On Thermal Effects in Deep Sub-Micron VLSI Interconnects

Kaustav Banerjee[1] Amit Mehrotra Alberto Sangiovanni-Vincentelli Chenming Hu

Department of Electrical Engineering and Computer Sciences
University of California, Berkeley, CA 94720
Email: {kaustav, mehrotra, alberto, hu} @eecs.berkeley.edu

Abstract

This paper presents a comprehensive analysis of the *thermal effects* in advanced high performance interconnect systems arising due to self-heating under various circuit conditions, including electrostatic discharge. Technology (Cu, low-k etc) and scaling effects on the thermal characteristics of the interconnects, and on their electromigration reliability has been analyzed simultaneously, which will have important implications for providing robust and aggressive deep sub-micron interconnect design guidelines. Furthermore, the impact of these thermal effects on the design (driver sizing) and optimization of the interconnect length between *repeaters* at the upper-level signal lines are investigated.

1 Introduction

The ever-increasing demand for speed and functionality of Si-based advanced high performance microprocessors, digital signal processing (DSP) chips, and application specific integrated circuits (ASICs), and the increasing density of memory (DRAM) devices has caused aggressive scaling of ICs beyond 0.25-μm minimum feature size [1], [2], [3]. These technology nodes, commonly referred to as *deep submicron*, (DSM) allow VLSI circuits to meet the required device density and various circuit performance specifications. This trend has resulted in a dramatic reduction of the interconnect metal pitch and increased the number of metallization levels. The aggressive interconnect scaling has resulted in increasing current densities [4], [5] and associated *thermal effects*.

Thermal effects are an inseparable aspect of electrical power distribution and signal transmission through the interconnects due to self-heating (or Joule heating) caused by the flow of current. Thermal effects impact interconnect design and reliability in the following ways. Firstly, they limit the maximum allowable RMS current density, $j_{RMS\text{-}max}$ (since the RMS value of the current density is responsible for heat generation) in the interconnects, in order to limit the temperature increase. Secondly, interconnect lifetime (reliability) which is limited by electromigration (EM) (transport of mass in metals under an applied current density), has an exponential dependence on the inverse metal temperature [6]. Hence, temperature rise of metal interconnects due to self-heating phenomenon can also limit the maximum allowed average current density, $j_{avg\text{-}max}$, since EM capability is dependent on the average current density [7]. Thirdly, thermally induced open circuit metal failure under short-duration high peak currents including electrostatic discharge (ESD) is also a reliability concern, [8] and can introduce latent EM damage [9] that has important reliability implications.

Chatterjee et al., [5] have argued that thermal effects will increasingly dominate interconnect design rules that specify maximum current densities for circuit designers. Recently, Hunter [10], solved the EM lifetime equation for Al-Cu, and the 1-D heat equation, in a self-consistent manner. In this approach, both EM and self-heating can be comprehended simultaneously. Rzepka et al., [11] have carried out detailed simulations of self-heating in multilevel interconnects using finite element analysis. They also concluded that in the near future, 3-D interconnect arrays will be affected by self-heating more severely than those of today's.

DAC 99, New Orleans, Louisiana

Furthermore, low dielectric constant (Low-k) materials are being introduced [12] as an alternative intra-level insulator to reduce interconnect capacitance (therefore delay) and cross-talk noise to enhance circuit performance. These materials can further exacerbate thermal effects owing to their lower thermal conductivity than silicon dioxide. Banerjee et al., [13] have recently shown that interconnect scaling using low-k as the intra-level (gap-fill) dielectric material can cause significant increase in thermal effects.

Presently, interconnect design rules are generated in a non self-consistent manner [14] as explained in section 2.1. As deep sub-micron interconnect technologies, like Cu and low-k dielectric materials are rapidly evolving, there is an increasing need to understand their effects simultaneously on the thermal characteristics of these interconnects and on their EM reliability, in order to provide robust design guidelines. Furthermore, it is not clear whether thermal constraints conflict with the performance optimization steps employed at the circuit level. Hence, a thorough analysis of thermal effects in DSM interconnects is necessary to comprehend their full impact on circuit design, accurately model their reliability, and provide thermally safe design guidelines for various technologies.

This paper presents a comprehensive analysis of the *thermal effects* in advanced high performance interconnect systems arising from self-heating under various circuit conditions, including ESD. The analysis examines the self-consistent solutions for allowed interconnect current density for technologies up to 0.1-μm involving Cu and various low-k materials in an eight-level metallization system as per the National Technology Roadmap for Semiconductors (NTRS) [15].

2 Interconnect Design Rules: Current Approach

2.1 Average, RMS, and Peak Current Density

Circuit designers are typically provided with the maximum allowable values for three interconnect current densities. These are the average current density, j_{avg}, the RMS current density, j_{rms}, and the peak current density, j_{peak}. These quantities are defined as follows:

The peak current density is simply the current density corresponding to the peak current of the waveform,

$$j_{peak} = \frac{I_{peak}}{A} \tag{1}$$

where A is the cross-sectional area of the interconnect.

The average current density is defined as,

$$j_{avg} = \frac{1}{T}\int_0^T j(t)dt \tag{2}$$

where T is the time period of the current waveform. The RMS current density is defined as,

$$j_{rms} = \sqrt{\frac{1}{T}\int_0^T j^2(t)dt} \tag{3}$$

For a fixed temperature, EM lifetime of interconnects is known to be determined by j_{avg} [7]. Self-heating is determined by j_{rms}. Presently, high performance interconnect design is based on the specified limits for the maximum values of the average, RMS, and peak current densities [14]. They are however not self-consistent, i.e., these values

[1]Now with the Center for Integrated Systems, Department of Electrical Engineering, Stanford University, Stanford, CA, 94305-4070.

do not simultaneously comprehend the two temperature dependent mechanisms: EM and self-heating.

Interconnects can be broadly classified into two categories: signal lines and power lines. They differ in that currents in signal lines are bi-directional (or bipolar) [7] while those in power lines are usually uni-directional (or unipolar).

We will now consider unipolar current waveforms for some illustrative analysis. Using the above definitions for the three current densities, one can easily show that

$$j_{avg} = r \cdot j_{peak} \tag{4}$$

and,

$$j_{rms} = \sqrt{r} \cdot j_{peak} \tag{5}$$

where r is the duty cycle defined as t_{on}/T in Fig. 1.

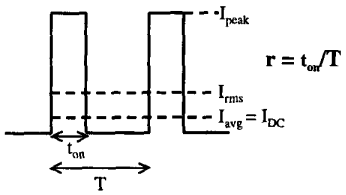

Figure 1. A unipolar pulsed waveform illustrating various current definitions.

2.2 Electromigration and Self-Heating

EM is the transport of mass in metals under an applied current density and is widely regarded as a major wear out or failure mechanism of VLSI interconnects [6]. When current flows through the interconnect metal, an *electronic wind* is set up opposite to the direction of current flow. These electrons upon colliding with the metal ions, impart sufficient momentum, and displace the metal ions from their lattice sites creating vacancies. These vacancies condense to form voids that result in increase of interconnect resistance or even open circuit conditions [16]. EM lifetime reliability of metal interconnects is modeled by the well known Black's equation [6] given by,

$$TTF = A^* \cdot j^{-n} \cdot exp(\frac{Q}{k_B T_m}) \tag{6}$$

where TTF is the time-to-fail (typically for 0.1% cumulative failure). A* is a constant that is dependent on the geometry and microstructure of the interconnect, j is the DC or average current density. The exponent n is typically 2 under normal use conditions, Q is the activation energy for grain-boundary diffusion and equals ~0.7 eV for Al-Cu, k_B is the Boltzmann's constant, and T_m is the metal temperature. The typical goal is to achieve 10 year lifetime at 100 °C, for which (6) and accelerated testing data produce a design rule value for the acceptable current density at T_{ref}, j_0. However, this design rule value does not comprehend self-heating.

The effect of self-heating can be analyzed from the following: The metal temperature, T_m in (6) is given by,

$$T_m = T_{ref} + \Delta T_{self-heating} \tag{7}$$

and,

$$\Delta T_{self-heating} = (T_m - T_{ref}) = I^2_{rms} \cdot R \cdot \theta_{int} \tag{8}$$

where T_{ref} is the reference chip (silicon junction) temperature and is typically taken as 100 °C, $\Delta T_{self-heating}$ is the temperature rise of the metal interconnect due to the flow of current, R is the interconnect resistance, and θ_{int} is the thermal impedance of the interconnect line to the chip. Thus, both EM and self-heating are temperature dependent effects, and as self-heating increases, EM lifetime decreases exponentially according to (6).

3 Thermal Effects in DSM Interconnects

3.1 Self-Consistent Interconnect Design Rules

In this section we will first introduce the formulation of the self-consistent solutions [10] for allowed interconnect current density, and then apply them to analyze Cu interconnects. The $\Delta T_{self-heating}$ in interconnects given by (8) can be written in terms of the RMS current density as,

$$j^2_{rms} = \frac{(T_m - T_{ref}) \cdot K_{ox} \cdot W_{eff}}{t_{ox} \cdot t_m \cdot W_m \cdot \rho_m(T_m)} \tag{9}$$

Here t_m and W_m are the thickness and width of interconnect metal line, and $\rho_m(T_m)$ is the metal resistivity at temperature T_m. Note that the thermal impedance θ_{int} in (8) has been expressed as,

$$\theta_{int} = \frac{t_{ox}}{K_{ox} \cdot L \cdot W_{eff}} \tag{10}$$

This expression for the thermal impedance is based on a quasi-1-D heat conduction model with $W_{eff} = W_m + 0.88t_{ox}$, valid for $W_m/t_{ox} > 0.4$ and is accurate to within 3% [17]. Here t_{ox} is the total thickness of the underlying dielectric, K_{ox} is the thermal conductivity normal to the plane of the dielectric, and L is the length of the interconnect.

Now, in order to achieve an EM reliability lifetime goal mentioned in section 2.2, we must have the lifetime at any (j_{avg}) current density and metal temperature T_m, equal to or larger than the lifetime value (eg. 10 year) under the design rule current density stress j_0, at the temperature T_{ref}. This value of j_0 is dependent on the specific interconnect metal technology. Therefore we have,

$$\frac{exp\left(\frac{Q}{k_B T_m}\right)}{j^2_{avg}} \geq \frac{exp\left(\frac{Q}{k_B T_{ref}}\right)}{j_0^2} \tag{11}$$

From (4) and (5) we have after eliminating j_{peak}

$$\frac{j^2_{avg}}{j^2_{rms}} = r \tag{12}$$

Substituting for j^2_{rms} from (9) and j^2_{avg} from (11) in (12) we get the self-consistent equation given by

$$r = j_0^2 \left(\frac{exp\left(\frac{Q}{k_B T_m}\right)}{exp\left(\frac{Q}{k_B T_{ref}}\right)} \right) \frac{t_{ox} \cdot t_m \cdot W_m \cdot \rho_m(T_m)}{(T_m - T_{ref}) \cdot K_{ox} \cdot W_{eff}} \tag{13}$$

Note that this is a single equation in the single unknown temperature T_m. Once this self-consistent temperature is obtained from (13) the corresponding maximum allowed j_{peak} and j_{rms} can be calculated from (5) and (9). The self-consistent equation given by (13) for unipolar pulses is also valid for more general time varying waveforms with an effective duty cycle r_{eff} [18]. One of the consequences of the self-consistent equation is that, for a certain j_0 as r decreases the self-consistent temperature and the maximum allowed j_{peak} increases. This effect is shown in Fig. 2 for Cu interconnects. Secondly, it can be observed that as r decreases the ratio j_{peak}(*self-consistent*)/j_{peak}(*without self-heating, i.e., the line labeled 1/r*) decreases monotonically. At $r = 10^{-2}$, the self-consistent j_{peak} is nearly 2 times smaller than the j_{peak} obtained from EM constraint only, i.e., without self-heating. From the EM lifetime relation given by (6) this implies that if a design used only the average (EM) current as the design guideline without comprehending self-heating it could have a lifetime nearly three times smaller than the reliability requirement. Thirdly, it can be observed from Fig. 3 that, as

j_0 is increased, the self-consistent interconnect metal temperature (T_m) increases as expected from (13). However, the maximum allowed j_{peak} does not increase much for values of $r < 10^{-3}$. That is j_0 becomes increasingly ineffective in increasing j_{peak} as the duty cycle r, decreases. From the above analysis, it apparently seems that EM capability might cease to be the dominant driver of interconnect technology evolution and thermal effects will limit the maximum allowed j_{peak}.

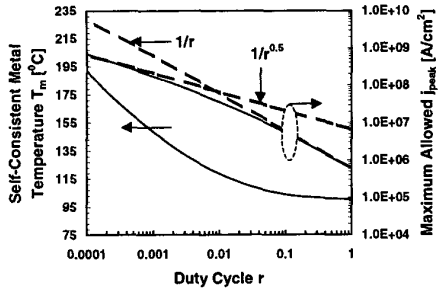

Figure 2. Self-consistent solutions for T_m and j_{peak}. Parameters used here are: j_0= 0.6 MA/cm², t_{ox} = 3 μm, t_m = 0.5 μm, and W_m = 3 μm. Interconnect metal is Cu with $\rho_m(T_m) = 1.67 \times 10^{-6}$ Ω-cm [1 + 6.8 x 10^{-3} $^0C^{-1}$ (T_m - T_{ref})]. The two dotted lines indicate j_{peak} values based on a) $j_{peak} = j_0/r$, and b) $j_{peak} = j_{rms}/r^{0.5}$.

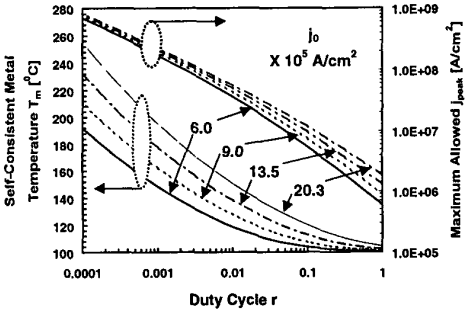

Figure 3. Self-consistent analysis showing dependence of T_m and j_{peak} on j_0. Parameters used here are: t_{ox} = 3.0 μm, t_m = 0.5 μm, and W_m = 3.0 μm. Interconnect metal is Cu.

The above analysis however motivates the following questions:

- Since the above analysis assumes a quasi-1D heat conduction model, what will be the effects on the self-consistent solutions if a more realistic model is used?
- Since the above analysis assumed that the intra-level and inter-level dielectric material is SiO_2, what will be the effects on the self-consistent solutions if low-k materials are considered as candidates for intra-level dielectric material?
- Since the above analysis assumed, simplistic values of interconnect geometry, and underlying oxide thickness, what will be the effects on the self-consistent solutions, if technology and interconnect scaling were considered?
- Since the waveform shape or duty cycle r plays an important role in determining the maximum allowed T_m and j_{peak}, what is a realistic value for r from a circuit designer's point of view?
- For optimized driver and interconnect signal length (between repeaters) design, how do the j_{peak} and j_{rms} values (from purely electrical considerations such as signal delay) compare with those generated using the self-consistent approach?
- Since the above analysis assumed an isolated interconnect line, what will be the effects on the self-consistent solutions if real 3-D interconnect arrays, with heat generation in all the lines, are considered?
- Finally, the question also arises that how do these j_{peak} values compare with peak current density values causing open circuit or latent metal damage in metal lines under ESD conditions?

3.2 Technology Scaling and Quasi 2-D Heat Conduction Effects on Self-Consistent Design Rules

In this section we will address the first three issues outlined above. We will first analyze the effects of line width scaling on the thermal impedance (or W_{eff}) defined in (10) which is accurate to within 3% for $W_m/t_{ox} > 0.4$ [17]. In current DSM technologies the ratio W_m/t_{ox} gets even smaller (especially for the top-level metal lines), and W_{eff}, which is meant to account for the extra heat conduction, (illustrated in Fig. 4) from the sides must be obtained for real DSM interconnect structures experimentally. We will first express W_{eff} in a more general form,

$$W_{eff} = W_m + \phi \cdot t_{ox} \qquad (14)$$

Here ϕ is introduced as the heat spreading parameter that must be extracted from data.

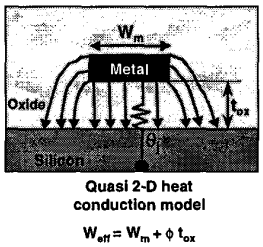

Figure 4. Schematic representation of quasi 2-D heat conduction scenario in DSM interconnects.

Furthermore, new intra-level insulating materials for DSM interconnects such as low-k dielectrics, which lower interconnect capacitance (and hence RC delay and cross-talk noise) and enhance circuit performance, are known to have poor thermal properties [13]. The thermal characteristics of these low-k dielectrics and their impact on the interconnect reliability is important.

For extracting the appropriate value of ϕ, the thermal impedance was experimentally determined using (8), for AlCu interconnects fabricated in a state-of-the-art 0.25-μm industrial CMOS process flow. Results are shown in Fig. 5 for a standard oxide process and a low-k (HSQ) intra-level (gap-fill) process. We can observe that for the narrowest line width (W_m = 0.35-μm) the thermal impedance for the low-k gap-fill process is around 20% higher than the value for standard oxide in agreement with the trends shown in [13]. Furthermore, since the ratio of W_m/t_{ox} (= 0.29) is the lowest for this line width, the heat spreading pattern for this case can be expected to be similar for top level metal lines in DSM technologies. Using the thermal impedance value for W_m = 0.35-μm and using (10) and (14), ϕ was extracted to be 2.45 (for the standard oxide case). The K_{ox} value used in the calculation is given in Table 1.

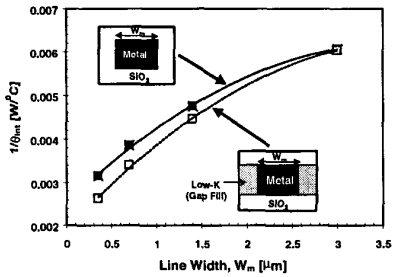

Figure 5. Experimentally obtained effective thermal impedance values for level 1 AlCu metal lines in a 0.25-μm CMOS process, with two different intra-level dielectrics. The t_{ox} was 1.2 μm, and interconnect length, L = 1000 μm.

Dielectric Material	Thermal Conductivity [W/m °K]
Oxide (PETEOS)	1.15*
HSQ	0.6**
Polyimide k	0.25**

Table 1. Thermal conductivity values (normal to the dielectric plane) for various dielectrics analyzed in this study. These values were obtained from [19] for * and [20] for **.

We now consider the technology specifications for a 0.25 μm and 0.1 μm process as per the NTRS [15] shown in Table 8 in the appendix. The interconnect metal is assumed to be Cu, and the effects of using various intra-level dielectrics (listed in Table 1) are analyzed.. We then solve for the self-consistent metal temperature, T_m and the corresponding maximum allowed j_{rms} and j_{peak} values. However, while solving the self-consistent equation (13), we replace the $(t_{ox}/K_{ox}W_{eff})$ term by the following generalized term:

$$\left(\frac{t_{ox}}{K_{ox}} + \frac{t_{low-k}}{K_{low-k}}\right)\frac{1}{W_m + \phi \cdot t_{total}} \qquad (15)$$

This is necessary to take care of the heat spreading effect in DSM structures and the use of low-k materials as intra-level gap-fill dielectric. The value of ϕ is kept unchanged even for the case of gap-fill interconnect structures, since the heat flow pattern is not significantly altered in the presence of gap-filled dielectric. Also, the duty cycle r in (13) is taken as 0.1, for all cases, due to reasons explained in the next section.

Before presenting the self-consistent solutions, we must point out an important caveat, which is the distinction between *thermally long* and *thermally short* metal lines. Since interconnect leads are typically connected to the diffusion (or another metal lead) through a contact (or via), the temperature at the end regions of these lines are in general lower than the temperature in regions far away from the ends. The solution of the differential equation for such 2-D heat conduction yields that the spatially dependent parts of the solution depend exponentially on a characteristic thermal length, λ [21] which is of the order of 25 μm - 200 μm. Metal lines which are >> λ are termed *thermally long* while those which are close to the value of λ are termed *thermally short*.

We will focus on the analysis of the thermally long lines which give rise to worst case scenario. We will not concentrate on thermal effects for small inter-block interconnects. Their lengths are usually of the same order of magnitude as the thermal characteristic length. Also these nets are typically routed on the first few levels of metal, which are closer to the silicon substrate, and hence the thermal problem is not as severe. Long interconnects which could be much larger than the thermal characteristic length will therefore form the main focus of the following analysis. These long lines usually realize the inter-block communication and hence for delay purposes, they are routed on higher, less resistive top few layers of metal. However, these metal layers are usually far away from the silicon substrate and hence can potentially experience thermal problems.

Therefore, self-consistent solutions were obtained for the top metal layers only in the 0.25-μm and 0.1-μm technologies. Table 2 gives the values obtained for the maximum allowed j_{peak} for a j_0 of 6 x 10^5 A/cm² and for two values of the duty cycle. Similar analysis was also done for a 300% higher value of j_0 (=1.8 x 10^6 A/cm²), to be more realistic for Cu EM, and is shown in Table 3.

We can observe from Table 2 and Table 3, that within a given technology node, as we go up the metallization level the maximum allowed value of j_{peak} reduces as expected due to increasing thermal effects. Secondly, the effect of the low-k materials with lower thermal conductivity lowers the maximum allowed j_{peak} even further.

Thirdly, as we can see from Table 3, increasing the value of j_0 (at r = 0.1) results in increases in the maximum j_{peak} values, indicating the positive result of using Cu as the interconnect metal, which has higher EM resistance compared to AlCu. In order to provide a direct comparison between Cu and AlCu, maximum allowed j_{peak} (self-consistent) values for AlCu are presented in Table 4.

	0.25 μm Cu Technology			0.1 μm Cu Technology		
Metal	Oxide	HSQ	Polyimide k	Oxide	HSQ	Polyimide k
M5	3.89	3.3	2.53	5.03	4.55	3.77
M6	3.82	3.22	2.46	4.98	4.49	3.71
M7	-	-	-	3.85	3.25	2.5
M8	-	-	-	3.80	3.20	2.45

Signal Lines (r = 0.1)

	0.25 μm Cu Technology			0.1 μm Cu Technology		
Metal	Oxide	HSQ	Polyimide k	Oxide	HSQ	Polyimide k
M5	0.551	0.519	0.457	0.586	0.574	0.546
M6	0.547	0.514	0.450	0.585	0.573	0.543
M7	-	-	-	0.549	0.516	0.454
M8	-	-	-	0.546	0.513	0.449

Power Lines (r = 1.0)

Table 2. Maximum allowed peak current density values (x 10^6 A/cm²) from the self-consistent approach, for two different technology nodes based on Cu and various low-k dielectrics. j_0 = 6 x 10^5 A/cm².

	0.25 μm Cu Technology			0.1 μm Cu Technology		
Metal	Oxide	HSQ	Polyimide k	Oxide	HSQ	Polyimide k
M5	6.06	4.80	3.43	9.47	7.77	5.78
M6	5.88	4.64	3.31	9.28	7.60	5.64
M7	-	-	-	5.94	4.72	3.38
M8	-	-	-	5.84	4.61	3.29

Signal Lines (r = 0.1)

	0.25 μm Cu Technology			0.1 μm Cu Technology		
Metal	Oxide	HSQ	Polyimide k	Oxide	HSQ	Polyimide k
M5	1.2	1.02	0.786	1.53	1.39	1.16
M6	1.17	0.994	0.765	1.52	1.37	1.14
M7	-	-	-	1.19	1.01	0.775
M8	-	-	-	1.17	0.989	0.760

Power Lines (r = 1.0)

Table 3. Maximum allowed peak current density values (x 10^6 A/cm²) from the self-consistent approach, for two different technology nodes based on Cu and various intra-level dielectrics. j_0= 1.8 x 10^6 A/cm².

	0.25 μm AlCu Technology			0.1 μm AlCu Technology		
Metal	Oxide	HSQ	Polyimide k	Oxide	HSQ	Polyimide k
M5	3.49	2.91	2.21	4.69	4.15	3.37
M6	3.41	2.84	2.15	4.63	4.09	3.30
M7	-	-	-	3.45	2.87	2.18
M8	-	-	-	3.39	2.82	2.13

Signal Lines (r = 0.1)

	0.25 μm AlCu Technology			0.1 μm AlCu Technology		
Metal	Oxide	HSQ	Polyimide k	Oxide	HSQ	Polyimide k
M5	0.529	0.488	0.420	0.578	0.561	0.522
M6	0.524	0.482	0.410	0.577	0.558	0.518
M7	-	-	-	0.527	0.486	0.415
M8	-	-	-	0.523	0.481	0.409

Power Lines (r = 1.0)

Table 4. Maximum allowed peak current density values (x 10^6 A/cm²) from the self-consistent approach, for two different technology nodes based on AlCu and various low-k dielectrics. j_0 = 6 x 10^5 A/cm².

4 Impact on Circuit Performance and Design

4.1 Effects on Signal Line Length, and Driver Size Optimization

In this section we will address the next two issues pointed out in section 3.1. As we have mentioned in Section 3.2, thermal effects are predominant in semi-global and global interconnects which are thermally long. These are used to connect functional blocks (*macros*) over relatively long distances and hence utilize the upper layers of interconnect. Furthermore, long interconnect lines are split into buffered segments [22]. Buffer insertion is also used to reduce the crosstalk noise in high performance circuits [23]. It has been shown [22] that for point to point connections, the optimal length of interconnect at which to insert repeaters is given by

$$l_{opt} = 1.3\sqrt{r_0 \frac{(c_0 + c_p)}{r \cdot c}} \qquad (16)$$

where r_0 is the effective driver resistance for a minimum sized driver, c_0 is the input capacitance, c_p is the output parasitic capacitance, r is the interconnect resistance per unit length and c is the interconnect capacitance per unit length (see Fig. 6). For DSM technologies, a significant fraction of c would be contributed by coupling capacitances to neighboring lines. For our simulations, we performed a full 3D-capacitance extraction using SPACE3D [24] for the signal lines to obtain the value of c for every metal layer for both technologies. The model assumes that the output V_{tr} makes a transition as soon as the input voltage crosses a threshold (say $0.5V_{dd}$). The optimum repeater size is given by

$$s_{opt} = \sqrt{\frac{r_0 \cdot c}{c_0 \cdot r}} \qquad (17)$$

which means that the width of the NMOS and PMOS for this inverter are obtained by scaling the width of the corresponding transistors in a minimum sized inverter by a factor s_{opt}. This also implies that the delay between any two optimally spaced and sized repeaters is independent of the layer [22]. This model also implies that it is not optimal to insert buffer in a line, which is of length less than l_{opt} for a given metal layer. Our simulations suggest that drivers and interconnects optimized with this model maintain good slew rates for rising and falling transitions for all the metal layers and across technologies. For lines of length less than l_{opt}, the buffer size can also be reduced to $\{s_{opt}(l/l_{opt})\}$ to reduce the power dissipation while still maintaining good slew rates. Hence, for a given metal level, the maximum RMS current occurs in an optimally buffered optimal length interconnect. Also, given the distributed nature of the interconnect, the maximum RMS current occurs close to the repeater output. Hence, we need to first verify whether this maximum current, which is dictated by delay considerations also meets the thermal specifications obtained in Section 3.2.

In the following, we perform some calculations using a well-reviewed technology file (see Appendix) based on [15] for 0.25 μm and 0.10-μm technology based on copper and low-k interconnect structures. We use these parameters to determine the peak and RMS current (and therefore current density) by SPICE simulation where we take into account all the device parasitics. We denote the maximum current density due to delay considerations as $j_{\text{-delay}}$ and due to thermal and EM considerations as $j_{\text{-self-consistent}}$.

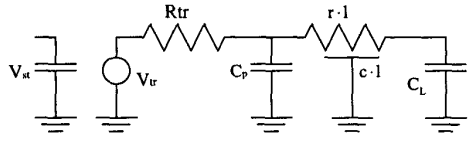

Figure 6. Equivalent distributed network for driver and interconnect. Vst is the voltage at the input capacitance that controls the voltage source. Rtr is the driver transistor resistance and C_L is the load capacitance of the next stage. l is the interconnect length.

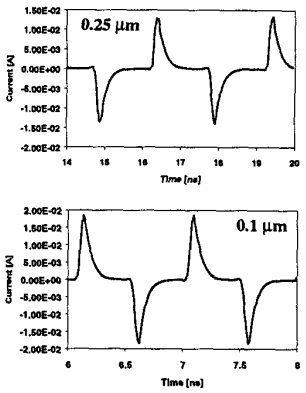

Figure 7. Current waveforms in the top layer metal lines obtained from SPICE simulations for the 0.25-μm and the 0.1-μm technologies.

As suggested earlier, we find that the relative slew rate defined as the 10%-90% rise time as a fraction of the clock period is almost constant across all metal layers and across technologies. This is also reflected by the fact that the effective duty cycle obtained for these simulations was 0.12±0.01 for every case considered. The interconnect current waveform in the top layer metal lines for 0.25 μm and 0.1 μm technologies, as obtained from transient SPICE simulations, is plotted in Fig. 7, which shows that the relative rise and fall skew is the same across both technologies. For lines which are not on critical path, the buffer size may be reduced to save power in which case the effective duty cycle will actually increase slightly, but, as suggested from Figures 2 and 3, this will not change $j_{\text{-self-consistent}}$ significantly. This justifies the selection of the duty cycle value as 0.1 in section 3.2 for the self-consistent analysis on various technologies.

Comparing the $j_{\text{peak-delay}}$ values (Table 5 and 6) with the values of the maximum allowed peak current density $j_{\text{peak-self-consistent}}$ as determined by the self consistent equations (Table 2), we find that the $j_{\text{peak-delay}} < j_{\text{peak-self-consistent}}$ for silicon dioxide as the dielectric. However, alternative dielectric materials with lower relative permittivity are being introduced to lower interconnect delay.

Layer	L_{opt} (mm)	S_{opt}	j_{rms} (x10^{5}A/cm^2)	j_{peak} (x10^{5}A/cm^2)
1	2.51	75	9.23	27.02
2	2.41	80	8.76	26.38
3	4.79	157	5.03	14.69
4	5.48	148	4.16	13.48
5	12.5	494	1.76	5.18
6	13.9	454	1.58	4.69

Table 5. Optimized interconnect and buffer parameters and corresponding RMS and peak current densities for a 0.25-μm Cu technology. Insulator dielectric constant = 3.3.

Layer	L_{opt} (mm)	S_{opt}	j_{rms} (x10^{5}A/cm^2)	j_{peak} (x10^{5}A/cm^2)
1	0.85	36	14.24	38.99
2	0.9	40	14.32	39.48
3	1.57	97	7.57	20.93
4	1.79	85	6.63	18.27
5	4.47	276	3.56	9.78
6	4.68	249	3.35	9.00
7	8.8	500	1.35	3.67
8	9.58	522	1.26	3.5

Table 6. Optimized interconnect and buffer parameters and corresponding RMS and peak current densities for a 0.1-μm Cu technology. Insulator dielectric constant = 2.0.

Also, the thermal conductivity of these low-k dielectrics is much lower than silicon dioxide (Table 1). This impacts in two ways. First, as suggested by (16) and (17) the optimum unbuffered interconnect length increases and the optimum repeater size decreases. It can be shown that s_{opt} and $c.l_{opt}$ decrease by the same factor and hence the RMS current density remains almost unchanged. Secondly, the lower thermal conductivity of these materials causes $j_{peak\text{-}self\text{-}consistent}$ to decrease by almost a factor of two for these low-k dielectrics and the margin between $j_{peak\text{-}self\text{-}consistent}$ and $j_{peak\text{-}delay}$ reduces. As we will see in Section 5, the thermal situation is exacerbated with the presence of other active metal lines at the same metal layer as well as other layers. This further reduces $j_{peak\text{-}self\text{-}consistent}$ for these configurations. Hence, self-heating needs to be considered in high performance deep sub-micron interconnect design that employs low-k dielectrics.

The above analysis assumes that the interconnects are always active. This is a valid assumption for global lines, because although the activity period of individual transistors is reducing with scaling and increasing transistor count, block size is also increasing and hence the activity period for the whole block is relatively unchanged. Hence the global wires which communicate with these blocks have a high activity rate. Therefore, duty cycle of 0.1 used in the study is valid. Also note that the circuit analysis of this section is for signal lines, which carry bi-directional currents. These lines are known to have much higher EM immunity, hence the self-consistent values of j_{rms} and j_{peak}, as determined in Table 2 for the unipolar case, are lower bounds.

5 Thermal Effects in Real 3-D Interconnect Arrays

In this section we will briefly address the penultimate issue pointed out in section 3.1. Our self-consistent analysis of self-heating and EM effects presented earlier in section 3.2 were based on single isolated interconnect lines. In a real IC there are densely packed layers of interconnect lines which form a 3-D array as illustrated in Fig. 8. The self-heating of interconnect lines within such array could be significantly more severe due to thermal coupling between neighboring lines. The heat flow analysis for such structures is complicated and must involve numerical simulation techniques. Rzepka et al., have recently done a very detailed analysis of self-heating effects in such structures using finite element simulations [11].

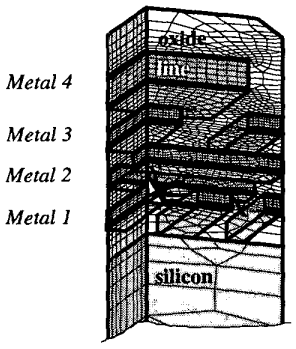

Figure 8. Densely packed 3-D interconnect array in a quadruple level metallization process. Self-heating of any line within such an array is strongly affected by the activity (*hot* or *cold*) of the neighboring lines.

The RMS current density can be empirically shown to obey the following relationship (see equation 9)

$$j^2_{rms} \propto \frac{T_m - T_{ref}}{\rho_m(T_m)} \qquad (18)$$

and the proportionality constant (which is independent of the interconnect material) can be obtained empirically from the finite element analysis presented in [11] for different interconnect configurations, which includes 2-D and 3-D effects. Substituting this equation and the EM lifetime equation (11) in (12), we can obtain another self-consistent equation similar to (13). From this we have obtained the maximum allowed j_{peak} for metal 4 in a 3-D configuration using Cu technology. We also compare this with the j_{peak} obtained from (13) for an isolated metal 4. We find that the maximum allowed j_{peak} reduces by nearly 40% for the 3-D case, where all the metal lines are heated. Hence we can conclude that the maximum allowed j_{peak} for real interconnects would also reduce significantly from the values calculated in Section 3.2.

	Maximum Allowed j_{peak} (MA/cm^2)
M1-M4 heated (3-D)	6.4
Isolated M4 heated (2-D)	10.6

Table 7. Comparison of maximum allowed peak current density values for a metal 4 line in a dense array (as in Fig. 8) with all neighboring lines heated and an isolated (as in section 3) metal 4 line.

6 Thermal Effects under ESD Conditions

In this section we will briefly address the last issue pointed out in section 3.1. A special case of thermally accelerated interconnect failure can also occur under ESD conditions. ESD is a high current [> 1 A], short-time scale [< 200 ns], phenomena that can lead to catastrophic open circuit failures, and to latent damage [25]. Integrated circuit failures due to ESD have significantly increased in importance as VLSI technologies continue to shrink towards the deep sub-micron regime [26]. Recently, Banerjee et al., [8] have shown that the critical current density for causing open circuit metal failure in AlCu interconnects is ~ 60 MA/cm^2, and have presented a short-timescale high-current failure model for designing robust interconnects to avoid thermal failure under high peak current conditions including those encountered in ESD protection circuits and I/O buffers. Interconnects have also been shown to suffer latent damage if the lines resolidify after melting [9] and this has been shown to degrade the EM lifetime. The model presented in [8] can be used to avoid this reliability hazard. This model can also be applied to design Cu interconnects as shown by Voldman et al. [27]. These interconnect design rules must be obeyed for high current robustness.

7 Conclusions

In conclusion, *thermal effects* in advanced high performance interconnect systems arising due to self-heating under various circuit conditions, including ESD has been examined in detail. A self-consistent approach, that simultaneously comprehends self-heating and EM, has been first modified using experimentally measured self-heating data to include quasi 2-D heat conduction in DSM metal lines to allow more aggressive design rules. This modified self-consistent formulation has been applied to analyze the implications of interconnect technology (Cu, low-k etc.) and scaling. Two different Cu based technology nodes (0.25-μm and 0.1-μm) as per NTRS has been used in this study.

Using circuit level simulations we have confirmed that the duty cycles for optimized global interconnects remain nearly invariant across metal layers and technologies. A value of 0.12±0.01 was obtained for all cases. This analysis has helped provide a more realistic value of the duty cycle for analyzing thermal and EM effects using the self-consistent approach. Secondly, the impact of these thermal effects on the design (driver sizing) and optimization of the interconnect length between *repeaters* at the upper-level signal lines have been investigated. We have shown that although the maximum allowed peak current density based on the self-consistent criteria is greater than the peak current density obtained from optimized driver and interconnect configurations (i.e., $j_{peak\text{-}self\text{-}consistent} > j_{peak\text{-}delay}$), for standard oxide, but they get closer as low-k dielectric materials are introduced. Furthermore, we have shown that in real 3-D interconnect arrays the greater self-heating due to thermal coupling effects could further lower the $j_{peak\text{-}self\text{-}consistent}$ by nearly 40%. Hence for DSM technologies thermal effects in interconnects need careful consideration.

Finally, we have pointed out that although the $j_{peak\text{-}self\text{-}consistent}$ values are much lower than the current densities causing open circuit metal failure under ESD type conditions, the interconnects in ESD protection circuits and I/O buffers must be designed separately

to meet high current robustness. These results will have important implications for providing robust and aggressive deep sub-micron interconnect design guidelines.

8 Acknowledgments

This work was supported by the SRC and the MICRO program. The authors would like to thank Dr. William Hunter, of Texas Instruments, Dallas, for several helpful discussions, and Phil Fischer of SEMATECH and Sunil Khatri of UC Berkeley for their help in generating the technology file in the appendix.

9 References

[1] C. R. Barrett, "Microprocessor evolution and technology impact," *Symp. VLSI Technol., Dig. Tech. Papers*, 1993, pp. 7-10.

[2] T. Makimoto, "Market and technology trends in the nomadic age," *Symp. VLSI Technol., Dig. Tech. Papers*, 1996, pp. 6-9.

[3] R. Whittier, "Push/Pull: PC technology/end user demand," *Symp. VLSI Technol., Dig. Tech. Papers*, 1996, pp. 2-5.

[4] R. H. Dennard, F. H. Gaensslen, H. Yu, V. L. Rideout, E. Bassous, and A. R. LeBank, "Design of ion-implanted MOSFETs with very small physical dimensions," *IEEE J. Solid-State Circuits*, Vol. SC-9, pp. 256-268, 1974.

[5] P. K. Chatterjee, W. R. Hunter, A. Amerasekera, S. Aur, C. Duvvury, P. E. Nicollian, L. M. Yang, and P. Yang, "Trends for deep submicron VLSI and their implications for reliability," *Proc. IRPS*, 1995, pp. 1-11.

[6] J. R. Black, "Electromigration – A brief survey and some recent results," *IEEE Trans. Electron Devices*, vol. ED-16, pp. 338-347, 1969.

[7] B. K. Liew, N. W. Cheung, and C. Hu, "Projecting interconnect electromigration lifetime for arbitrary current waveforms," *IEEE Trans. Electron Devices*, vol. 37, pp. 1343-50, 1990.

[8] K. Banerjee, A. Amerasekera, N. Cheung and C. Hu, "High-current failure model for VLSI interconnects under short-pulse stress conditions," *IEEE Electron Device Lett.*, vol. 18, No. 9, pp. 405-407, 1997.

[9] K. Banerjee, A. Amerasekera and C. Hu, "Characterization of VLSI circuit interconnect heating and failure under ESD conditions," *Proc. IRPS*, 1996, pp. 237-245.

[10] W. R. Hunter, "Self-consistent solutions for allowed interconnect current density – Part I: Implications for technology evolution," *IEEE Trans. Electron Devices*, vol. ED-44, pp. 304-309, 1997.

[11] S. Rzepka, K. Banerjee, E. Meusel, and C. Hu, "Characterization of self-heating in advanced VLSI interconnect lines based on thermal finite element simulation," *IEEE Trans. on Components, Packaging and Manufacturing Technology-Part A*, vol. 21, No. 3, pp. 1-6, 1998.

[12] J. Ida et al, "Reduction of wiring capacitance with new low dielectric SiOF interlayer film for high speed/low power sub-half micron CMOS," *Tech. Dig. VLSI Symp.*, pp. 59-60, 1994.

[13] K. Banerjee, A. Amerasekera, G. Dixit and C. Hu, "The effect of interconnect scaling and low-k dielectric on the thermal characteristics of the IC metal," in *Tech. Dig. IEDM*, 1996, pp. 65-68.

[14] NS Nagaraj, F. Cano, H. Haznedar, and D. Young, "A practical approach to static signal electromigration analysis," *Proc. 35th Design Automation Conf.*, 1998, pp. 572-577.

[15] National Technology Roadmap for Semiconductors (NTRS).

[16] J. R. Black, "Electromigration failure modes in aluminum metallization for semiconductor devices," *IEEE Trans. Electron Devices*, vol. 57, no. 9, pp. 1587-1594, 1969.

[17] A. A. Bilotti, "Static temperature distribution in IC chips with isothermal heat sources," *IEEE Trans. Electron Devices*, vol. ED-21, pp. 217-226, 1974.

[18] W. R. Hunter, "Self-consistent solutions for allowed Interconnect current density – Part II: Application to design guidelines," *IEEE Trans. Electron Devices*, vol. ED-44, pp. 310-316, 1997.

[19] C. Jin, L. Ting, K. Taylor, T. Seta, and J. D. Luttmer, "Thermal conductivity measurement of low dielectric constant films," in *Proc. Second International Dielectrics for VLSI/ULSI Multilevel Interconnection Conference (DUMIC)*, 1996, pp. 21-28.

[20] *Private Communications*, Professor Kenneth Goodson, Thermosciences Division, Mechanical Eng. Department, Stanford University.

[21] H. A. Schafft, "Thermal analysis of electromigration test structures," *IEEE Trans. Electron Devices*, vol. ED-34, pp. 664-672, 1987.

[22] R. H.J.M. Otten and R. K. Brayton, "Planning for performance," *Proc. 35th Design Automation Conf.*, 1998, pp. 122-127.

[23] J. Culetu, C. Amir, and J. McDonald, "A practical repeater insertion method in high speed VLSI circuits," *Proc. 35th Design Automation Conf.*, 1998, pp. 392-395.

[24] "Physical design modelling and verification project (SPACE Project)", http://cas.et.tudelft.nl/research/space.html

[25] C. Duvvury and A. Amerasekera, "State-of-the-art issues for technology and circuit design of ESD protection in CMOS ICs," *Semiconductor Science and Tech.*, pp. 833-850, 1996.

[26] A. Amerasekera and C. Duvvury, "The impact of technology scaling on ESD robustness and protection circuit design," in *EOS/ESD Symp. Proc.*, 1994, pp. 237-245.

[27] S. H. Voldman, "ESD robustness and scaling implications of aluminum and copper interconnects in advanced semiconductor technology," in *EOS/ESD Symp. Proc.*, 1997, pp. 316-329.

Appendix

Process (μ)		0.25	0.1
V_{DD} (V)		2.0	1.2
L_{eff} (nm)		160	50
T_{ox} (nm)		5.0	2.5
Metal Levels		6	8
Poly	H (μ)	0.2	0.1
	W (μ)	0.25	0.1
	Space (μ)	0.25	0.1
	Sheet ρ (Ω/square)	4	8
M1-2	H (μ)	0.5	0.26
	W (μ)	0.3	0.13
	Space (μ)	0.3	0.13
	Sheet ρ (Ω/square)	0.044	0.085
	t_{ins} (nm)	650	320
M3-4	H (μ)	0.9	0.55
	W (μ)	0.6	0.3
	Space (μ)	0.6	0.30
	Sheet ρ (Ω/square)	0.024	0.04
	t_{ins} (nm)	900	600
M5-6	H (μ)	2.5	1.2
	W (μ)	2.0	1.0
	Space (μ)	2.0	1.0
	Sheet ρ (Ω/square)	0.009	0.018
	t_{ins} (nm)	1400	800
M7-8	H (μ)	-	2.5
	W (μ)	-	2.0
	Space (μ)	-	2.0
	Sheet ρ (Ω/square)	-	0.009
	t_{ins} (nm)	-	1400

Table 8. Cu based NTRS interconnect technology file for the two processes used in this study.