# Thermal Analysis of the Fusion Limits of Metal Interconnect under Short Duration Current Pulses

*Kaustav Banerjee, Sven Rzepka, Ajith Amerasekera*, Nathan Cheung and Chenming Hu*
Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94720.
Tel: 510 643 7036 Fax: 510 642 2739 e-mail: kaustav@vivante.eecs.berkeley.edu

*Semiconductor Process and Device Center, Texas Instruments Inc., 13536 North Central Expressway, MS 461, Dallas, TX 75243. Tel: 214 995 7985 Fax: 214 995 2770 e-mail: ajith@spdc.ti.com

## Abstract

Thermal analysis of the fusion limits of the IC metal under short duration current pulses has been performed using a quadruple level TiN/AlCu/TiN metallization system. A finite element (FE) simulation program has been calibrated to analyze the thermal effects in detail. The program can be used to predict self heating under DC and transient current conditions for various metal levels, geometries and current loading conditions. It is shown both experimentally and using FE simulations that the metal temperatures rise past 1000 °C before open circuit failure under short duration current pulses. The critical failure current is strongly influenced by the metal thickness, thermal capacity and pulse width. Further, it is shown that the ratio of the critical energy causing open circuit conditions (fusion limit), to the theoretical melt energy increases with scaling. As a result, narrower metal lines can sustain higher current densities before failure.

## Introduction

Aggressive scaling of Si based IC devices motivated by the desire for faster speed and higher packing density has increased the functional complexity of VLSI circuits. This has in turn, reduced the metal pitch and increased the number of metallization levels. Recently it has been demonstrated that thermal effects, instead of electromigration (EM) itself will start to dominate interconnect design guidelines for advanced high performance interconnects [1,2]. Further, metal interconnects get heavily stressed due to high joule heating in field programmable gate arrays (FPGA) [3] wherein the interconnect metal is exposed to short duration, high current pulses. Also, metal heating and failure caused by short time current pulses as encountered during ESD/EOS testing and radiation testing for latchup has become a reliability issue [4]. It is desirable to comprehend the fusion limits of the interconnect under such current pulses since IC interconnects will soon encounter this limit as the ultimate limit of their scalability. We have recently presented a model for interconnect heating and failure under pulsed current stress [5]. The purpose of this paper is to extend our previous work, using experimental data and finite element simulations, to comprehend the implications of fusion limits of the IC metal under a short duration current pulse and to study the effect of interconnect scaling on this limit.

## Experiments/Finite Element Model

The test samples used in the experiments were prepared in a quadruple level metallization (QLM) wafer process. The test structures, were isolated, i.e. there were no metal layers on top of another. A 2-D finite element model was set up to describe the cross section in detail (Fig. 1). Fig. 1 also shows the 3-D model which has been generated to consider a cell of densely packed array structures including the case of lines crossing at different levels, as an example for real circuits. In both models it is possible to select any of the metallization levels and lines to be included in the actual simulation test.

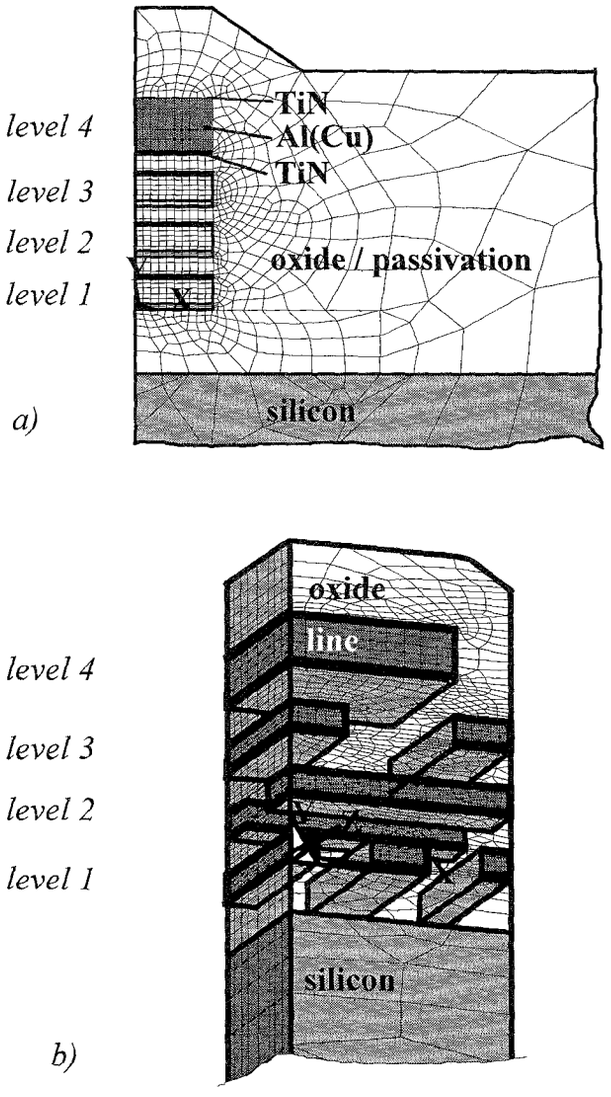



*Figure 1.* 2-D and 3-D finite element models. (a) test structure, W=3μm for all levels, example shown for metal 4, and (b) real structure in a chip, W= 0.5/0.5/1.0/3.0μm for metal 1/2/3/4.


0-7803-3598-8/97/$10.00 ©1997 IEEE

In our example, all test lines were 3 μm by 1000 μm NIST recommended structures. The metal system was multilayered with a stacking sequence TiN/AlCu(0.5%)/TiN. The material data for the models were mostly taken from the literature. The resistance rise under various DC loads was measured for the QLM wafer. One data point per level 1 and 4 was chosen to calibrate the finite element models with respect to thermal conductivity of the dielectric and the heat film coefficient between passivation and air. Fig. 2 depicts the excellent agreement between experimental and simulation results. Deviations are randomly distributed. They never did exceed 5 %.

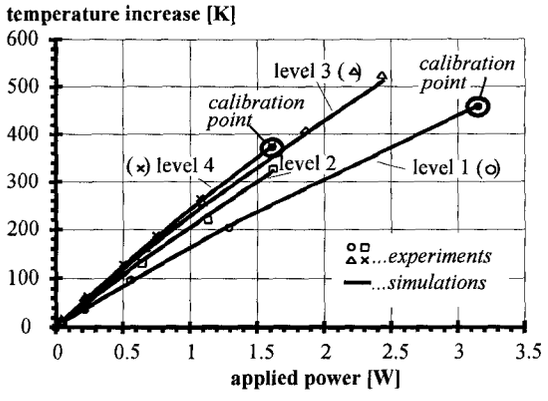
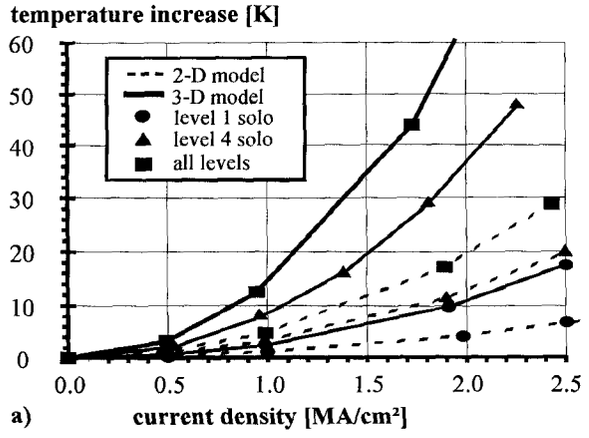



*Figure 2.* Calibration of the finite element model using DC current induced self heating in the QLM test structures. Excellent agreement between all experimental data and simulation results proves the accuracy of the model.

Applying the calibrated models temperature increase (ΔT), thermal impedance, and other characteristics for various design alternatives can be accurately extracted from simulations without further experiments. Fig. 3a gives simulation example for temperature increase in single line and densely packed structures under single line loads and complex load cases similar to real operational conditions. It can be observed that self heating is more severe in real structures (solid lines) than in typical test structures (broken lines). Fig. 3b plots simulated thermal impedance (of metal 4) for different current loads in levels 1, 2 and 3, shown as a fraction of current load in level 4, for a test structure. It is observed that the thermal impedance of metal 4 increases as other levels are loaded. Interestingly, all the lines cross over at 40%, indicating a optimum load fraction in the lower levels at which self heating of metal 4 remains unchanged.



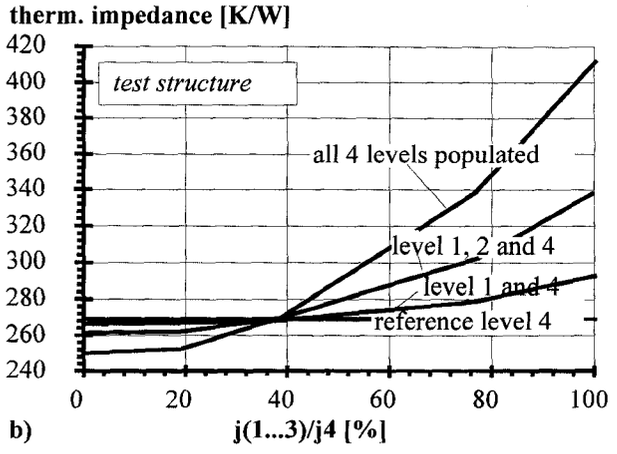




*Figure 3.* Simulation results using the FE method showing (a) simulated temperature rise under a DC current, test structures: 2-D model (broken lines), real structures: 3-D model (solid lines) and (b) thermal impedance of interconnects.

A standard transmission line pulsing technique [6] was then used to generate constant current pulses of varying widths (Δt = 100 ns, 200 ns, 400 ns and 500 ns) and amplitudes. The voltage, and hence the resistance of the metal lines increased roughly linearly with time during all the pulsing events as shown schematically in Fig. 4.

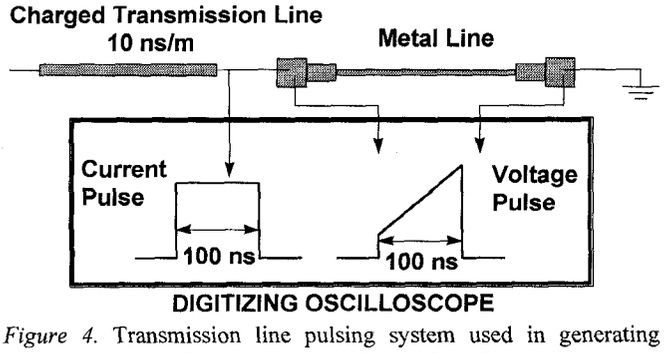



*Figure 4.* Transmission line pulsing system used in generating constant current pulses of varying widths and amplitudes.

In Fig. 5. the maximum resistance rise above initial room temperature value, ($\Delta R_{max}/R_0=\gamma$, where, $\Delta R_{max}= R_f - R_0$), for each metal level has been plotted against (a) the current density, and (b) the pulse energy, for a 200 ns pulse. From this figure it can be observed that for all the metal levels γ rises superlinearly with J, the current density, and they all fail when γ goes beyond a critical value ($=\gamma_{crit}$). It was shown that the metal lines were being heated past 1000 °C before the passivating dielectric cracked open due to the thermal stress causing an open circuit [6]. The melt temperature of the metal lines is ~660 °C which indicated that for open circuit condition the lines not only need to melt but the molten metal must expand to crack the passivating dielectric (in this case, layers of $SiO_2$ and $Si_3N_4$). Furthermore, it can be observed that metal lines with identical geometry and hence thermal capacity (level 1, 2 and 3) show similar temperature rise. This implies that the net underlying oxide thickness has no effect. This is expected since the pulse widths are much shorter than the thermal time constants (~2μs) of these interconnect structures, due to which steady state is not attained.

Finite element simulation have been conducted to determine whether or not the interconnection lines may really melt

during these pulses without cracking the passivating layer. Applying the calibrated models, a level 1 metal line was picked as the example for transient simulations. The simulated pulse had a constant current of 850 mA for 100/200/400/500 ns. The corresponding current density (J) was 47 MA/cm². The simulation results show that the lines can survive pulses (400 ns pulse in Fig. 7.) during which they are in the molten state. The longest pulse (500 ns) led to temperatures calculated above 1500 °C.

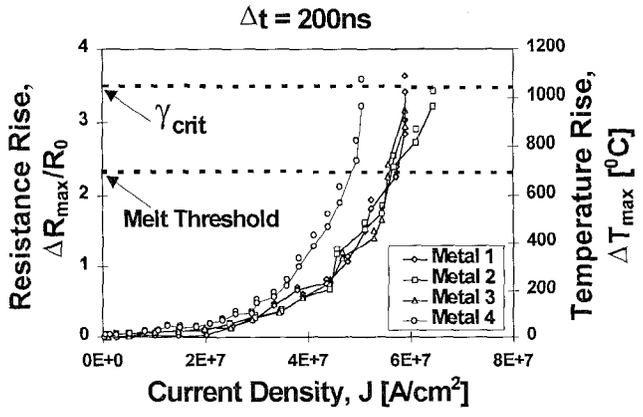



*Figure 5.* Interconnect metal heating under short duration current pulses showing resistance and temperature rise with current density. Note the fusion limit of 1000 °C.

In Fig. 6, the 500 ns curve never crosses the dotted line marking the simulated current density. That means, the line was blown up in the experiment even with a smaller current pulse. The calculated peak temperature increase for the 400 ns pulse was 990 °C, i.e. also 355 °C above the melting point of aluminum. Even when the rapid increase in the line's electrical resistivity during melting was not considered (dotted curve in Fig. 7), the temperature increased well beyond 660 °C. In this case the latent heat of fusion caused delays in temperature evolution at the melting point.

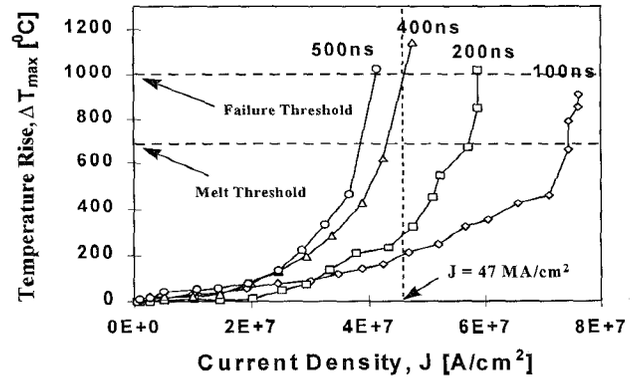



*Figure 6.* Experimental data summarizing the effect of pulse width on heating characteristics. The vertical dotted line shows the current density value (47MA/cm2) at which the transient heating during short pulses were simulated.

Fig. 8 shows a high resolution optical micrograph of interconnect failures caused by a short duration pulse. It can be observed that there are multiple failure spots. This arises due to the rupture of the overlying dielectric layers at localized weak spots that act as stress concentrators. Further, the failure spots are all located away from the bond pads in the region where temperature reaches a maximum value during all pulsing events.

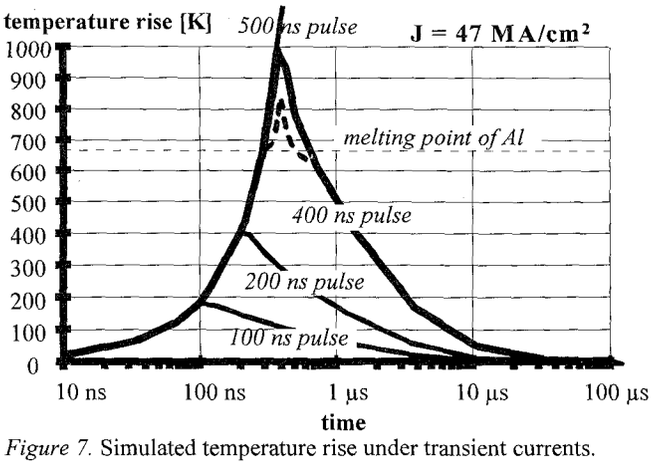



*Figure 7.* Simulated temperature rise under transient currents.

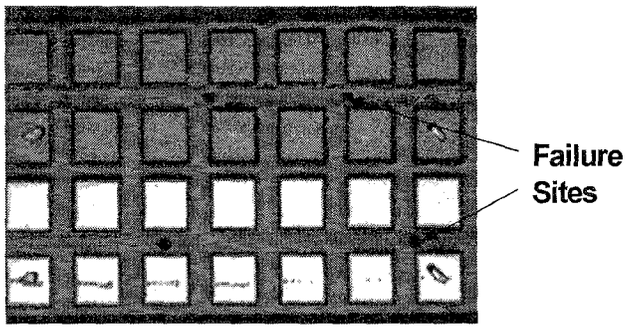



*Figure 8.* Optical micrograph showing open circuit interconnect failure at multiple spots under short duration pulsed current stress.

Before analyzing the effect of interconnect scaling on the fusion limits it is important to first illustrate the heating of a thin sheath of oxide around the metal lines that contribute to the net thermal capacity of the interconnect structures during these short pulses. Fig. 9 shows a plot of the resistance rise (or temperature rise) with pulse energy. The thermal capacities can be extracted from the slopes of these lines since the energy going into the lines can be expressed as the product of the temperature rise and the net thermal capacity of the interconnect structures.

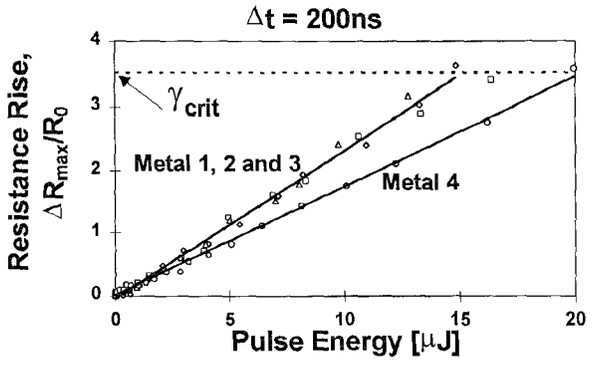



*Figure 9.* Interconnect resistance rise with pulse energy. Thermal capacities are inversely proportional to the slopes of these lines indicating that metal 4 has a bigger thermal capacity.

These thermal capacities are larger than the calculated thermal capacity of the stacked metal line (constant for a given line geometry) which indicates that the extra thermal capacity must be coming from a heated thin oxide sheath surrounding the metal lines. Fig. 10 shows this important concept wherein the extracted thermal capacity is shown to be increasing with pulse width.

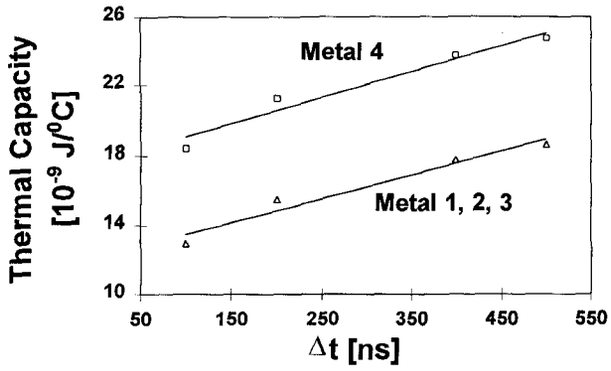



*Figure 10.* Thermal capacity increases with increasing pule width.

This is expected since the thickness of the oxide sheath must increase due to increasing heat diffusion time into the oxide. Note that metal 4 has a relatively bigger thermal capacity due to the greater AlCu thickness.

Fig. 11 shows the relative pulse energies required to cause melting and open circuit failure. These energies increase with pulse width as a result of greater heat diffusion into the surrounding oxide which results in the dissipation of a greater fraction of energy into the surrounding oxide. For pulse widths <<100 ns adiabatic conditions can be realised and nearly all the energy can be absorbed by the metal line itself.

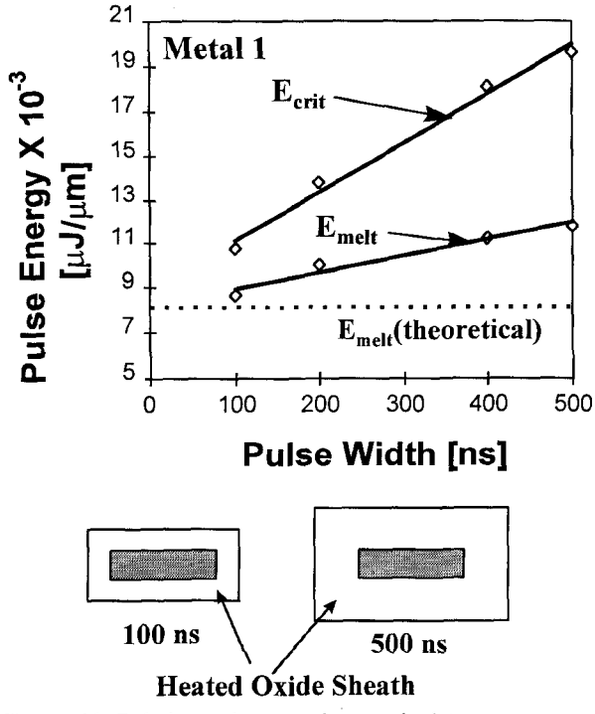



*Figure 11.* Relative pulse energies required to cause open circuit and melting in 3X1000 μm metal leads.

Furthermore, as pulse width increases the volume of surrounding oxide sheath increases, thereby increasing the difference between $E_{crit}$ and $E_{melt}$ as illustrated with the schematic below. It can also be observed that as pulse width approaches the near zero limit, $E_{crit}$ and $E_{melt}$ both approach a constant value as expected from theory. The $E_{crit}$ value will still be higher than the $E_{melt}$ value since the line must be heated past 1000 °C as compared to ~660 °C in the latter case.

Fig. 12 shows the variation of critical current density and current density to cause melting with pulse width. It is observed that the J values are quite close contrary to the energy values (Fig. 11), since ΔT rises superlinearly with J (Fig. 6).

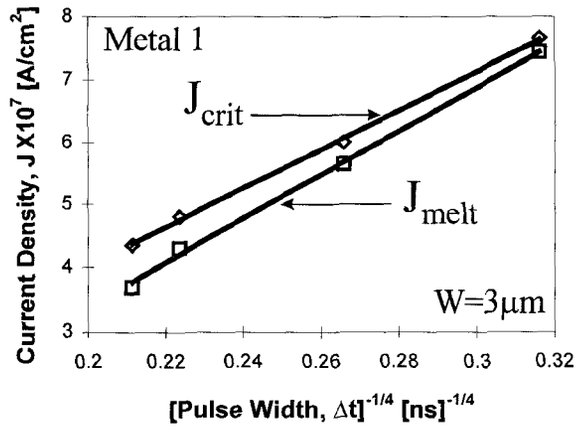
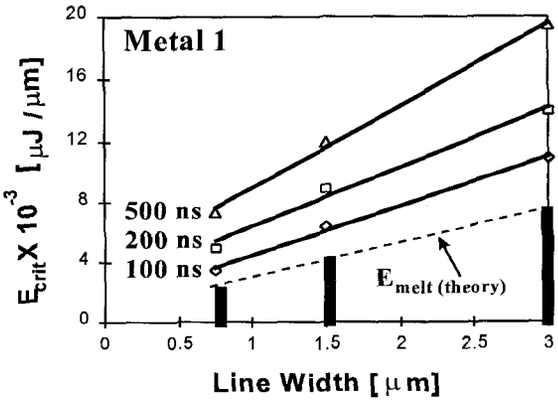


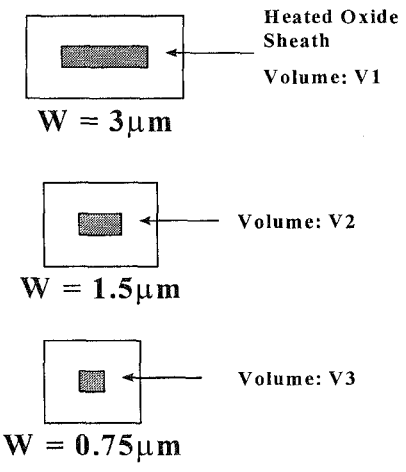

*Figure 12.* Critical and melt current densities for metal interconnects under short pulsed stress.

The effect of interconnect scaling on the fusion energies (causing open circuit or $E_{crit}$) is shown in Fig. 13a for various pulse widths. It can be observed that as line width approaches the infinitesimally thin limit a finite amount of energy is dissipated into the surrounding oxide (all three lines can be extrapolated to intercept the y axis above zero). Furthermore, for a given pulse width the thickness of the oxide sheath remains constant with scaling as illustrated in Fig. 13b. The volume of the oxide sheath decreases with scaling but does not scale with line width. This results in an increase in the ratio, $E_{crit}/E_{melt(theory)}$ . Here $E_{melt(theory)}$ is the calculated melt energy for the metal stack excluding the oxide sheath.




a)

*Figure 13a.* Effect of scaling on the fusion limits of metal lines.




b)

*Figure 13b.* As interconnects are scaled the thickness of the heated oxide sheath remains constant for a given pulse width. The volume of the oxide sheath decreases but does not scale down with line width resulting in increased fraction of energy dissipation in the surrounding oxide.

## Conclusions

In conclusion, thermal behavior of TiN/AlCu/TiN metal interconnects has been characterized and simulated under DC and short duration pulsed current stress. 2-D and 3-D finite element models has been calibrated using DC data and has been used to predict self heating for various geometry and current loading conditions for a multi-level metal system. It has been shown using experiments and these finite element simulations that temperature rise in metal leads can reach ~1000 $^0$C under short duration pulsed current stress before open circuit failure.

Furthermore, the effect of scaling on the fusion limits of interconnects has been analyzed and it has been shown that the ratio of $E_{crit}$ (fusion limit) to $E_{melt}$ increases with decreasing line width. This arises due to the unequal scaling down (with line width) of the volume of a thin oxide sheath around the metal lines as compared to the volume of the metal stack itself. The oxide sheath results from quasi adiabatic conditions during these short duration pulsing events. These fusion limits will have serious implications on the design of high density ICs and must be considered while developing design guidelines for EM, ESD/EOS and I/O buffer interconnects.

## Acknowledgments

This work was funded by the Semiconductor Research Corporation (SRC-96-IJ-148), Deutsche Forschungsgemeinschaft (Germany) and Texas Instruments Inc. The authors would also like to acknowledge W. Hunter of Texas Instruments for his support and encouragement.

## References

[1] P. K. Chatterjee et al, "Trends for Deep Submicron VLSI and Their Implications for Reliability", *Proc. IRPS*, pp. 1-11, 1995.

[2] W. R. Hunter, "The Implications of Self-Consistent Current Density Design Guidelines Comprehending Electromigration and Joule Heating for Interconnect Technology Evolution", *Tech. Dig. IEDM*, pp. 483-486, 1995.

[3] S. Chiang et al, "Antifuse structure comparison for field programmable gate arrays", *Tech. Dig. IEDM*, pp. 611-614, 1992.

[4] S. Ramaswamy et al, "EOS/ESD Reliability of Deep Sub-Micron NMOS Protection Devices", *Proc. IRPS*, pp. 284-291, 1995.

[5] K. Banerjee et al, "Characterization of VLSI Circuit Interconnect Heating and Failure under ESD Conditions", *Proc. IRPS*, pp. 237-245, 1996.

[6] T. J. Maloney et al, "Transmission line pulsing techniques for circuit modeling of ESD phenomena", *ESD/EOS Symp. Proc.*, pp. 49-54, 1985.

## Questions and Answers

**Q1:** In Figure 7, on the right side, temperature is decreasing. Is that because the power has stopped and it is just cooling off?

**A1:** Yes. After the pulse, the line cools down with a characteristic thermal time constant.

**Q2:** In figure 7 the temperature rise is shown in Kelvin while in figure 5 it is in $^0$C, is this O.K?

**A2:** It doesn't matter, since it is temperature difference.

**Q3:** Did you take into account the change in resistivity below the melting point of Aluminum?

**A3:** We compared the measured resistivity (for temperature up to 300 $^0$C) with values reported in the literature which documents resistivity up to the melting point. The resistivity rise with temperature was found to be quite linear with very little change in its slope. So we did not consider it to be important.

**Q4:** Your analysis used the current and voltage to determine the resistance change and the temperature rise and that is an average effect. But in figure 8, where you showed the fusion spots, they are local effects. Is there some way that you might be able to modify your analysis because you probably have problems occurring even before you think you do because you may have local transients which are higher than the average, in fact you won't have an average unless you had higher and lower values?

**A4:** It is true that there may be localized hot spots due to defects, impurities etc., where the temperature may exceed the average temperature. However, since the temperature at these spots will be higher, the average temperature is a good conservative indicator of this maximum allowable temperature.

**Q5:** I was wondering if you have compared your simulator results with some of the things that have appeared in the literature like some computer simulations done over 20 years ago on metallization burnout or there is an analytical model in the 1982 proceedings of the ESD/EOS symposium. Have you tried to see how your simulator agrees with that work?

**A5:** Simulation programs have been used in the past and we are aware of that. The most important thing in all simulations is their calibration with experimental data taken on structures "similar" to the ones being modeled (since technology difference will affect thermal parameters). We are trying to develop a physical understanding of high current effects and have used a finite element simulator, after calibrating it with DC data taken on similar samples, only to check the temperature rise obtained during a short pulse using the resistance of lines as the thermometer.