ELECTROTHERMAL ANALYSIS OF VLSI INTERCONNECTS

A DISSERTATION

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL ENGINEERING

AND THE COMMITTEE ON GRADUATE STUDIES

OF STANFORD UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Ting-Yen Chiang

June 2004

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy

_____

Professor Krishna C. Saraswat, Principal Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy

_____

Professor Robert W. Dutton

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy

_____

Dr. James P. McVittie

Approved for the University Committee on Graduate Studies:

_____

Abstract

The scaling of VLSI structures leads to continuous increase in current density that results in ever greater interconnect Joule heating. In addition, a variety of low-k materials have been introduced to reduce the RC delay, dynamic power consumption and crosstalk noise in advanced technology. Together with the poor thermal conductivity of such materials and more metal levels added, the increasing thermal impedance further exacerbates temperature rise in interconnects. As a result, not only will thermal effects be a major reliability concern, but also the increase of resistivity with temperature can degrade the expected speed performance. On the other hand, overly pessimistic estimation of the interconnect temperature will lead to overly conservative approach. Hence, performing a more realistic thermal modeling and analysis of interconnects is critical.

This research proposes both compact analytical models and fast SPICE based 3-D electro-thermal simulation methodology to characterize thermal effects due to Joule heating in high performance Cu/low-k interconnects under steady-state and transient stress conditions. The results demonstrate excellent agreement with experimental data and those using Finite Element (FE) thermal simulations (ANSYS). The effect of vias, as efficient heat transfer paths to alleviate the temperature rise in the metal wires, is included in our analysis for the first time to provide more accurate and realistic thermal diagnosis. It shows that the effectiveness of vias in reducing the temperature rise in interconnects is highly dependent on the via separation and the dielectric materials used. The simulation

methodology has also been applied to quantify the use of dummy thermal vias as additional heat sinking paths and possible solution to hot wires.

The impact of Joule heating on the scaling trends of advanced VLSI interconnects has been evaluated in detail. It shows the interconnect Joule heating can strongly affect the maximum operating temperature of the global wires which will, in turn, constrain the scaling of current density to mitigate electromigration and, thus, greatly degrade the expected speed improvement from the use of low-k dielectrics. Coupled analysis of reliability and delay, under the influence of thermal effects, is performed to optimize interconnect structures such as wire aspect ratio and ILD thickness. Finally, potential bottlenecks and opportunities of future heterogeneous three dimensional (3-D) ICs with various integration scenarios are identified from the thermal point of view. It is shown that under certain scenarios, 3-D ICs can actually lead to better thermal performance than planar (2-D) ICs. Tradeoffs among power, performance, chip real estate and thermal impact for 3-D ICs are evaluated.

Acknowledgements

I am deeply indebted to many people who made my Stanford years an enlightening, rewarding and memorable experience. It is a great pleasure to acknowledge the people who have supported me throughout these years.

First and foremost, I would like to thank my adviser Professor Krishna C. Saraswat for his guidance, insight and patience. I am forever grateful his unwavering confidence in me throughout my time at Stanford, both in my research and in my start-up dream. I feel very fortune to have worked with Professor Saraswat, who provides a creative academic atmosphere for the students and always shows his unconditional support to the students in all aspects. Being one of the most friendly, down-to-earth professors, Professor Saraswat is not only an admirable educator, but also a life-time mentor. The experience I have gained under his guidance is invaluable and has left an indelible mark on my self and persona.

I would like to thank members of my Oral and Reading Committees, Professor Robert W. Dutton, Professor Piero A. Pianetta, and Dr. Jim McVittie for their enthusiastic discussion and valuable feedback. I also thank Professor Michael A. Kelly for his continuous help and encouragement. My gratitude also goes out to Professor Kenneth E. Goodson for his time and effort acquainting me with thermal analysis and, to Dr. Pawan Kapur for numerous entertaining discussions and invaluable brain storming sessions, and to Dr. Shukri Souri, whose sense of out-of-box thinking and professionalism I thoroughly enjoyed, and who made possible the thermal analysis of 3-D ICs through his 3-D ICs performance study. I also thank Dr. Kaustav Banerjee and Sungjun Im whose generosity

of discussions and challenges elevated my intellectual awareness of my own work. In addition, I would like to thank MARCO Interconnect Focus Center for funding my Ph.D. work.

My colleagues from various research groups have been terrific and my work has benefited tremendously from interactions with the past and present members. I am very thankful for the assistance and friendship of Dr. Shahram Alibeik, Dr. David Bang, Dr. Richard Chang, Ray Chen, Tsu-Ju Chin, Hoon Cho, Hoyeol Cho, Chi-On Chui, Dr. Mar Hershenson, Yu-Li Hsueh, Dr. Nabeel Ibrahim, Dr. Amol Joshi, Dr. Mayur Joshi, Dr, Steve Jurichich, Dr. Pranav Kalavade, Dr. Kambiz Kaviani, Rohan D. Kekatpure, Donghyun Kim, Tejas Krishnamohan, Dr. Marci Liao, Dr. Alvin Loke, Yi-Chang Lu, Ammar Nayfeh, Ali Kemal Okyay, Abhijit Jayant Pethe, Dr. Xiao Ning Qi, Pinkesh Sachdev, Rohit Shenoy, Dr. Ben Shieh and Junfeng Xu.

Although it was short, the wonderful experience of writing business plan, presenting our ideas in front of venture capitalists and winning the $10,000 prize in the *Stanford Entrepreneur's challenge business plan competition* was thrilling. I would like to thank Ashish Chopra and Alper Tekin for all the dreams of riches and early retirement and for our TAAV team which aimed to be the leader in the physical analysis segment of electronic design automation with its breakthrough full chip thermal simulation technology.

I thank Irene Sweeney for great administrative support. I would also like to thank Dr. Richard Dasher for coordinating and coming along for a SPIE trip to Philips, Netherlands. I especially thank Professor Tom Lee, Carmen Miraflor and Maureen Rochford for working hard to get me back to U.S. when my green card got stolen in Amsterdam, Netherlands, during the SPIE trip.

Without question, none of this would have been possible without the support, moral and otherwise, of family and loved ones. To Aunt Chuling and her family, for all your support in a time of need, my deepest appreciation. To my brother and sister, Ting-Chun and Sher-Feng, I am forever grateful. For your friendship and taking care of our parents when I am far away from home. To my grandparents, I thank you for your endless love and encouragements. Most of all, my eternal gratitude goes to my mother and father, Min-Chuan and Ho-Chen, whose constant inspiration and unreserved support have made it all possible. This dissertation, its accomplishment through all the obstacles, is a manifestation of their dedication, sacrifices, commitment and hard work towards each one of us, their children, for all their years. To them, I am eternally indebted and dedicate this work.

Finally, for my beloved wife, Ching-Jung, the greatest blessing in my heart, for being always there for me. Without all her unparalleled love and support through the years, I would not go back to pursue Ph.D. degree and this thesis would never have been finished. Thanks to my precious son, Eric, for always greeting me with a smiling, loving face to go home to after a tough day. There will be another book. Ching-Jung and Eric, are the starting actors, and I am grateful to be part of the cast.

Contents

List of Tables

List of Figures

`

Chapter 1

Introduction

1.1 Background

The unprecedented prosperity of information technology has driven semiconductor industry to evolve at an incredible rate. The growing demand for higher performance in integrated circuits (ICs) results in faster device switching speed, greater number of transistors, increased functional density and larger chip size [1, 2]. Consequently, the communication, supported by on-chip interconnects, between devices and between circuit blocks is becoming more complex and a challenging problem. The connection of miniaturized and closely packed transistors requires reduced wire cross-sections in the local levels, while the rapid increase in functional density and chip size leads to longer-distance communication in the global levels. With the aggressive scaling of VLSI technology, the interconnect delay, due to longer wires and smaller wire pitch, now plays a key role and becomes bottleneck in the con-

tinued improvements in ICs density and speed [3, 4]. As illustrated by the International Technology Roadmap for Semiconductors (ITRS) [5], interconnect RC delay is dominating the chip performance in advanced technology nodes (Fig. 1.1).

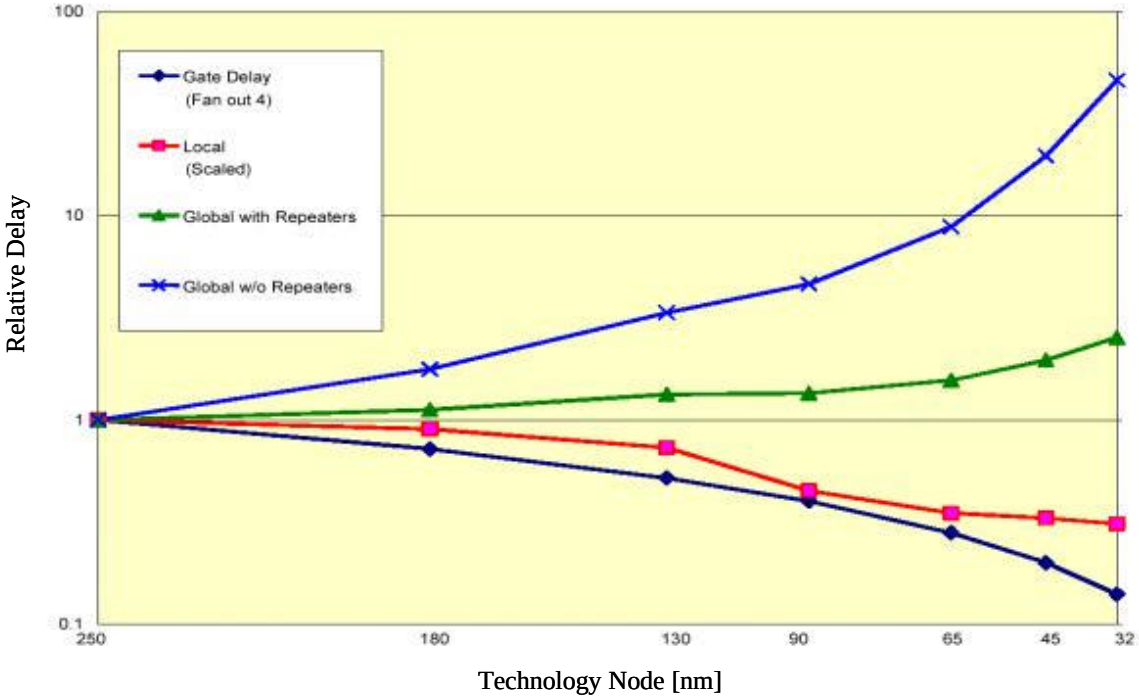


Fig. 1.1: Predictions for device and wire delays from ITRS 2001 [5].

To mitigate this adverse-scaling trend of signal propagation (RC) delay, considerable work has gone into overcoming the interconnect limitations. For instance, hierarchical interconnect structure has been adopted to incorporate several metal levels, so that long global interconnects can be routed to higher tier and maintain larger cross-sections to minimize wire resistance [6-8]. This interconnect architecture along with the

demand for more wiring has resulted in an increasing number of interconnect metal levels.

Year	1972	1982	1993	2002
Processor	8008	286	Pentium®	Itanium 2®
Technology Node	10 μm	1.5 μm	0.8 μm	0.18 μm
Frequency (MHz)	0.2	6	60	1000
Transistors	3500	120,000	3,100,000	221,000,000
Chip Area (mm^2)	15.2	68.7	217	421
Metal Layers	1	2	3	6

Table 1.1: Evolution of characteristics for Intel® commercial processors

(Table 1.1). In addition, Cu has replaced Al as the interconnect metal for its lower resistivity [9, 10] and low dielectric constant (low-k) materials has been pursued as an alternative to silicon dioxide to reduce interconnect capacitance [11, 12]. However, although all these tremendous efforts have salvaged the attention-getting interconnect delay problem, the interconnect reliability due to electromigration and thermal effects is quickly emerging as a serious integration issue. In addition, the fast increasing power density in ICs (Fig. 1.2) [13], together with the scaling trend toward lower power supply (Vdd) to reduce short channel effect, hot electrons, gate stress and leakage [14, 15], has significantly increased the current density in interconnects. Furthermore, interconnects are farther away from the substrate, where heat sink is usually attached. As a result, interconnects generally experience higher average temperature than transistors do.

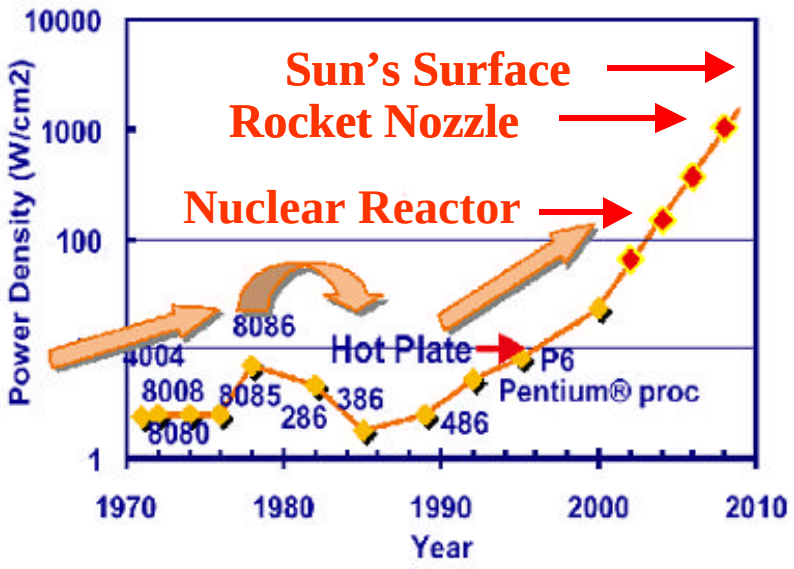


Fig. 1.2: Evolution of power density for Intel processor [13] (Courtesy of Intel).

It is not difficult to imagine that not only will interconnect delay be a major show-stopper for continuous ICs performance advancement, but also the concern of reliability and thermal effect of interconnect can potentially become another serious system design constraint. In short, as semiconductor technology strives to keep up with Moore's Law [1], it is imperative to understand and analyze the emerging VLSI "hot wire" phenomena to handle the impact of interconnect thermal effect in early design phase.

1.2 Thermal Implication of Interconnect Scaling

Thermal effects are an inseparable aspect of electrical power distribution and signal transmission through interconnect nets due to Joule heating (or self-heating) caused by

current flow in the wire. The temperature rise, on top of the substrate temperature, in the interconnects is determined by the product of Joule heating power dissipation and the thermal impedance from the wires to the substrate. Before proceeding to further discussion, it is instructive to be familiar with basic on-chip interconnect structure and terminology. The interconnects (or wires) in a modern IC chip can be broadly characterized into three groups according to the functions they perform. These are the signaling interconnects, the clock distribution interconnects and the power and ground supply distribution interconnects. Interconnects can be further subdivided into three categories according to the range of their lengths and their cross section dimension. These are the local, semiglobal/intermediate and the global interconnects. The global interconnects are responsible for long distance communication on a chip and have a larger cross sectional area to minimize resistance. Whereas, the local interconnects have the shortest range and the tightest cross sectional dimensions. Modern ICs have multiple levels of interconnects to accommodate their large numbers, starting from the local at the lowest level to the global at the top most levels. This is depicted in Fig. 1.3 [5]. The lower level interconnects are used in local routing and they connect transistors that are a few microns apart. The higher level interconnects are used for global routing and they can span across an entire chip and may be as long as 2-3 cm in modern high-performance chips. Vias are the metal fillings enabling inter-level wire connections. The wires are separated by inter-layer dielectrics (ILD) from level to level and isolated by inter-metal dielectrics (IMD) within the same level. Traditionally, silicon dioxide (SiO_2) has been

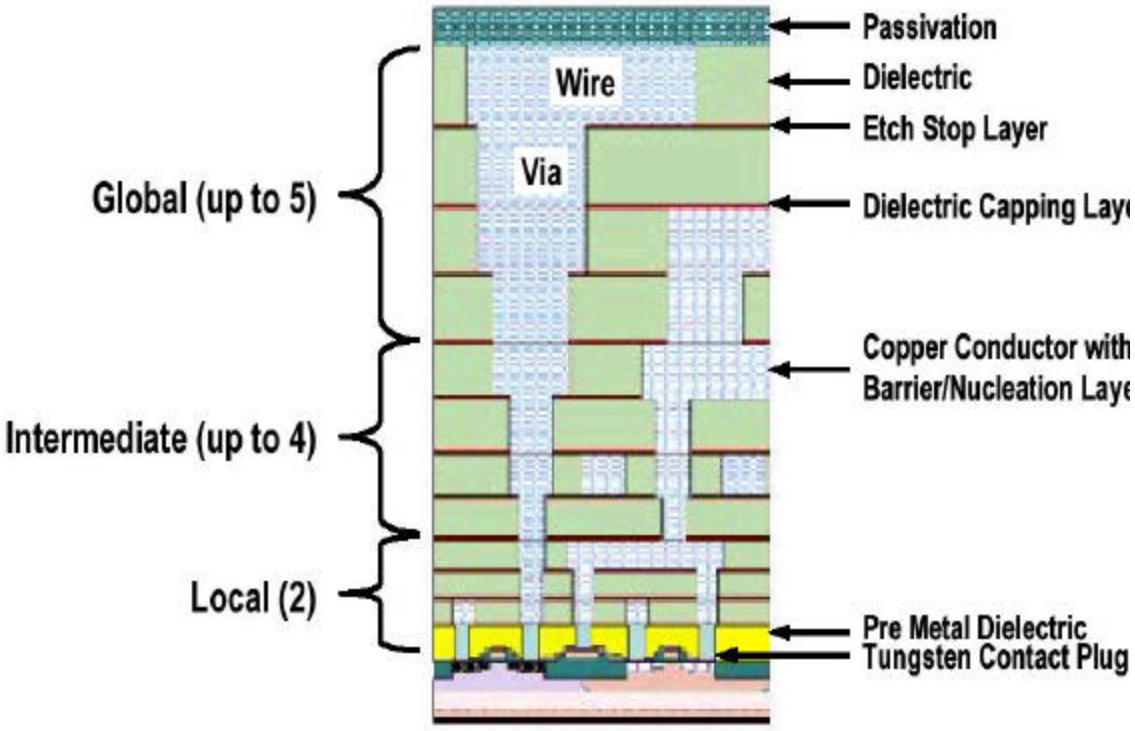


Fig. 1.3: Schematic showing the hierarchy of metal levels for distribution of interconnects in modern ICs. Figure taken from ITRS 2001 [5].

the choice of the dielectrics in ICs for its superior properties. However, a variety of low dielectric constant materials, including organic and inorganic films, as well as the use of porosity and air-gaps [16-18], have been fiercely pursued since a decade ago for the benefit of reducing wire capacitance, cross-talk noise and dynamic power consumption. Unfortunately, these low-k dielectrics inevitably have very poor thermal properties (Fig. 1.4 [5]). Compounded with the poor thermal conductivity of such materials and more metal levels added with more advanced technology nodes, the increasing thermal impedance further exacerbates temperature rise in interconnects.

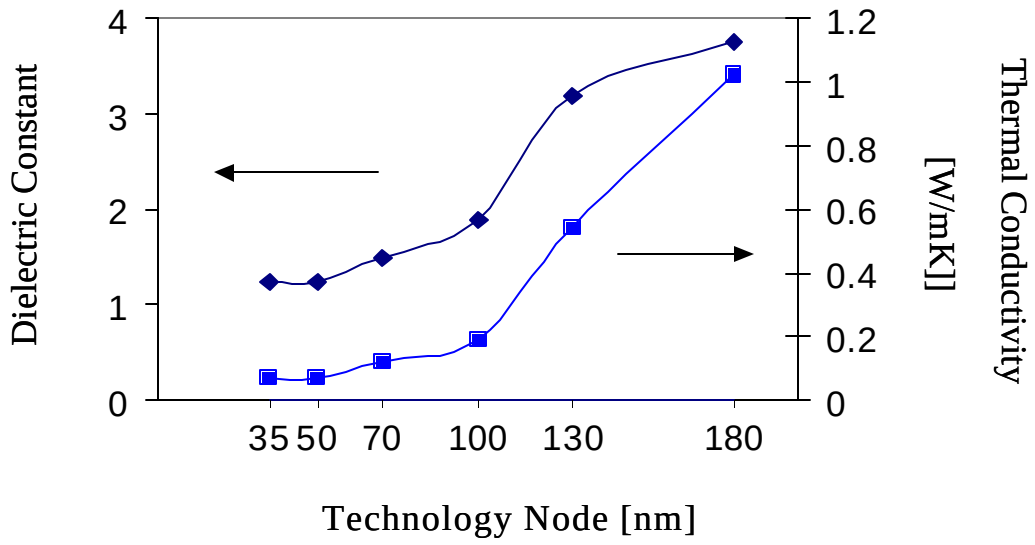


Fig. 1.4: Both dielectric constant and thermal conductivity of ILD materials decrease with advanced technology nodes [5].

On the other hand, the power consumption in the chips is rapidly becoming unmanageable as clock frequency continues to climb and chip temperature has been always a major concern to further chip advancement. The reduction in supply voltage slows down the power consumption at the transistor level; however, the resultant increased current requirement leads to greater interconnect Joule heating. The typical supply current is in the 100A range in modern chips and is expected to increase for future processors. Although the power dissipated by Joule heating is not a significant portion of the total chip power consumption yet, the interconnects are separated from the substrate by dielectrics with low thermal conductivity, and heat cannot be removed efficiently Therefore the temperature rise in interconnects can be non-negligible. The interconnect

structures with decreasing critical dimension will inevitably experience ever-higher current density and, the resultant greater temperature rise.

1.3 Via effect

Since the lifetime of interconnects has an exponential dependence on the inverse metal temperature owing to electromigration [19, 20], thermal effect will limit the maximum allowable current density in the wires to limit the temperature rise in interconnects. It can be envisioned that, to some extent, the thermal effect may not only be a major reliability concern, but also constrain the speed performance due to limited current drive capability. With the aggressive scaling trend, it is essential not to under-estimate the impact of interconnect temperature rise. On the contrary, overly pessimistic estimation of the interconnect temperature will lead to overly conservative design. Therefore, accurate temperature estimation is extremely important to regulate the design rules and, hence, performing a more realistic thermal analysis of interconnects is critical.

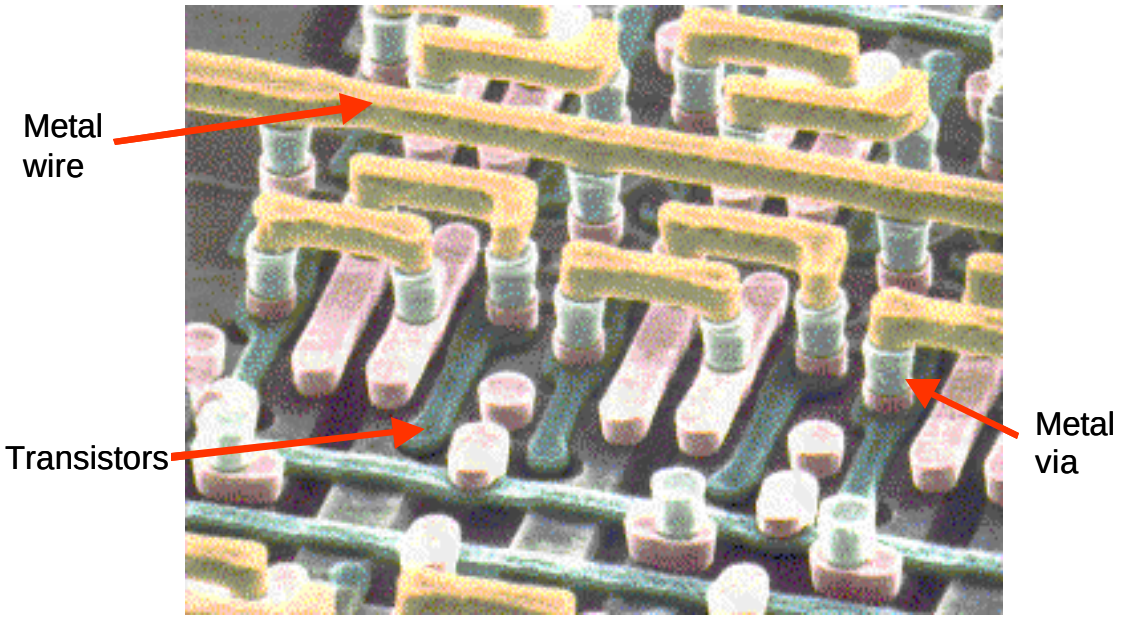


Fig. 1.5: SEM photo of local level interconnect

Several publications have addressed the issue of low-k dielectrics and their impact on interconnect temperature, reliability and performance [21, 22]. Poor thermal conductivity of low-k insulators has been a major concern to cause substantial rise in interconnect temperature. However, the effect of vias (Fig. 1.5), which have much higher thermal conductivity than the dielectrics and therefore can serve as efficient heat dissipation paths, has not been adequately addressed in those simplified thermal model [23, 24]. Consequently, if via effect is ignored, the predicted temperature rise of interconnects can be much higher than the reality. Neglecting via effect on interconnect temperature estimation can lead to intolerable prediction errors. This research proposes both compact analytical models and fast SPICE based 3-D electro-thermal simulation methodology to

characterize thermal effects due to Joule heating in high performance Cu/low-k interconnects under steady-state and transient stress conditions. The results demonstrate excellent agreement with experimental data and those using Finite Element (FE) thermal simulations (ANSYS). The effect of vias, as efficient heat transfer paths to alleviate the temperature rise in the metal wires, is included in our analysis for the first time to provide more accurate and realistic thermal diagnosis.

1.4 Dissertation Organization

With the pursuit of realistic evaluation in mind, via effect is included in all the thermal models and simulations. The impact of Joule heating on the scaling trends of advanced VLSI interconnects has been evaluated in detail.

Chapter 2 derives the analytical interconnect thermal models to provide building blocks, which will enable us to quantify the impact of via effect and various interconnect parameters on the interconnect temperature rise. The model is then applied to estimate the temperature rise of densely packed multi-level interconnects.

Chapter 3 discusses the impact of Joule heating on the scaling trends of advanced VLSI interconnects. Through a combination of extensive electrothermal simulation and 2D field solver for capacitance calculation, the thermal characteristics of various Cu/low-k schemes are quantified and their effects on electromigration reliability and interconnect delay is determined.

Chapter 4 describes a compact 3D electro-thermal simulation methodology to evaluate interconnect design options from a thermal point of view. Use of dummy thermal vias as additional heat sinking paths is evaluated to alleviate the thermal impact on interconnect reliability. Furthermore, this simulation methodology provides an efficient diagnosis tool for transient thermal stress analysis, which is usually extremely cumbersome, mostly unsolvable, in analytical forms.

Chapter 5 presents detailed thermal analysis of high performance three dimensional (3-D) ICs under various integration schemes. A complete thermal model including power consumption due to both transistors and interconnect joule heating from multiple strata is presented. Furthermore, tradeoffs among power, performance, chip real estate and thermal impact for 3-D ICs are evaluated.

Chapter 6 summarizes the main findings and concludes this dissertation with suggestion on possible future work.

Chapter 1: **Introduction**

Chapter 2

Analytical Interconnect Thermal Model

2.1 Motivation

The conventional aluminum(Al)/silicon dioxide (SiO_2) interconnect scheme reached its performance limitation a decade ago [3]. Hence, lower wire resistance and higher current density requirements have prompted an industry-wide shift from aluminum(Al) to copper(Cu). Similarly, the minimization of interconnect capacitance is driving the introduction of a variety of low-permittivity dielectric materials. Since these low-k dielectrics also have lower thermal conductivity than silicon dioxide, heat dissipation in the interconnect level has become more difficult. Furthermore, VLSI scaling trends, such as increasing number of interconnect metal levels and increasing current density, have caused ever-profound thermal effects in the interconnects. In a state-of –art chip, the interconnect temperature can rise more than 100 degrees above the ambient

temperature due to Joule heating and the heat dissipation from transistors in the substrate level.

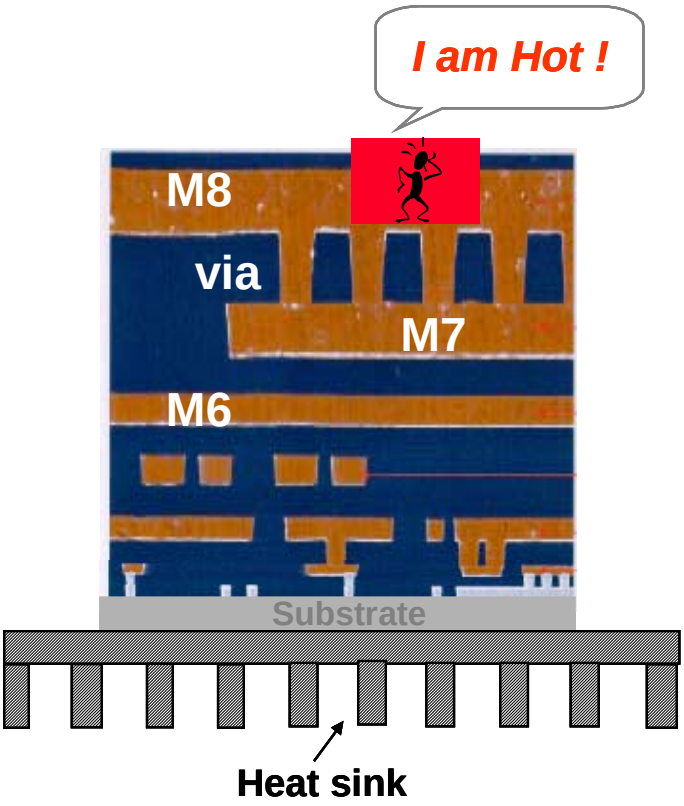


Fig. 2.1: Typical multilevel interconnect structure with heat sink attached to substrate.

Although knowing the interconnect temperature is the first step toward any kind of further thermal effect analysis, the temperature profile of a single interconnect is difficult enough to be obtained experimentally for its miniature scale [25], not to mention multi-level measurements. In general, interconnect thermometry based on temperature-dependent electrical resistivity of the wire is used to calculate a spatially averaged

temperature rise of the interconnects [26]. However, it can provide neither the local temperature rise profile on the wire nor the temperature distribution along the multilevel interconnects. On the other hand, various papers have been devoted to deriving analytical expressions for the temperature distribution in IC chips [27-29]. However, most of them only consider a single heat source without considering the effect of vias, which is not the realistic case for most interconnects. In addition, the complicated nature of these expressions makes it very difficult to easily apply them to multi-level metal layers. Recently, Im [24] and Hunter [30, 31] have proposed simplified formule, while Chen [23] and Rzepka [32] applied 3-D finite element analysis to estimate temperature rise on multi-level interconnects. However, the effect of vias, as an efficient path for conduction of heat has not been included in those simplified thermal model and simulation. Consequently, the predicted temperature rise of interconnects can be much higher than actually observed.

Analytical thermal models are desirable to facilitate quick estimation of temperature rise in order to provide better insight and thermal design guidelines for advanced interconnect structures. This chapter presents compact analytical thermal models for estimating the temperature rise of multilevel VLSI interconnect lines incorporating the effect of vias as efficient paths for conduction of heat. The impact of vias has been modeled using i) a characteristic thermal length representing the distance along a metal line through which vias effectively remove the heat, and ii) an effective thermal conductivity of inter-layer dielectrics (ILD), $k_{ILD,eff}$, with $k_{ILD,eff} = k_{ILD}/\eta$, where k_{ILD} is the

nominal thermal conductivity of the ILD and η is a physical correction factor representing the heat removal by vias.. Both the spatial temperature profile along the metal lines and their average temperature rise can be easily obtained using these models. The predicted temperature profiles are shown to be in excellent agreement with the 3-D finite element thermal simulation results. The model is then applied to estimate the temperature rise of densely packed multi-level interconnects. It is shown that in multi-level interconnect arrays, via density along the lines can significantly affect the temperature rise of such interconnect structures.. The influence of via self heating and thermal impedance of vias is considered here, providing more comprehensive and accurate results.

The remainder of this chapter is organized as follows. To provide a baseline, in Section 2.2, we validate that it is a legitimate approach to consider via effect. Sections 2.3 and 2.4 describe the derivation of our analytical model and discuss the impact of via effect. In Section 2.5, we apply the model to multilevel interconnects and discuss the temperature distribution along metal layers. The main results are summarized in Section 2.6.

2.2 Preliminary Observations

The fact that the interconnect temperature does not increase inversely proportional to the nominal thermal conductivity of dielectrics suggests that the vias, which have

much higher thermal conductivity than ILD, serve as efficient heat dissipation paths. If the via effect is not properly considered, the predicted temperature will be significantly higher than the temperature in practical situations. As a result, there is danger of significant errors in using overly simplistic assumptions to assess thermal problems in advanced interconnect structures with low-k materials. In this work, the via effect is evaluated in details and is incorporated into our multilevel temperature distribution model. Furthermore, due to the interconnect architecture that, in general, lower metal levels have much higher via density than the higher levels, the temperature rise among metal layers is expected to be quite different from what is predicted by the simplistic model.

2.2.1 First glance at via effect

Since the cross sectional area of vias is generally smaller than that of the wires, it may raise a concern that the resultant higher current density in the vias may generate significant heat, deteriorating their effectiveness in heat conduction. However, as shown in [33], the temperature rise in the via is not as high as that in the wire. A simulation validation was obtained by using ANSYS, a three dimensional finite element thermal simulation package. The configuration simulated, as shown in Fig. 2.2, is Cu wire/via embedded in a low-k dielectric with via separation, L, varied as 250 μm, 100 μm and 5 μm. The metal width is taken to be 0.3 μm and both of the metal height and via height are 0.6 μm. Polymer is chosen as the low-k insulator with nominal thermal conductivity of 0.3 W/m-K. The electrical resistivity and thermal conductivity of Cu are assumed to be 2.2 μΩ-cm and 400 W/m-K. The RMS current density of 2.1e6 A/cm^2 is flowing in

the wire. A uniform temperature, T_0 , is assigned to the underlying layer of the block. The temperature rise of wires on top of substrate temperature is represented by $\theta_m(x) \equiv T_m(x)-T_0$, where $T_m(x)$ is the wire temperature. Adiabatic boundary conditions are applied to the top and four sides of the block where symmetry boundary conditions are satisfied.

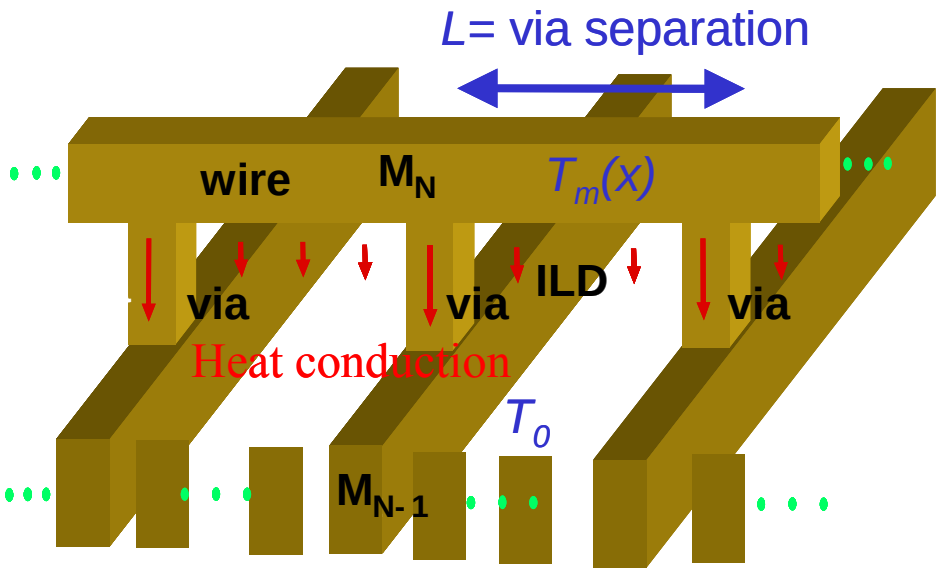


Fig. 2.2: Heat generated by *Joule–heating* is conducted downwards through both ILD and vias. L represents the separation between vias.

Two extreme simulation conditions have been carried out in Fig. 2.3, with no current flowing through the via, condition (a); and with all the current flowing through via, condition (b). Therefore, only the thermal impedance caused by vias, $R_{th,\,v,}$ is included in the analysis in condition (a), while the extra heat generated in the via due to current

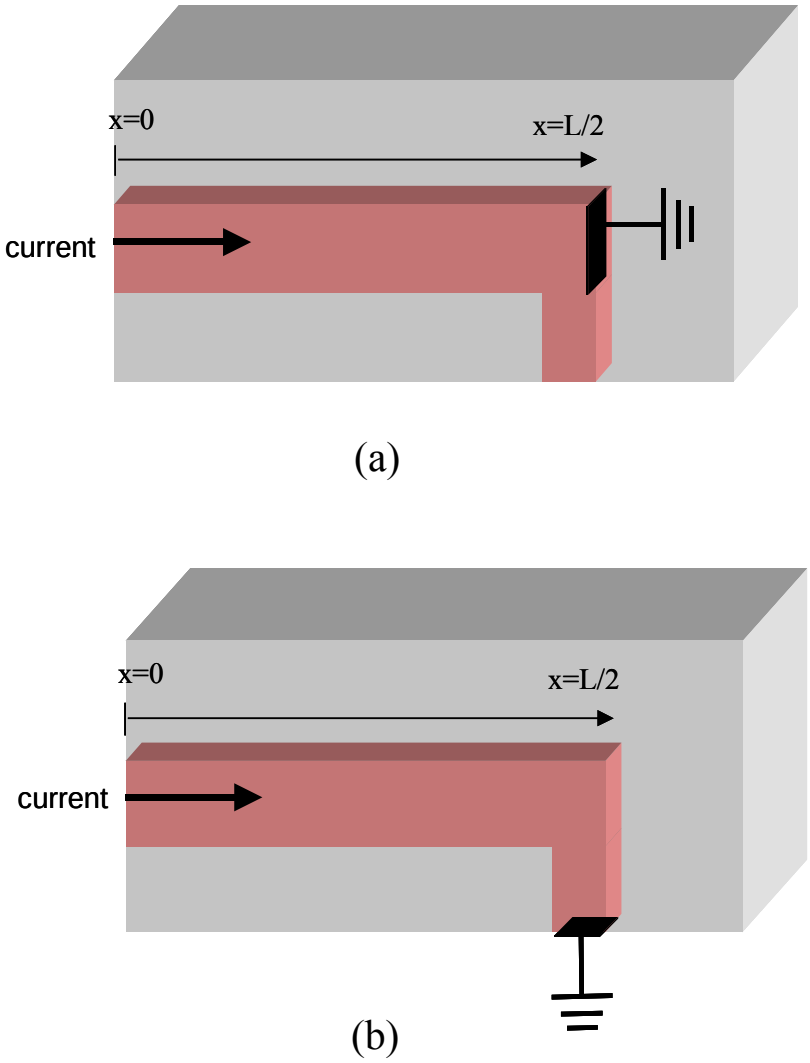


Fig. 2.3: Two extreme simulation conditions: (a) no current flows through via, so there is no heat generation in the via; (b) all the current flowing through the wire will also flow through the via and cause heat generation. It should be noted that, in both cases, the via helps transfer heat to underlying layer and reduce the temperature rise on the wire.

density flowing in the via, J_v, is also taken into account in condition (b). In practice, all the vias do not carry the same amount of current as in the wire.

The comparison of temperature distributions along the metal wires are shown in Fig. 2.4 (a), (b) and (c) for via separation of 250 μm, 100 μm and 5 μm, respectively. The temperature profiles of the simplified 1-D thermal model, constant temperature value, with via effect completely ignored are also plotted in Fig. 2.4 (a) and (b). It has the same value as well for the case of 5 μm via separation, although it is not plotted in Fig. 2.4 (c) to keep other curves more readable. It should be noted that, with shorter via separation, the average temperature as well as the peak temperature in the wire is reduced. It appears that within a certain range from via, the temperature reduction with the help of vias is particularly consumptious. In addition, the simulations show that the temperature rise in the vias is minimal. Even though percentage wise, the via contact exhibits a major contribution to the temperature rise in the wire for short via separation, the absolute value is quite small. The vias are simply too short to produce much heat. Furthermore, under normal operating conditions, all the vias are not conducting current at the same time but they always help dissipate heat. Consequently, it is legitimate to consider vias as efficient thermal sink paths. Finally, as long as the vias are connected to lower-temperature underlying layers, they will help dissipate heat no matter how far they are away from substrate. With the above preliminary observations in mind, in the following sections, we will quantify via effect and evaluate if vias are always helpful.

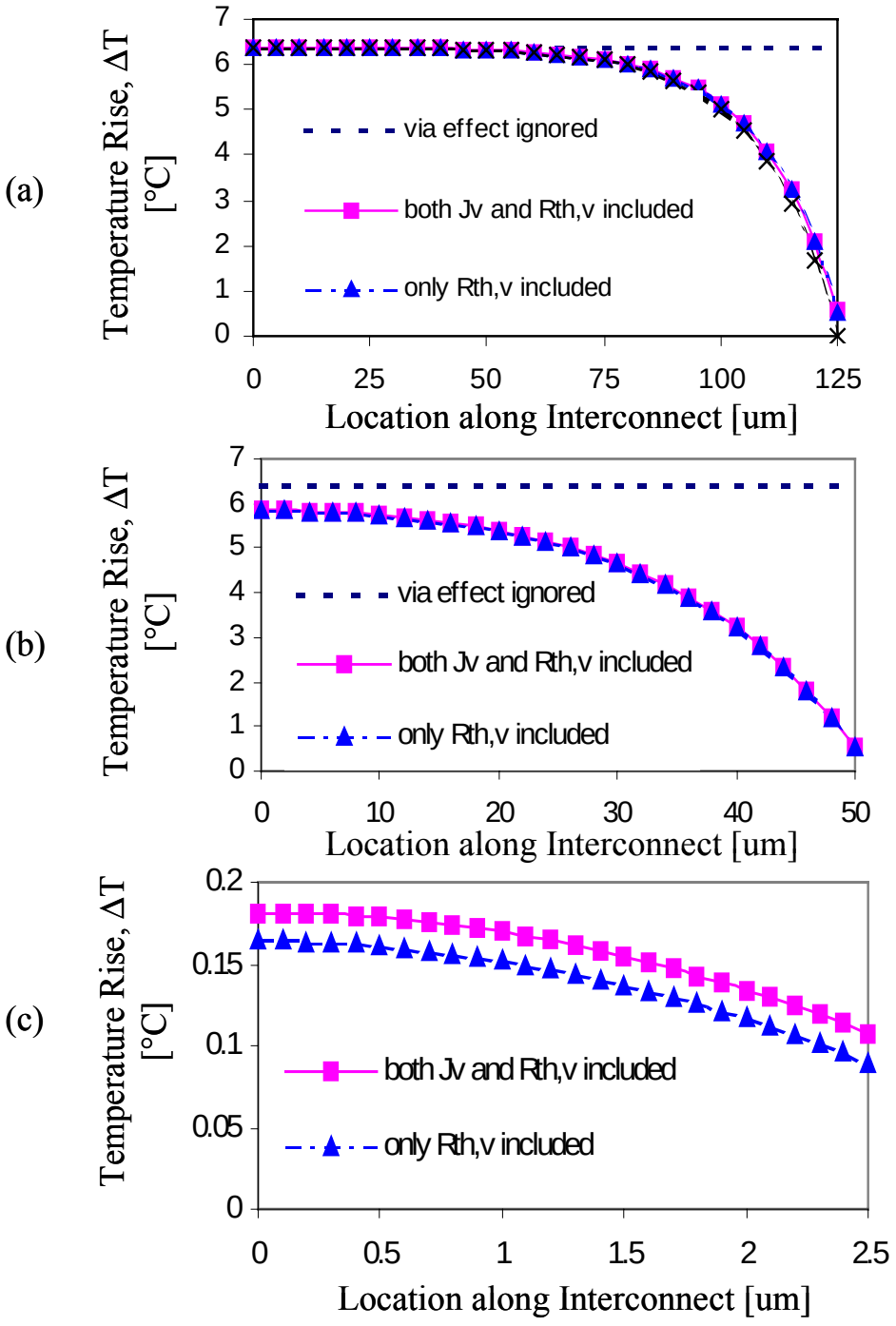


Fig. 2.4: Temperature distribution along the metal wire with via separation of 250 µm (top), 100 µm (middle) and 5 µm (bottom) for case (a) and (b) of Fig. 2.3. It can be observed that there is no significant difference between the two simulation conditions and temperature rise in the via is minimum.

2.3 Interconnect Thermal Model and Assumptions

2.3.1 Energy conservation law

In either steady state or transient condition, the first principle of *energy conservation law* must be satisfied at any instance,

$$\dot{E}_{in} + \dot{E}_{gen} - \dot{E}_{out} = \frac{dE_{st}}{dt} \tag{2.1}$$

where the thermal energy entering, leaving, and generated in the control volume are $\dot{E}_{in}$, $\dot{E}_{out}$, and $\dot{E}_{gen}$, respectively. The rate of change of energy stored within the control volume is designated as dE_{st}/dt. Consider a rectangular metal wire with cross-sectional area A_c, and periphery p embedded in a dielectric medium as shown in Fig. 2.5. The temperature variation can be modeled as a one-dimensional situation which refers to the fact that only one coordinate (x) is needed to describe the spatial temperature variation. The temperature gradient along y and z directions is negligible for a long metal wire with relatively small cross section and high thermal conductivity. Garden has proven that the error due to this simplification is less than 1% even for larger cross sections [34]. The conduction heat rate perpendicular to the control surface at x is indicated by the term q_x.

$$q_x = k_m \frac{\partial T}{\partial x} A_c \tag{2.2}$$

where k_m is the thermal conductivity of the metal wire and A_c is the cross-sectional area of the wire. The conduction heat rate at the opposite surface can then be expressed as a Taylor series expansion where, neglecting higher order terms,

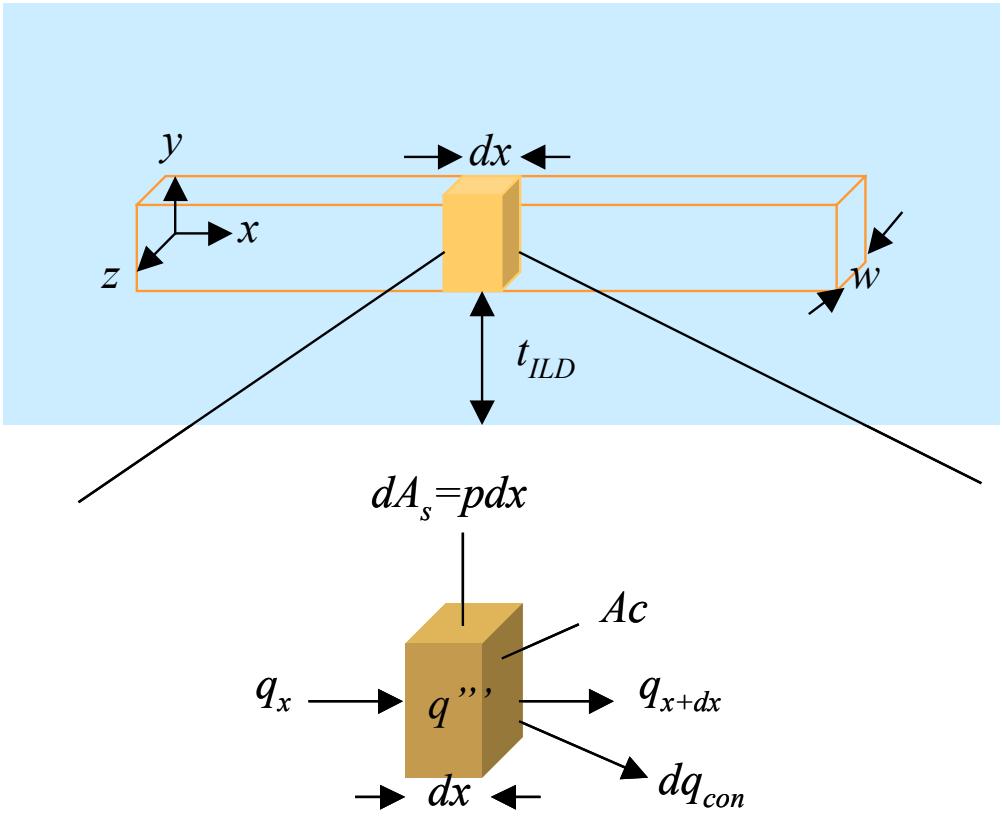


Fig. 2.5: A differential control volume, from a conducting metal wire embedded in a dielectric medium, shows energy balance.

$$q_{x+dx} = q_x + \frac{\partial q_x}{\partial x} dx \tag{2.3}$$

dq_{con} represents the heat conducted to the surrounding medium. It is proportional to the difference between the surface and the surrounding temperatures and the contiguous surface area.

$$dq_{con} = h(S'_m \, dA_s)(T_m(x) - T_\infty) \tag{2.4}$$

where the proportionality constant, h, is termed as *heat transfer coefficient.* dA_s is the surface area where heat in the control volume can be transferred into the medium. S'_m is the spreading factor to model the multi-dimensional heat spreading phenomena. $T_m(x)$ is the metal temperature at x and T_∞ is the temperature of the medium far from the metal surface.

As required by *energy conservation law* Eq. 2.1, the heat entering the differential control volume (q_x-q_{x+dx}) plus the volumetric rate of thermal energy generation in the control volume (q''') minus the heat dissipated to the medium (dq_{con}) must be equal to the rate of change of thermal energy stored within the differential control volume. Therefore, the governing heat diffusion equation with temperature variation dominating in x direction can be concluded as

$$\frac{\partial}{\partial x}[k_m A_c \frac{\partial T}{\partial x}] - hS'_m\, p(T - T_\infty) + q''' A_c = \rho\, c_p A_c \frac{\partial T}{\partial t} \qquad (2.5)$$

where ρ is the metal density and c_p is the metal specific heat. The product, $\rho\, c_p$ [J/m^{3}K], commonly termed the *volumetric heat capacity*, measures the ability of a material to store thermal energy.

2.3.2 Conduction dominant heat transfer

Since, in general, thermally insulated package materials encapsulate the IC chip, heat convection to the ambient air is ignored by the application of the adiabatic boundary condition on the four side walls and top of the chip. In addition, heat radiation is simply too small to be taken into account. Therefore, only heat conduction will be considered as

the prevailing heat transfer mechanism. Therefore, the net heat transfer processes can be quantified by the rate equation known as *Fourier's law*,

$$q'' = -k\nabla T \tag{2.6}$$

where q'' [W/m^2] is the heat flux and k [W/m-K] is the thermal conductivity. Furthermore, the substrate, to which a heat sink is usually attached, is assumed to be the sole heat dissipation path to the outside ambient, hence only heat conduction toward the substrate will be considered in our interconnect thermal modeling. This assumption is validated in Fig. 2.6, as Zhou has shown that, with heat sink attached, the majority of heat will transfer through the substrate [35].

Eq. 2.4 can then be further simplified if we consider heat is only transferred downwards and only heat conduction mechanism takes place, i.e.

$$dq_{con} = (\frac{k_d}{t_{ILD}})(S'_m\, w dx)(T_m(x) - T_0) \tag{2.7}$$

where k_d is the thermal conductivity of the dielectric medium and t_{ILD} is the underlying dielectric thickness. k_d / t_{ILD} acting as the proportionality constant and can be treated as an equivalent *conduction heat transfer coefficient*. w is the width of the metal wire and T_0 is the temperature of underlying layer. Accordingly, Eq. 2.5 results in

$$\frac{\partial}{\partial x}[k_m A_c \frac{\partial T}{\partial x}] - (\frac{k_D}{t_{ILD}})(S'_m\, w)(T - T_\infty) + q''' A_c = \rho\, c_p A_c \frac{\partial T}{\partial t} \tag{2.8}$$

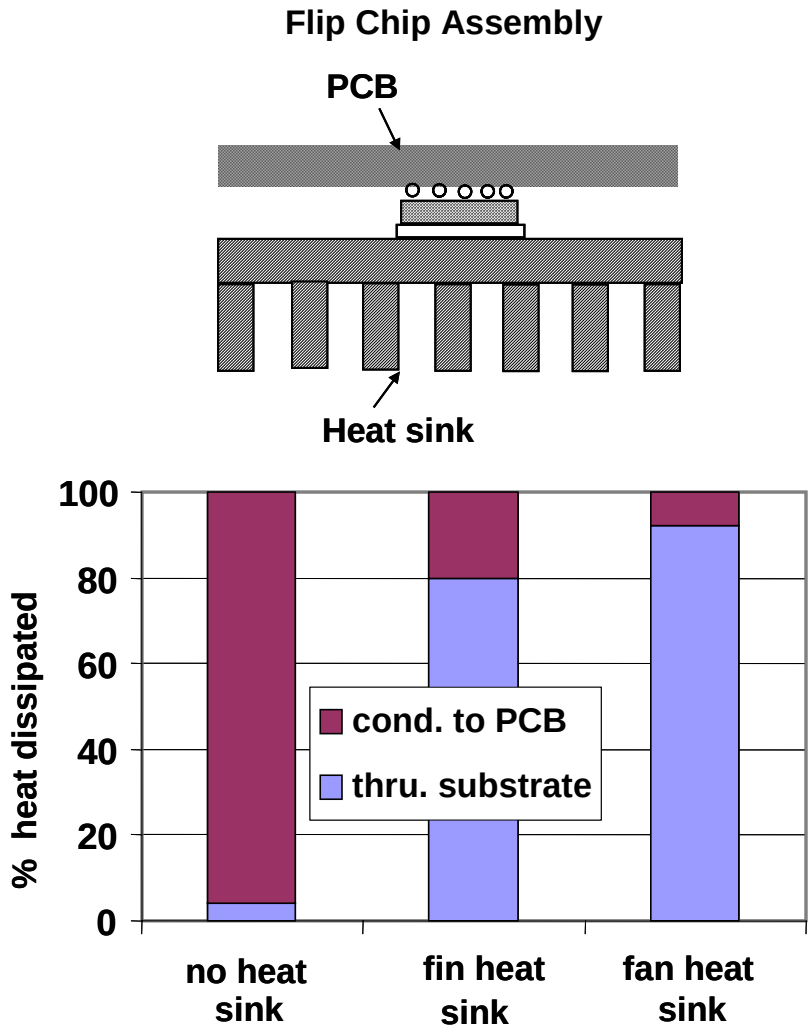


Fig. 2.6: With the help of heat sink, particularly with fan, the majority of heat will conduct to outside ambient through substrate rather than to printed circuit board (PCB) [35].

2.3.3 Steady state thermal modeling

In this section, we will develop the interconnect thermal modeling under normal chip operation circumstances. The thermal environment on the interconnects is considered to be under steady-state condition. Transient stress analysis, such as under electrostatic discharge (ESD), will be discussed in chapter 4. To begin with, consider a rectangular metal wire with thickness H, width w, length L, electrical resistivity ρ_m and thermal conductivity k_m, separated from the underlying layer by inter-layer dielectrics (ILD) of thickness t_{ILD} and thermal conductivity k_{ILD}. The two ends, at $x=\pm L/2$, of the wire are connected to the underlying layer through vias (Fig. 2.7). The vias are modeled

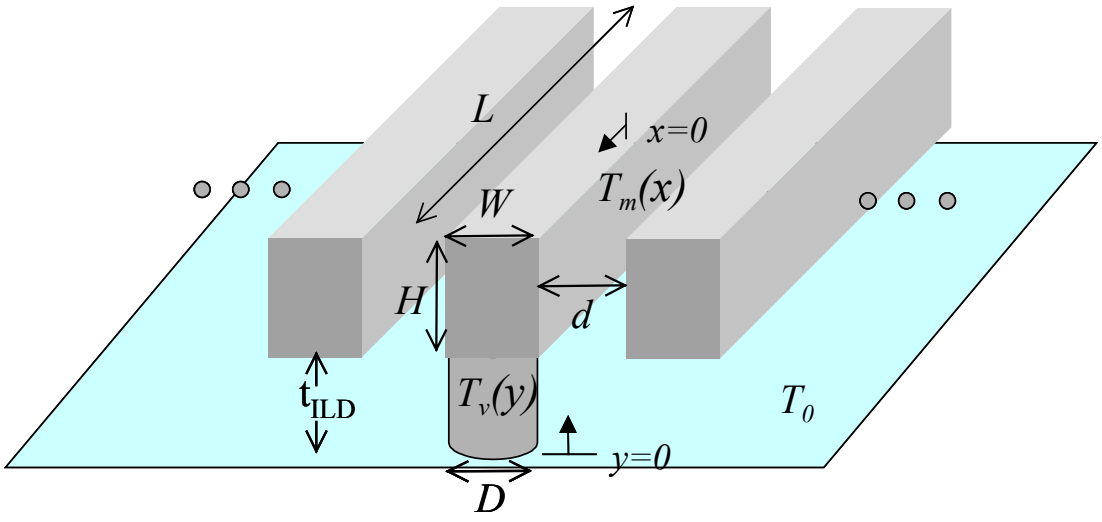


Fig. 2.7: Geometry used to derive interconnect thermal model, representing parallel metal wire array.

as round pillars, with diameter D and height t_{ILD}. The electrical resistivity and thermal conductivity of vias are ρ_v and k_v. The underlying layer temperature set to be T_0, and therefore temperature at the via bottom $T_v(0)=T_0$. The temperature rise of wires and vias is represented by $\theta_m(x) \equiv T_m(x)-T_0$ and $\theta_v(y) \equiv T_v(y)-T_0$, respectively. It is assumed that heat only flows downwards toward silicon substrate which is usually attached to a heat sink. Following Eq. 2.8, under steady state conditions, with uniform root-mean-square current, j_{rms}, flowing in the conductors, the governing heat equations for the wire and the via are

$$\frac{d^2\theta_m(x)}{dx^2} - \frac{\theta_m(x)}{L_{H,m}^2} = -\frac{q_m'''}{k_m}, \tag{2.9}$$

and

$$\frac{d^2\theta_v(y)}{dy^2} - \frac{\theta_v(y)}{L_{H,v}^2} = -\frac{q_v'''}{k_v}, \tag{2.10}$$

where $q_m''' (=j^2_{rms,m}\rho_m)$ and $q_v''' (=j^2_{rms,v}\rho_v)$ are the volumetric heat generation in the wire and the via due to Joule heating. $L_{H,m}$ and $L_{H,v}$ are the *thermal healing length* of the wire and the via, defined as

$$L_{H,m} \equiv \left[\frac{k_m A_m}{k_{ILD} S_m}\right]^{1/2} = \left[\frac{k_m wH}{k_{ILD} S_m}\right]^{1/2}, \tag{2.11}$$

and

$$L_{H,v} \equiv \left[\frac{k_v A_v}{k_{ILD} S_v}\right]^{1/2} = \left[\frac{k_v (\pi D^2/4)}{k_{ILD} S_v}\right]^{1/2}, \tag{2.12}$$

where S_m and S_v are the shape factor per unit length of the wire and the via, respectively, to represent heat spreading and, thus, to correct the deviation from one-dimensional heat

flow. It should be noted that the shape factor is only a function of geometric parameters. As it will be seen later, shape factor S_m and spreading factor S'_m are interchangeable and both describe the multi-dimension heat conduction flow from metal surface to surrounding medium. The physical meaning of $L_{H,m}$ is that within the range of thermal healing length, $L_{H,m}$, from vias, heat generated will flow through vias to the underlying layer. Beyond $L_{H,m}$, heat will flow through ILD and the via effect is diminished.

The two boundary conditions used to solve Eq. 2.9 are the adiabatic condition at the middle of the wire (x=0) due to symmetry, and the junction temperature θ_J at the two ends of the wire (x=±L/2), with $\theta_J=\theta_m(x=\pm L/2)=\theta_v(y=t_{ILD})$.

$$\left.\frac{d\theta_m(x)}{dx}\right|_{x=0} = 0, \qquad \theta_m(x=\pm L/2) = \theta_J, \tag{2.13}$$

In addition, the two boundary conditions to solve Eq. 2.10 are the temperature rise θ_J at the top of the via ($y=t_{ILD}$) and zero at the bottom of the via (y=0).

$$\theta_v(y=t_{ILD}) = \theta_J, \qquad \theta_v(y=0)=0. \tag{2.14}$$

The temperature rise along the wire and within the via can then be solved from Eq. 2.9 and Eq. 2.10 as

$$\theta_m(x) = \theta_J \frac{\cosh\left(x/L_{H,m}\right)}{\cosh\left(L/2L_{H,m}\right)} + \frac{q'''_m L_{H,m}^2}{k_m}\left(1 - \frac{\cosh\left(x/L_{H,m}\right)}{\cosh\left(L/2L_{H,m}\right)}\right), \tag{2.15}$$

$$\theta_v(y) = \theta_J \frac{\sinh\left(y/L_{H,v}\right)}{\sinh\left(t_{ILD}/L_{H,v}\right)} + \frac{q'''_v L_{H,v}^2}{k_v}\left(1 - \frac{\sinh\left(y/L_{H,v}\right) + \sinh[(t_{ILD}-y)/L_{H,v}]}{\sinh\left(t_{ILD}/L_{H,v}\right)}\right), \tag{2.16}$$

with $-L/2 \leq x \leq L/2$ and $0 \leq y \leq t_{ILD}$. It is instructive to note that we can define $q_m'''L_{H,m}^2/k_m$ $(= j_{rms,m}^2 \rho_m L_{H,m}^2/k_m = j_{rms,m}^2 \rho_m Hw/k_{ILD}S_m)$ in the second term of Eq. 2.15 as $\Delta T_{1\text{-}D}$, which represents the temperature rise in the wire obtained from simplified 1-D heat conduction model where via effect is not included. Thereafter, the wire temperature rise, $\theta_m(x)$, from Eq. 2.15 can be interpreted as the combined contribution from via self-heating and the thermal impedance caused by the via (first term) and the Joule heating in the wire corrected with the help of via effect (second term). The ratio of $\theta_J/\Delta T_{1\text{-}D}$ has other important indications as will be explained in a later section. To determine θ_J, we need to realize the continuous flow of heat at the junction of the wire and the via, which can be shown as the continuity equation

$$-k_m A_m \frac{d\theta_m(x)}{dx}\Big|_{x=\pm L/2} = k_v A_v \frac{d\theta_v(y)}{dy}\Big|_{y=t_{ILD}} \quad . \tag{2.17}$$

Substituting Eq. 2.15 and Eq. 2.16 into Eq. 2.17, θ_J can then be solved as

$$\theta_J = \frac{A_m L_{H,m} q_m''' \tanh(L/2L_{H,m})}{\frac{k_m A_m}{L_{H,m}} \tanh(L/2L_{H,m}) + \frac{k_v A_v}{L_{H,v}} \coth(t_{ILD}/L_{H,v})}$$

$$+ \frac{A_v L_{H,v} q_v''' [\coth(t_{ILD}/L_{H,v}) - \operatorname{csch}(t_{ILD}/L_{H,v})]}{\frac{k_m A_m}{L_{H,m}} \tanh(L/2L_{H,m}) + \frac{k_v A_v}{L_{H,v}} \coth(t_{ILD}/L_{H,v})} . \tag{2.18}$$

Effect of variation in electrical resistivity ρ with temperature was found to be small for practical situations and is ignored in this model. It should be noted that we should certainly include the temperature coefficient of resistivity (TCR) for electromigration tests where very high current densities are employed. It is simple to do so because

electromigration tests are usually conducted on isolated wires and not multilevel structures. Several papers have been devoted to study the case under electromigration test condition [36, 37]. The focus of our work, however, is to incorporate the via effect and to predict temperature distributions along multilevel interconnects under normal operation conditions. Therefore, the resistivity used in this work is assumed to be a constant at $\rho(T_{Die})$ to a first order approximation.

The last pieces of the puzzle in this model are the shape factors S_m and S_v, which accommodate 3-D heat spreading and different boundary conditions. The commonly used Bilotti's equation [29] is not adopted in this work to account for the deviations from 1-D heat flow. This is due to the fact that it assumes a single heat source, whereas, in typical IC layout, there are multiple heat sources due to parallel metal wire array. A new expression of heat spreading factor, S'_m, is therefore derived here for uniformly separated infinite number of parallel wires and will be used in the following analysis throughout this work. For the worst case scenario, all metal wires are assumed to carry the maximum RMS current density and separated by spacing d. As shown in Fig. 2.8, the Joule heat transfers downward as well as spreads laterally in the ILD. Assuming the lateral spreading to increase linearly with vertical coordinate, the spreading thermal resistance, R_{spr}, can be derived as

$$R_{spr} = \int_0^{d/2} \frac{1}{k_{ILD}} \frac{dy}{w + 2y} = \frac{1}{2k_{ILD}} \ln\left(\frac{w+d}{w}\right). \tag{2.19}$$

where d is the spacing between wires and is shared by the two wires for heat dissipation. Then, the total thermal resistance of ILD, $R_{th,ILD}$, can be calculated by combing the spreading resistance with the volume resistance,

$$R_{th,ILD} = \frac{1}{2k_{ILD}} \ln\left(\frac{w+d}{d}\right) + \frac{1}{k_{ILD}} \frac{t_{ILD} - d/2}{w+d} \quad . \tag{2.20}$$

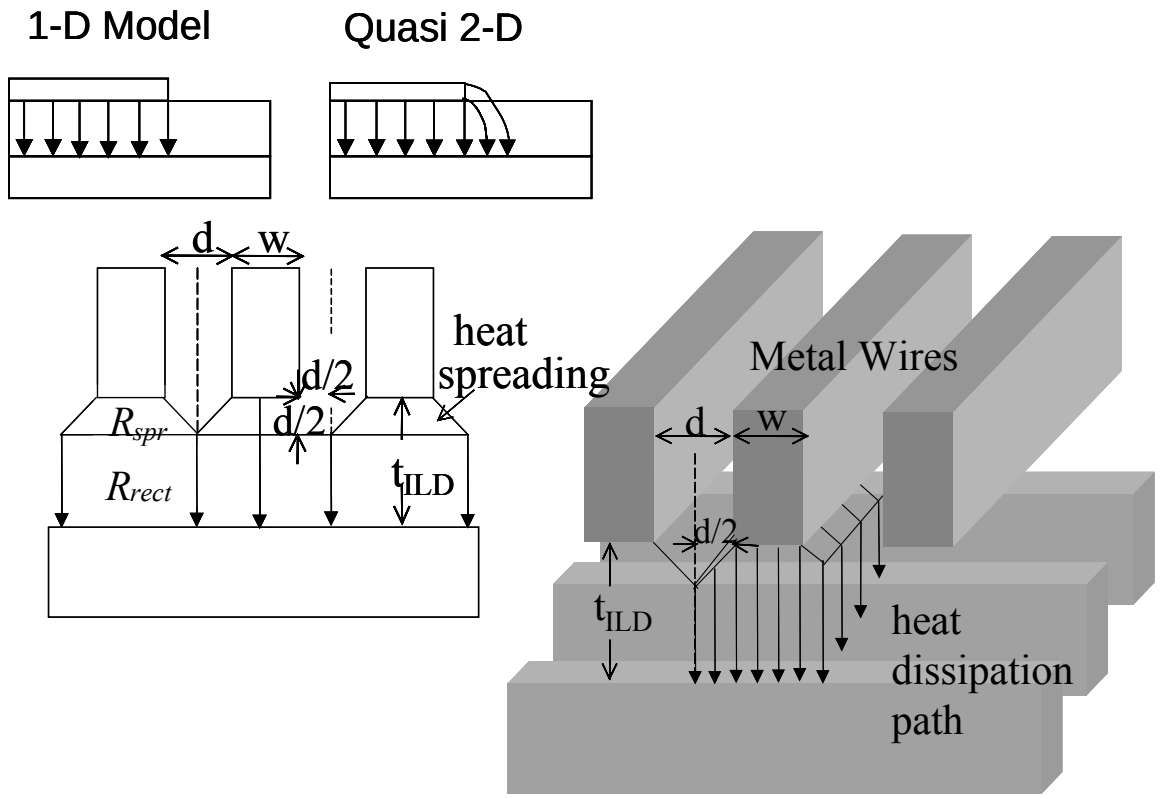


Fig. 2.8: Geometry used for calculating R_{spr}, R_{rect} and the spreading factor S'_m. Cross sectional view is shown on the left hand side. Quasi 2-D heat conduction is used to correct 1-D heat conduction model. Space between any two wires is shared for heat dissipation.

On the other hand, $R_{th,ILD}$ can be also expressed as

$$R_{th,ILD} = \frac{t_{ILD}}{k_{ILD} w_{effective}} = \frac{t_{ILD}}{k_{ILD} w S'_m}. \tag{2.21}$$

By comparing Eq. 2.20 and Eq. 2.21, , S'_m can be obtained as

$$S'_m = \left(\frac{w}{t_{ILD}} \frac{1}{2} \left(\frac{w+d}{w} \right) + \frac{w}{t_{ILD}} \frac{t_{ILD} - d/2}{w+d} \right)^{-1}. \tag{2.22}$$

The next step is to convert spreading factor S'_m to shape factor S_m. By the definition of shape factor [38],

$$q(x) = S_m k_{ILD} \theta_m(x) = \frac{\theta_m(x)}{R_{th,ILD}} \tag{2.23}$$

where $q(x)$ is the heat transfer rate from the metal wire to the dielectric medium. Hence, , $R_{th,ILD}$ can be manipulated as,

$$R_{th,ILD} = \frac{1}{S_m k_{ILD}} \tag{2.24}$$

Again, by comparing Eq. 2.21 and Eq. 2.24, the conversion between S_m and S'_m can be readily obtained as,

$$S_m = (\frac{w}{t_{ILD}}) S'_m \tag{2.25}$$

and thus,

$$S_m = \left([\frac{1}{2} \ln \left(\frac{w+d}{w} \right) + \frac{t_{ILD} - d/2}{w+d}] \right)^{-1} \tag{2.26}$$

It should be noted that, depending on different layout and operating conditions, S_m could have different expressions than Eq. 2.26. However, all the equations derived here can still be valid as long as the appropriate S_m is determined by either an appropriate analytical expression or extracted from simulation. The S_v is adopted from [38] for a vertical cylinder in a semi-infinite medium attached to a constant temperature surface,

$$S_v = 2\pi / [\ln(4t_{ILD} / D)] . \tag{2.27}$$

After the shape factors are installed in Eq. 2.11 and Eq. 2.12, the effect of the via separa-

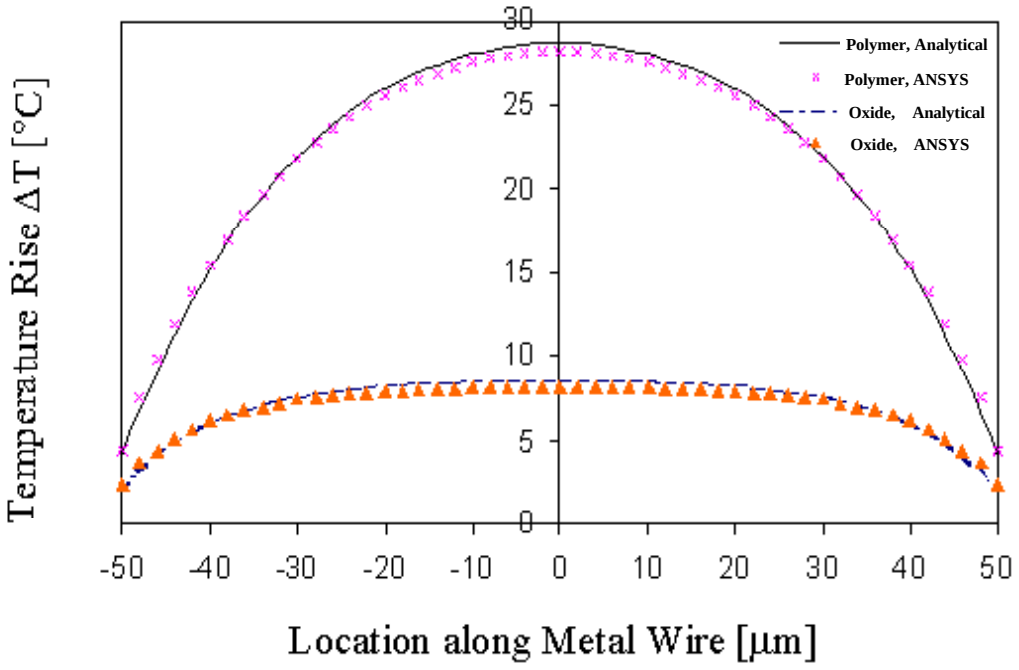


Fig. 2.9: Temperature profile along the Cu wires with 100 μm via separation. $H = tILD$ = 0.8 μm, $w = d$ = 0.3 μm, and $jrms,m$ = 3.7E6 A/cm^2. $k_{polymer}$ = 0.3 W/mK and k_{oxide} =1.2 W/mK are assumed here.

tion and heat spreading on the temperature profile along metal wire can be captured completely by Eq. 2.15. As can be observed form Fig. 2.9, the result from analytical expression is shown to be within 3% agreement with the 3-D finite element thermal simulation using ANSYS.

2.4 Impact of Via Effect on Effective k_{ILD}

2.4.1 Via correction factor

It should be noted that as predicted from Eq. 2.11 and validated from Fig. 2.9, the thermal healing length, $L_{H,m}$, in a wire is longer if the ILD has lower thermal conductivity, ($k_{polymer}$ = 0.3 W/mK v.s. k_{oxide} =1.2 W/mK). Consequently, the via effect is more important for low-k insulators. By defining a via correction factor (η), the via effect can be incorporated into the effective thermal conductivity of ILD, $k_{ILD,eff}$, which can then be used in place of the nominal k_{ILD} in the conventional thermal equations. An analytical expression for $k_{ILD,eff}$ incorporating the via effect is now derived here. ΔT_{ave} is defined as the average temperature rise in one metal layer and it can be expressed as

$$\Delta T_{ave} = q_m R_{th,ILD} = (I_m^2 R_{elec,m}) R_{th,ILD} = j_{rms,m}^2 \rho_m \frac{Hw}{k_{ILD,eff} S_m}. \tag{2.28}$$

On the other hand, ΔT_{ave} can also be obtained from Eq. 2.15 as

$$\Delta T_{ave} = \theta_{m,ave} = \frac{1}{L}\int_{-L/2}^{L/2} (T(x) - T_0)dx = \frac{1}{L}\int_{-L/2}^{L/2} \theta_m(x)dx$$

$$= j_{rms,m}^2 \rho_m \frac{Hw}{k_{ILD} S_m} \left[\frac{\theta_J}{(\frac{j_{rms,m}^2 \rho_m Hw}{k_{ILD} S_m})} \frac{\tanh\left(\frac{L/2}{L_{H,m}}\right)}{\frac{L/2}{L_{H,m}}} + (1 - \frac{\tanh\left(\frac{L/2}{L_{H,m}}\right)}{\frac{L/2}{L_{H,m}}}) \right] \tag{2.29}$$

Comparing Eq. 2.28 and Eq. 2.29, via correction factor, η, can be deduced as

$$\eta = 1 - (1 - \frac{\theta_J}{\Delta T_{1-D}}) \frac{\tanh\left(L/2L_{H,m}\right)}{L/2L_{H,m}} \ , \tag{2.30}$$

which yields two important results.

$$k_{ILD,eff} = k_{ILD} / \eta \tag{2.31}$$

and

$$\Delta T_{ave} = \Delta T_{1-D} \ \eta \ . \tag{2.32}$$

2.4.2 Hot spot location

As implied by Eq. 2.31 and Eq. 2.32, via effect will help increase the effective thermal conductivity of the ILD and, therefore, alleviate interconnect temperature rise if $\eta<1$. On the other hand, if via self heating and thermal impedance caused by the via is excessive, $\eta>1$ may occur and via effect will be detrimental to the wire. In addition, hot spots can occur within the vias if $\eta>1$. The essential criterion to differentiate these two domains is to evaluate the ratio of $\theta_J/\Delta T_{1\text{-}D}$, as indicated by Eq. 2.30. With $\theta_J/\Delta T_{1\text{-}D}<1$, a beneficial via effect is guaranteed. This critical condition can be further shown to be

$$\left[1 - (\frac{q_m'''}{q_v'''})(\frac{k_v}{k_m})(\frac{L_{H,m}}{L_{H,v}})^2 \right] \cosh(\frac{t_{ILD}}{L_{H,v}}) < 1 , \tag{2.33}$$

or

$$\left[1 - (\frac{j_{rms,m}^2}{j_{rms,v}^2})(\frac{\rho_m}{\rho_v})(\frac{S_v}{S_m})(\frac{A_m}{A_v}) \right] \cosh(\frac{t_{ILD}}{L_{H,v}}) < 1 . \tag{2.34}$$

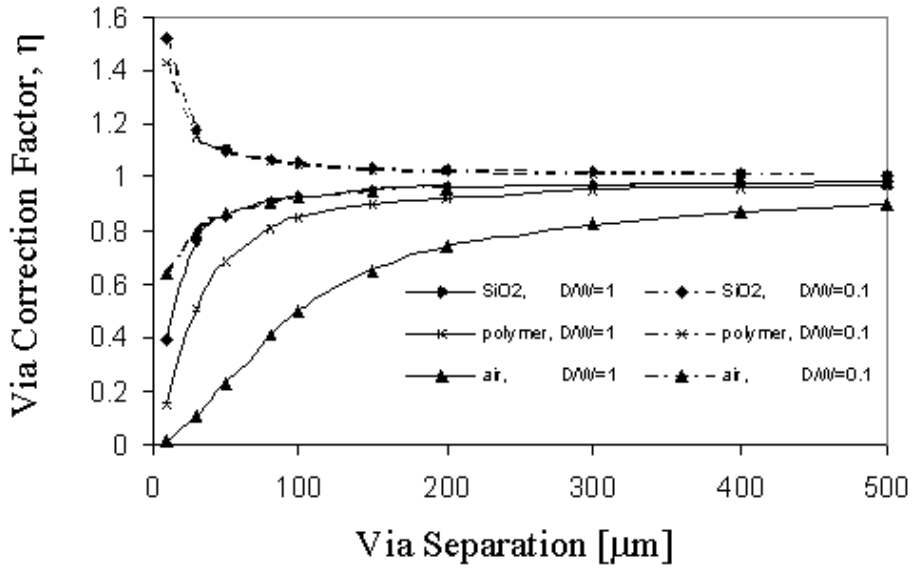


Fig. 2.10: Via correction factor, η, with two different D/W ratios, are plotted for three types of dielectrics. Dielectrics with lower nominal k_{ILD} experience lower η even with longer via separation. Wires with shorter via separation feel stronger via effect.

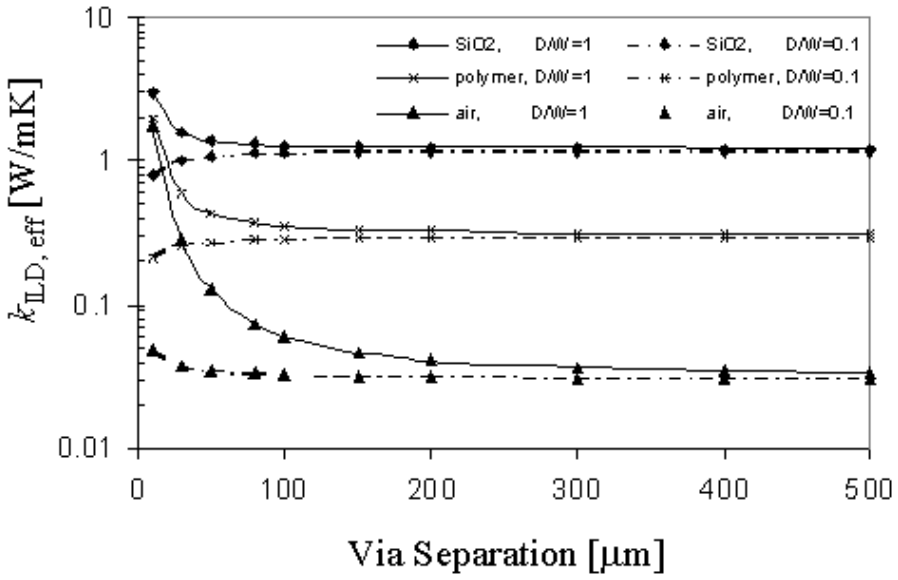


Fig. 2.11: Effective ILD thermal conductivity increases with decreasing via separation. The lower the nominal k_{ILD}, the longer the $L_{H,m}$, and hence, the stronger is the via effect.

From Eq. 2.33 and Eq. 2.34, various interconnect parameters can be quickly evaluated to determine the hot spot location (in the wire or within via) and the nature of via effect (alleviate or enhance wire temperature rise). To illustrate the effect of the via, η and $k_{ILD,eff}$ are plotted against via separation with two via diameter-to-wire width ratios, *D/w*, in Fig. 2.10 and Fig. 2.11, respectively. Geometries of the interconnect structure are taken from ITRS [5] 65nm technology node for global wires. It can be observed that, under the condition $\eta<1$, incorporation of the via effect results in increased $k_{ILD,eff}$ especially for ILD materials with lower nominal thermal conductivity. This fact can explain why the interconnect temperature is not as high as commonly assumed when low-k ILD is implemented in the advanced interconnect structure. Fig. 2.11 also shows that the via effect diminishes rapidly for the portion of wire beyond thermal healing length from each end, the $L_{H,m}$'s for oxide, polymer and air are about 5μm, 10um and 30um as calculated from Eq. 2.11. On the other hand, as shown in Fig. 2.10, for the case of *D/w*=0.1, the cross section of the via is much smaller than that of the wire which results in substantial via heating. The temperature profile along the interconnect is shown in Fig. 2.12. In this case, short via separation suffers more heat backflow from the vias. However, in the typical VLSI interconnect structure, *D/w* is generally greater than 0.5, so that the hot spot locates in the middle of the wire and the via effect is always beneficial.

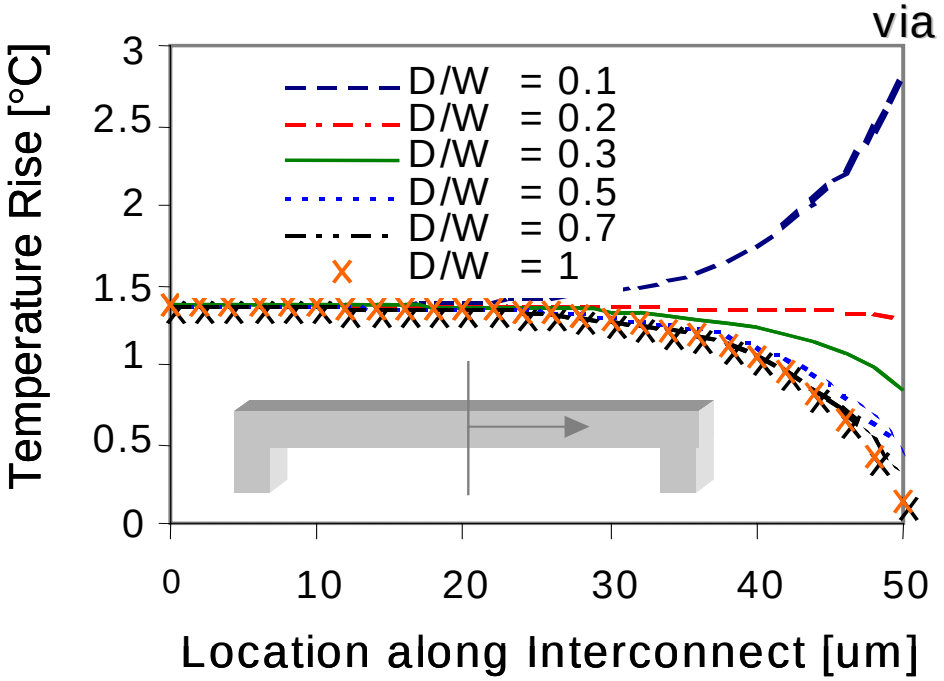


Fig. 2.12: The figure shows the temperature profile along the wire from the middle of the interconnect to the via at the right end. The hot spots usually locate in the wire except for extremely small via dimensions.

2.5 Temperature in Multilevel Metal Layers

2.5.1 Multilevel interconnect formula

Following the previous argument, Joule heat generated in metal wires is considered to dissipate only through the heat sink attached to the Si substrate. This assumption is fairly legitimate due to the fact that the chip is usually encapsulated with insulation materials. Therefore, all the heat generated in the upper metal levels has to transfer through the lower metal levels to the substrate. With $\Delta T_{i-1,\ i}$ defined as the average

temperature rise between metal layers *i-1* and *i*, the temperature rise at the top layer for an N-level metal interconnect can be obtained as

$$\Delta T_N = T_N - T_{substrate} = \sum_{i=1}^{N} \Delta T_{i-1,i} = \sum_{i=1}^{N} R_{th,i} Q_i \tag{2.35}$$

$$\cong \sum_{i=1}^{N} \frac{w_i \eta_i}{k_{ILD,i} S_{m,i} \alpha_i} \sum_{j=i}^{N} j_{rms,j}^2 \rho_{m,j} \alpha_j H_j \tag{2.36}$$

where $Q_{i,}$ is the accumulated heat dissipated through the i^{th} level interconnects and α_i is the metal coverage of the i^{th} level metal layer. For the case when the via effect is neglected, η_i is set to be 1. As can been seen from Eq. 2.35, there is more heat flowing through the lower levels since Q_i represents the sum of all the heat generated from i^{th} layer to N^{th} layer. As a result, a substantial temperature rise will occur in local wires if the effect of the dense via population is not taken into account. For the following demonstration of the importance of the via effect, some reasonable values of via separation are assigned to each of the 10 metal layers with polymer used as the ILD in a 65 nm technology node structure. In addition, a worst case current density, j_{rms}, of 2.1e6 A/cm^2 is assumed for all wires and 50% metal coverage is assumed for each metal layer. First, as shown in Fig. 2.13, the overall temperature rise is much lower with the help of the vias. Second, it can be observed that the temperature distribution among metal layers is quite different in these two cases. Ignoring the via effect results in large temperature jump in the lower layers and then the temperature rise levels off. On the other hand, with the via effect considered, there is hardly any temperature rise in the lower levels even when

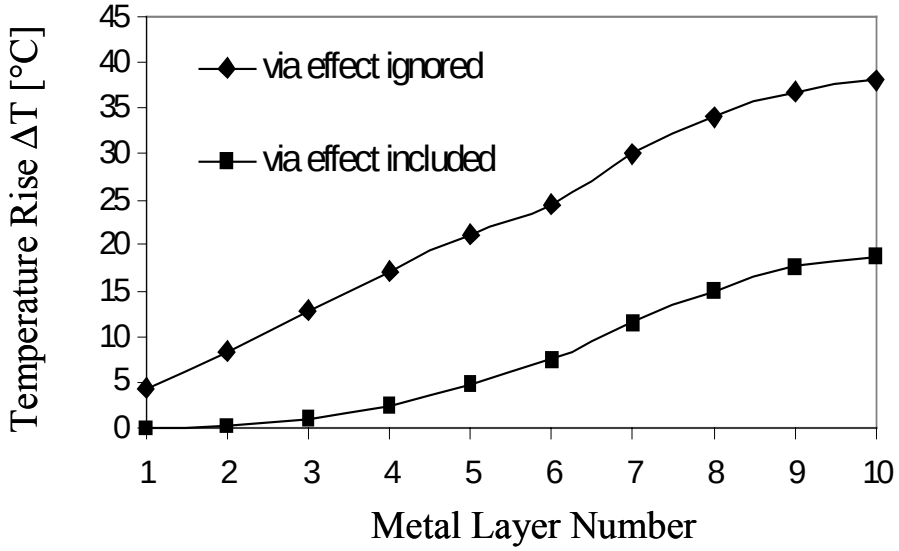


Fig. 2.13: Temperature rise distribution along metal layers from substrate to top metal level. In the case of via effect included, the via separations assigned to the metal layers, from 1st to 10th levels, are 5, 10, 30, 50, 100, 150, 200, 300, 500 and 1000 metal pitches,respectively, based on 65nm technology node.

$k_{polymer}$ is only one fourth of k_{oxide}. Most of the temperature rise is attributed to the upper metal layers with long via separation. Therefore, from the thermal design point of view, global interconnects are more problematic . The concern of RC delay in the global wires may get worse with this additional temperature effect.

2.5.2 Interconnect temperature rise trend

The trend of interconnect temperature rise due to Joule heating is investigated in Fig. 2.14, curve (a). It is interesting to see that even with increasing current density and lower thermal conductivity from the low-k ILDs used for the more advanced technology nodes, the interconnect temperature rise will reach a plateau and then drop. This phenomena can

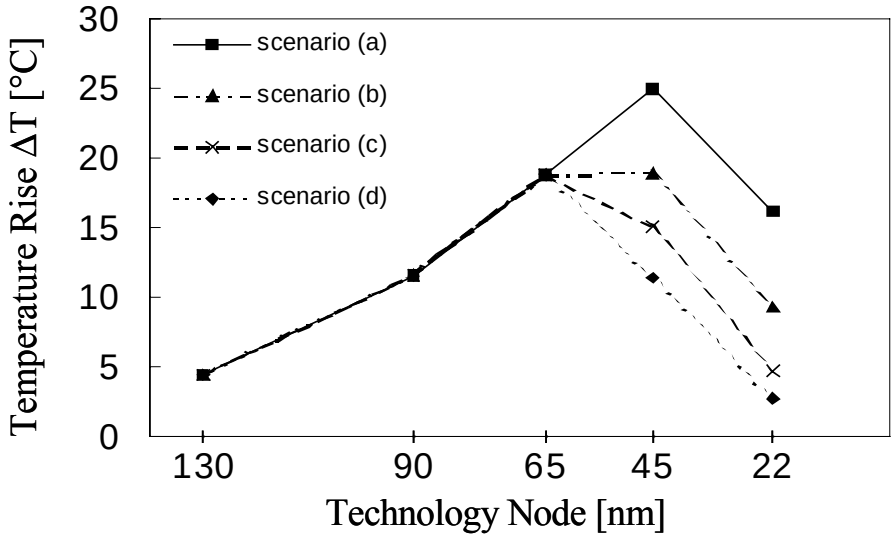


Fig. 2.14: Trend of interconnect temperature rise along technology nodes for different scenarios. (a) both j and k_{ILD} scale according to ITRS. (b) j scales with ITRS, but k_{ILD} stops scaling at 65nm node. (c) j stops scaling at 65nm node, while k_{ILD} continues scaling. (d) both j and k_{ILD} stops scaling at 65 nm node

be understood by realizing that the via separation keeps reducing due to scaling. The same functionality can be reached within a shorter distance. On the other hand, the magnitude of current density and the dielectric constant of the low-k insulators are believed to encounter a bottleneck due to reliability concern at 65nm technology node. Several scenarios are studied here to identify their impact on interconnect temperature, as shown in Fig. 2.14. Overall, the interconnect temperature rise will be less if either the trend of increasing current density or lower-k materials stops at 65nm nodes, while all the other aspects of scaling continue.

2.6 Summary

In conclusion, a compact analytical thermal model has been presented to evaluate the spatial thermal distribution and the average interconnect temperature rise under the influence of vias. Both the via effect and heat spreading have been taken into account to ensure accurate predictions. This model provides a quick and accurate interconnect temperature rise estimation as well as a comprehensive thermal design guideline. It will be essential for high performance chip designer being able to comprehend the thermal impact on the circuit performance and reliability implication in the early design stage. It has been shown that with the help of vias as efficient thermal paths, the effective thermal conductivity of the ILD materials can be significantly higher than their nominal values if via separation is comparable to the characteristic thermal length. The interconnect temperature can be substantially lower than that predicted from overly simplified 1-D thermal model. Therefore, the thermal problem associated with low-k insulators is not as bad as it appears.

Additionally, a closed form thermal model incorporating the via effect has been formulated to estimate the temperature rise of interconnects in multi-level metal arrays. It is shown that via effect must be considered in the thermal analysis of interconnect structures. It is observed that global interconnects would suffer much higher temperature rises than local interconnects due to the much longer via separation. Beyond the 45nm node closer packing of vias will alleviate the temperature rise problem.

Chapter 3

Thermal Impact on Interconnect Design

3.1 Introduction

Thermal effects are very important in determining both reliability and performance of interconnects. Accurate integrity estimation of on-chip interconnect temperature rises caused by joule heating is essential for high performance chip design because the wire current-carrying capability is stringently limited by interconnect temperature. Thermal effects impact the interconnect design in the following ways. First, the wire slows down with higher temperature due to increasing metal electrical resistivity. For the temperature coefficient of resistivity (TCR) of Cu given in [39], wire resistance goes up by 5-10% as interconnect temperature rises by 10-20°C. As a result, the delay reduction expected from the introduction of low-k dielectrics, which have invariably poor thermal conductivity, will be discounted to some extent. Second, most of interconnect failure mechanisms are temperature related including electromigration (EM). EM lifetime, mean-time to-fail

(MTF), decreased exponentially with inverse interconnect temperature according to Black's equation [19]. It is reported that MTF could be reduced by 90% when interconnect temperature rises from 25°C to 52.5°C [32]. Therefore, traditionally, the current density design limits developed by reliability engineers for circuit designers are very conservative. The interconnect temperature is commonly limited to ~105°C for electromigration life time considerations [40, 41].

Having realized the importance of the temperature rise on interconnects, many efforts have been devoted to estimate the interconnect temperature [24, 31] and propose "self-consistent" analysis for allowed current density [30, 42]. However, without considering via effect properly, the estimated interconnect temperature rise can be much higher than twice as high compared to the realistic case [43, 44]. The temperature and consequent performance/reliability predictions with those overly simplified thermal analysis will deviate even more profoundly with more advanced technology nodes as the via effect is more effective for lower-k dielectrics and the via separation is shortened with scaling [45]. Extending the work from Chapter 2, our multi-level interconnect thermal model incorporating the via effect enables more realistic circuit timing simulation and reliability assessment without being excessively conservative. In this chapter, the impact of Joule heating on the scaling of deep sub-micron Cu/low-k interconnects will be investigated in detail. [46]. In addition, traditional wisdom requires the on-chip interconnect temperature to be no more than 5°C above the silicon temperature (100°C for a typical high performance microprocessor) for EM lifetime consideration. However, the prospects of rising

interconnect temperature and the need for greater current carrying capability are not compatible. Although it has been proposed to consider EM lifetime and wire self-heating simultaneously in generating EM guidelines [23, 30, 42], the lack of realistic interconnect temperature estimation makes the proposals fruitless. In a later section of this chapter, combined with our new interconnect thermal model, a comprehensive EM evaluation methodology is proposed from the view of coupled performance and reliability analysis [45].

The remainder of this chapter is organized as follows. In Section 3.2, we discuss how the interconnect metrics are influenced by thermal effect. Section 3.3 addresses the foreseeable impact from interconnect scaling trend on wire temperature due to Joule heating. Section 3.4 describes the definition of a 'reasonable' worst-case scenario of interconnect thermal analysis and applies it to evaluate the thermal effect on Cu/low-k interconnects. In Section 3.5, we perform a coupled assessment of EM reliability and current drivability for global interconnects for 22-130 nm technology nodes. Section 3.6 investigates the trend of interconnect temperature rise under various scaling scenarios. Finally, we summarize and conclude in Section 3.7.

3.2 Thermal Effect on Interconnect Metrics

It suffices to mention here that a non-negligible and increasing resistance of the wires leads to a $j^2\rho$ (j is the current density) power dissipation in the form of heat. This

raises the temperature of the interconnects above the device temperature, especially if lower dielectric constant materials, which are also invariably poorer heat conductors, are used.

3.2.1 RC delay

The delay of the wires can be well approximated by the product of resistance (R) and the capacitance of the wire (C), if inductive effects are not important. The wire capacitance typically has three components: inter-level, C_{ILD}, inter-metal (within the same level between metal lines), C_{IMD}, (Fig. 3.1) and the fringe component. To model the RC delay of the wires it is imperative to accurately model both the resistance and the capacitance accurately.

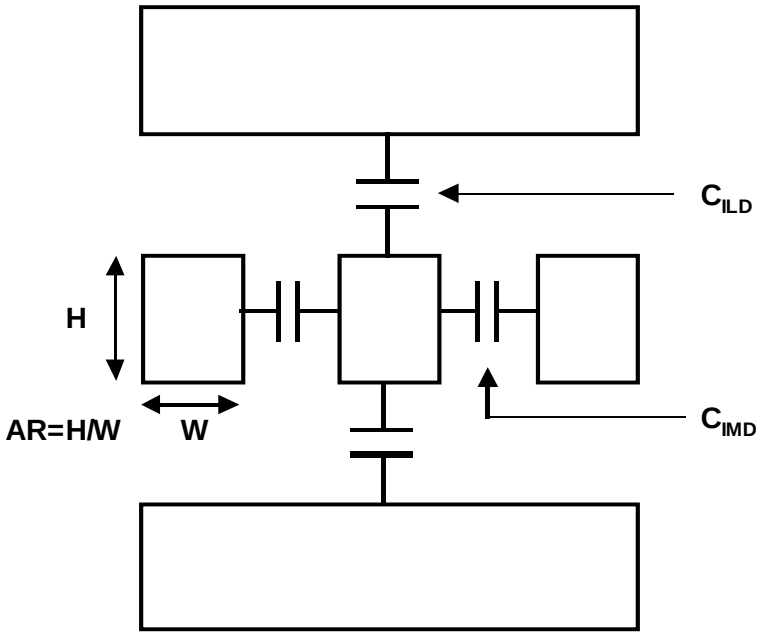


Fig. 3.1: Schematic showing the inter-metal and the inter-level components of capacitance. Also showing aspect ratio (AR).

It is well known that the electrical resistivity, ρ_m,is a function of wire temperature, and can be modeled as,

$$\rho_m(T) = \rho_m(T_0)[1 + \beta(T - T_0)] \tag{3.1}$$

here β being the *temperature coefficient of resistivity* (TCR), and has the value of 4.3e-3/K for Cu and 4.5e-3/K for Al. In terms of interconnect capacitance, to gain better insight of the role of C_{ILD} and C_{IMD}, let us ignore any fringe capacitance for this moment. The total interconnect capacitance per unit length can then be simply expressed as:

$$C_{total}=2(C_{ILD}+C_{IMD}), \tag{3.2}$$

where $C_{ILD}=\varepsilon_{ILD}/2AR$ and $C_{IMD}=\varepsilon_{IMD}AR,$ (3.3)

with *AR* denotes wire *aspect ratio*, defined by *H/w*, where *H* is the metal thickness (height) and *w* is the width of the wire. ε represents the respective premittivity for IMD and ILD. Thickness of the ILD is approximately the same as the height of the metal wire. The factor of 2 in the denominator for C_{ILD} accounts for the overlap with orthogonal wires on adjacent levels. The length of overlap is taken to be half the length of the interconnect based on the assumption that wire width is half the pitch. First, we notice that, due to the high wire *AR*, the intra-level (line-to-line) capacitance, $C_{IMD,}$ is dominant. Second, since most of the heat is flowing downwards toward the heat sink, the ILD assumes the major contribution to the thermal impedance and IMD being a minimal factor. Therefore, heterogeneous dielectric schemes (with different dielectric materials for ILD and IMD), should be exploited to optimize RC delay and subside thermal effect, such as use low-k dielectrics for IMD and keep silicon dioxide for ILD.

3.2.2 Dynamic power consumption

The second metric of importance, the power dissipation due to interconnects, is a result of charging and discharging its capacitance and is given by the dynamic power dissipation formula

$$P_{int} = S_w C_{int} V^2 f \tag{3.4}$$

Here, S_w is the switching activity factor representing the probability of a particular interconnect switching during a clock cycle, C_{int} is the total interconnect capacitance, V is the voltage to which the interconnect charges and f is clock frequency. Thus, at a given technology node, the interconnect power is heavily dependent on its total capacitance.

3.2.3 Cross talk noise

. Cross talk is proportional to the ratio of the inter-metal to the total capacitance [47] of the wire. Thus, from the cross talk perspective it is more beneficial to lower just the inter-metal capacitance. Hence, again, an heterogeneous dielectric approach will be appropriate to lower coupling noise while contain interconnect temperature rise.

3.2.4 Electromigration reliablity

Electromigration is wildly regarded a major failure mechanism of VLSI interconnects[28, 48]. The current through metal wires leads to metal atom migration due to momentum exchange between electrons and metal atoms [19]. This migration, over time, leads to depletion of enough material so as to initially increase the wire resistance and finally cause an open circuit [49]. On the other hand, it also causes excess metal atoms to

accumulate at a different location along the wire, which in extreme cases can cause a short to the adjacent wire through metal hillocks. The primary factors, which influence electromigration can be divided into those related to the physical structure of the metal wire and those related to the conditions of operation. Certain crystal orientations of thin films (metal wire) are more conducive to preventing electromigration. Among the conditions of operation, temperature and the current density play the most important role in dictating electromigration, as modeled by the well known Black's equation, given by Eq. 3.5,

$$MTF = A\, j^{-n} \exp(Q / k_B T) \tag{3.5}$$

where *MTF* is the meat-time-to-failure (typically for 0.1% of accumulative failure or 10 years lifetime), 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 operation conductions. The activation energy Q for Cu has been reported in the range of 0.5-1eV [50-52], k_B is the Boltzmann's constant, and T_m is the metal wire temperature. For a middle of the range Q value of 0.75eV, *MTF* will drop by 50% when interconnect temperature rises from 100°C to 110°C. Usually, the maximum allowable current density is limited by the goal that the electromigration lifetime will achieve 10 years with interconnect temperature maintained at or below 105°C.

3.3 Scaling Trend of Joule Heating

As VLSI technology advances, interconnects have become the limiting factor to IC chip performance [3]. Aggressive interconnect scaling has resulted in increasing current density, more metal levels, and introduction of low dielectric constant (low-k) materials. The growing demand of higher current driving capability and the aggressive shrinking of metal pitch has resulted in significant current density rise in the wires (Table 3.1). Wire width is half the wire pitch for all cases if not defined specifically, n is the number of layers in each tier, $\varepsilon_{r,ILD}$ and k_{ILD} are the relative dielectric constant and thermal conductivity of the inter-level dielectrics (ILD) for each technology node.

Tech. Node	$\varepsilon_{r,ILD}$	k_{ILD} [W/m-K]	Local Tier				Semiglobal Tier				Global Tier			
			n	Wire Pitch [nm]	Wire AR	Via AR	n	Wire Pitch [nm]	Wire AR	Via AR	n	Wire Pitch [nm]	Wire AR	Via AR
130	3.3	0.7	2	350	1.6	1.6	4	450	1.6	1.4	2	670	2.0	1.8
90	2.8	0.45	2	210	1.7	1.7	4	265	1.7	1.5	3	460	2.1	1.9
65	2.5	0.36	2	150	1.7	1.7	4	195	1.8	1.6	4	290	2.2	2.0
45	2.1	0.25	2	105	1.8	1.8	4	135	1.8	1.6	4	205	2.3	2.1
22	1.8	0.17	2	50	2.0	2.0	4	65	2.0	1.8	5	100	2.5	2.3

Table 3.1: Interconnect parameters for 130 nm to 22nm technology nodes based on ITRS[5].

So explosive has the growth been, according to ITRS [5], the rise of interconnect current density will outpace the average chip power density by a factor of two for high performance microprocessor throughout the technology nodes (Fig. 3.2).

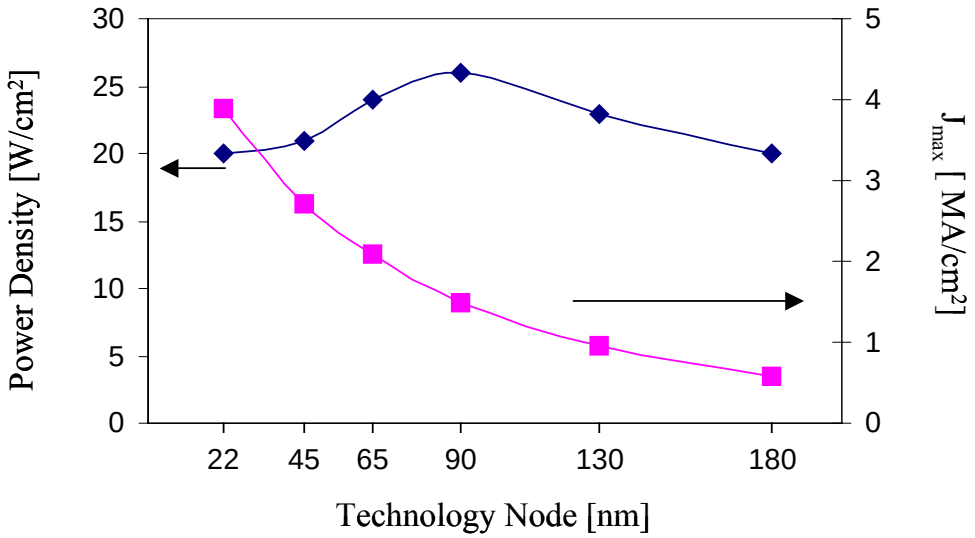


Fig. 3.2: Trends of chip power density and interconnect J_{max} along technology nodes suggested by ITRS [5]. Chip power density is calculated by total power of the chip divided by chip size.

Furthermore, Cu resistivity will increase due to barriers, surface scattering and skin effect [53, 54]. As a result, interconnect joule heating is becoming non-negligible. In addition, thermal conductivity of low-k dielectrics is decreasing rapidly with the reduction of dielectric constant (Table 3.1). The combination of greater heat generation and thermal impedance is leading to a continuous increasing interconnect temperature and consequent impact on wire delay and reliability wires is fast emerging as an urgent issue. Therefore, the impact of the interconnect "thermal"scaling trend on wire temperature due to Joule heating requires immediate attention. The work presented in this chapter is for the 65nm technology node with all parameters quoted from the newly updated ITRS '01[5]. We assign the first two metal layers as local tier, the following four layers as semi-global tier and the remaining layers as global tier. The substrate temperature, T_{ref}, is

assumed to be fixed at 100°C (which will be true if two-phase microchannel cooler is employed in the future [55]), and the temperature of the top global wire is $T_m = T_{ref} + \Delta T_{joule\ heating}$.

3.4 Temperature Effect on Cu/low-k Interconnects

3.4.1 Definition of worst case condition

To provide robust thermal analysis for interconnects, it is important to identify a reasonable worst case scenario. Previous work has attempted to evaluate the thermal characteristics of interconnects neglecting vias as an effective heat conduction path [4]. However, ignoring vias in heat transfer predicts unrealistically high temperature rises even with moderate current densities. This is especially true for dielectrics with lower thermal conductivities [44, 46]. Therefore, this work includes the via effect in the analysis to obtain more realistic results. Case (1) in Fig. 3.3 shows the widely used condition, i.e., all wires flowing the same current density and the via effect ignored. Case (2) represents an isolated global wire. Although there is no additional heat source between this wire and the substrate, the thermal impedance is higher in the absence of lower level metal. Case (3) has the same current condition as case (1) but with vias taken into account for all metal layers. Reasonable via separations are assigned for each metal level from level one to ten: 5, 20, 50, 100, 150, 200, 300 ,500, 1000 and 2500 interconnect pitches. As shown in Fig. 3.4, case (1) has the highest temperature rise ($\Delta T_{joule\ heating}$).

But we will disregard this condition as unrealistic as vias are ignored. Case (3) is worse than case (2) and we will use this condition for the following analysis.

Fig. 3.3: Configurations of the three thermally worst case scenarios. Current flows through all wires at all metal layers in cases (1) and (3).

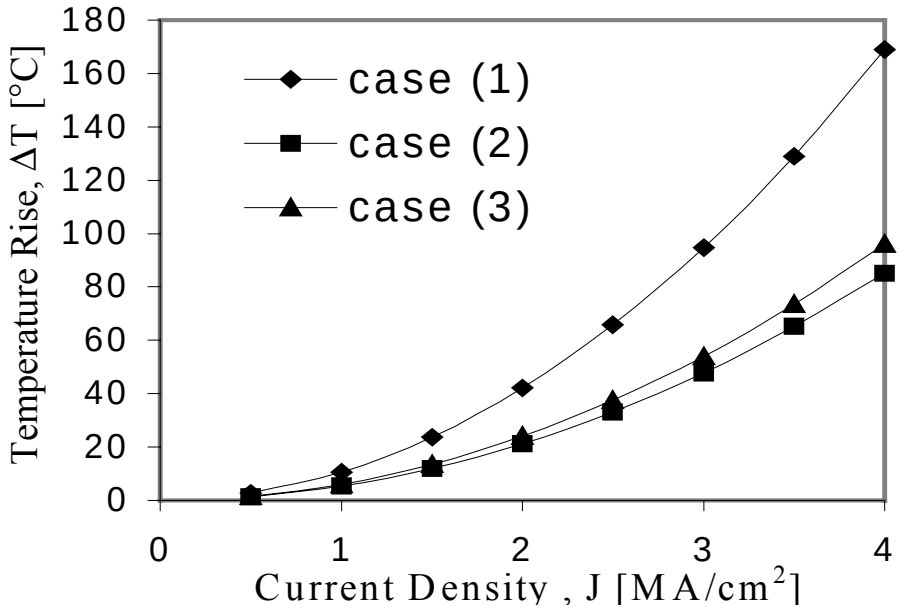


Fig. 3.4: Temperature on the top-level interconnects rises rapidly with increasing current density.

3.4.1 Thermal impact on Cu/low-k Interconnects

Fig.3.5 compares the thermal performance of several low-k materials. The materials properties are shown in Table 3.2. As the figure shows, interconnects with lower dielectric constant materials exhibit significantly higher temperature rises, thus higher

Parameter	Conductor			Dielectrics					
	Cu	Al	W	Air	Aerogel	Polyimide	HSQ	FSG	Oxide
k [W/m-K]	400	240	180	0.03	0.17	0.3	0.54	1.05	1.2
ρ [μΩ-cm]	2.2	3.6	10	-	-	-	-	-	-
εr	-	-	-	1.0	1.7	2.5	3.1	3.7	4.0

Table 3.2: Materials properties of various conductors and dielectrics used in this study. *k* [W/m-K] is thermal conductivity, ρ [μΩ-cm] is electrical resistivity near 100ºC and ε_r is the relative dielectric constant

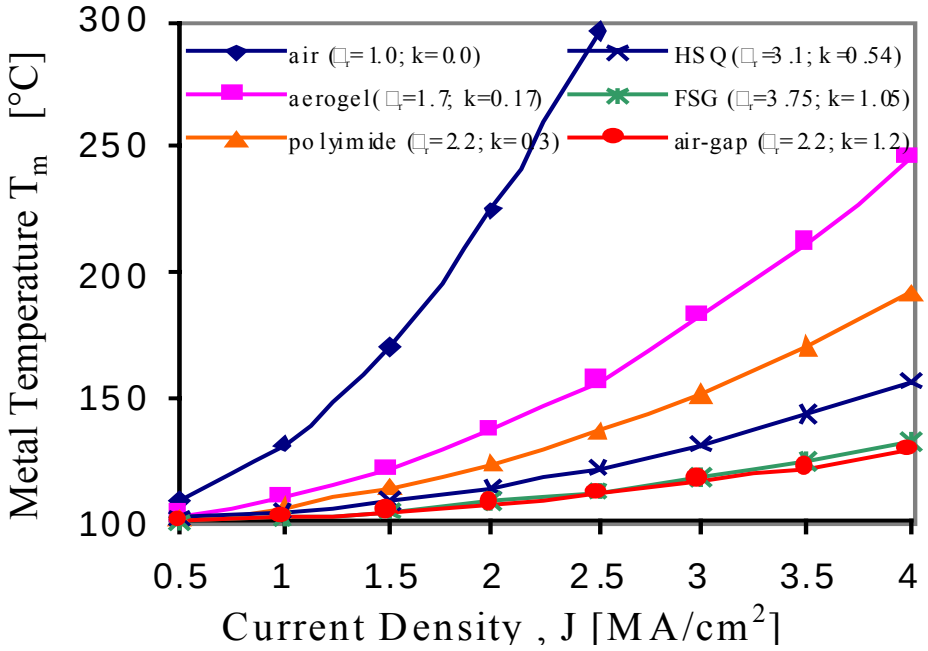


Fig. 3.5: Temperature of top global interconnects rises sharply for Low-k dielectrics. ε_r: relative dielectric constant, k: thermal conductivity [W/Km].

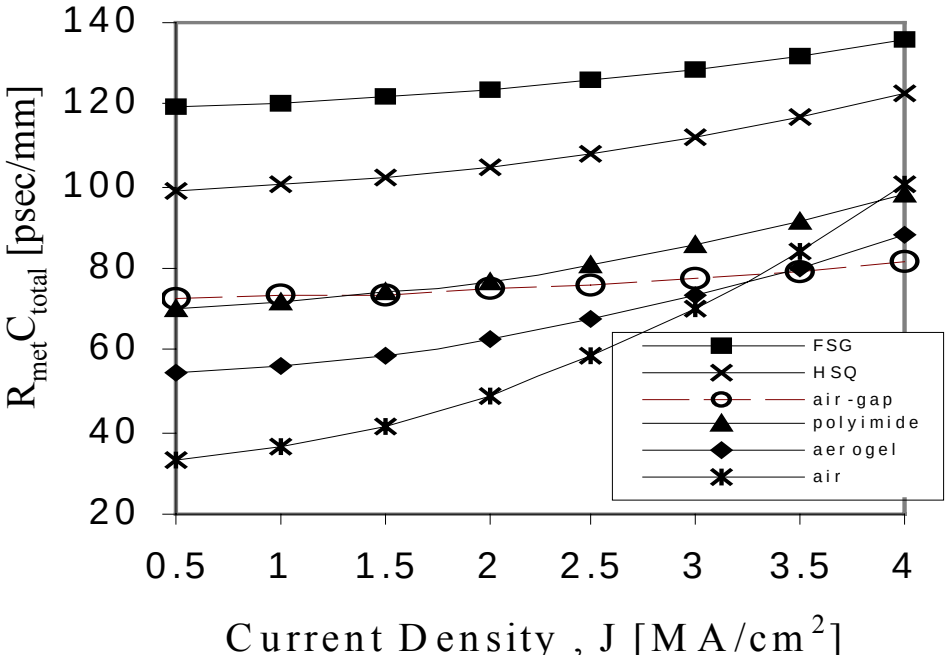


Fig. 3.6: RC delay is strong function of current density on the wires because of Joule heating. The lower the dielectric constant, the stronger the Joule heating and greater RC degradation.

interconnect temperature, T_m. The product of $R_{met}C_{total}$ per unit length for these dielectrics is shown in Fig. 3.6 as a function of current density. As the figure shows, the low-k advantage of reduced capacitance can be offset by the increased temperature rise due to poor thermal conductivity. It should be noted that for the case of air-gap scheme (ILD: SiO_2 and IMD: Air), the $R_{met}C_{total}$ is relatively constant through the range of current density and it is even better than polyimide. This is because air-gaps reduce the dominating line-to-line capacitance while leaving the SiO_2 ILD intact for better thermal conductivity. Since electromigration mean time to failure (MTF) is exponentially de-

pendent on wire temperature, the maximum allowable current density for a given temperature rise is evaluated in Fig. 3.7 for different low-k materials.

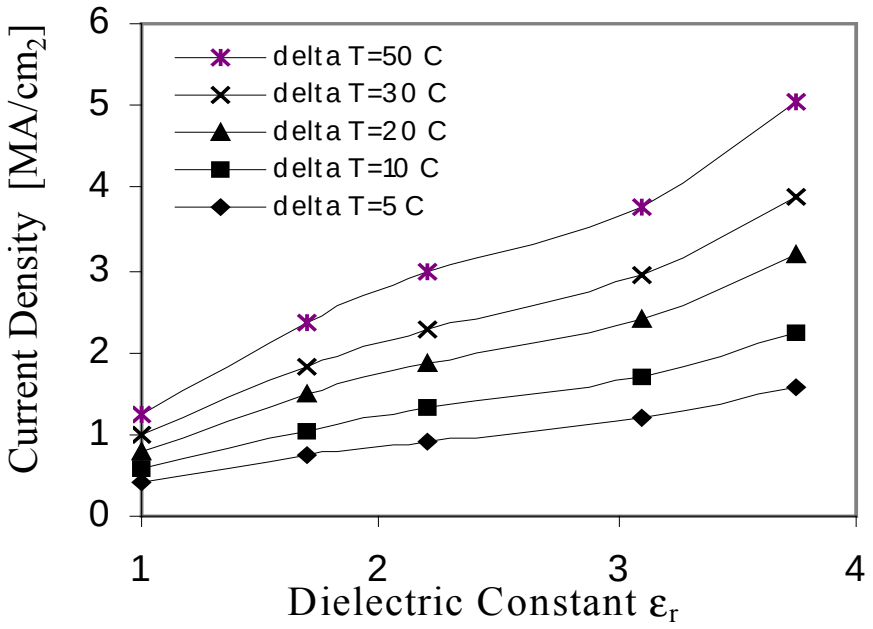


Fig. 3.7: Maximum allowed current density is limited by the maximum allowed ΔT on the metal wires. The constraint is more stringent for low-k dielectrics.

3.5 Delay and Reliability Optimization

With the possibility that much fatter wires may be used in the global tier to reduce IR drop, it is important to know the thermal impact of using low aspect ratio lines on the expected performance and reliability. In addition, ILD thickness should be evaluated to assess the trade-offs of thermal impedance and capacitance. The contours of constant $R_{met}C_{total}$, in Fig.3.8 show the optimization of the wire aspect ratio (AR) and the ILD

thickness for delay consideration. The value of H/S (=0.32μm/0.15μm) is fixed, where H is metal height and S is wire spacing. It demonstrates that Joule heating can increase delay of the interconnect by as much as 15%. Fig. 3.9 shows constant MTF contours, with the temperature effect included. The MTF can never reach 50% of MTF at 105°C which is the wire temperature specified in ITRS. We can conclude from Fig. 3.8 and 3.9 that although fat wires can provide better speed performance, but they are also more subject to electromigration failure since the temperature is higher.

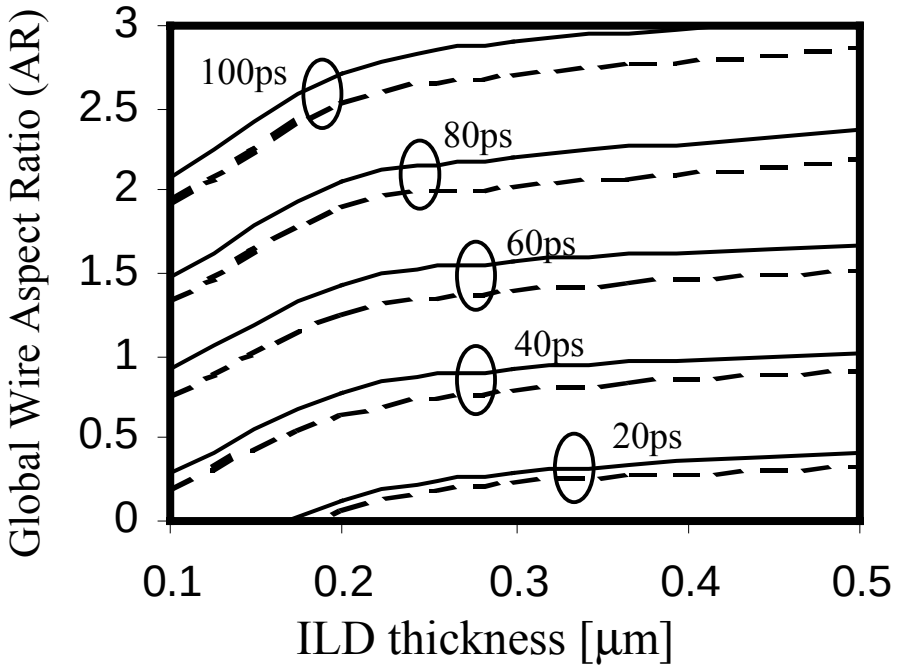


Fig. 3.8: Constant $R_{met}C_{total}$ contour plots of global level wiring as functions of wire aspect ratio (AR) and ILD thickness. Solid curves represent the case thermal effect neglected. Dash-line curves include the influence of thermal effect. (H/S=0.32/0.15 μm).

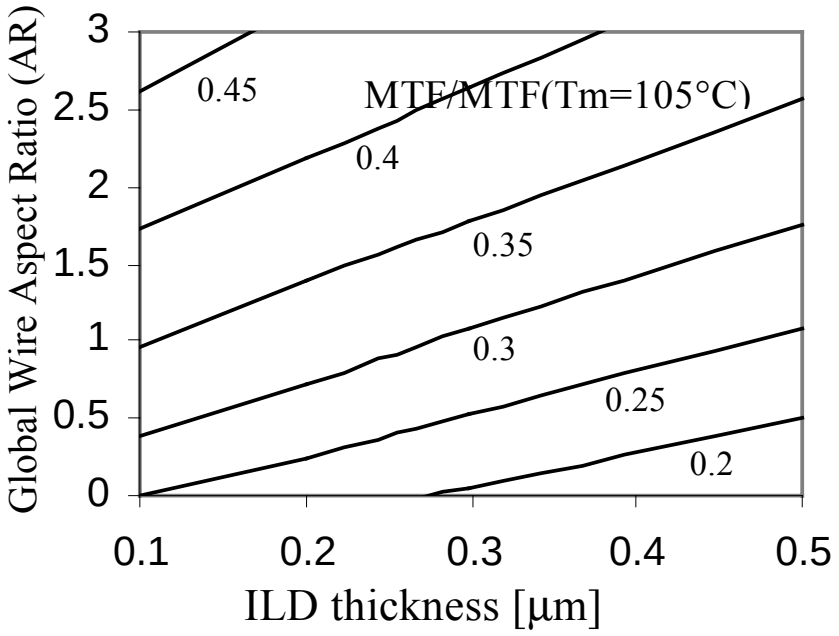


Fig. 3.9: Constant normalized electromigration MTF, MTF(Tm)/MTF(105°C), contour plots of global level wiring as functions of wire aspect ratio (AR) and ILD thickness. (H/S=0.32/0.15 μm).

3.6 Impact of Joule Heating on Scaling Trend

Table 3.3 shows the coupled analysis of delay and electromigration MTF for various technology nodes. Column 2 shows the maximum current density specified in ITRS and column 3 shows the resultant temperature of the top global wires. The $R_{met}C_{total}$ delay is shown in column 4 and the corresponding MTF is shown in Fig. 3.10 as the solid curve. The achievable MTF is about 90% at 130nm node, but drops sharply for the following technology nodes. The MTF(T_m) is compared to the MTF(105°C). To confine the wire temperature at 105°C, the current density has to be reduced as shown in column

5 and resultant current drive, compared to the maximum current density specified in ITRS, drops drastically as can be observed in column 6. To relax the temperature rise to 10°C above substrate, the current drive can be improved as shown in column 7. We notice that in both cases trying to match certain specific wire temperatures (105°C and 110°C), the corresponding MTF rises above the required MTF, as shown in Fig.3.10. It implies that this approach is overly conservative. Since the MTF of electromigration is also inversely proportional to the square of current density, as modeled by the well-known Black's equation, we can achieve the expected MTF by optimizing both the current density and wire temperature, which are related by Joule heating. The result is shown in the last three columns of the table. The optimal current density is much closer to the value specified in ITRS. On the other hand, although the wire temperature is higher than 105°C, as shown in the last column, the MTF is on target, as shown in the flat line in Fig.3.10.

Tech. Node	Max. J_{rms} (ITRS) [MA/cm^2]	T_m [°C]	$R_{wire}C_{wire}(T_m)$ normalized to $R_{wire}C_{wire}(T_m=105°C)$	J_{rms} $T_m=105°C$ [MA/cm^2]	Current drivability normalized to Max. J_{rms}	J_{rms} $T_m=110°C$ [MA/cm^2]	Current drivability normalized to Max. J_{rms}	$J_{rms,optimal}$ MTF=MTF($T_m=105°C$) [MA/cm^2]	Current drivability normalized to Max. J_{rms}	T_m [°C]
130	0.96	107	1.01	0.79	0.82	1.12	1.16	0.91	0.95	106
90	1.5	117	1.05	0.82	0.55	1.16	0.78	1.24	0.83	111
65	2.1	122	1.09	1.0	0.48	1.41	0.67	1.64	0.78	113
45	2.7	126	1.09	1.2	0.44	1.69	0.63	2.04	0.76	114
22	3.9	128	1.10	1.6	0.41	2.33	0.60	2.84	0.73	116

Table 3.3: Coupled evaluation of electromigration reliability and performance for global interconnects for 22-130 nm technology nodes. T_m is the top global wire temperature with all the heat sources, including substrate and Joule heating from all metal levels underneath, taken into account.

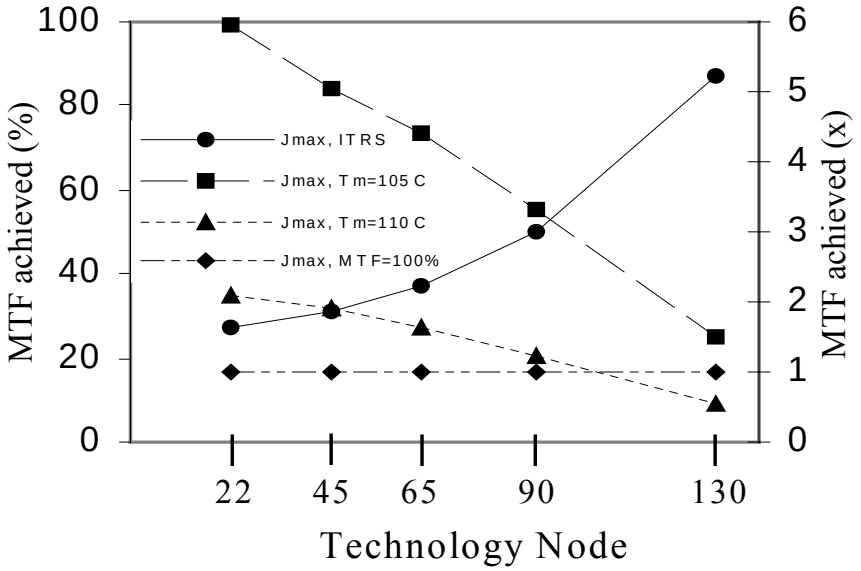


Fig. 3.10: The solid curve shows the MTF (%) that can be achieved under the current density specified in ITRS. The broken-line curves show the MTF under various wire temperature (*Tm*) criterions and the values refer to the right axis.

3.7 Summary

In conclusion, a detailed analysis of the impact of Joule heating on the characteristics of future Cu/low-k interconnects is presented using a realistic full chip model with via effect included. Thermal effects can severely degrade both reliability and speed performance. Optimization with various interconnect parameters is provided. Joule heating will limit scaling of current density and use of low-k materials. Global wires will be more problematic with higher operating temperature and careful consideration is imperative.

Chapter 4

SPICE-Based Electro-Thermal Simulation Methodology

4.1 Motivation

In Chapter 2, we analyzed the thermal effects in interconnects using analytical thermal models and demonstrated the strong influence of vias on the temperature distribution in metal lines. Nevertheless, without solving complex differential equations, the electro-thermal simulation approach is desirable to facilitate quick estimation of temperature rises and to investigate thermal coupling effects between wires in order to provide thermal design guidelines for advanced interconnect structures [56, 57].

Teng [58] has used lumped thermal circuit model to predict interconnect temperature. However, without including thermal capacitance, the model is not capable of transient analysis. This chapter presents a fast SPICE based 3-D electro-thermal

simulation methodology to characterize thermal effects due to Joule heating in high performance Cu/low-k interconnects under steady-state and transient stress conditions. The results demonstrate excellent agreement with experimental data and those using Finite Element (FE) thermal simulations (ANSYS). The effect of vias, as efficient heat sinking paths to alleviate the temperature rise in the metal wires, is included in our analysis to provide more accurate and realistic thermal diagnosis. The simulation methodology has also been applied to quantify the use of *dummy thermal vias* as additional heat transfer paths to lower the temperature rise in the metal wires for the first time. Furthermore, thermal coupling between wires is evaluated and found to be significant. Finally, the impact of metal wire aspect ratio on interconnect thermal characteristics is discussed. All the dimensions of the interconnect structure in this paper were taken from the 100 nm technology node based on the ITRS [59]. The remainder of this chapter is organized as follows. To serve as a starting point, in Section 4.2, we give an overview of our thermal simulation methodology. Sections 4.3 & 4.4 describe the impact of via effect of interconnect thermal characteristics under steady-state and transient conditions, respectively. Finally, we summarize and conclude in Section 4.5.

4.2 SPICE-Based Thermal Modeling

In order to provide robust thermal analysis for deep sub-micron Cu/low-k interconnects, it is very desirable to have an efficient 3-D simulation methodology to

estimate the temperature profiles in the metal wires and evaluate the thermal coupling between them. Based on the thermal-electrical analogy (Fig. 4.1), a 3-D RC distributed thermal circuit model has been developed, as shown in Fig. 4.2. This thermal network can be easily implemented and simulated using SPICE in the same manner as an electrical circuit network by simply employing the proper counterparts as illustrated in Fig. 4.1. For example, the thermal resistance along the metal wire per unit length, R_m, and the thermal resistance of the insulator, R_i, per segment can be computed as,

$$R_m = \frac{\Delta \ell}{k_M w H} \tag{4.1}$$

Thermal		Electrical
Temperature T [K]	$\Leftrightarrow$	Voltage V [V]
Heat Q [J]	$\Leftrightarrow$	Charge Q [C]
Heat transfer rate q [W]	$\Leftrightarrow$	Current i [A]
Thermal resistance R_T [K/W]	$\Leftrightarrow$	Electrical resistance R [V/A]
Thermal capacitance C_T [J/K]	$\Leftrightarrow$	Electrical capacitance C [C/V]

Governing equations

Steady-State condition

Temperature Rise		Voltage Difference
$\Delta T = q R_T$	$\Leftrightarrow$	$\Delta V = i R$

Transient condition

Heat diffusion		RC transmission line
$\nabla^2 T = R_T C_T \frac{\partial T}{\partial t}$	$\Leftrightarrow$	$\nabla^2 V = RC \frac{\partial V}{\partial t}$

Fig. 4.1: Thermal-Electrical analogous parameters.

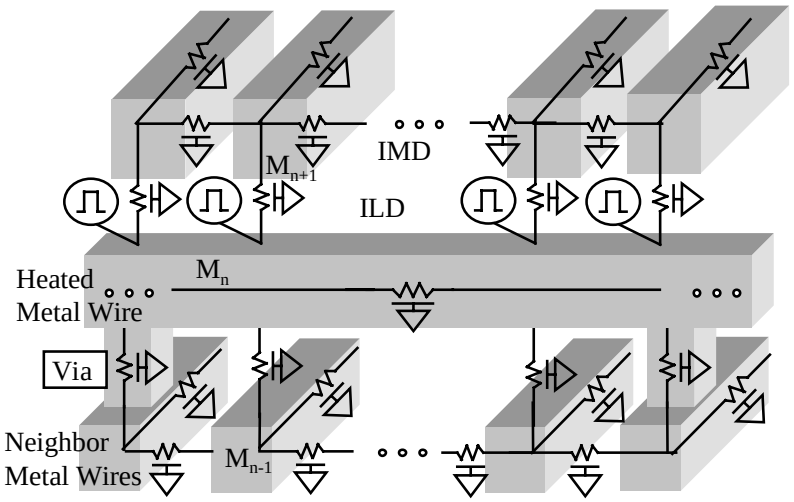


Fig. 4.2: 3-D thermal circuit RC transmission line model for transient thermal analysis of interconnect structures.

and
$$R_i = \frac{t_I}{k_I \Delta\ell \Delta w} \tag{4.2}$$

where $\Delta\ell$ is the unit length of metal wire, w is the width of metal wire, H is metal height, t_I is the thickness of insulator, and Δw is the width of the chosen insulator segment. k_M and k_I are the thermal conductivity of metal wire and insulator, respectively. Furthermore, the thermal capacitance along the metal wire per unit length, C_m, and the thermal capacitance of the insulator per segment, C_i, can be shown as,

$$C_m = c_{p,M} \rho_M (wH\Delta\ell) \tag{4.3}$$

and
$$C_I = c_{p,I} \rho_I (t_I \Delta\ell \Delta w) \tag{4.4}$$

where $c_{p,M,}$ ρ_M and $c_{p,I,}$ ρ_I are the specific heat and density of metal wire and insulator, respectively. Inclusion of the lateral thermal impedance in the model captures the heat spreading effect and thermal coupling from nearby interconnects. Therefore, there is no need to add any data-fitting modification in the circuit model as in [58]. In addition, RC transmission lines are used to model the heat diffusion due to the similarity of the governing equations as illustrated in Fig. 4.1 and transient thermal effects in interconnects can, thus, also be conveniently analyzed. To account for the temperature dependence of the metal resistivity, the heat generation (q) in each segment of the wire has been modeled as a voltage (temperature) controlled current source,

$$q = (j_{rms} wH)^2 (\rho(T) \frac{\Delta \ell}{wH}) \tag{4.5}$$

where j_{rms} is the root-mean-square current density flowing through the wire and $\rho(T)(=\rho_0(1+\beta T))$ is the temperature dependent metal resistivity. ρ_0 is the metal resistivity at 0°C and β is the temperature coefficient of the resistivity (TCR). The main advantage of this SPICE-based methodology is that once layout data is available, the thermal analysis can be quickly done. HSPICE was used for the 3-D simulations of the distributed thermal RC circuits in this work. This technique was validated by comparing with both experimental data [60] and simulation results carried out by ANSYS, a finite element (FE) simulation package. It will be shown in the following sections that it exhibits excellent accuracy within 5% agreement.

4.3 Steady-State Analysis

4.3.1 Impact of Via Separation on Effective k_{ILD}

In the steady-state analysis, thermal capacitance can be removed from our RC thermal circuit model for the obvious reason. The resultant simplified model is shown in Fig. 4.3. The temperature profiles along the embedded Cu wires with 100μm via separation for two different dielectrics are shown in Fig. 4.4. Simulated temperature profile along Cu interconnect with current density J=3.7x10^6 A/cm^2 and via separation of 100 μm. Current is flowing in alternate wires, i.e. current is not flowing in the nearest, third

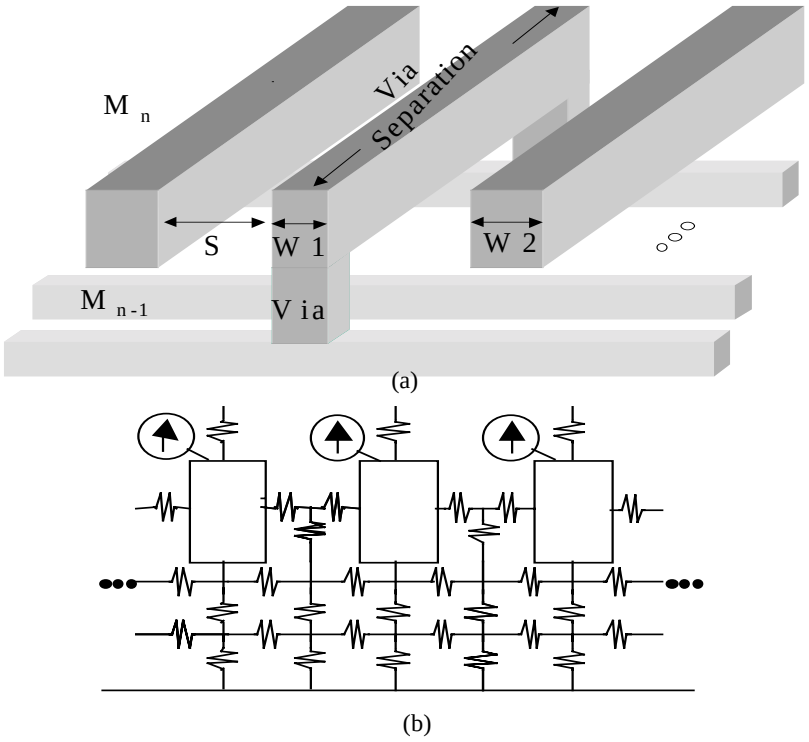


Fig. 4.3: (a) Interconnect configuration. (b) Correspondent distributed thermal circuitry.

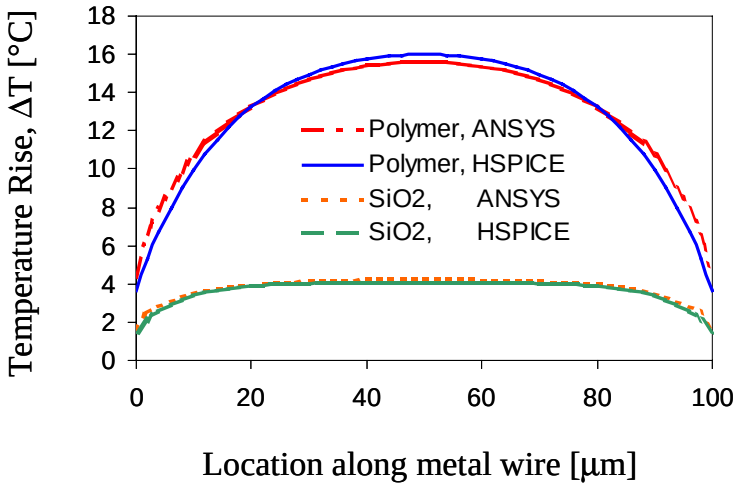


Fig. 4.4: Validation of HSPICE thermal simulation with Finite Element Analysis (ANSYS) data.

nearest, fifth nearest and so on neighboring wires but current is flowing in the second nearest, forth nearest and so on neighboring wires. Geometries used in the simulations are metal height H=0.8 μm, wire width (w)=wire spacing (S)=0.3μm, and ILD thickness t_{ILD}=0.8 μm. The interconnect geometries used in the simulations are stated in the caption and the materials properties are listed in Table 4.1. The diameter of the vias is assumed to be the same as the width of the wire for the damascene process. Since the

Parameters	Conductor		Dielectrics		
	Cu	Al	Oxide	Polymer	Air
k [W/m-K]	400	220	1.2	0.3	0.03
ρ [kg/m3]	8933	2720	2220	1380	0.87
Cp [J/kg-K]	385	900	745	1195	1014
ρ [μΩ-cm] @ 0°C	1.67	2.66	-	-	-
TCR × 103/K	4.3	4.5	-	-	-
εr	-	-	4.0	2.5	1.0

Table 4.1: Materials properties used in this work

substrate, to which a heat sink is usually attached, is assumed to be the sole heat dissipation path to the outside ambient, only heat conduction downwards is considered in this interconnect thermal modeling. Heat convection to the ambient air is ignored by the application of adiabatic boundary condition on the four side walls and top of the chip. This is a reasonable assumption because, in general, the chip is enclosed by thermally insulated package materials. Owing to the lack of available experimental data for such small spatial resolution under steady-state current, the results from SPICE simulations are compared against with those using finite element thermal simulation (ANSYS), and it demonstrates excellent agreement (Fig. 4.4). However, because of the wide familiarity of SPICE to the circuit design community, this approach, as compared to traditional finite element simulation, can be adopted more easily and it requires much less effort to set up as well as less CPU time to run. It can be observed in Fig. 4.4 that the temperature profiles along the wire for dielectrics with lower thermal conductivity ($k_{polymer}$ = 0.3 W/m-K compared to k_{SiO2} =1.2 W/m-K) is more curved. This is due to the larger ratio between the thermal conductivity of Cu and the dielectric and hence a stronger influence from vias is expected. The distance over which via effect is important can be roughly estimated by thermal healing length $L_{H,}$ ($=k_M t_{ILD} H / k_{ILD} s)^{1/2}$ [45], which is 14 μm and 30 μm for SiO_2 and polymer respectively, where k_M, is the metal thermal conductivity, H is the metal wire thickness, separated from the underlying layer by ILD of thickness t_{ILD} and thermal conductivity k_{ILD} and s is the heat spreading factor to accommodate the deviation from 1-D heat flow between a metal wire and the underlying layer. It can be thought that within

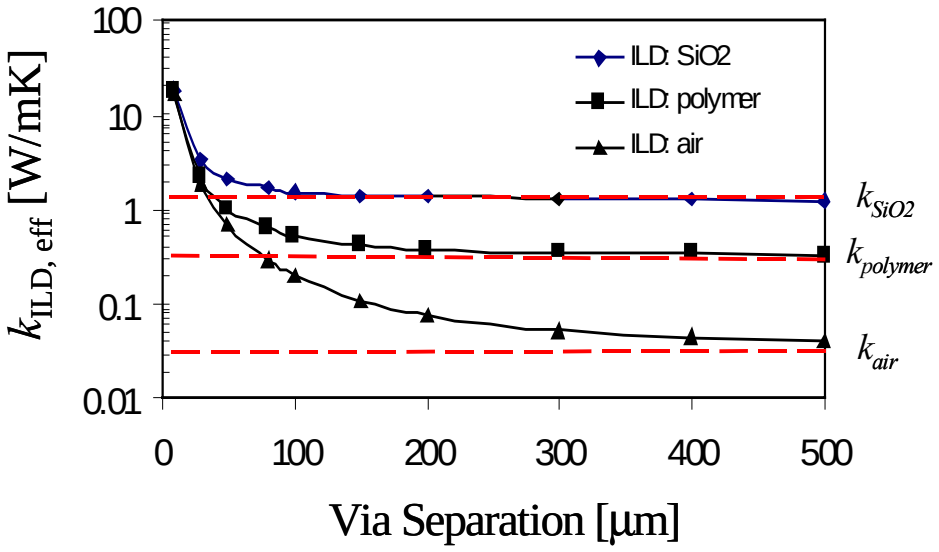


Fig. 4.5: Effective ILD thermal conductivity increases with decreasing via separation. The lower the nominal k_{ILD}, the longer the L_H, and hence, the stronger is the via effect.

the range of L_H from vias, heat generated will flow through the vias to the underlying layer. Beyond L_H, heat will flow through the ILD and the via effect is diminished. It can be observed that the lower the nominal $k_{ILD,}$ the longer the L_H and the stronger the via effect. Since the via separation for the local interconnects is generally much shorter than L_H, via effect can not be ignored in local interconnects. Even at the global level, large portions of the wires are still under the influence of vias. Therefore, the effective thermal conductivity of low-k dielectrics is considerably higher than the nominal values with the help of vias, as shown in Fig. 4.5, where the $k_{ILD,eff}$ is plotted against via separation for three different ILD materials. Knowing the temperature profile (e.g. Fig. 4.4), the $k_{ILD,eff}$ is obtained by taking the average temperature of the wire and adjusting from the nominal

k_{ILD} accordingly. It can be observed that incorporation of via effect results in increased $k_{ILD,eff}$ especially for ILD materials with lower nominal thermal conductivity. This fact can explain why the interconnect temperature is not as high as commonly assumed when low-k ILD is implemented in advanced interconnect structures and, thus, via effect has to be included for accurate interconnect temperature estimation. As concluded from Fig. 4.4 and 4.5, significant difference in temperature profiles along the wires and in the maximum temperature rise can arise between the realistic situation of heat dissipation in the presence of vias and the overly simplified case that ignores via effect. In addition, it should be noted that the steeper temperature profile for the low-k dielectrics can exacerbate electromigration due to the larger temperature gradient.

4.3.2 Thermal Coupling Effects

Thermal coupling between wires in the same layer and different layers has become an important issue due to the scaling of wire spacing. It is very difficult to solve the coupling effect analytically. However, our SPICE-based electrothermal methodology provides a very convenient solution. To investigate the parallel thermal coupling effect between wires, the configuration shown in Fig. 4.3(a) is examined and symmetrical boundary conditions applied. The interconnect structure used here is the same as that used for Fig. 4.4, but with wire width fixed at 0.3 μm, and wire spacing varied from 0.1 μm to 2 μm. Furthermore, polymer is used as the insulation material. In Fig. 4.6(a), the middle wire is carrying a current density of J_0(=1.4x10^6 A/cm^2), which is the maximum current density specified for the 100 nm technology node from ITRS [58]. The tempera-

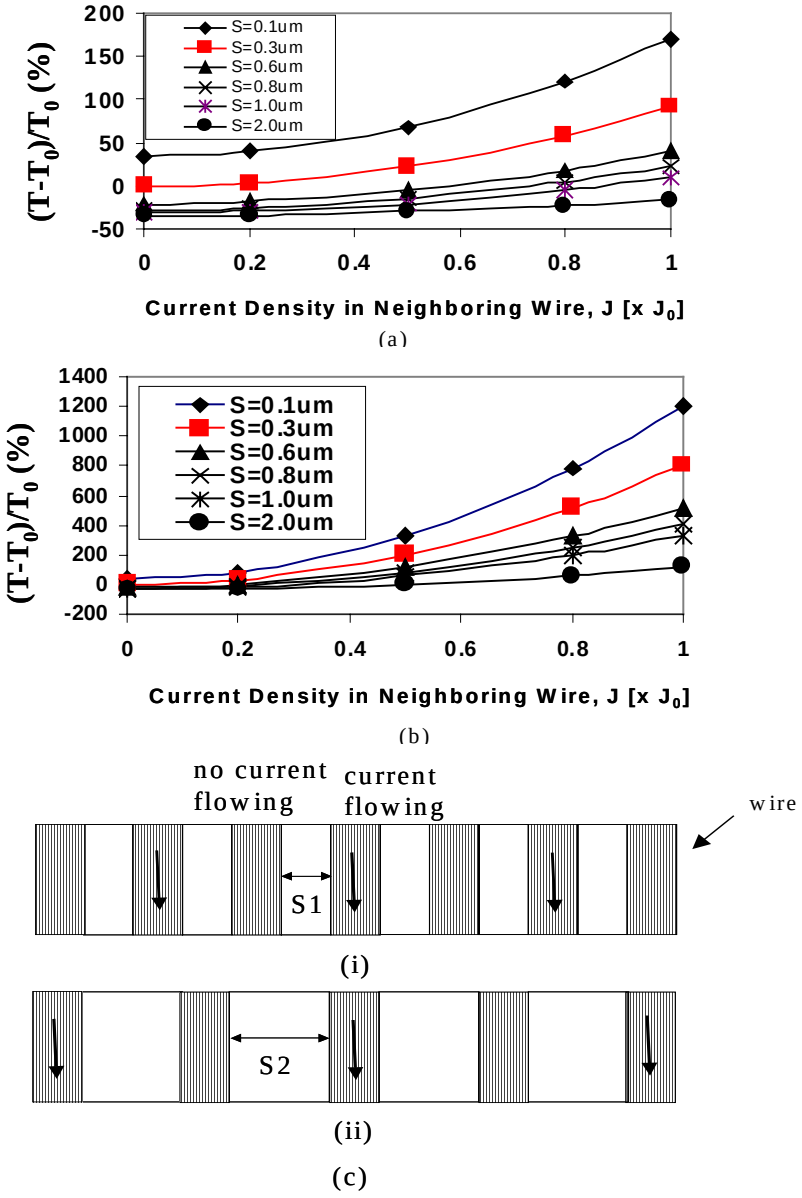


Fig. 4.6: (a) T is defined as the maximum temperature in the middle wire. Reference temperature T_0 is the maximum temperature in the middle wire with no current flowing in the nearest neighboring wires and with equal line width and spacing (S). The current density in the middle wire is J_0, 1.4×10^6 A/cm^2. Line width is kept at 0.3µm in this simulation. (b) The current density in the middle wire is one third of J_0. Line width is 0.3µm. The temperature of wires carrying lower current density is strongly affected by the heat coupled from the neighboring wires. (c) Top view of interconnect structures, with different spacing, used in this simulation and current flow patterns showing current flowing on alternate wires.

ture variation due to thermal coupling is normalized to T_0, which is the maximum temperature in the middle wire when there is no current flowing in the nearest neighboring wires and with equal wire width and wire spacing, S (=0.3μm in this simulation). It can be seen that both the current density of the neighboring wire and the line spacing greatly affect the temperature of the middle wire. In the case of equal line width and spacing (0.3μm), the temperature of the middle wire can be increased by nearly 100%. The negative values are due to the symmetric boundary condition chosen in this simulation and the temperature variation is normalized to $T_{0,}$ which is with S equal to w. Fig. 4.6(c) depicts the top view of two current patterns and can be used to understand the results. With S1 < S2, more current is flowing in Fig. 4.6(c)(i) for a given area, thus causing higher thermal coupling effect to the middle wire. The coupling effect can be even more drastic when the wire is carrying a low current density, which represents a typical signal line. In this case, only one third of J_0 is flowing in the middle wire and the thermal coupling effect is shown in Fig. 4.6(b). The strong thermal coupling may cause resistance variations and further enhance crosstalk problem.

4.4. Transient Stress Analysis

Apart from normal circuit operating conditions, ICs also experience high-current stress events, the most crucial of them being electrostatic discharge (ESD), which causes accelerates thermal failures [61]. Semiconductor industry survey has indicated that ESD

is one of the major causes of the failure of ICs [62]. Furthermore, I/O circuitry also exposes to high current events since it interfaces between multiple power supplies [63]. Therefore, it is important to be able to analyze interconnects heating under transient stress conditions.

4.4.1 Analytical Model vs. SPICE-Based Simulation Methodology

In general, a high-current short-pulse (J > 10MA/cm^2, and pulse width t_{pulse} < 200 ns) usually causes much higher ΔT_{max}, the maximum temperature rise on the wire, due to more severe self-heating than under normal operating conditions and the heat diffusion is limited to the immediate materials in contact with the metal line. Again, only heat conduction is considered in this analysis. The governing heat diffusion equation of temperature rise during brief transients can be written down based on energy conservation law as [60, 64],

$$\frac{\partial}{\partial x}(k_M \frac{\partial T}{\partial x}) + \frac{\partial}{\partial y}(k_{ILD} \frac{\partial T}{\partial y}) + \frac{\partial}{\partial z}(k_{IMD} \frac{\partial T}{\partial z}) + \frac{i_{rms}^2 \rho(T)}{wH} = C_{eff} \frac{\partial T}{\partial t} \qquad (4.6)$$

where

$$C_{eff} \cong [c_M \rho_M wH + 2c_{ILD} \rho_{ILD} s_V w\sqrt{\alpha_{ILD} t} + 2c_{IMD} \rho_{IMD} s_l H \sqrt{\alpha_{IMD} t}]^{-1} \qquad (4.7)$$

and c_M, c_{ILD}, c_{IMD} and ρ_M, ρ_{ILD}, ρ_{IMD} are the specific heat and density of metal, ILD and IMD materials respectively. S_v and S_l are shape factors in vertical and lateral directions to account for three-dimensional heat spreading. α_{ILD} and α_{IMD} are the thermal diffusion coefficients of ILD and IMD materials respectively. C_{eff} can be understood as the effective thermal capacitance per unit length to absorb the heat. Due to the nature of heat

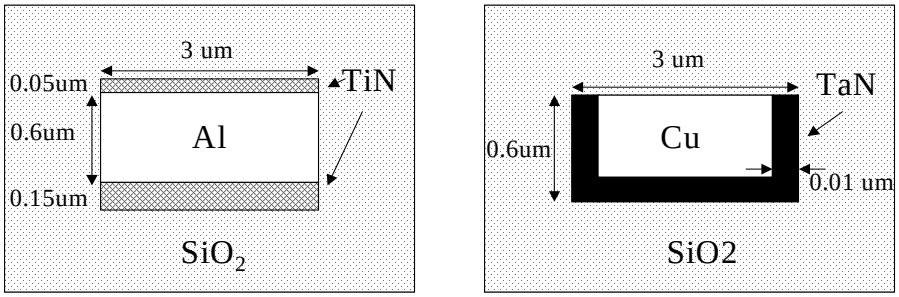
diffusion process, C_{eff} is increasing with time as heat diffuses through larger volume. Therefore, temperature rise on the wire under current pulse i_{rms} with duration t_{pulse} can be shown as,

$$\Delta T = \frac{E}{C_{eff}} \tag{4.8}$$

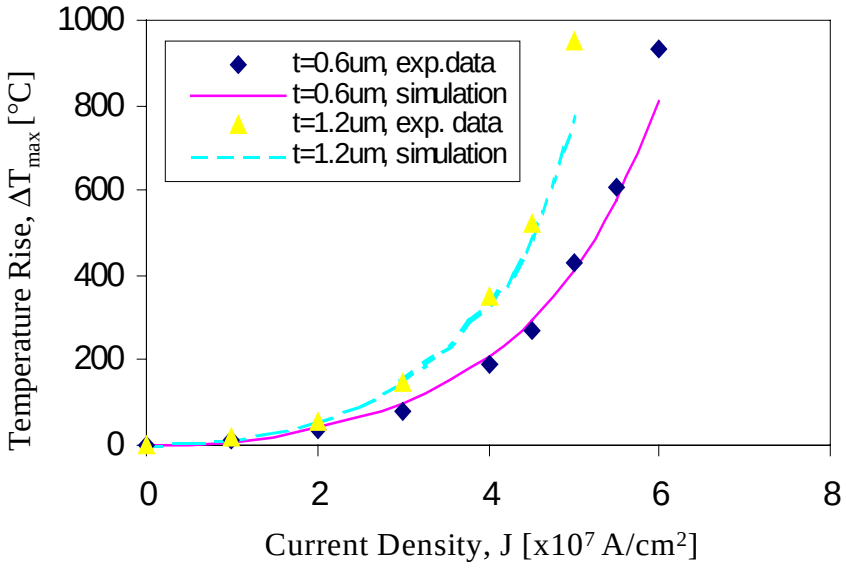
$$\cong \frac{i_{rms}^2 \rho(T) t_{pulse}}{wH} [c_M \rho_M wH + 2c_{ILD}\rho_{ILD} s_V w\sqrt{\alpha_{ILD} t} + 2c_{IMD}\rho_{IMD} s_l H\sqrt{\alpha_{IMD} t}]^{-1} \tag{4.9}$$

where E is the pulse energy. It should be noted that, unlike in the case of normal steady-state operation condition, the temperature dependence of $\rho(T)$ absolutely can not be ignored because of the large temperature rise. Together with the time-dependent C_{eff}, there is no easy analytical solution available for Eq. 4.6. Therefore, an efficient simulation methodology is desirable and will be discussed in details in the following. In contrast with the unsolvable nature of the analytical analysis, SPICE-based simulation methodology provides a convenient and efficient transient thermal analysis. RC transmission lines are used to model the heat diffusion (Fig. 4.2) due to the similarity of the respective governing equations as illustrated in Fig. 4.1. The distributed thermal resistance and thermal capacitance components are constructed, using Eq 4.1-Eq. 4.4, to form thermal RC network. Heat generation, calculated by Eq. 4.5, is realized as voltage (temperature) controlled current sources as shown in Fig. 4.2. After substituting these thermal quantities into the correspondent electrical identities in the thermal circuit model, the temperature along the interconnect, manifested as node voltage, can be easily

obtained through SPICE simulation. This technique was validated by comparing with experimental data from [60] as shown in Fig. 4.7, where it shows excellent agreement.



(a)



(b)

Fig. 4.7: (a) Schematic cross section of Al and Cu interconnects with cladding layers used in model validation Fig. 4.7(b) and for simulations in Fig. 4.8.
(b) Validation of HSPICE thermal simulation with experimental data from [60]. Al interconnect test structure shown in Fig. 4.7(a) is used. Two different wire thickness , *t*, are under examination.

4.4.2. Al vs. Cu Interconnects

With the interconnect structure of Al and Cu wires described in Fig. 4.7(a), temperature rise is simulated and the comparison is shown in Fig. 4.8. It should be noted that the simulation is done for a 3-D parallel metal array, but only the wire under examination has a current pulse flowing through it because event like ESD usually strikes very few wires at a time. The materials properties again are listed in Table 4.1. As can be observed from Fig. 4.8, for the same cross section, current density, and surrounding dielectric, Cu wires would experience lower ΔT_{max} than Al wires, due to their higher

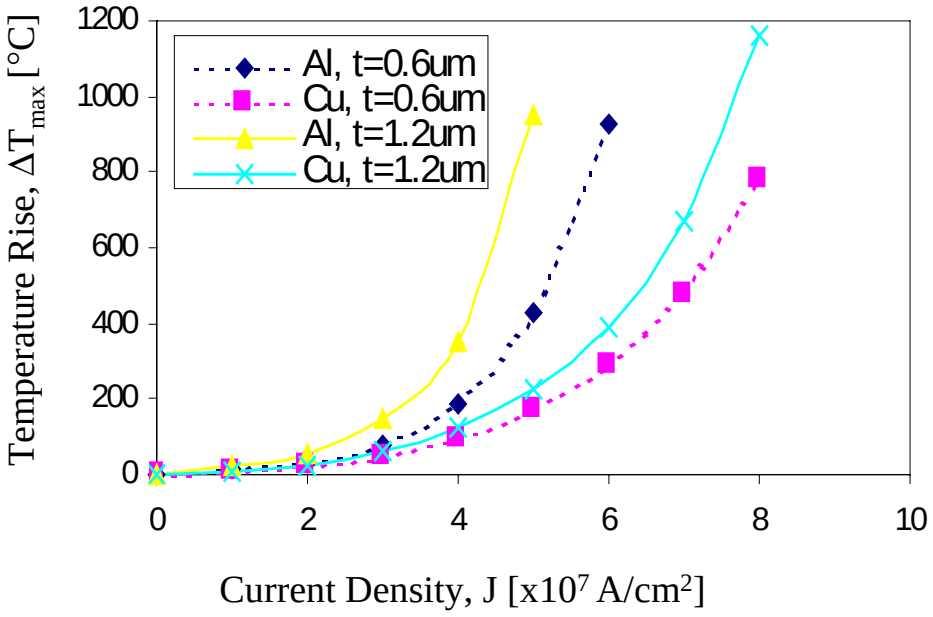


Fig. 4.8: For the same wire cross sections, shown in Fig. 4.7(a), Cu wire shows much lower temperature rise compared to Al wire under a 200 ns pulse stress. Symbols represent simulation data.

thermal conductivity and thermal capacity, and most importantly, due to their lower resistivity. This along with higher melting point (~1100 ^{0}C) than that of Al (~660 ^{0}C) would provide more thermal margin to Cu interconnects. However, ΔT_{max} can still be high when metal dimensions are scaled down and low-k dielectrics are incorporated. Therefore, it is prudent to study the effect of vias on the thermal characteristics of Cu wires to make reliable use of them in deep sub-micron designs.

4.4.3 Impact of Via Separation and Low-k Dielectrics

Before continuing our discussion, we would like to remind the readers again that, if it is not otherwise mentioned, the interconnect dimensions used throughout the following sections are taken from global interconnect levels from the 100 nm technology node based on the ITRS [58] with metal height H=0.8 μm, wire width w=wire spacing S=0.3μm, and ILD thickness t_{ILD}=0.8 μm. If a local interconnect is simulated, then the dimensions are metal height H=0.26 μm, wire width w=wire spacing S=0.13μm, and ILD thickness t_{ILD}=0.32 μm.

Fig. 4.9 compares the normalized spatial temperature distribution along an interconnect line with polymer as insulation material and a via separation of 100 μm, subjected to transient current pulses of 200 ns and 2 μs duration. It can be observed that the temperature rise profile for the 2 μs pulse is more gradual due to the increased influence of vias for longer diffusion time, which results in longer diffusion lengths. The heat diffusion length, $L_D \propto (\alpha t)^{1/2}$, is 5 μm and 16 μm for the 200 ns and 2 μs pulse durations respectively, where α is the thermal diffusivity of the interconnect materials. Thermal diffusion length L_D can be interpreted as the distance over which via effect is prominent, which is a function of time and should not be confused with the thermal healing length L_H under steady-state condition as mentioned in Chapter 2 and Section 4.3.

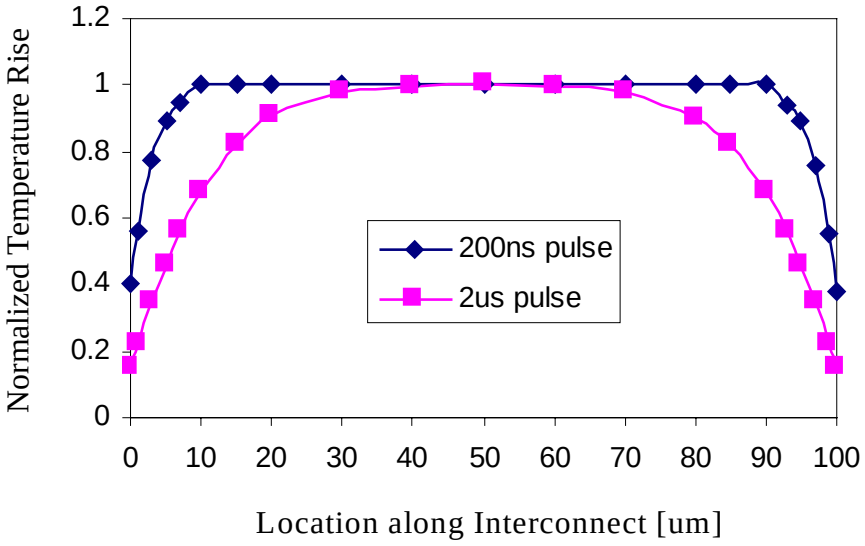


Fig. 4.9: Simulated normalized temperature ($\Delta T/\Delta T_{max}$) profile along global Cu/low-k (polymer) interconnect for two pulse durations with current density J=4x10^{7}A/cm^2.

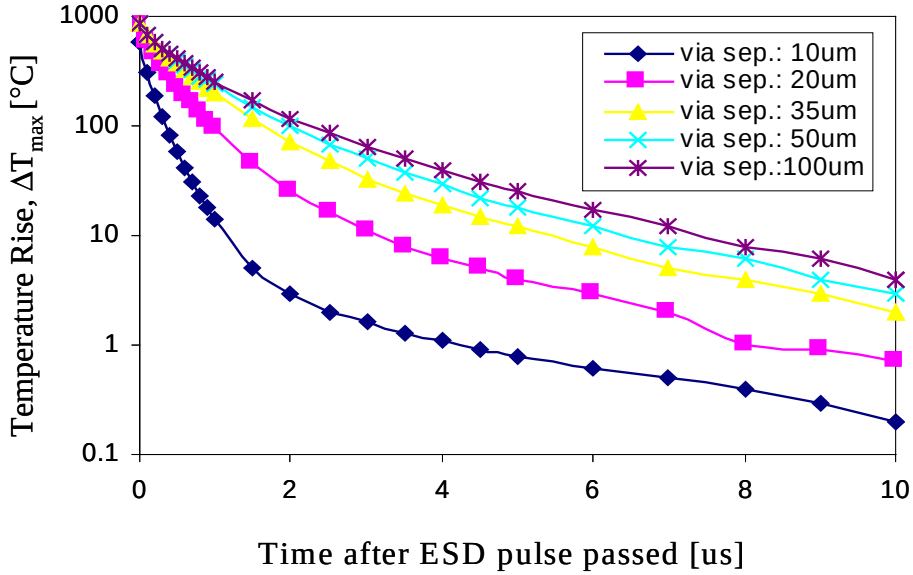
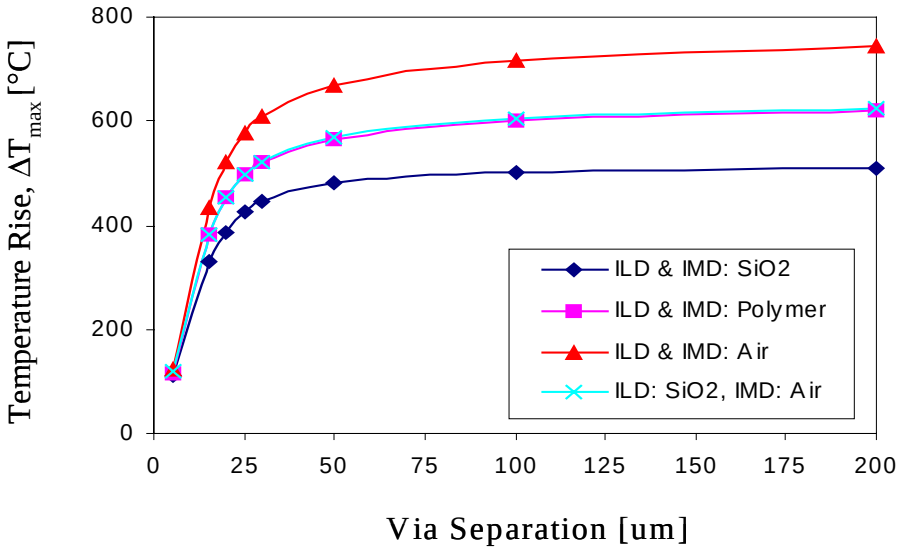


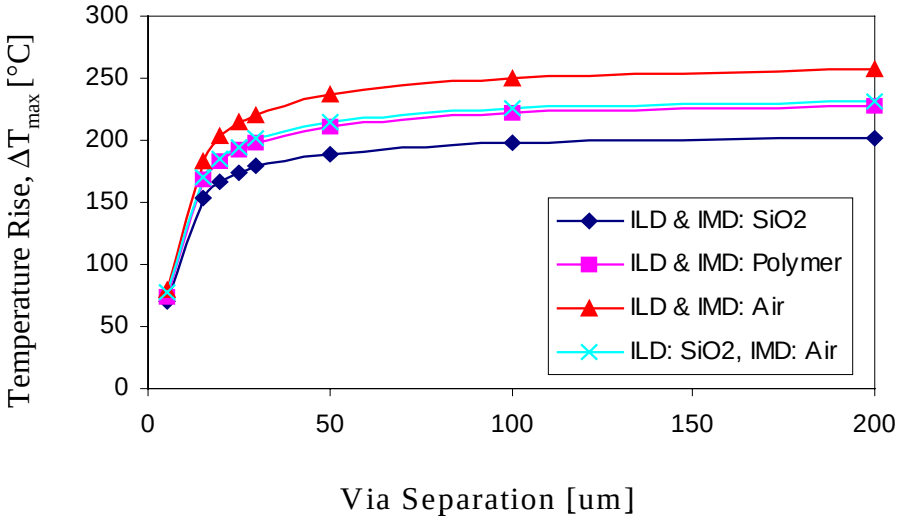
Fig. 4.10: The temperature decay after a 200 ns current pulse with J=8x10^7 A/cm^2, for Cu/low-k (polymer) global wires. The decay is facilitated by the presence of vias.

L_D is a function of time under transient condition and L_H is only a function of materials properties and is not varying with time. Fig. 4.10 plots the temperature decay after a 200 ns pulse. It can be observed that the temperature decays more rapidly with vias placed closer. This shortens the high temperature span that the interconnect would experience and thus reduces thermal problems. ΔT_{max} vs. via separation for different dielectrics is shown for high current pulse duration of 200 ns in Fig. 4.11(a) and 100 ns in Fig. 4.11(b). ΔT_{max} is higher for the 200 ns pulse due to higher pulse energy, which results in greater Joule heating. However, both Fig. 4.11(a) and Fig. 4.11(b) indicate that ΔT_{max} would be reduced dramatically for smaller via separation. For larger via separation ΔT_{max} saturates since the effect of vias diminishes. This is dictated by thermal diffusion length (L_D), which can be interpreted as the distance over which heat generated in the wire flows through the via. It can also be observed in Fig. 4.11(b) that for shorter pulse duration the differences between ΔT_{max} for various dielectrics is much smaller. This is due to the fact that heat does not have sufficient time to diffuse through the surrounding dielectrics.

The maximum temperature rise, at the end of a 200 ns pulse, ΔT_{max}, is shown for global (Fig. 12(a)) and local interconnects (Fig. 12(b)) for different dielectrics as a function of the current density. It can be observed that while the temperature rises sharply and low-k dielectrics show worse situation for a typical global line with via separation of 100 μm, the temperature rise is significantly alleviated for local interconnects due to a smaller via separation of 1 μm. Even if air is used as both the ILD and the IMD dielectric (worst case thermal scenario), no significantly higher ΔT_{max} is observed for the local interconnects. This observation is further validated in Fig. 13 with 1 μm via separation, which shows that the spatial temperature distributions for various dielectrics are similar and ΔT_{max} is nearly independent of the dielectric material. This suggests that they are all within L_D and via effect dominates the thermal characteristics.

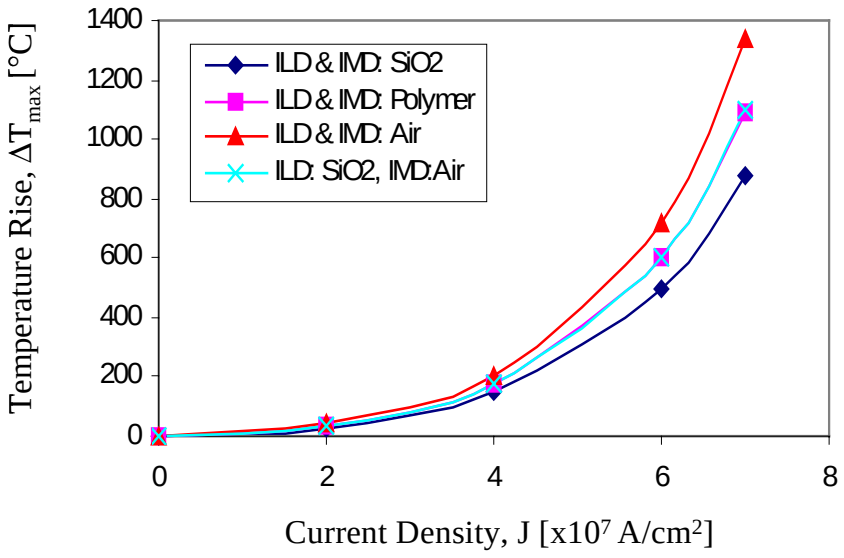


(a)

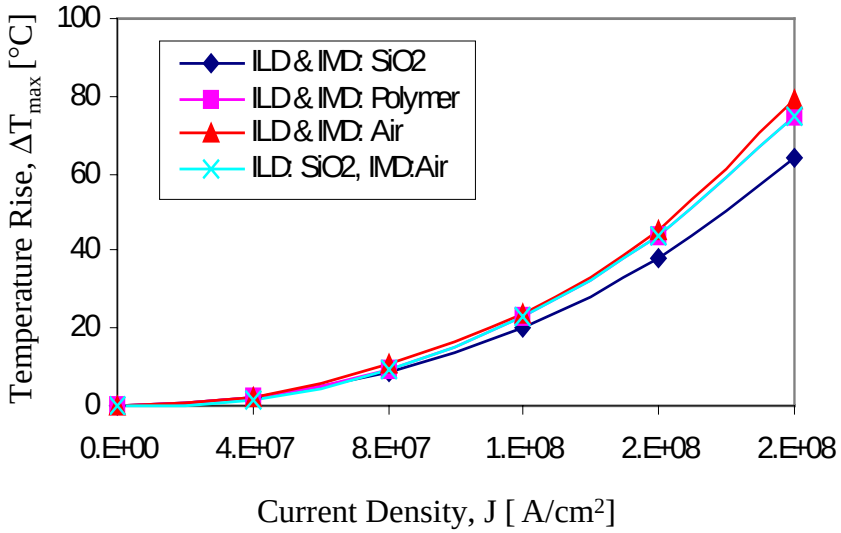


(b)

Fig. 4.11: (a) ΔT_{max} of global interconnects for a 200 ns current pulse with J=6x10^7 A/cm^2. ΔT_{max} increases with via separation and saturates when via effect has diminished. (b) For a short pulse of 100 ns, with the same current density as in Fig. 4.11(a), ΔT_{max} of the global Cu interconnect is much less dependent on the dielectric materials.



(a)



(b)

Fig. 4.12: (a) ΔT_{max} of Cu global interconnect with 100 μm via separation under a 200 ns current pulse. Temperature rises sharply with current density. (b) ΔT_{max} of Cu local interconnect with 1 μm via separation under a 200 ns current pulse. Temperature rises are much lower and nearly independent of the surrounding dielectric materials.

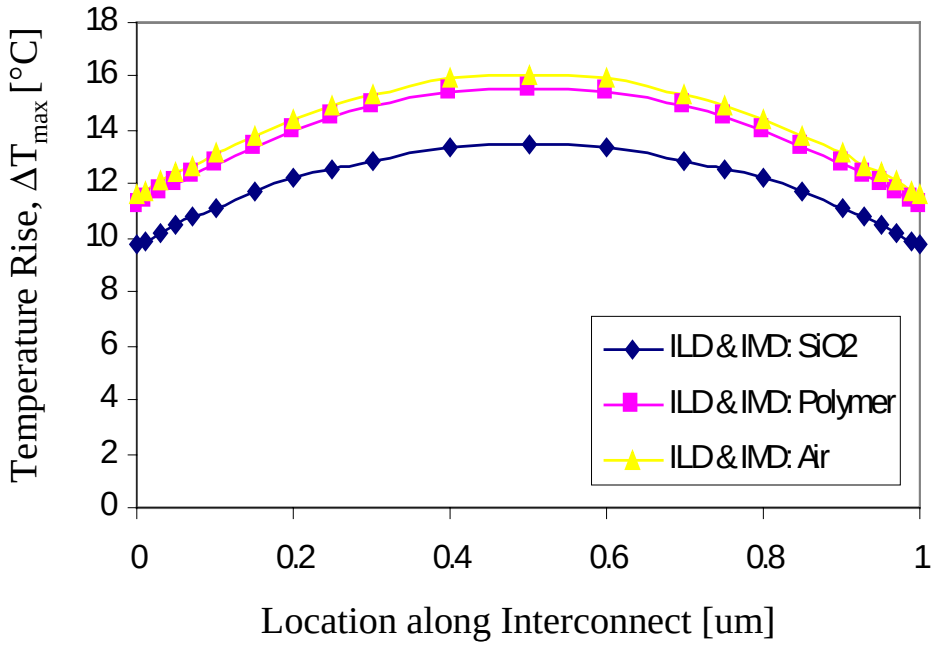
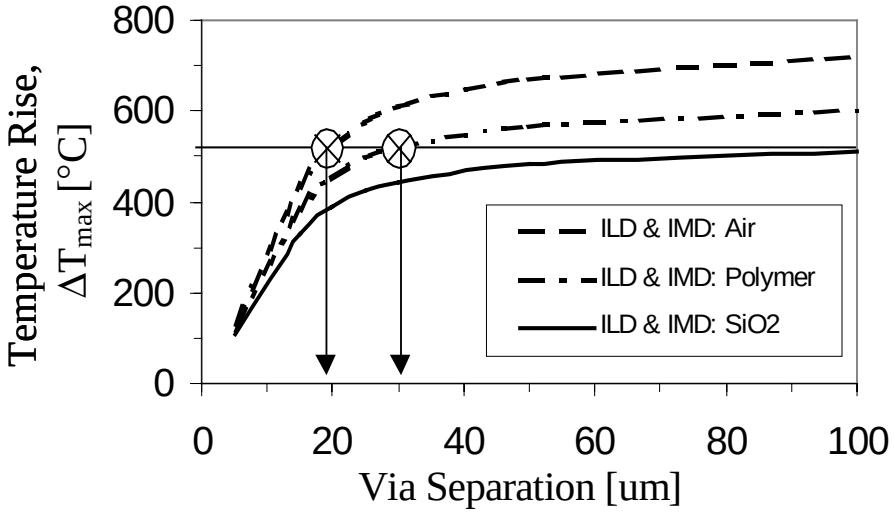


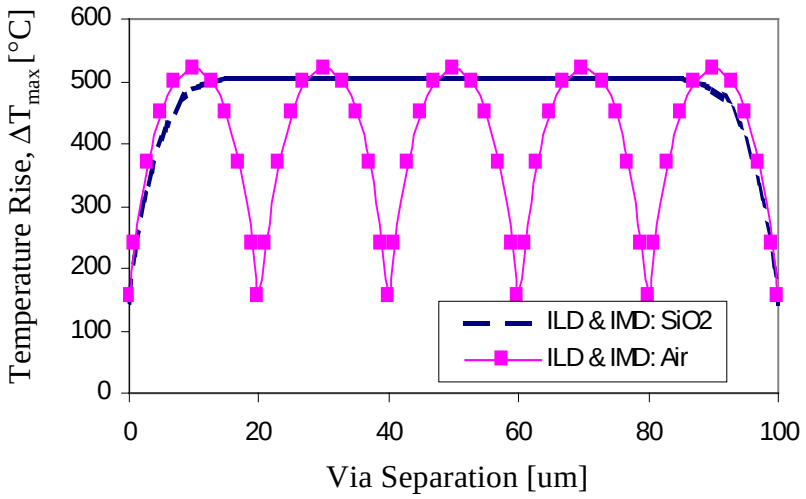
Fig. 4.13: Temperature profiles along local Cu interconnect under a 200 ns pulse are gradual for all the dielectric materials and ΔT_{max} are about the same with J=1x10^8 A/cm^2.

4.4.4 Impact of Dummy Thermal Vias

Dummy thermal vias, which conduct heat but are electrically isolated, can be installed in Cu/low-k structure to lower the temperature rise. Since a high-current short-pulse (J > 10MA/cm^2, and t_{pulse} < 200 ns) usually causes much higher ΔT_{max} due to more severe self-heating than under normal operating conditions and the heat diffusion is limited to the immediate materials in contact with the metal line, the transient condition will define the most stringent via separation requirement for lowering the temperature of the metal wires. It should be noted that the simulation methodology developed here is quite general, and can be easily extended to study steady-state stress conditions by using longer pulses. The advantage of the dummy via effect can be demonstrated by using the



(a)



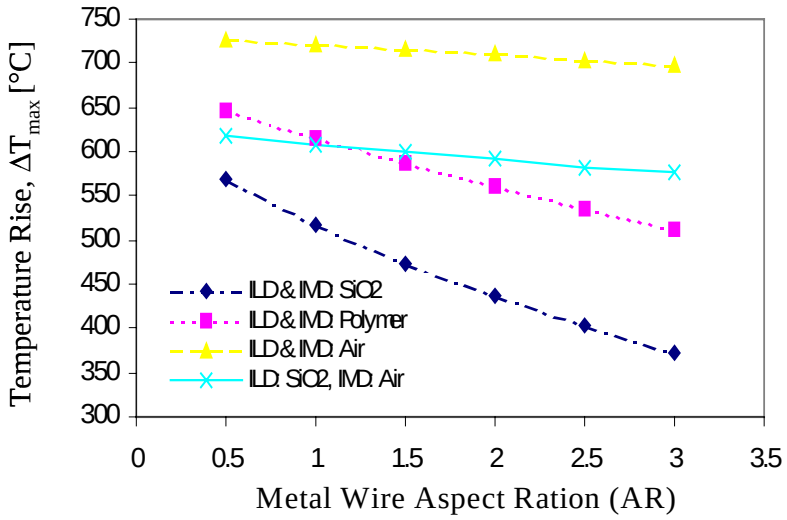
(b)

Fig. 4.14: (a) ΔT_{max} of Cu/air and Cu/polymer can match ΔT_{max} of Cu/oxide global interconnect if dummy thermal vias are added every 20 μm and 30 μm, respectively, with t_{pulse} = 200 ns and J=6x10^7 A/cm^2. (b) Cu/air with thermal vias every 20 um shows nearly the same ΔT_{max} as that of Cu/oxide global interconnect with 100 μm via separation under a 200 ns and J=6x10^7 A/cm^2 current pulse.

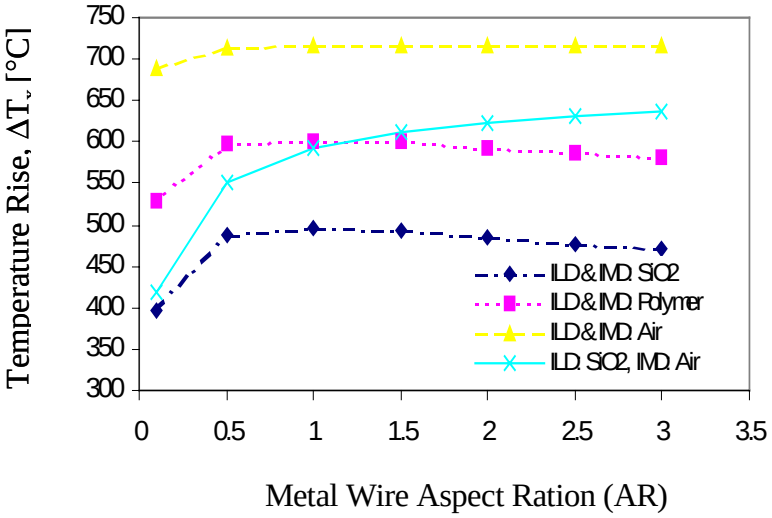
lowest-k dielectric, air, in ULSI interconnects despite its poor thermal properties. In the case of global interconnect, Fig. 4.14(a) shows that thermal vias would be required approximately every 20 μm in Cu/air and every 30 μm in Cu/polymer interconnect structures to match the temperature rise of Cu/SiO_2. For the purpose of comparison, under normal steady state operating condition with the J_{max} specified in the ITRS [17], the thermal via separation can be much more relaxed based on a simple analytical evaluation. The temperature profiles for Cu/SiO_2 with via separation of 100 μm, and for Cu/air with via separation of 20 μm, experiencing the same ΔT_{max} under a 200 ns pulse are shown in Fig. 4.14(b). It can be observed that the temperature rise along the Cu/air wire varies spatially and that the average temperature is much lower than that of Cu/SiO_2 wire, resulting in reduced thermal problems and may relax the requirement for thermal via separation.

4.4.5. Impact of Interconnect Aspect Ratio

Finally, the effect of wire aspect ratio on the thermal characteristics is evaluated. For the same metal thickness, wires with smaller aspect ratio would suffer higher ΔT_{max} because of the smaller surface area-to-volume ratio (Fig. 4.15(a)). For the same cross section area, indicating same current capability, a larger perimeter would result in lower ΔT_{max} by offering larger area for heat to diffuse out of metal wires, as shown in Fig. 4.15(b). However, for embedded air gap (ILD: SiO_2 and IMD: Air), ΔT_{max} increases slightly with aspect ratio. This can be explained by the fact that with higher aspect ratio, embedded air gap interconnect structure would have increasing area contacted by air. Fig. 15(a) and (b) can be used to provide thermal design guidelines for interconnect.



(a)



(b)

Fig. 4.15: (a) For the same metal thickness (2.5 μm), global interconnects with lower aspect ratio (AR), shows higher ΔT_{max} due to lower surface area-to-volume ratio. t_{pulse} = 200 ns and J=6x10^7 A/cm^2. (b) For the same metal wire cross section area (5 μm^2), the impact of aspect ratio (AR) and dielectric structure strategy is shown. ΔT_{max} peaks at aspect ratio=1 due to the smallest perimeters. t_{pulse} = 200 ns and J=6x10^7 A/cm^2.

4.5 Conclusions

In this chapter, a simple SPICE-based 3-D thermal circuit simulation methodology for steady-state and transient stress conditions is presented which allows quick evaluation of various Cu/low-k interconnect structures. The SPICE simulation results show excellent agreement with rigorous finite element simulations and experimental data and can be employed for accurate reliability and performance analysis. It is demonstrated that via density strongly affects the spatial temperature distribution as well as the maximum temperature rise in interconnects and should be considered in interconnect design. In fact, the effective thermal conductivity of ILDs can be significantly higher than the nominal value if via separation is comparable to the thermal characteristic length. Additionally, the impact of dummy thermal vias on the thermal characteristics of interconnects has been presented. It is shown that by optimal spacing of dummy thermal vias the Cu/low-k structures can have the same thermal capability as Cu/SiO_2. Furthermore, it has been demonstrated that thermal coupling can be significant for wires in densely packed structure. The impact of wire aspect ratio (AR) on the thermal characteristics has also been shown to be important for the thermal design of deep sub-micron interconnect structures. Finally, as technology keeps scaling, thermal effects should be carefully evaluated not only to address reliability concerns, but also for accurate interconnect performance analysis. It has been shown that, to accurately estimate the temperature rise and the thermal characteristics of interconnect, the effect of via as efficient thermal paths must be considered properly.

Chapter 5

Thermal Analysis of 3-D ICs

5.1 3-D Integration: Background and Motivation

Interconnect RC delay is increasingly becoming the dominant factor determining the performance of advanced ICs. On the other hand, System-on-a-Chip (SoC) designs are often driven by the ever-growing demand for increased system functionality, heading to dramatic increase in chip area and the use of numerous digital blocks with many long global wires. The three dimensional (3-D) IC architecture has emerged as a unique solution to alleviate the interconnect delay problem [65-67]. Additionally, 3-D ICs provide a vehicle for heterogeneous integration to realize SoC designs, e.g., to incorporate such diverse combinations as embedded DRAM, high-performance logic, analog, RF ICs and optical interconnects [68-72], as schematically illustrated in Fig. 5.1. Furthermore, the noise generated by the interference between different embedded circuit blocks containing digital and analog circuits can be minimized.

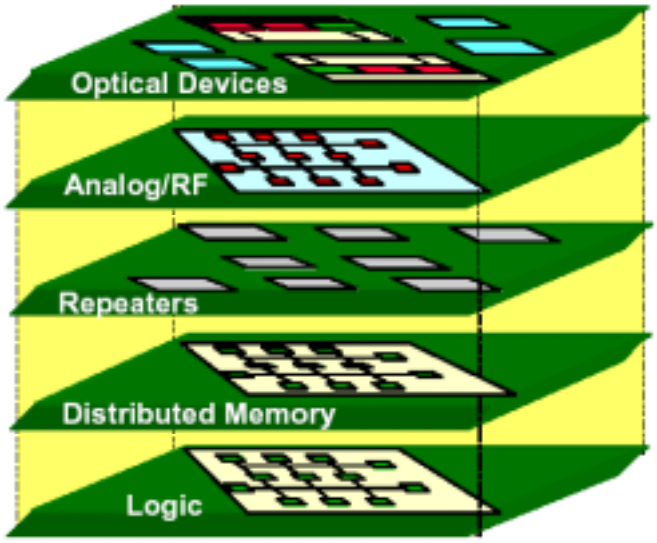


Fig. 5.1: Schematic of a 3-D chip showing integrated heterogeneous technologies.

3-D integration (schematically illustrated in Fig. 5.2) to create the vertical stacking of multilevel active layers of Si ICs is a concept that may significantly improve deep submicron interconnect performance, increase transistor packing density, and reduce chip area and power dissipation. Each Si layer in the 3-D structure may have its own dedicated or shared interconnect network. Each of these layers are connected together through Vertical Inter-Layer Interconnects (VILICs) and common global interconnects as shown in Fig. 5.2. Souri [73] has demonstrated the basic concept for this 3-D analysis, as depicted in Fig. 5.3. A general representation of a wire-pitch limited 2-D IC is considered as consisting of a number of logic blocks. By migrating to a 3-D structure it is assumed that logic blocks can be rearranged in some fashion so as to occupy any number

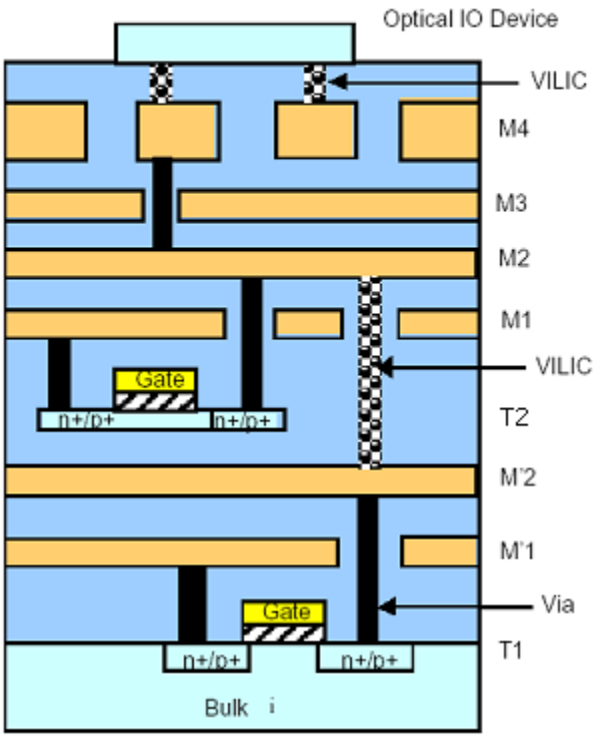


Fig. 5.2: Schematic representation of 3-D integration with multilevel wiring network and VILICs. T1: first active layer device, T2: second active layer device, Optical I/O device: third active layer I/O device. M'1 and M'2 are for T1, M1 and M2 are for T2. M3 and M4 are shared by T1, T2, and the I/O de-

of active layers of Si. Such an arrangement makes the vertical dimension available for logic block interconnectivity. For instance, a global wire in the 2-D IC connecting two logic blocks across the chip and contributing to the chip size can now be replaced with a VILIC connecting the same two blocks, which can be arranged vertically stacked on top of another. This VILIC is characterized by its much shorter length and smaller contribution to chip size as compared to the original global wire. By performing such

replacements across the entire interconnect network, a significant fraction of lateral wires can thus be replaced with VILICs which ultimately reduces the horizontal wiring requirement and chip size and prevents the interconnect delay problem from dominating IC performance.

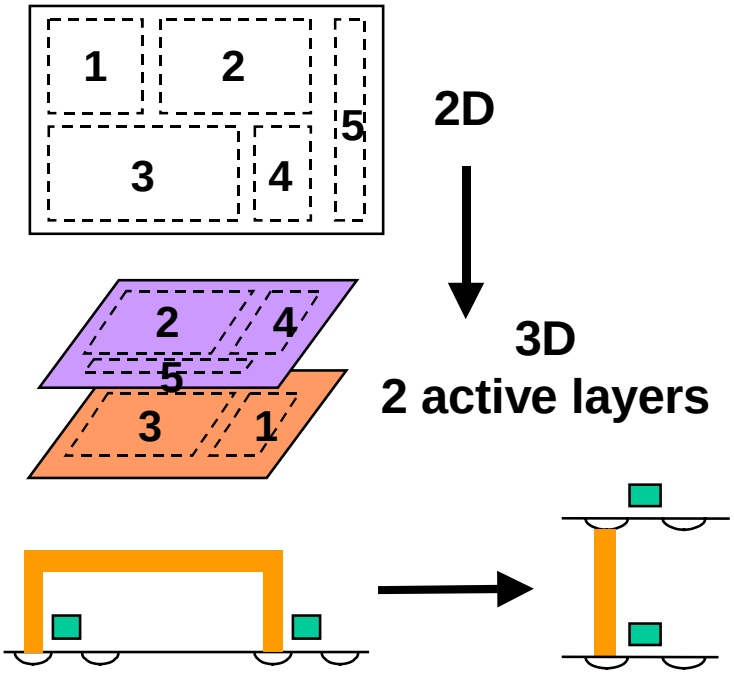


Fig. 5.3: Horizontal interconnects are replaced with VILICs, reducing wiring requirement, chip area and interconnect delays.

However, there is concern of poor heat dissipation and consequent chip temperature rise in 3-D ICs due to increased power density and lower thermal conductivity of inter layer dielectrics (ILD) [24]. The majority of the thermal power dissipated in integrated circuits arises due to signal switching. This heat is typically conducted

through the silicon substrate to the package and then to the ambient by a heat sink. With multi-layer device designs, devices in the upper layers will also generate a significant fraction of the heat. Furthermore, all the active layers will be insulated from each other by layers of dielectrics (LTO, HSQ, polyimide etc.) which typically have much lower thermal conductivity than does Si [74, 75]. Assuming low thermal conductivity of the ILDs, the heat dissipation issue may become even more acute for 3-D ICs and may cause degradation in device performance, and reduction in chip reliability due to increased junction leakage, electromigration failures, and by accelerating other failure mechanisms [76].

In this chapter, a detailed thermal analysis of high performance 3-D ICs is presented under various integration schemes [77]. The analysis presented here is the culmination of collaborative work with Shukri Souri at Stanford University. In Section 5.2, a complete thermal model including power consumption due to both transistors and interconnect joule heating from multiple strata is presented. The vertical links (VILICs) and vias have much higher thermal conductivity and hence can effectively reduce the thermal resistance caused by the ILD layers. Ignoring the effect of these structures can result in overly pessimistic estimations predicting unacceptably high 3-D chip temperatures. Section 5.3 shows schematically the power analysis of 3-D ICs. Tradeoffs between power, performance, chip area and thermal impact are evaluated. Section 5.4 discusses the thermal impact on heterogeneous 3-D integration and the main results are summarized in Section 5.5.

5.2 3-D Thermal Modeling

The significant increase in device switching speed and current density in the inter-connect layers, together with increasing Cu resistivity, due to barriers, surface scattering and skin effect, have augmented heat generation in IC chips. Exacerbated by the low-k dielectrics with poor thermal conductivity, thermal effects will not only lead to higher interconnect temperature in 2-D ICs but also impact the device temperature in various active layers in 3-D ICs (Fig. 5.4).

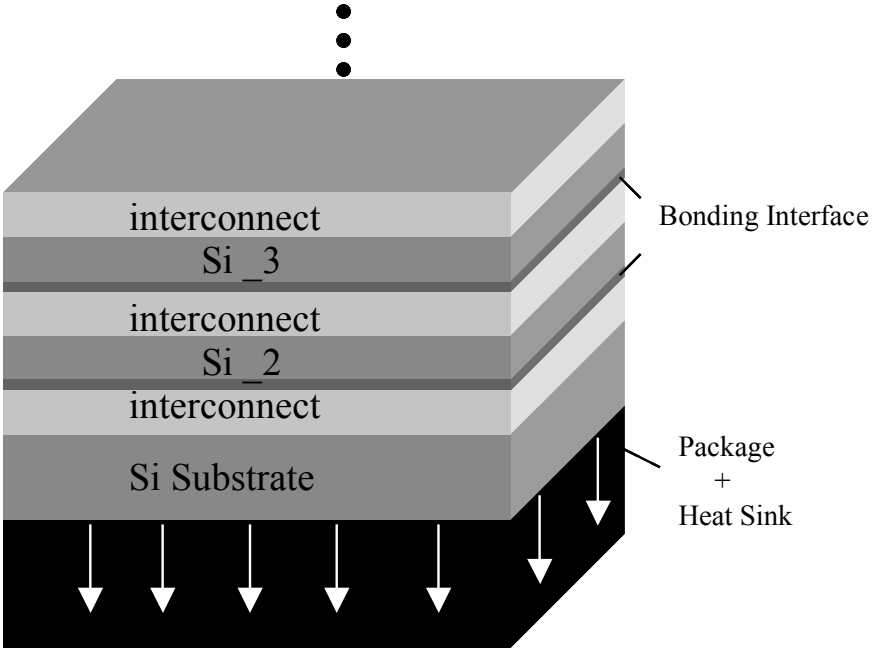


Fig. 5.4: Schematic of multi-level 3-D IC with a heat sink attached to Si substrate.

As seen in Fig. 5.5, even for the case of two active layers, the ratio of thermal resistance caused by the ILD layers (R_{ILD}) to the required package (including glue layers, heat sink) thermal resistance (R_{pkg}) increases rapidly for future technology nodes. The required R_{pkg} is the maximum allowed value which gives the maximum junction temperature specified in the ITRS [58]. With multiple active layers, R_{ILD} will become the dominant factor to determine temperature rise in 3-D ICs.

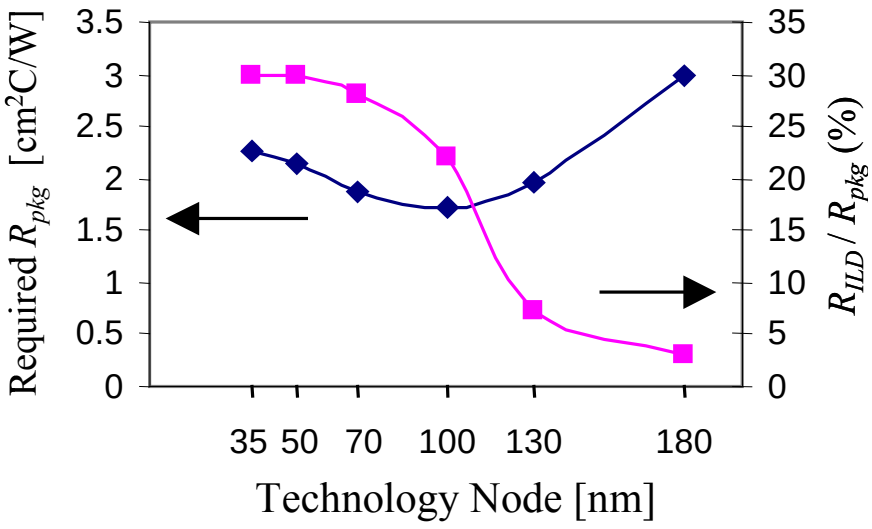


Fig. 5.5: The required package thermal resistance, R_{pkg}, to achieve the maximum junction temperature specified in ITRS and the ratio of R_{ILD} and R_{pkg} vs. technology nodes.

The analytical expression derived based on first principles, to evaluate temperature rise in 3-D structure is given below [77]:

$$\overline{T_{Si\_N}} = T_{amb} + \sum_{m=1}^{N} \{ [\underbrace{\sum_{n=1}^{N_m} \frac{t_{ILD,mn}}{k_{ILD,mn}\, s_{mn}} \eta_{mn} (\sum_{i=n}^{N_m} j_{rms,mn}^2 \, \rho H_{mn} + \sum_{j=m+1}^{M} \Phi_j) }_{\text{Temperature rise caused by ILDs}}] + \underbrace{R_m (\sum_{k=m}^{M} \Phi_k)}_{\substack{\text{Temp. Rise caused by} \\ \text{PKG, glue layer, Si sub.}}} \}$$

where:

- T_{amb} : ambient temperature.
- M : number of strata.
- N_m : number of metal levels in the mth stratum.
- mn : the nth interconnect level in the mth stratum.
- t_{ILD} : thickness of ILD.
- k_{ILD} : thermal conductivity of ILD materials.
- s : heat spreading factor [45].
- η : via correction factor, $0 \leq \eta \leq 1$ [45].
- j_{rms} : root-mean-square value of current density flowing in the wires.
- ρ : electrical resistivity of metal wires.
- H : thickness of metal wires.
- Φ : total power consumption of mth stratum, including power consumed by active layer and interconnect joule heating.
- R : thermal resistance of glue layer and Si layer for each of the stratum, with R_1 represents the total thermal resistance of package, heat sink and Si substrate.

Via effect is incorporated in the expression by the via correction factor η ($0 \leq \eta \leq 1$), with $k_{ILD,eff} = k_{ILD}/\eta$, where $k_{ILD,eff}$ is the effective thermal conductivity of ILD with the

help of via effect and k_{ILD} is the nominal thermal conductivity with via effect ignored [45]. Power consumption due to both active (device) layers and interconnect joule heating are included. This expression can be better understood by comparing it with the Elmore-delay model following an electrical-thermal analogy (Fig. 5.6).

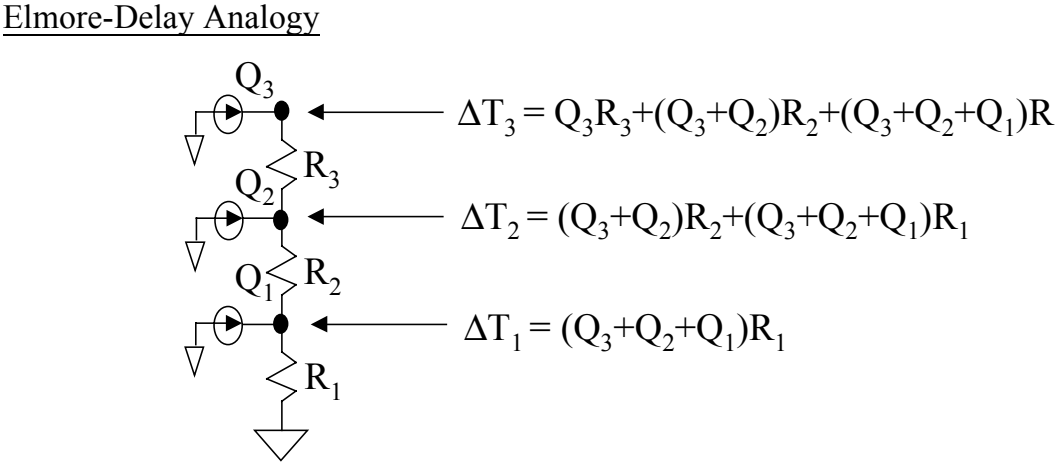


Fig. 5.6: Elmore-Delay model – electrical analogy for the thermal mode employed [45].

The model has been validated by comparing it with full chip thermal simulations done using ANSYS [24]. The two-layer 3-D structure with wafer bonding technique used for the validation is shown in Fig. 5.7. Since it is too complicated to construct the 3-D structure with thousands of vias in ANSYS, the validation is done for the case where via effect is ignored ($\eta = 1$). The results obtained from analytical expressions show excellent agreement with ANSYS (Fig. 5.8), where J_{rms} is 4.85×10^5 A/cm^2 and power

density of each active layer is assumed to be 0.615 W/cm^2. Interconnect parameters are quoted from the ITRS 100 nm technology node. For the case where via effect is included, the via separations assigned to the metal layers, from 1st to 8th levels, are 1, 5, 15, 30, 50, 80, 200 and 500 μm, respectively. However, the analytical model takes much less computation time and provides key design insight. Furthermore, as shown in Fig. 5.8, via effect greatly helps with heat dissipation and the resultant temperatures are much lower as compared to previous works [24, 67, 78]. Therefore, it is crucial to include via effect for thermal analysis of 3-D ICs.

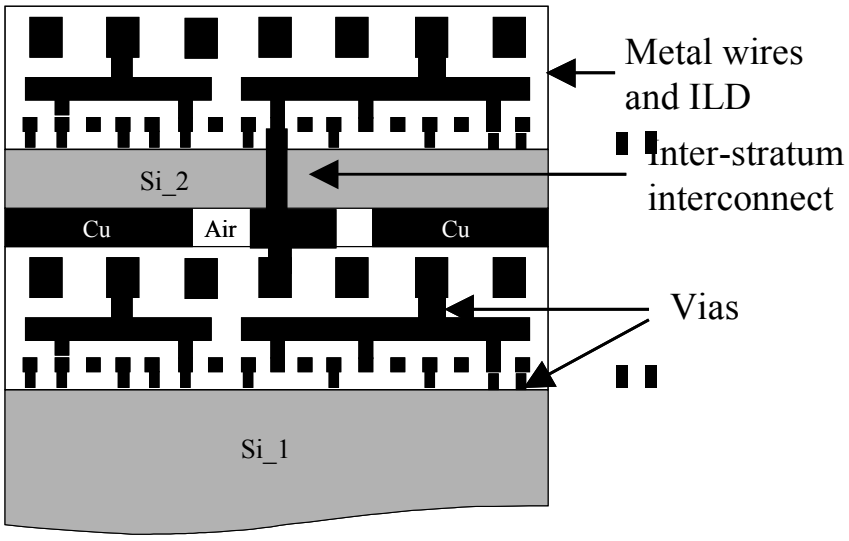


Fig. 5.7: Schematic of 3-D structure fabricated by wafer bonding using Cu pad thermocompression.

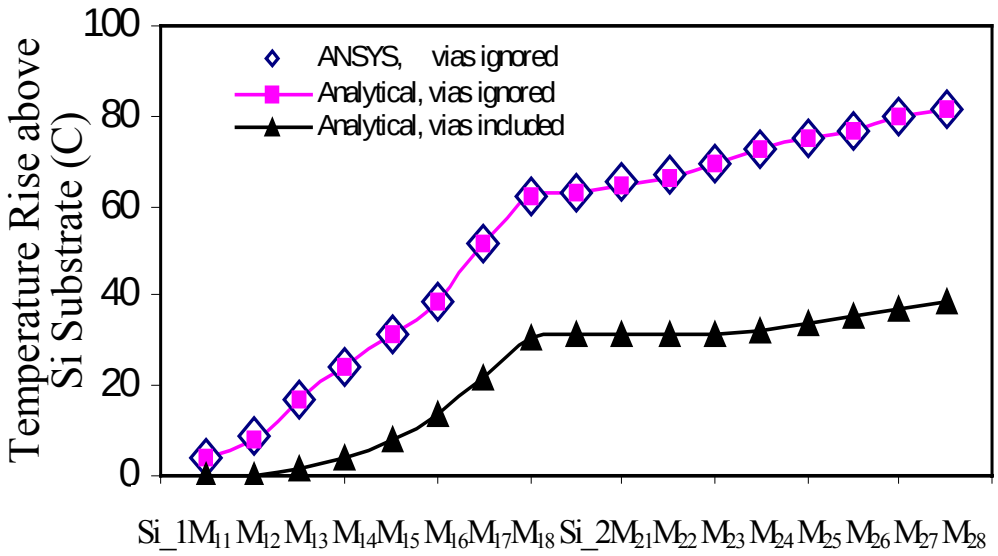


Fig. 5.8: Temperature distribution along the vertical layers from the Si substrate surface (Si_1) to top metal level of the second stratum. The 3-D structure is shown in Fig. 6. M_{11} represents the first metal level in stratum 1, etc.

5.3 Power Analysis of 3-D ICs

In general, performing any comparison between 2-D and 3-D configurations of the same IC is problematic. At issue is the lack of common ground on which to perform an objective comparison. For instance, the operating frequency can be maintained constant, yet the chip area and hence power densities must change. Conversely, if the chip areas are assumed invariant, then the operating frequencies are necessarily different. Such an assumption serves well for a performance comparison. Comparing power dissipation and temperature rises as a result of migration from 2-D to 3-D, on the other

hand, is a different matter. To address this thermal comparative issue, the analysis in this section is again focused on the ITRS projection at the 50nm technology node. However, several 3-D integration cases and scenarios are explored to cover an adequate space for the comparison.

A summary of this wire-pitch limited 2-D case and different 3-D integration cases used in this comparative study is listed in Table 5.1. All the data in this table are calculated based on the 50nm technology node and the thermal resistance of the package is assumed to be 2.15cm^2 °C/W for a constant supply voltage of 0.6V from ITRS projections for 2-D ICs at the 50 nm node. The data in the 2-D column represents the standard 2-D IC. 3-D, Case 1, is a special 3-D integration case in that memory and logic from the 2-D are each dedicated to separate active layers without any modifications to the wiring. The resulting chip area, A_c, is determined by the larger logic area and power dissipation is unchanged relative to the 2-D case. The remaining four 3-D cases are obtained, and compared to 2-D, by modifying the chip wiring and assuming the memory is interspaced with logic. Their characteristics are summarized below:

- 3-D, Case 2: Equal f_c and decreased A_c;
- 3-D, Case 3: Equal f_c and A_c;
- 3-D, Case 4: $2f_c$ and equal A_c;
- 3-D, Case 5: Equal A_c with f_c determined by maintaining 2-D P_{Total}.

where A_c is the chip area, f_c is the operating frequency and P_{total} is the total power dissipation. The different characteristics of these 3-D integration cases give rise to different

power dissipation results as summarized in Table 5.1. The dynamic power dissipation components considered are due to logic, interconnect (local, semi-global and global), clock distribution and repeaters and are calculated using $P_{Dynamic}=1/2\alpha CV_{dd}^2 f_c$ where α is the activity factor (assumed to be 0.1), V_{dd} is the supply voltage obtained from ITRS, f_c is the operating frequency and C is the capacitance. Other power dissipating components include memory, I/O pads and static components, such as leakage and short-circuit currents, are all combined under P_{Other}.

	2-D	3-D, Case 1	3-D, Case 2	3-D, Case 3	3-D, Case 4	3-D, Case 5
Active Layers	1	2	2	2	2	2
f_c (MHz)	3000	3000	3000	3000	6000	3559
Feature Size (nm)	50	50	50	50	50	50
Chip Area (cm^2)	8.17	4.25	4.51	8.17	8.17	8.17
Memory Area (cm^2)	3.92	3.92	3.92	3.92	3.92	3.92
Logic Area (cm^2)	4.25	4.25	5.1	12.42	12.42	12.42
P_{Logic} (W)	34.8	34.8	34.8	34.8	69.6	41.28
P_{Local} (W)	17.4	17.4	17.44	20.66	10.44	6.19
$P_{Semi\text{-}Global}$ (W)	14.63	14.63	6.89	8.16	30.68	18.2
P_{Global} (W)	6.96	6.96	4.18	5.63	11.78	6.99
P_{Clock} (W)	34.8	34.8	22.97	27.21	56.93	33.76
$P_{Repeaters}$ (W)	45.24	45.24	29.7	35.19	73.6	43.65
P_{Other} (W)	20.17	20.17	20.17	20.17	40.34	23.93
P_{Total} (W)	174	174	136.15	151.82	293.37	174
Power Density Per Active Layer (Wcm^{-2})	21.30	20.47	15.09	9.29	17.95	10.65

Table 5.1: Comparison of power dissipation due to logic, interconnect, clock distribution and repeaters for 2-D and 3-D ICs with 2 active layers for ITRS 50nm technology node. 3-D IC cases are presented for comparison by varying the chip area, A_c, and operating frequency, f_c, and represent the same 2-D IC (conserving feature size, number of transistors and functionality) converted to 3-D with 2 active layers

The capacitance, C, is calculated for each component to determine the associated power dissipated. For P_{Logic}, the device capacitance is calculated by considering gate oxide capacitance, overlap capacitance and junction capacitance all of which are calculated from ITRS data [58]. Interconnect capacitances for the local, semi-global and global tiers are found from the wire-length distribution and the dimensions of the wire pitches for each tier [73]. Clock distribution capacitances are calculated using the BACPAC model proposed in [79] by considering a buffered H-Tree model. Power dissipated by repeaters is calculated based on the driver capacitances and the number of repeaters. P_{Other} is determined in the 2-D case to be the sum of remaining components to achieve the ITRS projected total power dissipation for this generation. Since this component is assumed dominated by dynamic dissipation, it is considered linearly dependent on the operating frequency for all 3-D cases.

In 3-D Case 2, the total power dissipation is seen to decrease primarily due to the reduction in the wiring requirement thus reducing the interconnect power dissipation, reducing number of required repeaters and reducing the clock distribution network. 3-D Case 3 is associated with a larger chip area which requires longer interconnect lines, a larger number of repeaters and clock-distribution network all of which increase the power dissipation. However, lower power density is achieved due to larger chip area. 3-D Case 4 shows a dramatic increase in the power dissipated primarily due to the significant increase in operating frequency. 3-D Case 5 illustrates the increase in the operating

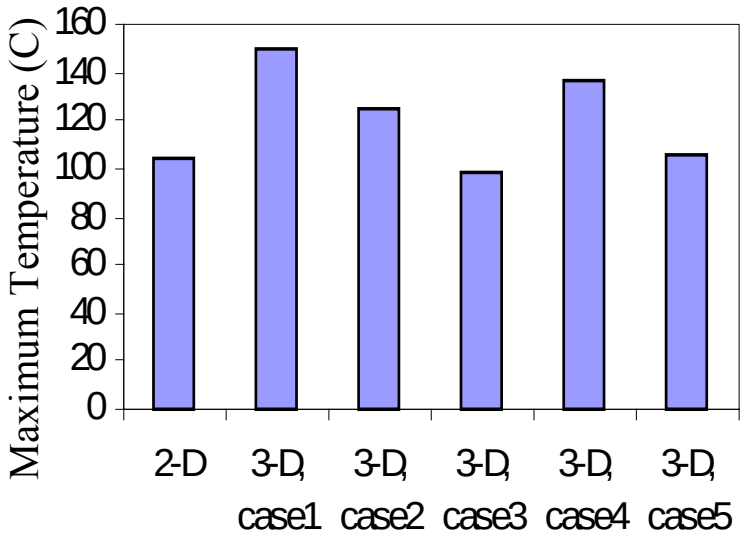


Fig. 5.9: Comparison of temperature performance among 2-D ICs and five different two-active-layer 3-D ICs scenarios

frequency if the chip area and the power dissipation requirements are maintained constant to 2-D.

Although the total power consumption as shown in Table 5.1 is reduced by going from 2-D to 3-D ICs due to the reduction in the interconnect and the clock network related capacitance, the heat removal capability could deteriorate as the upper active layers experience a longer heat dissipation path to the heat sink. Fig. 5.9 compares the temperature rise for different 3-D integration scenarios. The lowest temperature achieved is that for 3-D Case 3. Here, although the wiring has been reduced by migrating to 3-D, the operating frequency and the chip area have been maintained constant as compared to 2-D. The power density, then, is significantly lower than in the 2-D case giving rise to a lower die temperature. It should be noted that, in most of the cases, 3-D ICs have similar

temperature rise but have the advantage of either reduced chip area (Case 2) or increased operating frequency (Case 5). In the case of equal chip area and operating frequency (Case 3), lower temperature than 2-D ICs can be achieved. Note that to double the operating frequency (Case 4), temperature will invariably increase. Nevertheless. the temperature is still much lower than that estimated with via effect ignored [24, 67, 78].

Furthermore, it is desirable to put memory in close proximity to the logic circuitry in order to reduce latency in high performance microprocessors. 3-D IC technology provides an excellent opportunity to stack memory and logic. The power consumption in on-chip memory is generally less than 10% of total power consumption and the area occupied by memory and logic are comparable at 50 nm node. With these assumptions, several schemes are developed, and applied in the following analysis to 3-D Case 4. These schemes are illustrated in Fig. 5.10 where four 3-D stack schemes are shown, each with a different configuration of memory and logic. Fig. 5.10 shows four 3-D stack schemes along with their temperature performance. It can be observed that with logic in the bottom active layer and memory in the upper layer (Scheme 1), the resultant temperature rise is the lowest. On the other hand, stacking logic parts back to back will experience much higher temperature rise.

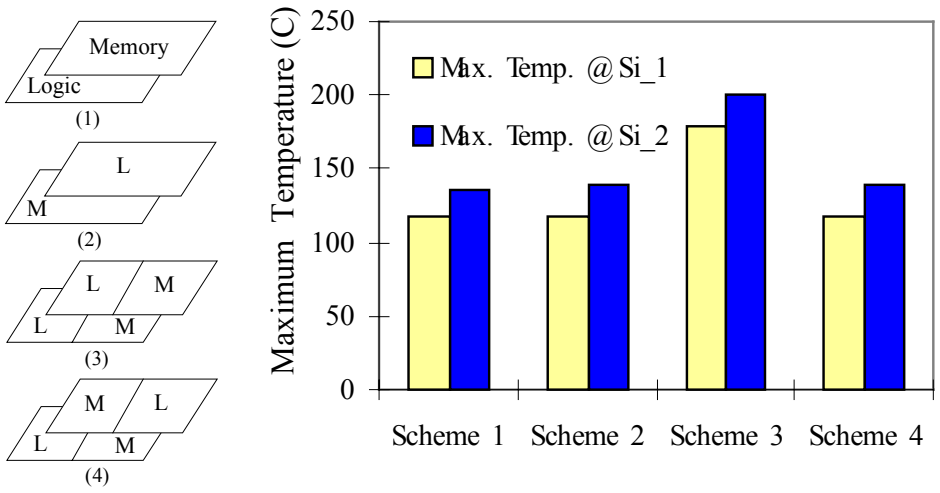


Fig. 5.10: Thermal capability of high performance 3-D ICs (Case 4) for four different 3-D logic-memory integration schemes.

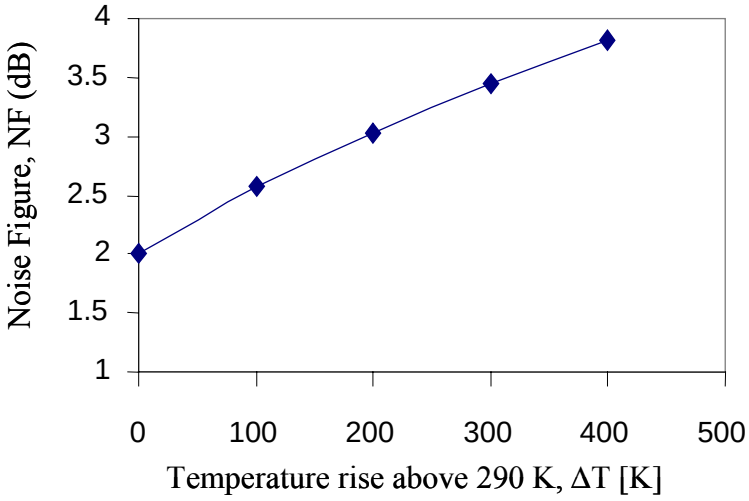


Fig. 5.11: The noise performance of a typical low noise amplifier (LNA) under the temperature effect is evaluated. NF is assumed to be 2dB @ 290K.

5.4 Thermal Impact on Heterogeneous 3-D Integration

Heterogeneous integration of RF circuit in 3-D ICs, e.g., with a microprocessor to facilitate wireless communication, has attracted major interest with the intent to isolate substrate-coupled noise between digital and analog components [67]. Evaluated in Fig. 5.11 is the noise figure, indexing RF performance, that deteriorates rapidly with temperature rise, imposing more stringent design requirements for 3-D integration.

Employing optical interconnects in the global signaling or clocking could eliminate many problems associated with large multi-GHz chips like reducing timing skew, power and area for clock distribution [80]. An initial investigation to look at the thermal aspects of the required optical components is shown below. In an on-chip optical interconnect system, photodetectors are responsible for optical-to-electrical signal conversion. Usually, the detector photocurrent is ~100μA. Even the dark current (noise) increases rapidly with temperature (Fig. 5.12), a photo-to-dark current ratio of 100 is still retained, which is enough for reliable communications. In a dense optical interconnect, the receivers are not noise-limited, but indeed gain-limited aiming at low power.

Focusing on the transmitting components like lasers and modulator diodes, the three major temperature-induced changes in performance are drop in quantum efficiency, degradation of laser threshold current, and shift in transmitter wavelength. Employing strained multiple quantum well (MQW) active regions with effective facet coatings [81],

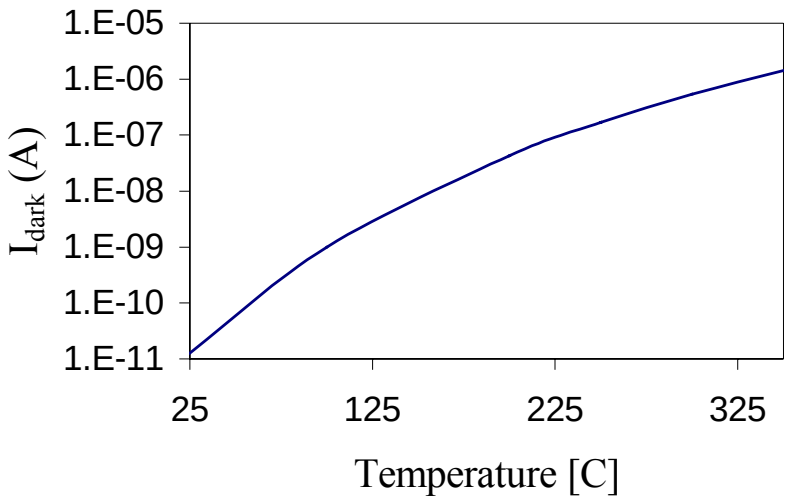


Fig. 5.12: Temperature effect on optical receiver performance is evaluated. A typical example of Ti-Si metal-semiconductor-metal (MSM) PD dark current vs. ambience temperature is shown here.

[83] and a bimetallic heatsink [84] were explored to get around the detrimental wave-length shift. Recently, a quantum-dot laser with a maximum operating temperature of 160°C in pulse mode [85] has been demonstrated.

The last major components are the waveguide-related devices like modulators, optical attenuators, and electric field sensors, etc. For a waveguide (attenuator) with an InGaAsP MQW active layer and InP claddings, a negligible change in absorption coefficient (an index for fiber loss) has been demonstrated from about 20 to 120°C [86]. Even up to 170°C, the absorption loss is still tolerable.. From the above analysis, the incorporation

of optical interconnect into the 3-D architecture may require some more device design efforts for higher temperature operation.

5.5 Conclusions

In conclusion, a compact analytical thermal model to evaluate temperature distribution in 3-D ICs including via effect is presented. It is demonstrated that via effect must be considered to evaluate the thermal capability of 3-D ICs. Temperature performance of various 3-D integration schemes has been examined thoroughly. It is shown that with careful thermal designs, 3-D ICs can have the same thermal capability as that of 2-D ICs. For the high performance case where higher temperature is not avoidable, better circuit design and advanced package solution will be necessary. Finally, examples of 3-D heterogeneous integration of RF circuits and optical interconnects have been explored from the thermal view point.

Chapter 6

Conclusion

6.1 Summary

Due to scaling trends of VLSI technology, Joule heating at the interconnect levels is becoming non-negligible. As a result, not only will thermal effects be a major reliability concern, but also the increase of wire electrical resistivity with temperature can degrade the expected speed performance. In this thesis, both a compact analytical formula and a convenient SPICE-based simulation methodology are developed to evaluate the temperature rise on interconnects under steady-state and transient stress conditions. The results demonstrate excellent agreement with experimental data and with Finite Element (FE) thermal simulations (ANSYS). With the effects of vias, as efficient heat transfer paths, being taken into account, more realistic and accurate interconnect thermal analysis is finally possible. Significant differences in temperature distribution and maximum temperature rise are observed between the realistic situation of heat

dissipation with vias and the overly simplified case that ignores the via effect. Furthermore, it shows that the effectiveness of vias in reducing the temperature rise is highly dependent on the via separation and the dielectric materials used. The effective thermal conductivity of ILDs can be significantly higher than the nominal value if via separation is comparable to the thermal characteristic length. Therefore, the thermal problem associated with low-k insulators is not as bad as it might appear. It should be noted that global interconnects would suffer much higher temperature rises than local interconnects due to the much longer via separation and further away from heat sink. Additionally, the thermal advantage gained by using dummy thermal vias in advanced Cu/low-k interconnects is quantified, which may offers a solution to alleviate hot interconnect phenomena.

The impact of Joule heating on the scaling trends of advanced VLSI interconnects has been evaluated in detail. Coupled analysis of delay and electromigration MTF for various technology nodes suggests that Joule heating will limit scaling of current density and use of low-k materials. It shows that the interconnect Joule heating can strongly affect the maximum operating temperature of the global wires which will, in turn, constrain the scaling of current density to mitigate electromigration and, thus, greatly degrade the expected speed improvement from the use of low-k dielectrics. Optimization with various interconnect parameters is performed to provide thermal design guideline/insight in the early design phase.

Finally, potential bottlenecks and opportunities of future heterogeneous three dimensional (3-D) ICs with various integration scenarios are identified from the thermal

point of view. It is shown that under certain scenarios, 3-D ICs can actually lead to better thermal performance than planar (2-D) ICs. Tradeoffs among power, performance, chip real estate and thermal impact for 3-D ICs is evaluated.

6.2 Future Work

In our analytical thermal model under steady-state normal operation condition, we assume the wire resistivity is at a constant value. Although errors introduced because of this simplification are minimal, more accurate results can be obtained by including the temperature dependency, carrier scattering and skin effect etc. This will lead to a non-closed form formula, and require a few iterations. However, with the help of computer, this process should not be too difficult.

On the other hand, we consider the heat sink as the sole thermal dissipation path to the outside ambient and assume adiabatic boundary conditions on the top and on all the four side walls of the chip, with the reasoning that the chip is encapulated by some insulating materials. These assumptions are fairly accurate with wire bonding package and still quite reasonable with the current flip-chip technology. However, with the increasing number of bumps for each generation, more heat will be dissipated through these thousands of metal balls. The effect of this additional heat transfer path should be investigated to ensure valid thermal analysis.

Lastly, the thermal modeling and insight developed in this work should not be limited to interconnect level; it can be easily extended to device level and chip level with certain modification. The thermal effects will continue to be a major issue for years to come in chip design, hopefully, more accurate and convenient methodologies can be built upon this work to facilitate the continued advancement of semiconductor industry

Bibliography

[1] G. E. Moore, “Cramming More Components onto Integrated Circuits,” *Electronics*, vol.38, no.8, April 19th 1965.

[2] C. Hu, “MOSFET Scaling in the Next Decade and Beyond,” *Semiconductor International*, pp. 105-114, 1994.

[3] M. T. Bohr, “Interconnect Scaling-The Real Limiter to High Performance ULSI,” *IEDM Tech. Dig.*, pp. 241-244, 1995.

[4] K. C. Saraswat and F. Mohammadi, “Effect of Interconnection Scaling on Time Delay of VLSI Circuits,” *IEEE Trans. On Electron Devices*, vol. ED-29, pp. 645-650,1982.

[5] *The International Technology Roadmap for Semiconductors* (ITRS), 2001

[6] M. T. Bohr and Y. A. El-Mansy, "Technology for Advanced High-Performance Microprocessors," *IEEE Trans. Electron Devices*, vol 45, no. 3, pp. 620-625, 1998.

[7] T. N. Theis, “The Future of Interconnection Technology”, *IBM Journal of R&D*, Vol. 44, 2000.

[8] R. Ho, K. Mai, and M. Horowitz, “The Future of Wires,” *Proceedings of the IEEE*, vol.89, no.4, pp. 490-504, April 2001.

[9] S. Venkatesan, A. V. Gelatos, S. Hisra, B. Smith, R. Islam, J. Cope, B. Wilson, D. Tuttle, R. Cardwell, S. Anderson, M. Angyal, R. Bajaj, C. Capasso, P. Crabtree, S. Das, J. Farkas, S. Filipiak, B. Fiordalice, M. Freeman, P. V. Gilbert, “A High Performance 1.8V, 0.20 μm CMOS Technology with Copper Metallization,” *IEDM Tech. Dig.*, pp. 769-772, 1997.

[10] N. Rohrer, C. Akrout, M. Canada, D. Cawthron, B. Davari, R. Floyd, S. Geissler, R. Goldblatt, R. Houle, P. Kartschoke, D. Kramer, P. McCormick, G. Salem, R. Schulz, L. Su, L. Whitney, “A 480MHz RISC microprocessor in a 0.12 μm L_{eff}

CMOS Technology with Copper Interconnects," *Int. Solid-State Circuits Conf., Tech. Digest*, pp. 240-241,1998.

[11] E. M. Zielinski, S. W. Russell, R. S. List, A. M. Wilson, C. Jin, K. J. Newton, J. P. Lu, T. Hurd, W. Y. Hsu, V. Cordasco, M. Gopikanth, V. Korthuis, W. Lee, G. Cerny, N. M. Russell, P. B. Smith, S. O'Brien, R. H. Havemann, "Damascene Integration of Copper and Ultra-Low-k Xerogel for High Performance Interconnects," *Tech. Dig. IEDM*, pp. 936-938,1997.

[12] B. Zhao, D. Feiler, V. Ramanathan, Q. Z. Liu, M. Brongo, J. Wu, H. Zhang, J. C. Kuei, D. Young, J. Brown, C. Vo, W. Xia, C. Chu, J. Zhou, C. Nguyen, L. Tsau, D. Dornish, L. Camilletti, P. Ding, G. Lai, B. Chin, M. Johnson, J. Turner, "A Cu/low-k Dual Damascene Interconnect for High Performance and Low Cost Integrated Circuits," *Symp. VLSI Technology, Tech. Digest*, pp. 28-29,1998.

[13] F. Pollack, "New Challenges in Microarchitecture and Compiler Design", *PACT 2000*, Oct 2000.

[14] P. K. Ko, "Hot-Electron Effects in MOSFET's," Ph.D. dissertation, Univ. of California, Berkeley, 1982.

[15] C. Fiegna, H. Iwai, T. Wada, T. Saito, E. Sangiorgi, and B. Ricco, "Scaling the MOS Transistor below 0.1 μm: Methodology, Device Structure, and Technology Requirements," *IEEE Trans. Electron Devices*, vol. 41, no. 6, p. 941, June 1994.

[16] J. Ida, M. Yoshimaru, T. Usami, A. Ohtomo, K.Shimokawa, A. Kita and M. Ino, "Reduction of Wiring Capacitance with New Low Dielectric SiOF Interlayer Film for High Speed/Low Power Sub-half Micron CMOS," *VLSI Technology Symposium, Dig. Tech. Papers*, pp. 59-60,1994.

[17] B. P. Shieh, B. Shieh, K. C. Saraswat, J.P. McVittie, S. List, S. Nag, M. Islamraja, R. H. Havemann, "Air-Gap Formation During ILD Deposition to Lower Interconnect Capacitance," *IEEE Electron Device Lett.*, vol. 19, no. 1, January 1998, pp. 16~18.

[18] D. H. Cho, Y. S. Eo, M. H. Seung, N. H. Kim, J. K. Wee, O. K. Kwon, and H.S. Park, “Interconnect Capacitance, Crosstalk, and Signal Propagation Delay for 0.35 μm CMOS Technology,” *IEDM Tech. Dig.*, pp. 619-622,1996.

[19] J. R. Black, “Electromigration – A Brief Survey and Some Recent Results,” *IEEE Transactions on Electron Devices*, vol. ED-16, pp. 338-347,1969.

[20] G. Yoh and F. N. Najm, "A Statistical Model for Electromigration Failures," *IEEE 2000 1st International Symposium on Quality Electronic Design*, San Jose, CA, March 20-22, pp. 45-50, 2000.

[21] 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,” *Tech. Dig. IEDM*, pp. 65-68,1996.

[22] K. Banerjee, A. Mehrotra, A. Sangiovanni-Vincentelli, and C. Hu, “On Thermal Effects in Deep Sub-micron VLSI Interconnects,” *36th ACM Design Automation Conference*, pp. 885-891,1999.

[23] Danqing Chen, Erhong Li, Elyse Rosenbaum and Sung-Mo Kang, “Interconnect Thermal Modeling for Determining Design Limits on Current Density,” *Computer-Aided Design of Integrated Circuits and Systems, IEEE transactions on*, vol. 19, issue 2, pp. 197-205 Feb. 2000.

[24] S. Im and K. Banerjee, “Full Chip Thermal Analysis of Planar (2-D) and Vertically Integrated (3-D) High Performance ICs,” *Tech. Dig. IEDM*, pp. 727-730, 2000.

[25] Y. S. Ju and K.E. Goodson, “Short-timescale Thermal Mapping of Interconnects,” *Proc. IEEE I.R.P.S*, pp. 320-324,1997.

[26] S. Ramaswamy, C. Duvvury, and S. Kang, “EOS/ESD Reliability of Deep Sub-micron NMOS Protection Devices,” *Proc. IEEE IRPS*, pp. 284-291, 1995.

[27] R. V. Andrews, “Solving Conductive Heat Transfer Problems with Electrical-analogue Shape Factor,” *Chemical Engineering Process*, vol. 51, No. 2, pp. 67-71, 1955.

[28] P. R. Gray and D. J. Hamilton, "Analysis of electrothermal integrated circuits," *IEEE J. Solid-State Circuits*, vol. SC-6, pp. 8-14, Feb. 1971.

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

[30] 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.

[31] 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.

[32] 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. 406-411, 1998.

[33] J. T. Trattles, A. G. O'Neill, and B.C. Mecrow, "Three-dimensional finite-element investigation of current crowding and peak temperatures in VLSI multilevel interconnections," *IEEE Trans. Electron Devices*, vol. 40, No. 7, pp. 1344-1347, 1993.

[34] K. A. Gardner, "Efficiency of extended surfaces," *Trans. ASME*, vol. 67, pp. 621-631, 1945.

[35] John H. Lau, *Low Cost Flip Chip Technologies,* McGraw-Hill, Inc., NY, pp. 314, 2000.

[36] H. Katto, M. Harada, and Y. Higuchi, "Wafer-level Jramp and J-constant electromigration testing of conventional & SWEAT patterns assisted by a thermal and electrical simulator," *Proc. of the IRPS*, p298-305, 1991.

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

[38] F. P. Incropera and D. P. DeWitt, *Introduction of heat transfer,* New York: John Wiley & Sons, pp. 192-194, 2002.

[39] Virgil A. Sandborn, Resistance temperature transducers, *Metrology Press*. Fort Collins, Colorado, pp. 511, 1972.

[40] W. R. Hunter, “The implications of self-consistent current density design guidelines: comprehending electromigration and joule heating for interconnect technology evolution,” *IEDM Dig. of Tech. Papers*, pp. 483-486, 1995.

[41] W. Y. Shin, M.C. Chang, R.H. Havemann, and J. Levine, “Implications and solutions for joule heating in high performance interconnects incorporating low-k dielectrics,” *Digest of 1996 Symposium on VLSI Technology*, June, pp. 83-84, 1997.

[42] K. Banerjee and A. Mehrotra, "Coupled Analysis of Electromigration Reliability and Performance in ULSI Signal Nets," *IEEE International Conference on Computer-Aided Design* (ICCAD), San Jose, CA, November 4-8, pp. 158-164, 2001

[43] T. Y. Chiang, K. Banerjee, and K. C. Saraswat, “Compact modeling and SPICE-based simulation for electrothermal analysis of multilevel ULSI,” *ICCAD*, pp 165-172, 2001.

[44] T. Y. Chiang, B. Shieh and K. Saraswat, "Impact of Joule Heating on Scaling of Deep Sub-Micron Cu/Low-K Interconnects," *Digest of 2002 Symposium on VLSI Technology*, Honolulu, Hawaii, June, pp. 141-142, 2002.

[45] T. Y. Chiang, K. Banerjee, and K. C. Saraswat, "Analytical Thermal Model for Multilevel VLSI Interconnects Incorporating Via Effect," *IEEE Electron Device Letters*, Vol. 23, No. 1, pp. 31-33, January, 2002.

[46] T. Y. Chiang, and K. Saraswat, "Closed-form Analytical Thermal Model for Accurate Temperature Estimatin of Multilevel ULSI Interconnects " *Digest of 2003 Symposium on VLSI Circuit*, Kyoto, Japan. June, pp. 275-278, 2003.

[47] W. J. Dally and J. W. Poulton, *Digital Systems Engineering*, Cambridge University Press, NY, 1998.

[48] G. Yoh and F. N. Najm, "A Statistical Model for Electromigration Failures," *IEEE 2000 1st International Symposium on Quality Electronic Design*, San Jose, CA, March 20-22, pp. 45-50, 2000.

[49] B. K. Liew, N. Cheung, and C. Hu, "Projecting interconnect electromogration lifetime for arbrtrary current waveforms," *IEEE Trans. Electron Devices*, vol. 37, pp. 1343-1351, May 1990.

[50] C. -K. Hu and J.M.E. Harper, "Copper interconnect: fabrication and reliability," *VLSI Technology, Systems, and Applications, 1997. Proceedings of Technical Papers, 1997 International Symposium on*, pp.18 – 22, June 1997.

[51] C. -K. Hu, B. Luther, F.B. Kaufman, J. Hummel, C. Uzoh, D. J. Pearson, "*Copper interconnection integration and reliability," Thin solid Films,* 262, pp. 84-92, 1995.

[52] Jiang Tao, Nathan W. Cheung, and C. Hu, "Electromigration characteristics of copper interconnects," *IEEE Electron Device Lett*., vol. 14, no. 5, pp. 249-251, May 1993.

[53] P. Kapur, G. Chandra and K.C. Saraswat, "Technology and reliability constrained future copper interconnects-Part I: Resistance Modeling," *IEEE Trans. Electron Devices*, vol. 49, no. 4, pp. 590-597, April 2002.

[54] P. Kapur, "Scaling induced performance challenges/limitations of on-chip metal interconnects and comparisons with optical interconnects," Ph.D. dissertation, Stanford, 2002.

[55] Linan Jiang; Jae-Mo Koo; Shulin Zeng; Mikkelsen, J.C.; Lian Zhang; Peng Zhou; Santiago, J.G.; Kenny, T.W.; Goodson, K.E.; Maveety, J.G.and Tran, Q.A., **"**Two-phase microchannel heat sinks for an electrokinetic VLSI chip cooling system," *Semiconductor Thermal Measurement and Management, Seventeenth Annual IEEE Symposium*, pp.153 – 157, March 2001

[56] T. Y. Chiang, K. Banerjee, and K. C. Saraswat, "Effect of via separation and low-k dielectric materials on the thermal characteristics of Cu interconnects," *Tech. Dig. IEDM*, pp. 261-264, 2000.

[57] T. Y. Chiang, and K. C. Saraswat, "Impact of vias on the thermal characteristics of deep sub-micron Cu/low-k interconnects," *Tech. Dig., Symp. VLSI Technology,* pp. 141-142, 2001

[58] C. C. Teng, Y.K. Cheng, E. Rosenhaum and S. M. Kang, "iTEM: a temperature dependent electromigration reliability diagnosis tool," *IEEE Trans. Computer-Aided Design*, vol. 16, No. 8, pp. 882-893, Aug. 1997.

[59] The International Technology Roadmap for Semiconductors, 1999.

[60] 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.

[61] C. Duvvury and A. Amerasekera, "ESD: A pervasive reliability concern for IC technologies," *Proc. IEEE*, vol. 81, pp. 690-702, May 1998.

[62] T. Green, "A review of EOS/ESD field failures in military equipment," in *Proc. 10th, EOS/EDS symp.*, pp. 7-14, 1998.

[63] M. J. Pelgrom and E. C. Dijkmans, "A 3/5 V compatible I/O buffer," *IEEE J. Solid-Statae Circuits*, vol. 30, no. 7, pp. 823-825, 1995.

[64] F. P. Incropera and D. P. DeWitt, *Introduction to heat transfer.* New York: John Wiley & Sons, pp. 61-65, 2002.

[65] S. J. Souri and K. C. Saraswat, "Interconnect performance modeling for 3D integrated circuits with multiple Si layers," *Int. Interconnect Technology Conf. Proceedings*, pp. 24-26, 1999.

[66] S. J. Souri, K. Banerjee, A. Mehrotra, and K. C. Saraswat, "Multiple Si layer ICs: motivation, performance analysis, and design implications," *Proc. 37th ACM Design Automation Conf.*, pp. 873-880, 2000.

[67] K. Banerjee, S. J. Souri, P. Kapur, and K. C. Saraswat, "3-D ICs: A Novel Chip Design for Improving Deep Submicron Interconnect Performance and Systems-on-Chip Integration," *Proceedings of the IEEE, Special Issue, Interconnections- Addressing The Next Challenge of IC Technology*, Vol. 89, No. 5, pp. 602-633, May 2001.

[68] Akasaki, "Three-Dimensional IC Trends", *Proceedings of the IEEE*, Dec. 1986

[69] N. Takahashi, N. Senba, Y. Shimada, I. Morisaki, K. Tokuno, "Three-dimensional memory module", *IEEE Trans. on Advanced Packaging*, Feb. 1998

[70] Y. Akasaka and T. Nishimura, "Concept and basic technologies for 3-D IC structure," *IEDM Tech. Dig.*, pp. 488-491,1986.

[71] T. Nishimura, Y. Inoue, K. Sugahara, S. Kusunoki, T. Kumamoto, S. Nakagawa, M. Nakaya, Y. Horiba, and Y. Akasaka, "Three dimensional IC for high performance image signal processor," *IEDM Tech. Dig.*, pp. 111-114, 1987.

[72] A. V. Krishnamoorthy, A. L. Lentine, K. W. Goossen, J. A. Walker, T. K. Woodward, J. E. Ford, G. F. Aplin, L. A. D'Asaro, S. P. Hui, B. Tseng, R. Leibenguth, D. Kossives, D. Dahringer, M. F. Chirovsky, D. A. B. Miller, "3-D integration of MQW modulators over active submicron CMOS circuits: 375 Mb/s transimpedance receiver-transmitter circuit," *IEEE Photonics Technology Letters*, vol. 7, no. 11, pp. 1288-1290, 1995.

[73] S. Souri, "3D ICs interconnect performance modeling and analysis," Ph.D. dissertation, Stanford, 2003.

[74] K. E. Goodson and Y. S. Ju, "Heat conduction in novel electronic films," *Annu. Rev. Mater. Sci.*, 29: pp. 261-293, 1999.

[75] K. Banerjee, A. Mehrotra, A. Sangiovanni-Vincentelli, and C. Hu, "On thermal effects in deep sub-micron VLSI interconnects," *Proc. 36th ACM Design Automation Conference*, pp. 885-891, 1999.

[76] K. Banerjee, Ph.D. Dissertation, University of California at Berkeley, 1999.

[77] Ting-Yen Chiang, S. J. Souri, Chi On Chui, K. C. Saraswat, "Thermal analysis of heterogeneous 3D ICs with various integration scenarios," *Electron Devices Meeting, 2001. IEDM Technical Digest*. 2001

[78] S. A. Kuhn, M. B. Kleiner, P. Ramm, and W. Weber, "Thermal analysis of vertically integrated circuits," in *IEDM Technical Digest*, pp.487-490, 1995

[79] D. Sylvester, Chenming Hu, "Analytical modeling and characterization of deep-submicrometer interconnect," *Proc. IEEE*, pp. 634-664, May 2001.

[80] D. A. B. Miller, "Optical interconnects to silicon;" *Selected Topics in Quantum Electronics, IEEE Journal of,* Volume: 6, Issue: 6 ,Nov.-Dec. pp.1312 – 1317, 2000.

[81] A. Kasukawa, N. Iwai, N. Yamanaka, N. Yokouchi, "Very high characteristic temperature and constant differential quantum efficiency:1.3-µm GaInAsP-InP strained-layer quantum-well lasers by use of temperature dependent reflectivity (TDR) mirror,"*Selected Topics in Quantum Electronics, IEEE Journal of,* Volume: 1, Issue: 2, pp. 293-300, Jun 1995.

[82] C-E Zah, R. Bhat, B. N. Pathak, F. Favire, Wei Lin Wang, M.C. Andreadakis, N.C. Hwang, D.M. Koza, M.A. Tein-Pei Lee, Zheng Wang D. Darby, D. Flanders, J. J. Heieh, "High-performance uncooled 1.3-µm AlxGayIn 1-x-yAs/InP strained-layerquantum-well lasers for subscriber loop applications," *Quantum Electronics, IEEE Journal of,* Volume: 30, Issue: 2, pp. 511-523, Feb. 1994.

[83] D. A. Cohen, Daniel A. Cohen, Mark E. Heimbuch, and Larry A. Coldrene, "Reduced temperature sensitivity of the wavelength of a diode laser in a stress-engineered hydrostatic package," *APL*, vol. 69, no. 4, pp. 455-457, 1996

[84] D. A. Cohen *et al.*, *OFC '97 Tech. Dig.*, pp. 303-304, 1997.

[85] F. Klopf, R. Krebs, J. P. Reithmaier and A. Forchel, "High-temperature operating 1.3-µm quantum-dot lasers for telecommunication," *IEEE PTL*, vol. 13, no. 8, pp. 764-766, 2001.

[86] M. Allard, R. A. Masut, M. Boudreau, "Temperature determination in optoelectronic waveguide modulators," *IEEE JLT*, vol. 18, no. 6, pp. 813-818, 2000.