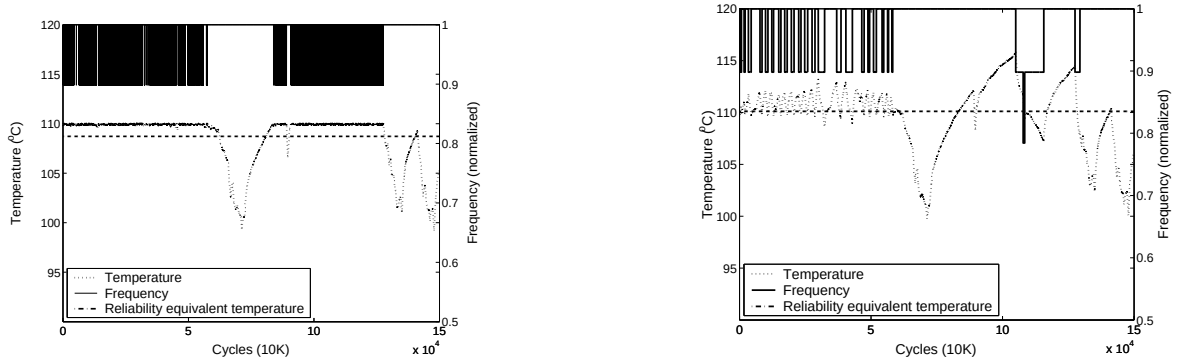


Fig. 16. Temperature and clock frequency profiles in different thermal management techniques for benchmark *gcc*. (a) Conventional DTM (threshold temperature =110^oC). (b) Reliability banking based DRM (reliability target temperature =110^oC).

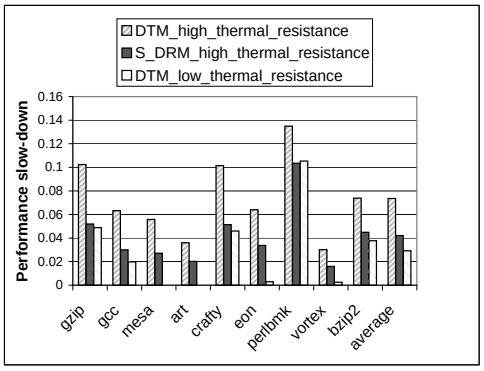


Fig. 17. Performance comparison of DTM and the proposed S_DRM. The results for S_DRM are based on high convection thermal resistance configuration. The results for DTM include two different thermal configurations. [14]

In addition, using the S_DRM technique, one can explicitly trade-off reliability with performance by targeting different lifetime budgets. That is one can increase the nominal lifetime consumption rate when lifetime target is allowed to be reduced. Figure 18 plots the performance of S_DRM averaging over all benchmarks at different lifetime budgets, with shorter expected lifetimes enabling faster execution. However, reducing lifetime by 10% only increases the performance by about 1%.

When compared with the conventional thermal threshold-based DTM, a distinct feature of S_DRM is its ability to "remember" the effects of previous behaviors. If the lifetime balance is high due to previous deposits, S_DRM

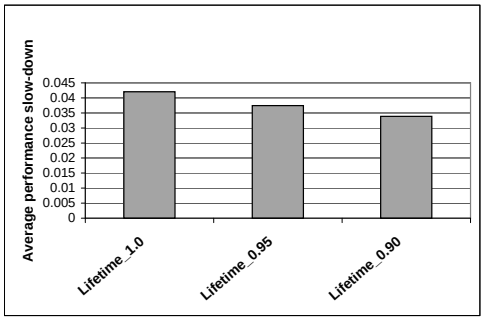


Fig. 18. S_DRM performance at different targeted lifetimes. [14]

will be more tolerant of higher operating temperatures for longer time intervals, thus reducing the performance penalties due to conventional DTM slow-down mechanisms. In summary, the advantage of S_DRM over DTM is largely dependent on the inherent variations in the temperature profile of the workload.

3) Multi-program workload: Another interesting program execution scenario is a workload of multiple programs with context-switching between them. When a hot benchmark and a cold benchmark are executed together, the average operating temperature should be between the individual benchmarks' operating temperatures. For example, *gcc*'s own operating temperature is around 115^oC and *applu*'s is around 70^oC. Figure 14(b) plots the temperature profile of a hybrid workload composed of *gcc* and *applu*, with different context-switch time intervals. Note that, in our simulation, the multi-program workload is constructed using the power trace of the individual program, and the overhead of context-switching is not modeled and simulated. Since we are only interested in the relative performance of different runtime management technique, such simplification should not affect the final conclusions.

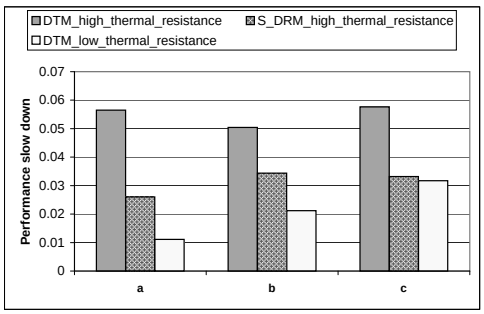


Fig. 19. Average performance comparison of DTM and DRM on a multi-program workload with different context-switch intervals ((a) 50μs, (b) 5ms, and (c) 25ms). [14]

As one expects, the smaller the context-switch interval, the less temperature fluctuation, with the thermal package of the chip working as if a low-pass filter. When the context-switch interval is increased, individual benchmarks can show their hot/cold properties, and the temperature variation in the workload becomes obvious. In order to investigate how multi-program workloads affect the performance of DTM and DRM, we reduced the temperature threshold of the targeted lifetime from $110°C$ to $90°C$. Figure 19 shows the performance penalties of DTM and S_DRM for this multi-program workload with different context-switch intervals. We observe a similar trend shown in the single-program workload. S_DRM outperforms DTM with the same thermal package configurations. As the context-switch interval increases, the performance of S_DRM becomes closer to that of DTM with a much smaller convection thermal resistance (three-fold smaller).

D. Dynamic Reliability Management for Server Workloads

We have investigated the application of banking based DRM in workloads for general-purpose computing. As we will see, the disadvantages of this technique become more obvious in server systems such as web servers, in which hot phases usually imply an increased number of service requests while the engagement of active cooling mechanisms then exacerbate the QoS provided by the server. In the following, we first discuss some distinct characteristics of server workloads in terms of both thermal behaviors and performance requirements. We then propose a profile-based dynamic reliability management (P_DRM) technique that can extract more benefits from lifetime banking for those server workloads.

1) Characteristics of server workloads: In general-purpose computing, the temperature variations of workloads are largely due to the inherent phased behaviors (i.e. phased activities or context-switches). These variations usually occur in a very small-scale time interval, which is comparable to the thermal constant of chip thermal package. Server workloads, in contrast, are dependent on user requests, which vary with a much larger time scale with tens of hours [40], [41]. There are several distinct characteristics in the server workloads like those presented in [40], [41]. First, there is clearly a cool phase (lower request rate) and a hot phase (higher request rate) in the workload distribution. For example, in a workload trace for the 1998 Winter Olympic Games, the request rate increases from around 50 (req./s) in the cool phase to above 400 (req./s) in the hot phase, an eight-fold difference. Second, as a consequence of the workload and associated processor utilization variation, the power consumption of the processor varies greatly (a two-fold difference in the Olympic Games servers), which implies a large variation in temperature. Third, each phase sustains for a very long time interval. Thus, each phase reaches its steady-state temperature and stays at that temperature for most time of the phase interval. This is quite different from general-purpose computing, where the interval for each thermal phase is very short and the steady-state temperature is seldom reached before the next phase arrives. These distinct thermal characteristics make our lifetime-banking-based reliability management promising for server workloads.

In the hot phase, conventional thermal threshold-based DTM clamps the maximum temperature to a predefined threshold by slowing down the processor, thus possibly exacerbating the situation. In contrast, banking-based runtime management can exploit the banking effects of the long cool phase and delay or reduce the performance loss due to engagement of a cooling mechanism. From an average user's point of view, the QoS provided by the server is largely dependent on its performance in the hot phase, as most requests are made during that time. Therefore, in our study, *we use the performance of the hot phase as our performance metric for comparison*.

2) Dynamic reliability management for server workloads: In order to evaluate our runtime management technique on server workloads, we construct a hybrid workload in a way similar to that of the multi-program workload, but with a much longer context-switch interval. This synthetic workload is composed of a cool phase and a hot phase, running Spec2000 benchmarks *applu* and *gcc* respectively. Figure 20 shows the temperature profile of the synthetic workload we use to mimic the thermal behavior of server workloads. From various experiments, we find that the thermal time constants of our simulated system are in the range of tens of milliseconds. Therefore, by simulating workloads in a time scale of several seconds, we can ensure that the portion of time in the profile spent on the transient behaviors from one phase to another is minimized, just like one may see in a temperature profile for server workloads. Although the total simulated time is short (i.e., about one second) compared to a real server workload, Figure 20 indicates that the time interval for each phase is long enough to reach the steady-state operating temperature of the individual program. The temperature variations within each program also mimic the workload variations in both the cool and hot phases of a real server workload. Therefore, the time units shown in Figure 20 could be interpreted as scaled down from a much longer time interval (e.g. several hours). One disadvantage of our synthetic workload is that power peaks due to individual requests in the cool phase are not modeled. However, the effect of those intermittent power peaks on reliability banking is not significant, because of the filtering effect of the thermal package on temperature.

In our synthetic workloads, the cool phase is followed by the hot phase, and lifetime will be banked first and then withdrawn. In other workloads where the hot phase is followed by the cool phase, DTM can be applied in the hot phase if there is no previous lifetime banking, and lifetime will be banked in the following cool phase and prepared for withdrawal in the future hot phase. Thus, our lifetime banking based approach is effective in spite of the detail of workloads (i.e. the order of cool and hot phases). We define the *duty cycle* of the cool phase as the portion of time the cool phase occupies in the whole length of simulation. For example, in Figure 20, the duty cycle of the cool phase is equal to 0.5. In our experiments, we also construct workloads with different duty cycles of the cool phase (e.g., 0.6 and 0.75), and in all of these workloads, individual programs reach their own steady-state temperatures.

The application of the S_DRM technique to server workloads is straightforward, just like in the context-switched multi-program workload studied previously. However, our simulation results reveal that S_DRM is not the best choice for server workloads. In the workloads of general-purpose computing, since each phase is very short, the lifetime balance deposited in the previous cool phases can support the subsequent over-consumption of lifetime for an interval comparable to that of the hot phase. S_DRM can minimize the impacts on the phase within these workloads. However, in the server workloads, the interval of the hot phase is much longer, and temperature rises steadily towards the hot phase steady-state temperature. At the same time, due to the exponential dependence of lifetime consumption rate on temperature, the lifetime balance is consumed more and more rapidly, despite a previous long cool phase. Figure 21 demonstrates such a process in the time interval $[0.6s, 0.68s]$. After 0.68s, the lifetime balance becomes *zero*. S_DRM performs during the rest of the hot phase just as it behaves in the single program workload. Therefore, only a small portion of the execution in the hot phase benefits from the lifetime banking by the cool phase.

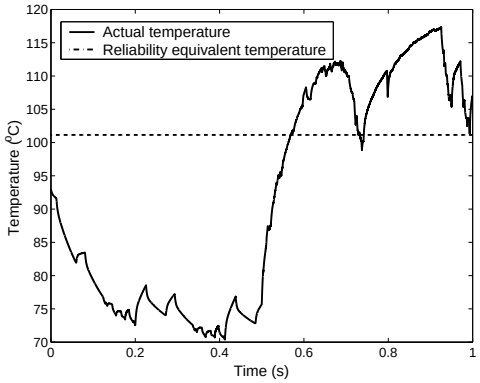


Fig. 20. A constructed workload used to mimic the thermal behavior of server workloads. [14]

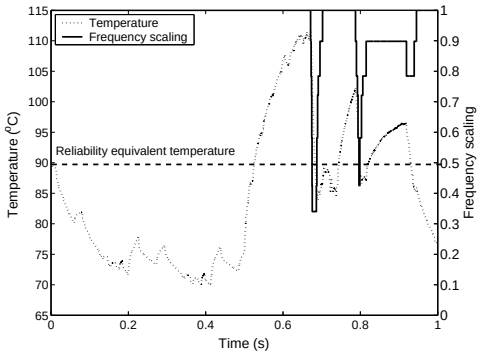


Fig. 21. S_DRM (simple dynamic reliability management) on the synthetic workload shown in Figure 20.

Due to the above reason, one should find a more strategic way to spend the lifetime balance in order to maximize the performance in the hot phase. Since in steady state, temperature can be modeled as a function of the operating frequency, one can find the relationship between lifetime consumption rate and operating frequency. Let $f(t)$ denotes the operating frequency curve in the hot phase, and $r(f(t))$ be the corresponding lifetime consumption rate. The problem to find the maximum performance operating scheduling while satisfying the reliability constraint can be formulated as a constrained optimization problem as follows.

$$Max(E[f(t)]), \text{ subject to } E[r(f(t)] = R,\ t \in \text{hot phase}$$

where $E[\,]$ is the expected-value function, and R is a constant in the hot phase that is determined by the lifetime balance deposited during the cool phase as well as the nominal lifetime consumption rate. We assume that, in the hot phase, system performance is proportional to the clock speed.

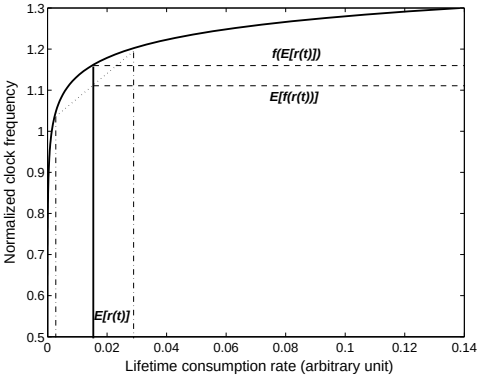


Fig. 22. Relationship between clock frequency and lifetime consumption rate.

Figure 22 plots clock frequency as a function of the lifetime consumption rate. It is obvious that the relationship

between clock speed and lifetime consumption rate forms a convex curve. According to Jensen's inequality, it follows that (as shown in Figure 22) $f(E[r(t)]) \geq E[f(r(t))]$, which implies that, in order to obtain the best performance, one should operate with a constant consumption rate. In other words, one should distribute the lifetime balance evenly across the hot phase. In order to calculate the desired consumption rate in the hot phase, one has to know the duration of the hot phase. Currently we assume that this information can be obtained through profiling technique thanks to the high regularity of the workload distribution for servers.

With the optimal operating condition in mind, we introduce our (P_DRM, profile-based dynamic reliability management) technique, which is a natural extension of our S_DRM with the awareness of the optimal operating points in the hot phase. When the server is running in the cool phase, P_DRM works the same way as S_DRM with lifetime balance banked. When the server enters the hot phase, P_DRM calculates a new nominal lifetime consumption rate based on the lifetime balance and the duration of the hot phase (obtained through profiling). Then P_DRM acts just like S_DRM, with the new calculated nominal consumption rate, which can further exploit some banking opportunities due to temperature variations within the hot phase.

The profiling only provides a *prediction* that allows the CPU to jump to the best operating point during a hot phase. In some cases we might not be able to obtain accurate workload profiles. However, with our P_DRM technique, the inaccuracy of workload profiles only affects the the performance optimality, and does not result in violations to the lifetime budget. That is because our technique always tracks the actual reliability consumption rate and compares it with the nominal lifetime consumption.

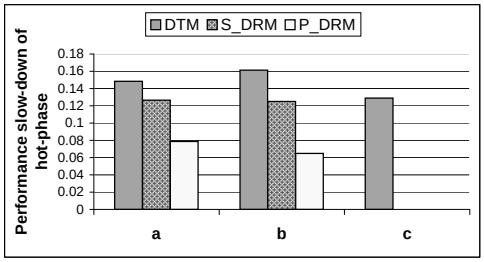


Fig. 23. Performance comparison of different runtime management techniques on the synthetic workload shown in Figure 20 with different duty cycles of the cool phase: (a) 0.5, (b) 0.6 and (c) 0.75. [14]

3) Simulation results for server workloads: We simulate the synthetic workload shown in Figure 20, which mimics the thermal behaviors of the real server workload, with different runtime management techniques. We change the program switching time so that we can test on 3 workloads with different duty cycles of the cool phase. We compare the performance slow-down in the hot phase and the results are presented in Figure 23. Both DRM techniques outperform DTM, and P_DRM performs the best. The performance of S_DRM is slightly better than that of DTM and much worse than P_DRM due to the reasons discussed in above. On the other hand, P_DRM can fully exploit the banking benefits of the cool phase. For example, when the cool phase occupies 60% of the total time (i.e. as indicated by (b) in Figure 23), P_DRM can reduce the performance penalty from 16%(DTM) to only 6% (or equivalently, the execution speed of the hot phase is increased by P_DRM by about 9.5% over DTM). Interestingly, for the case when the cool phase occupies 75% of the total time (i.e., (c) in Figure 23), no performance slow-down is incurred for both DRM techniques, because the reliability equivalent temperature for that workload is less than the reliability nominal temperature. Thus, in that case, the lifetime balance banked in the cool phase is enough to support the full speed execution in the hot phase, while DTM clamps the hot phase temperature to the reliability temperature, resulting in about a 13% performance penalty in the hot phase.

4) An analytical model for P_DRM for server workloads: In order to fully understand the potential benefits of p_DRM on server workloads, we present a first order analytical model, providing some insights of our proposed runtime techniques. In this model, we approximate server workloads using square waveforms as shown in Figure 24. The solid blue line represents the temperature/performance profile with DTM. The temperature profile with P_DRM in the cool phase overlaps with that of DTM. And P_DRM allows operating points above the reliability temperature in the hot phase, as presented by the dotted green line in the figure. We want to find out what is the allowable performance difference between the dotted green line and the solid blue line (i.e., the performance gain of P_DRM over DTM), subject to a fixed lifetime budget. Here we make an assumption that the processor can operate at a clock frequency higher than that clamped by the thermal threshold. There are two aspects to this. First, temperature excursions will require a reduction in frequency, thus reducing performance somewhat, but should still outperform a strict, temperature-limited form of DTM because the temperature dependence of frequency is mild [4]. Second, many ICs are actually under-clocked due to thermal limitations. In both cases, there exist possibilities that we can over-drive the processor in the hot phase to meet the QoS requirements without sacrificing reliability lifetime.

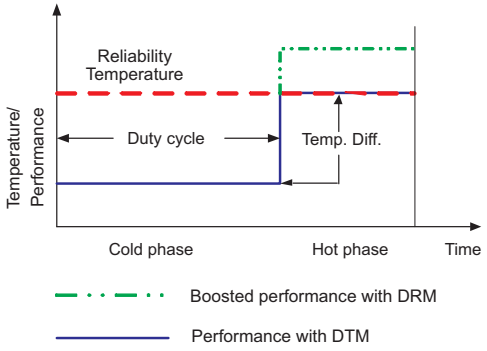


Fig. 24. Modeling thermal behaviors of server workloads using square waveforms. [14]

As one can see from Figure 24, two factors might affect the potential performance boost by P_DRM over DTM: 1. the difference between the steady-state temperatures in both the hot phase and the cool phase, and 2. the duty cycle of the cool phase. We set the reliability temperature $T_n = 105^oC$, associated with the clock frequency $f_n = 3.0GHz$. This setting means that in the hot phase, the maximum performance achieved by DTM is to operate at $3.0GHz$. If we can assume that the dynamic power consumption of the processor is proportional to the cubic of clock frequency, the steady state temperature can be denoted by $T(f) = K_f f^3 + T_0$, where K_f is a constant and T_0 represents the ambient temperature of the thermal package. Accounting for the contribution of static power consumption to temperature, we set a higher ambient temperature $T_0 = 55^oC$, and obtain $K_f = 1.85K/GHz^3$. Let ΔT denote the temperature difference between the hot phase and the cool phase, f_2 the allowable operating clock frequency in the hot phase by P_DRM, and α the duty cycle of the cool phase. The following equation should be satisfied to retain the same lifetime budget with P_DRM:

$$[r_n(T_n) - r_1(T_n - \Delta T)]\alpha = [r_2(f_2, T(f_2)) - r_n(T_n)](1 - \alpha) \quad (13)$$

where r_n is the nominal reliability consumption rate at temperature T_n, r_1 is the consumption rate in the cool phase, and r_2 is the consumption rate in the hot phase with clock frequency f_2 and temperature $T(f_2)$. The left hand side of the above equation represents the reliability balance banked during the cool phase and the right hand side represents the banking deposits to be consumed in the hot phase. Although the temperature dependence of static power is not taken into account in this model, we feel that it captures the key relationships between performance, operating temperature and reliability consumption rate, and is thus sufficient for our purposes.

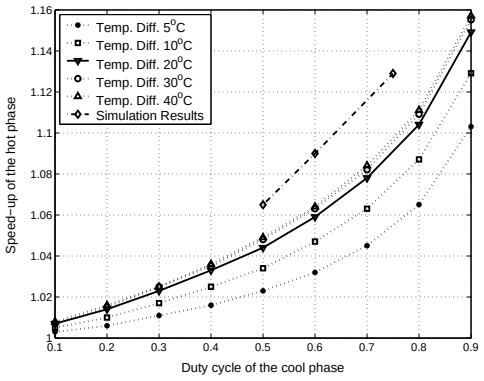


Fig. 25. Performance speed-up due to lifetime banking on different workload characteristics. [14]

Using the above analytical model (i.e. Equation (13), we can calculate the performance speed-up by P_DRM (i.e. $\frac{f_2}{f_n}$ in the hot phase) as a function of ΔT and the duty cycle of the cool phase. The results are presented in Figure 25, which shows that the performance speed-up is highly dependent on the duty cycle of the cool phase. When the duty cycle of the cool phase is fixed, the increase of temperature difference will also increase the speed-up. However, after some value (e.g. about 20^oC), the temperature difference has a minor effect on the speed-up, due to the exponential dependence of the reliability consumption rate on temperature. Because the extra reliability balance brought by further lowering the temperature in the cool phase is negligible when compared to the very

high consumption rate in the hot phase. This figure suggests that the "sweet spot" for performance speed-up with P_DRM lies in the case when the duty cycle of the cool phase is more than 50% and the temperature difference is more than 20^oC, and we can expect more than 5% of performance speed-up. Fortunately, as shown before, many server workloads satisfy these requirements.

The simulation results of DTM and P_DRM from Figure 20 are re-plotted in Figure 25. The workloads for these data are similar to that shown in Figure 20, with the cool phase duty cycle equal to 0.5, 0.6 and 0.75 respectively. The reliability temperature is set to 90^oC, while the temperature of the cool phase in these workloads is about 70^oC. These simulation results show a similar trend to that predicted by our simple analytical model, though our analytical model is not calibrated against any specific simulation data. Therefore, these simulation results confirm the applicability of our analytical model. Compared with the simulation results, it seems that the analytical model underestimates the performance speed-up by P_DRM. Two major reasons might help explain the discrepancy. First, in our analytical model, we use a cubic relationship between power and operating frequency, which exaggerates the effect of clock frequency on the temperature, leading to a more conservative estimate of the performance speed-up. Second, in the simulations, we include the idle penalties for frequency/voltage transitions due to dynamic frequency/voltage scaling, while in the analytical model, we do not assume any extra performance penalty for DTM.

VIII. Conclusions

This paper presented an analysis of interconnect EM failures subject to temporal and spatial thermal gradients. For EM under time-varying stresses (temperature/current), we proposed a dynamic reliability model, which returns reliability equivalent temperatures/currents. For EM under non-uniform temperature distributions, we obtained close bounding temperatures to estimate the actual lifetime. Therefore, the commonly used Black's equation is still applicable by using our constant reliability equivalent temperatures. Our analysis reveals that blindly using the maximum or average temperature to evaluate EM lifetime is inappropriate. Our results not only increase the accuracy of reliability estimates but enable designers to more aggressively explore the design space and to reclaim the design margin imposed by less accurate, more pessimistic models. Existing constant-temperature models require designers to observe a static worst-case temperature limit, but the analysis presented here enables temperature-aware designers to evaluate the system reliability using runtime information, thus increasing the confidence about the actual behavior of the system. The dynamic nature of our reliability model also makes it suitable for DTM.

As an application example, we detailed the use of the temperature variability and lifetime resource models to develop novel DRM techniques that reduce the performance penalties associated with existing DTM techniques while maintaining the required expected IC reliability lifetime. When the operating temperature is below a nominal temperature (i.e., the threshold temperature used in DTM techniques), lifetime is being consumed at a slower than nominal rate, effectively banking lifetime for future consumption. A positive lifetime balance allows the nominal temperature to be exceeded for some time (thus consuming lifetime at a faster than nominal rate) instead of automatically engaging DTM and unnecessarily suffering the associated performance penalties. For general-purpose computing, simulation results revealed that S_DRM provides performance improvements over traditional threshold-based DTM without sacrificing expected lifetime, or allows the usage of cheaper thermal package without sacrificing performance. For server workloads, simulations on synthetic workloads demonstrate the possibility to increase server QoS by using P_DRM when service requests are aggregated. A conservative analytical model further identifies the "sweet spots" of server workloads that benefit from our P_DRM. Although the DRM experiments presented in this paper do not explicitly study the scenarios with long periods of inactivity, which are commonly seen in non-server, user-driven workloads, our lifetime banking techniques can be applied in a straightforward way, because our dynamic reliability model (Equation (12)) also holds true in these situations. Consequently, much better performance gains would be expected compared to those obtained in the server style workloads presented here, because more lifetime banking opportunities are available during those inactivity periods.

In the future, we will compare our analysis with experimental data. We will also investigate other dynamic reliability models by considering such failure mechanisms as fast thermal cycling, stress-migration, and dielectric/gate oxide breakdown. Finally, we will integrate the dynamic reliability models into a reliability-aware design flow.

IX. Acknowledgements

This work is supported in part by NSF grants CCR-0133634 (Career), CCR-0306404, CCF-0105626, CCF-0429765, and CNS-0410526; Army Research Office grant W911NF-04-1-0288; a University of Virginia FEST Award; the Woodrow W. Everett Jr., SCEEE Development Fund in cooperation with the Southeastern Association of Electrical Engineering Department Heads; two research grants from Intel MRL; and an IBM Faculty Partnership Award. We thank Karthik Sankaranarayanan for his help with power trace extraction.

APPENDIX

Consider an EM process with zero stress at both ends of the metal line and with non-uniform temperature distribution across the metal line. In the steady state, the mechanical stress along the line reaches its steady distribution, or $\frac{\partial \sigma}{\partial t} = 0$. And the following equation is hold (from Equation (3)):

$$\nabla J = \frac{\partial}{\partial x}\left(\beta\left(T(x)\right)\left(\frac{\partial \sigma(x)}{\partial x} - \alpha(T(x))\right)\right) = 0$$

with the boundary conditions: $\sigma(0,t) = \sigma(l,t) = 0$, where $T(x)$ is the temperature profile.

Therefore, $\beta\left(T(x)\right)\left(\frac{\partial \sigma(x)}{\partial x} - \alpha(T(x))\right) = J$, where J is a constant and represents the steady state atom flux. Thus, in the steady state of EM process, there exists a constant atomic flux from one end (cathode) of the metal line to the other (anode). It follows that $\left(\frac{\partial \sigma(x)}{\partial x} - \alpha(T(x))\right) = \frac{J}{\beta(T(x))}$. By integrating both sides of this equation along the metal line, and noticing that $\sigma(0) = \sigma(l) = 0$, we obtain $-\int_0^l \alpha(T(x)dx = J\int_0^l \frac{1}{\beta(T(x))}dx$. And finally the steady state atomic flux can be expressed as:

$$J = -\frac{\int_0^l \alpha(T(x))dx}{\int_0^l \frac{1}{\beta(T(x))}dx}$$

REFERENCES

[1] K. Banerjee and A. Mehrotra, "Global (interconnect) warnning," *IEEE Circuits and Device Magazine*, pp. 16–32, September 2001.

[2] K. Skadron, M. R. Stan, W. Huang, S. Velusamy, K. Sankanarayanan, and D. Tarjan, "Temperature-aware microarchitecture," in *Proc. of the 30th International Symposium on Computer Architecture*, June 2003, pp. 2–13.

[3] J. Srinivasan, S. V. Adve, P. Bose, and J. A. Rivers, "The case for lifetime reliability-aware microprocessors," in *Proc. of the 31st International Symposium on Computer Architecture*, June 2004, pp. 276–287.

[4] K. Skadron, K. Sankaranarayanan, S. Velusamy, D. Tarjan, M. Stan, and W. Huang, "Temperature-aware microarchitecture: Modeling and implementation," *ACM Trans. on Architecture and Code Optimization*, vol. 1, no. 1, pp. 94–125, March 2004.

[5] Y.-K. Cheng and S.-M. Kang, "A temperature-aware simulation environment for reliable ULSI chip design," *IEEE Trans. Computer-Aided Design*, vol. 19, no. 10, pp. 1211–20, October 2000.

[6] T.-Y. Wang and C. C.-P. Chen, "3-D thermal-ADI: A linear-time chip level transient thermal simulator," *IEEE Trans. Computer-Aided Design*, vol. 21, no. 12, pp. 1434–45, December 2002.

[7] M. Hauschildt, M. Gall, S. Thraasher, P. Justison, L. Michaelson, H. Kawasaki, and P. S. Ho, "Statistical analysis of electromigration lifetimes and void evolution for cu interconnects," in *Materials, Technology and Reliability of Advanced Interconnects – 2004*, ser. Mater. Res. Soc. Symp. Proc., vol. 812, 2004, pp. 379–384.

[8] J. R. Black, "Mass transport of aluminum by momentum exchange with conducting electrons," in *IEEE Int. Rel. Phys. Symp.*, 1967, pp. 148–159.

[9] Y.-J. Park, V. K. Andleigh, and C. V. Thompson, "Simulations of stress evolution and the current density scaling of electromigration-induced failure times in pure and alloyed interconnects," *J. Appl. Phys.*, vol. 85, no. 7, pp. 3546–3555, April 1999.

[10] Z. Lu, W. Huang, J. Lach, M. Stan, and K. Skadron, "Interconnect lifetime prediction under dynamic stress for reliability-aware design," in *Proc. of International Conference on Computer Aided Design*, November 2004, pp. 327–334.

[11] S. M. Alam, G. C. Lip, C. V. Thompson, and D. E. Troxel, "Circuit level reliability analysis of Cu interconnects," in *Proc. of the 5th International Symposium on Quality Electronic Design (ISQED'04)*, 2004, pp. 238–243.

[12] D. Brooks and M. Martonosi, "Dynamic thermal management for high-performance microprocessors," in *Proc. of the 7th International Symposium on High-Performance Computer Architecture (HPCA-7)*, January 2001, pp. 171–182.

[13] J. Srinivasan and S. V. Adve, "Predictive dynamic thermal management for multimedia applications," in *Proc. of the 17th Annual ACM International Conference on Supercomputing*, June 2003, pp. 109–120.

[14] Z. Lu, J. Lach, M. Stan, and K. Skadron, "Improved thermal management with reliability banking," *IEEE Micro.*, vol. 25, no. 6, pp. 40–49, November/December 2005, special issue on Reliability-Aware Microarchitectures.

[15] J. Srinivasan, S. V. Adve, P. Bose, and J. A. Rivers, "Exploiting structural duplication for lifetime reliability enhancement," in *Proc. of the 32st International Symposium on Computer Architecture*, June 2005, pp. 520–531.

[16] A. Ramakrishnan and M. G. Pecht, "A life consumption monitoring methodology for electronic systems," *IEEE Trans. Comp. Packag. Technol.*, vol. 26, no. 3, pp. 625–634, September 2003.

[17] J. Clement, "Electromigration modeling for integrated circuit interconnect reliability analysis," *IEEE Trans. Device Mat. Rel.*, vol. 1, no. 1, pp. 33–42, March 2001.

[18] M. E. Sarychev, Y. V. Zhitnikov, L. Borucki, C.-L. Liu, and T. M. Markhviladze, "General model for mechanical stress evolution during electromigration," *J. Appl. Phys.*, vol. 86, no. 6, pp. 3068–75, September 1999.

[19] J. Clement, "Reliability analysis for encapsulated interconnect lines under DC and pulsed DC current using a continuum electromigration transport model," *J. Appl. Phys.*, vol. 82, no. 12, p. 5991, December 1997.
[20] M. A. Korhonen and P. Bøgesen, "Stress evolution due to electromigration in confined metal lines," *J. Appl. Phys.*, vol. 73, no. 8, p. 3790, April 1993.
[21] H. Ye, C. Basaran, and D. C. Hopkins, "Numerical simulation of stress evolution during electromigration in IC interconnect lines," *IEEE Trans. Comp. Packag. Technol.*, vol. 26, no. 3, pp. 673–681, September 2003.
[22] R. C. McOwen, *Partial differential equations: methods and applications*. Prentice-Hall, 1995.
[23] W. R. Hunter, "Self-consistent solutions for allowed interconnect current density-part II: Application to design guidelines," *IEEE Trans. Electron Devices*, vol. 44, no. 2, pp. 310–316, February 1997.
[24] L. M. Ting, J. S. May, W. R. Hunter, and J. W. McPherson, "AC electromigration characterization and modeling of multilayered interconnects," in *Proc. Int. Reliability Physics Symposium*, March 1993, pp. 311–316.
[25] T.-Y. Chiang, K. Banerjee, and K. C. Saraswat, "Analytical thermal model for multilevel VLSI interconnects incorporating via effect," *IEEE Electron Device Letters*, vol. 23, no. 1, pp. 31–33, January 2002.
[26] A. H. Ajami, K. Banerjee, and M. Pedram, "Modeling and analysis of non-uniform substrate temperature effects on global ulsi interconnects," *IEEE Trans. Computer-Aided Design*, vol. 24, no. 6, pp. 849–861, 2005.
[27] W. Guo, Z. Li, H. Zhu, W. Zhang, Y. Ji, Y. Sun, and G. Shen, "Temperature gradient impact on electromigration failure in vlsi metalization," in *Proceedings of 14th IEEE SEMI-THERM Symposium*, 1998, pp. 122–127.
[28] H. V. Nguyen, C. Salm, B. Krabbenborg, K. Weide-Zaage, J. Bisschop, A. J. Mouthaan, and F. G. Kuper, "Effect of thermal gradients on the electromigration lifetime in power electronics," in *Proc. of the 42nd Annual International Reliability Physics Symposium (IRPS)*, 2004, pp. 619–620.
[29] K. Jonggook, V. C. Tyree, and C. R. Crowell, "Temperature gradient effects in electromigration using an extended transition probability model and temperature gradient free tests," in *Proceedings of the IEEE Integrated Reliability Workshop (IRW)*, 1999, pp. 24–40.
[30] E. T. Ogawa, K.-D. Lee, V. A. Blaschke, and P. S. Ho, "Electromigration reliability issues in dual-damascene Cu interconnections," *IEEE Trans. Rel.*, vol. 51, no. 4, pp. 403–419, December 2002.
[31] Z.-S. Choi, C. L. Gan, F. Wei, C. V. Thompson, J. H. Lee, K. L. Pey, and W. K. Choi, "Fatal void size comparisons in via-below and via-above Cu dual-damascene interconnects," in *Materials, Technology and Reliability of Advanced Interconnects – 2004*, ser. Mater. Res. Soc. Symp. Proc., vol. 812, 2004, pp. 373–378.
[32] R. A. Wachnik, R. G. Filippi, T. M. Shaw, and P. C. Lin, "Practical benefits of the electromigration short-length effect, including a new design rule methodology and an electromigration resistant power grid with enhanced wireability," in *2000 Symposium on VLSI Technology Digest*, 2000, pp. 220–221.
[33] W. K. Meyer, R. Solanki, and D. Evans, "Near-threshold electromigration of damascene copper on tin barrier," in *Materials, Technology and Reliability of Advanced Interconnects – 2003*, ser. Mater. Res. Soc. Symp. Proc., vol. 766, 2003, pp. 145–150.
[34] S. Rochel, G. Steele, J. R. Lloyd, S. Z. Hussain, and D. Overhauser, "Full-chip reliability analysis," in *Proc. Int. Reliability Physics Symposium*, 1998, pp. 356–362.
[35] T.-Y. Wang, J.-L. Tsai, and C. C.-P. Chen, "Thermal and power integrity based power/ground networks optimization," in *Proc. of the Design, Automation and Test in Europe Conference and Exhibition*, vol. 2, February 2004, pp. 830–835.
[36] J. A. Davis and J. D. Meindl, Eds., *Interconnect Technology and Design for Gigascale Integration*. Kluwer, 2003.
[37] D. T. Blaauw, C. Oh, V. Zolotov, and A. Dasgupta, "Static electromigration analysis for on-chip signal interconnects," *IEEE Trans. Computer-Aided Design*, vol. 22, no. 1, pp. 39–48, January 2003.
[38] D. Burger and T. M. Austin, "The simplescalar toolset, version 2.0," *Computer Architecture News*, vol. 25, no. 3, pp. 13–25, June 1997.
[39] Y. Zhang, D. Parikh, K. Sankaranarayanan, K. Skadron, and M. Stan, "Hotleakage: A temperature-aware model of subthreshold and gate leakage for architects," Department of Computer Science, University of Virginia, Tech. Rep. CS-2003-05, March 2003.
[40] M. Elnozahy, M. Kistler, and R. Rajamony, "Energy conservation policies for web servers," in *Proc. of the 4th USENIX Symposium on Internet Technologies and Systems*, March 2003.
[41] P. Bohrer, E. Elnozahy, T. Keller, M. Kistler, C. Lefurgy, and R. Rajamony, "The case for power management in web servers," in *Power Aware Computing*, R. Graybill and R. Melhem, Eds. Kluwer Academic Publications, 2002.