

Interconnect Lifetime Prediction under Dynamic Stress for Reliability-Aware Design

Zhijian Lu, Wei Huang, John Lach, Mircea Stan, Kevin Skadron[§]

Department of Electrical and Computer Engineering, [§]Department of Computer Science

University of Virginia

Charlottesville, VA 22904

{z14j, wh6p, jlach, mircea}@virginia.edu, skadron@cs.virginia.edu

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 presents a physics-based model for estimating interconnect lifetime for any time-varying temperature/current profile. This model is verified with numerical solutions.

With this model, we show that designers may be more aggressive with the temperature profiles that are allowed on a chip. In fact, our model reveals that when the temperature magnitude variation is small, average temperature (instead of worst-case temperature) can be used to accurately predict interconnect lifetime, allowing for significant design margin reclamation in reliability-aware design. Even when the variation of temperature magnitude is large, our model shows that using the maximum temperature is still too conservative for interconnect lifetime prediction. Therefore, our model not only increases the accuracy of reliability estimates, but also enables designers to consider more aggressive designs. This model is similarly useful for temperature-aware dynamic runtime management.

Keywords

Electromigration, reliability-aware design, dynamic stress.

1. 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) [1]. In the paper, we focus on temperature-related EM. 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 [13, 14] are being developed to best manage this consumption. While the dynamic temperature profile of a system is workload-dependent [13, 14], several efficient and accurate techniques have been proposed to simulate transient chip-wide temperature distribu-

tion [3, 13, 16], 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. Historically, Black [2] proposed a semi-empirical temperature-dependent model for EM failures:

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

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 [10]. Our techniques is applicable for any value of n . For simplicity, in the following discussion, we use $n = 2$. Black's model is widely used in thermal reliability analysis and design.

However, Black's model assumes a constant temperature. Thus, a worst-case temperature profile is usually used when applying this model, resulting in pessimistic estimations and unnecessarily restricted design spaces. As an example, we use the *Hotspot* toolset [13], 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 $110^\circ C$ and $114^\circ C$, 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.

Recently, Srinivasan *et al.* [14] proposed an architecture-level dynamic reliability model, but their model does not consider the impact of time-varying stresses from a physics perspective. In this paper, we present a simple physics-based model to estimate interconnect lifetime due to EM failures under time-varying temperature and current distributions. While designers are currently constrained by constant, worst-case temperature assumptions, the model presented here 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. This also allows

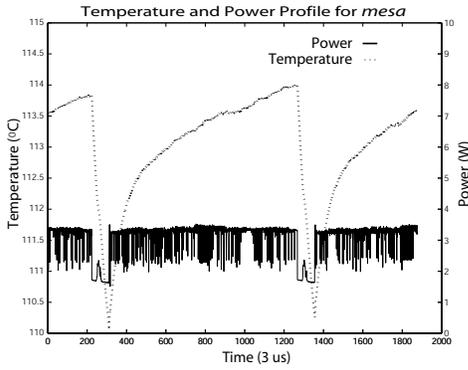


Figure 1. A simulated temperature/power profile for an integer unit running the *mesa* Spec2000 benchmark.

more aggressive runtime DTM techniques.

This paper is organized as follows. Section 2 introduces a constant temperature analytic model for EM. In Section 3, 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 4. We illustrate some possible applications of this model in Section 5. Finally, we summarize the paper in Section 6.

2. ANALYTIC MODEL FOR EM WITH CONSTANT TEMPERATURE

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 [6] provides a review of 1-D analytic EM models. Several more sophisticated EM models are also available [10, 12]. In this paper, we only discuss the EM-induced stress build-up model of Clement and Korhonen [5, 8], which has been widely used in EM analysis and agrees well with simulation results using a more advanced model by Ye *et al.* [18].

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 [5, 8]:

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

where $\sigma(x, t)$ is the stress function, and an interconnect failure is considered to happen when $\sigma(x, t)$ reaches a threshold 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 $\frac{\rho j}{\Omega}$, 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 backflow flux created by the stress gradient to counter-balance the EM flux. This equation assumes that the temperature is uniform across the characteristic length. If we 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:

$$\frac{\partial \sigma}{\partial t} - \beta(T) \frac{\partial}{\partial x} \left(\frac{\partial \sigma}{\partial x} - \alpha(j) \right) = 0 \quad (3)$$

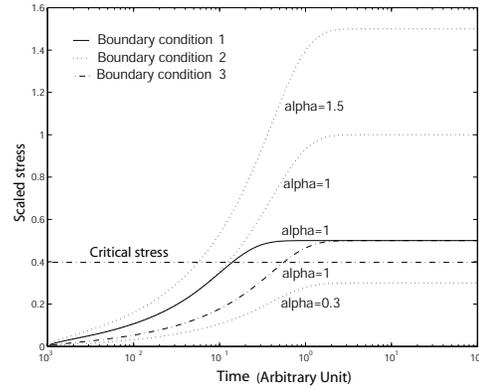


Figure 2. EM stress build-up for different boundary conditions and α values. All processes have $\beta = 1$. (α and β are defined in Equation (3).)

Clement [5] investigated the effect of current density on stress build-up using Equation (3), 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: $D_a = D_{a0} \exp\left(\frac{-Q}{kT}\right)$.

Applying the parabolic maximum principles [9] to Equation (3), 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 solutions for Equation (3) 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 [5] 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.

3. 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.

Time-dependent current stress

Clement [5] 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 [7, 15]. 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.

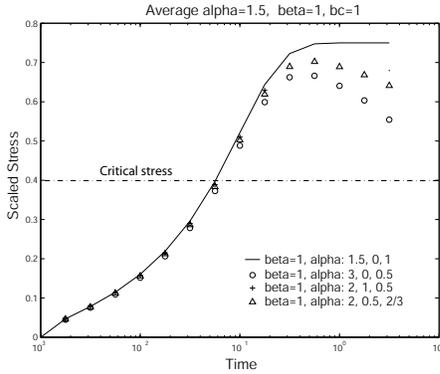


Figure 3. EM stress build-up under time-dependent current stress. In each EM process, α (defined in Equation (3)) oscillates between two values with different duty cycles. The time dependence of α is given in the legend.² All curves have the same average value of α . The solid line is the stress build-up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

We numerically solve Equation (3) 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.

²For 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.

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 (3) 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 (3) 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. \square

Theorem 1 tells us that the stress build-up processes in the interconnect are independent of the value of β in Equation (3). 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 (3) is a segment-wise function, described as:

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

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 β_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 i th 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.

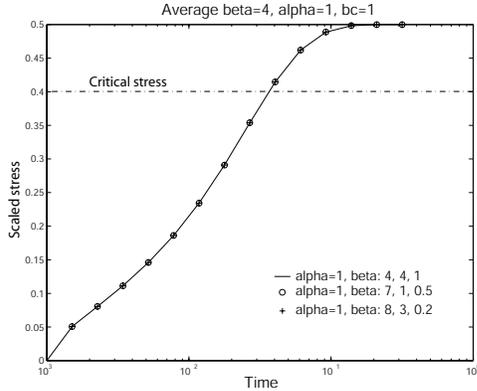


Figure 4. EM stress build-up at one end of the interconnect with different time-dependent β functions (square waveform). The solid line is the case with a constant value of β equal to the average value of β in other curve.

It follows that at the end of the i th time segment, the stress in $M0$ is specified as: $\sigma_0(\sum_{k=1}^i \Delta t_k) = \sigma_1(\sum_{k=1}^i (\frac{\beta_k}{\beta_1}) \Delta t_k)$. As $\Delta t_i \rightarrow dt$, $\beta_i \rightarrow \beta(T(t))$, we obtain the integral version for the stress build-up function:

$$\sigma_0(t) = \sigma_1 \left(\left(\frac{1}{\beta_1} \right) \int_0^t \beta(T(t)) dt \right) \quad (4)$$

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

$$\int_0^{t_{failure}} \beta(T(t)) dt = \varphi_{th} \quad (5)$$

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:

$$t_{failure} = \frac{\varphi_{th}}{E(\beta(T(t)))} \quad (6)$$

where $E(\beta(t))$ is the expected value for $\beta(t)$, and $\beta(t)$ is the temperature factor, as defined in Equation (3), 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 (6) indicates that the average of temperature factor β should be used.

One way to interpret Equation (5) 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 (6). Substitute this relation in Equation (6) again and consider

the time-varying temperature, and we obtain an alternative form for Equation (6):

$$t_{failure} = \frac{1}{E(1/MTF(T))} \quad (7)$$

Equation (7) 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 (6), 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 .

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 i th 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:

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

Since the stress evolution in $M1$ is under constant thermal stress, we may apply the "average current model". As

$\Delta t_i \rightarrow dt$, $\beta_i \rightarrow \beta(T(t))$ and $\alpha_i \rightarrow \alpha(t)$, we derive the EM reliability equivalent current for $M0$ (or the average current for $M1$) as:

$$j_{equivalent} = \frac{\int_0^T j(t)\beta(t)dt}{\int_0^T \beta(t)dt} = \frac{E[j(t)\beta(t)]}{E[\beta(t)]} \quad (8)$$

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 (4), (5), and (6), combined with Black's equation, we obtain the time to failure for $M0$:

$$t_{failure} = \frac{C}{j_{equivalent}^2 E(\beta(T(t)))} \quad (9)$$

where $j_{equivalent}$ is defined by Equation (8), and C is a constant.

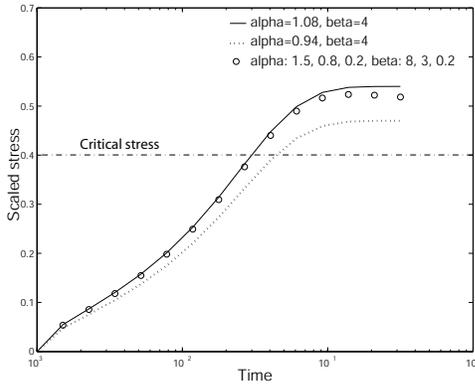


Figure 5. EM stress build-up at one end of the interconnect with time-varying α (current) and β (temperature) functions (i.e., square waveforms). The circles represent the numerical solution for time-varying α and β . The solid line is with a constant value of α calculated according to Equation (8) and 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).

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 (8)). 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.

As a matter of fact, if the current and the temperature are statistically independent, we have $\frac{E[j(t)\beta(t)]}{E[\beta(t)]} = E[j(t)]$ in Equation (8). 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 (8) and (9) will lead us to Equation (6).

4. ANALYSIS OF THE PROPOSED MODEL

Equations (8) and (9) 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.

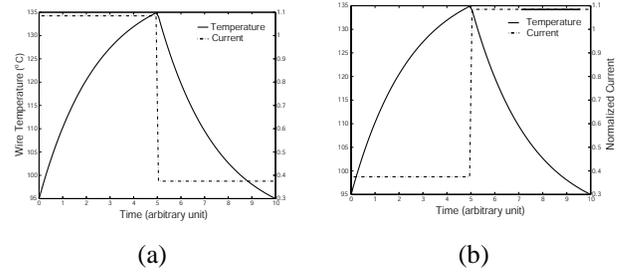


Figure 6. Temperature and current waveforms analyzed in the paper: (a) in phase current/temperature, (b) out of phase current/temperature.

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

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

where MTF_1 is the time to failure under time-varying temperature profile $T_1(t)$ and electric current profile $j_1(t)$. Figure 6 shows two extreme relations between temperature and current profiles. 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., $40^\circ C$). The temperature difference between the substrate and the interconnects is fixed to be $21^\circ C$, which is a reasonable assumption for high-layer interconnects [4]. Using the data from Figure 1, the maximum temperature of the substrate is assumed to be $114^\circ C$ (i.e., $135^\circ C$ 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). According to our proposed model, the temperature will also impact the reliability equiv-

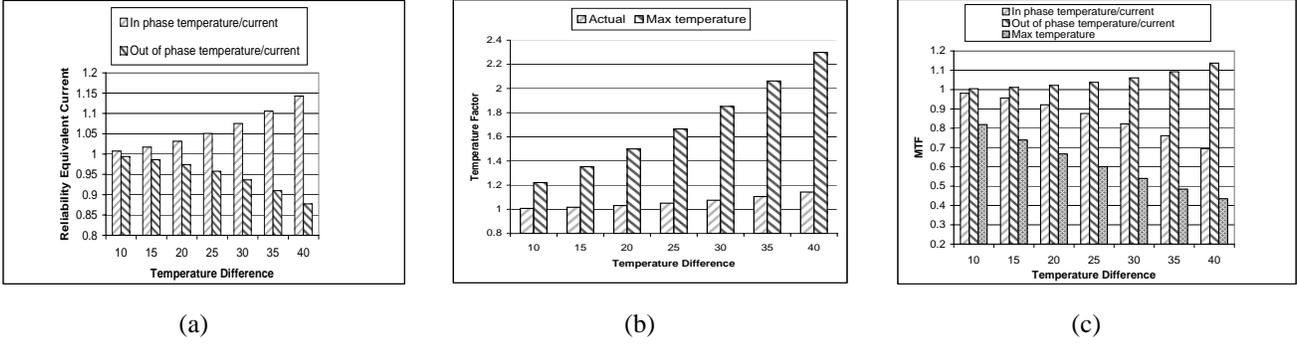


Figure 7. Comparison of electric current, temperature factor (β) and MTF for different peak to peak temperature cycles. All results are normalized to the average current and/or temperature case. (a) Ratio of reliability equivalent current (our model) to average current. Both cases of current variation (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.

alent current, so we also investigate the case in which current is out of phase with temperature as shown in Figure 6(b). 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 $10^{\circ}C$), 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).

5. APPLICATION OF THE PROPOSED MODEL

In this section, we discuss some possible applications of our proposed dynamic stress model. When the thermal profile of a circuit can be predicted at design time, designers can use our model to reclaim some design margin that is hidden by the conservative nature of traditional worst-case design assumptions. Even if the runtime thermal profile is not predictable, our dynamic reliability model can be used to guide dynamic thermal management at runtime to achieve customized reliability goals.

Design time optimization

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 [11]. However, as we have shown above, using worst-case temperature is too conservative and could result in wasteful excessive design margins. Here we propose a design flow incorporating runtime stress information as shown Figure 8. 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 reliability projection from these models will

generally enable more relaxed design constraints and provide a wider design space.

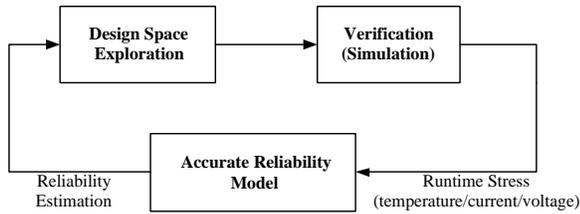


Figure 8. A proposed design flow incorporating runtime stress information.

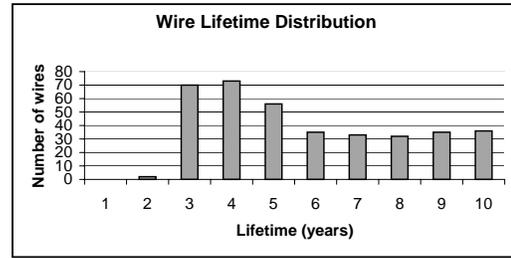
For instance, when temperature fluctuates within a relatively small range (e.g., $10^{\circ}C$), our model predicts that using average temperature is good enough for reliability evaluation. Therefore, we could potentially reduce the number of design points falsely flagged for design rule violations when using the worst-case temperature. One example is illustrated in Figure 9 using data from a power grid design [17]. In this example, the worst-case temperature of a design is $135^{\circ}C$, and Wang *et al.* [17] 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 $135^{\circ}C$ 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 9(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.* [17].

Runtime management

Another advantage of our proposed model is its inherent suitability for runtime management. Using our model, the MTF of an integrated circuit can be formulated as a resource to be consumed over time, as expressed in Equation (5), and be incorporated into a runtime dynamic management framework, such as [13, 14]. For example, using this model, we can continuously monitor the reliability consumption by the system, compare it to the reliability budget, and decide the correct operation strategy to maximize performance subject to reliability. If the reliability budget is in surplus, the system may operate in a more aggressive way to boost performance. Otherwise, the system must operate in a more conservative way to maintain lifetime.

To demonstrate the benefits of considering temporal temperature variations in runtime management, we present a case



(a)

Temperature variation ($^{\circ}C$)	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)

Figure 9. (a) Distribution of wires violating the MTF specification using maximum temperature (data extracted from [18]) with a total of 372 wires. (b) Reduction of the number of wires violating the MTF specification under different temperature variations (maximum temperature: $135^{\circ}C$).

study using temperature values obtained from simulating a microprocessor with characteristics similar to a $0.13\mu m$ Alpha 21364. Using *Hotspot* [13], a compact thermal model, we can obtain the steady-state and transient temperature responses of the substrate and interconnect. For example, the transient thermal behavior of interconnect above the floating point register file is shown in Figure 10 for the Spec2000 benchmark program *applu*, revealing obvious variations. The thermal-package characteristics we used in the simulation were derived so that the reliability-equivalent temperature value, according to our model, that yields the same expected lifetime as the pattern in Figure 10 is $110^{\circ}C$ (a common limit). This temperature is plotted in Figure 10 as a straight line.

These results illustrate the potential benefits of accounting for temporal variation. If the lifetime budget is used to dictate only some fixed worst-case temperature (e.g., $110^{\circ}C$), then a more expensive cooling solution is required to bring *applu*'s actual behavior within specification while achieving the same performance. The alternative is to reduce the voltage and clock speed, or a DTM technique must be engaged to reduce processor activity and enforce the $110^{\circ}C$ limit whenever the operating temperature exceeds the threshold. Using microarchitecture simulation techniques described in [13], we estimate that selecting a lower design point for voltage and frequency would require a 13% reduction in clock fre-

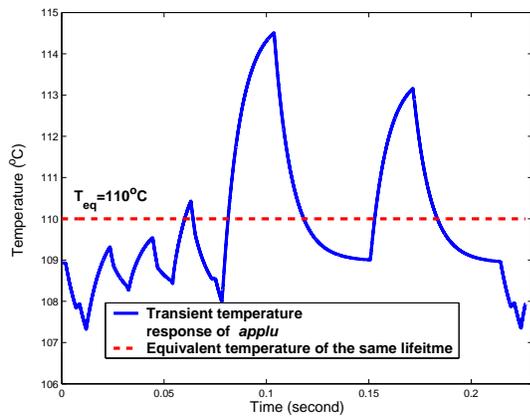


Figure 10. Transient temperature response of floating point register interconnects. Constant operating temperature for the same interconnect lifetime is also shown.

quency. Dynamic thermal management would reduce performance by about 10% using dynamic voltage scaling and 50% using fine-grained fetch gating. If temporal temperature variations are taken into account using our model, none of these costly solutions are needed and there is no impact on system performance.

6. CONCLUSIONS AND FUTURE WORK

This paper presented a dynamic reliability model for interconnect EM failures with time-varying temperature and current stress. This model will not only increase the accuracy of reliability estimates but will enable designers to more aggressively explore the design space and to reclaim the design margin imposed by less accurate, more pessimistic models. To our knowledge, this is the first physics-based EM model dealing with dynamic stresses.

Existing constant-temperature models require designers to observe a static worst-case temperature limit, but the model 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 model also makes it suitable for DTM — an important approach for post-manufacture optimization. In the future, we will compare our model predictions 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.

ACKNOWLEDGEMENTS

This work is supported in part by the National Science Foundation under grant Nos. CCR-0105626, CCR-0133634, and a grant from Intel MRL. We would also like to thank the anonymous reviewers for their helpful comments.

REFERENCES

- [1] Banerjee, K., and Mehrotra, A. Global (interconnect) warning. *IEEE Circuits and Device Magazine* (September 2001), pp. 16.
- [2] Black, J. R. Mass transport of aluminum by momentum exchange with conducting electrons. In *IEEE Int. Rel. Phys. Symp.* (1967), pp. 148–159.
- [3] Cheng, Y.-K., and Kang, S.-M. A temperature-aware simulation environment for reliable ULSI chip design. *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems* 19, 10 (October 2000), pp. 1211–20.
- [4] Chiang, T.-Y., Banerjee, K., and Saraswat, K. C. Analytical thermal model for multilevel VLSI interconnects incorporating via effect. *IEEE Electron Device Letters* 23, 1 (January 2002), pp. 31–33.
- [5] Clement, J. Reliability analysis for encapsulated interconnect lines under DC and pulsed DC current using a continuum electromigration transport model. *J. Appl. Phys.* 82, 12 (December 1997), pp. 5991.
- [6] Clement, J. Electromigration modeling for integrated circuit interconnect reliability analysis. *IEEE Trans. on Device and Materials Reliability* 1, 1 (March 2001), pp. 33–42.
- [7] Hunter, W. R. Self-consistent solutions for allowed interconnect current density-part II: Application to design guidelines. *IEEE Trans. on Electron Devices* 44, 2 (February 1997), pp. 310–316.
- [8] Korhonen, M. A., and Bøgesen, P. Stress evolution due to electromigration in confined metal lines. *J. Appl. Phys.* 73, 8 (April 1993), pp. 3790.
- [9] McOwen, R. C. *Partial differential equations: methods and applications*. Prentice-Hall, 1995.
- [10] Park, Y.-J., Andleigh, V. K., and Thompson, C. V. Simulations of stress evolution and the current density scaling of electromigration-induced failure times in pure and alloyed interconnects. *J. Appl. Phys.* 85, 7 (April 1999), pp. 3546–3555.
- [11] Rochel, S., Steele, G., Lloyd, J. R., Hussain, S. Z., and Overhauser, D. Full-chip reliability analysis. In *Proc. Int. Reliability Physics Symposium* (1998), pp. 356–362.
- [12] Sarychev, M. E., Zhitnikov, Y. V., Borucki, L., Liu, C.-L., and Markhviladze, T. M. General model for mechanical stress evolution during electromigration. *J. Appl. Phys.* 86, 6 (September 1999), pp. 3068–75.
- [13] Skadron, K., Stan, M. R., Huang, W., Velusamy, S., Sankaranarayanan, K., and Tarjan, D. Temperature-aware microarchitecture. In *Proc. of the 30th International Symposium on Computer Architecture* (June 2003), pp. 2–13.
- [14] Srinivasan, J., Adve, S. V., Bose, P., and Rivers, J. A. The case for lifetime reliability-aware microprocessors. In *Proc. of the 31st International Symposium on Computer Architecture* (June 2004), pp. 276–287.
- [15] Ting, L. M., May, J. S., Hunter, W. R., and McPherson, J. W. AC electromigration characterization and modeling of multilayered interconnects. In *Proc. Int. Reliability Physics Symposium* (March 1993), pp. 311–316.
- [16] Wang, T.-Y., and Chen, C. C.-P. 3-D thermal-ADI: A linear-time chip level transient thermal simulator. *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems* 21, 12 (December 2002), pp. 1434–45.
- [17] Wang, T.-Y., Tsai, J.-L., and Chen, C. C.-P. Thermal and power integrity based power/ground networks optimization. In *Proc. of the Design, Automation and Test in Europe Conference and Exhibition* (February 2004), vol. 2, pp. 830–835.
- [18] Ye, H., Basaran, C., and Hopkins, D. C. Numerical simulation of stress evolution during electromigration in IC interconnect lines. *IEEE Trans. on Components and Packaging Technologies* 26, 3 (September 2003), pp. 673–681.