An Accurate Interconnect Thermal Model Using Equivalent Transmission Line Circuit

Baohua Wang and Pinaki Mazumder
Dept. of Electrical Engineering and Computer Science
University of Michigan, Ann Arbor, MI USA
{baohuaw@umich.edu, mazum@eecs.umich.edu}

Abstract—This paper presents an accurate interconnect thermal model for analyzing the temperature distribution of an on-chip interconnect wire. The model addresses the ambient temperatures and the heat transfer rates of the packaging materials. Particularly, the model considers the effect of the interconnect temperature gradients. The paper employs an equivalent transmission line circuit to obtain the temperature distribution solution from the model. Then an $O(n)$ algorithm is introduced to compute the interconnect temperatures. Experimental results demonstrate the accuracy of the thermal model, by comparisons with the computational fluid dynamics tool FLUENT.

I. INTRODUCTION

Rapid technology scaling enables hundreds of millions of transistors to be densely packed into a single chip. However, the enormous heat dissipation can cause high temperatures and large temperature gradients inside the chip. High temperatures significantly exacerbate interconnect electromigration and reduce the mean-time-to-fail (MTTF) of a metal wire [1]. On the other hand, large temperature gradients induce severe circuit timing failures because of the spatial variation of interconnect conductivities [2]. Particularly, temperature gradients make the optimization of on-chip clock distribution network challenging in order to reduce the clock skews [3].

To alleviate the timing and reliability issues, it is necessary to accurately analyze the temperatures of on-chip interconnect wires. Interconnect thermal analysis is usually separated into two steps: first, partition the entire interconnect network into individual wire segments based on a library of routing patterns; second, estimate the temperature distribution of each wire segment using the thermal characteristic data for the specific routing pattern [4], [5]. The thermal characteristic data typically give the thermal impedances of the interconnect wires in the specific routing pattern based on the heat conductivities of the dielectric materials and the interconnect geometries [6], [7], [8], [3]. The latter step usually employs 1-D models to estimate the longitudinal temperatures of each wire [7], [3].

Temperatures inside an interconnect wire in a 3-D chip are influenced by several factors, such as the ambient temperatures, substrate temperatures, heat transfer rates of packaging materials, and thermal impedances of vias. However, these factors were often inadequately addressed, and approximations were made in traditional 1-D interconnect models.

Traditional interconnect thermal models are often inflexible to handle distributed heat sources, which can be located at a typical global interconnect wire due to the leakage of current through distributed interconnect capacitances [9].

To address the aforementioned issues, this paper introduces a new interconnect thermal model that adapts to non-uniform substrate temperatures and distributed heat sources. In addition, the new model addresses the effect of interconnect temperature gradients, neglecting which can overestimate the temperature differences along an interconnect wire. The paper obtains the interconnect temperature distribution from the model by solving the voltages in an equivalent transmission line (TL) circuit. Then an efficient $O(n)$ algorithm is introduced to compute the interconnect temperatures.

The rest of the paper is organized as follows. Section II introduces the new interconnect thermal model, derives the temperature distribution solution from the equivalent TL circuit, and introduces the $O(n)$ algorithm to compute interconnect temperatures. Section III presents the experimental results to demonstrate the accuracy of the new model and the efficiency of the $O(n)$ algorithm, by comparisons with FLUENT 3-D thermal simulation.

II. AN ACCURATE INTERCONNECT THERMAL MODEL

A. Modeling Vertical Heat Dissipation

Fig. 1 shows an interconnect wire in a multilayer chip that may include the active region, thermal adhesive layer, heat sink, and packaging materials. As in [7], assume that heat in the interconnect wire flows in two forms. One is called vertical heat dissipation, i.e., heat dissipates vertically through the dielectric materials surrounding the wire. The other is called longitudinal heat conduction, i.e., heat conducts along the interconnect longitudinal direction.

First, consider the vertical heat dissipation. The thermal impedances between the interconnect wire and the surrounding dielectric materials can be estimated by 2-D thermal simulation of the chip cross-sectional area, as shown in Fig.1b. Then the amount of vertical heat dissipation at the longitudinal location $y$ of the wire, denoted by $p_v(y)$, through the surrounding dielectric materials, can be given by

$$p_v(y) = \frac{T(a)}{R} - \frac{\bar{T}}{R} - \frac{T(ta)}{R}$$

where $\bar{T}$ is the respective ambient temperatures on the top and bottom surfaces of the multilayer chip.
for the location \( y \) of the interconnect wire, as shown in Fig.1a and Fig.1b. Determined by 2-D thermal simulation of the chip cross-sectional area, \( R, \bar{R}, \text{ and } R \) are the self thermal impedance of the interconnect wire, thermal impedances from the interconnect wire to the top and bottom surfaces of the multilayer chip, respectively.

**B. Modeling the Effect of Interconnect Temperature Gradients**

Take the three identical heat conduction plates in Fig.2, for example. If the three plates have the same temperature distribution at their boundaries, their interior temperature distributions must be the same because they satisfy the same 2-D Laplace’s equation. Therefore, in this case there is no heat flow among the three plates when they are attached together (Fig.2a). This is an assumption underlying traditional interconnect thermal models. However, when the three plates have different temperatures at their boundaries, e.g., \( T_l < T_r < T_m \), they must have similar interior temperature gradients. That is, at an interior location, the temperature of the left plate is the lowest and that of the middle one is the highest. Therefore, when the three plates are attached together, heat flows from the middle one to the left and right ones, and the vertical heat dissipation of the middle one increases (Fig.2b). The example demonstrates that interconnect temperature gradients affect the amount of vertical heat dissipation. Neglecting the effect can overestimate the temperature gradients in the interconnect wire.

Therefore, introduce an augmented vertical heat dissipation model, \( p^{+}_v(y) \), which linearly approximates the effect of interconnect temperature gradients:

\[
p^{+}_v(y) = p_v(y) - \beta_1 \frac{\partial^2 T(y)}{\partial y^2} + \beta_2 \frac{\partial^2 T^a(y)}{\partial y^2} + \beta_3 \frac{\partial^2 T^u(y)}{\partial y^2}
\]

where coefficients \( \beta_1, \beta_2, \text{ and } \beta_3 \) can be determined experimentally.

**C. The New Interconnect Thermal Model**

Denote the interconnect width by \( w \), thickness by \( t \), length by \( L \), thermal conductivity by \( k \), and the power density at location \( y \) by \( p(y) \). Consider an incremental interconnect segment of length \( \Delta y \), as shown in Fig.1c. The total amount of heat entering the box specified by \([y, y + \Delta y]\) through its left end, denoted by \( q_l \), is given by

\[
q_l = -wtk \frac{dT(y)}{dy}.
\]

By first-order approximation, the total amount of heat leaving the box from its right side, denoted by \( q_r \), is given by

\[
q_r = q_l + \frac{d}{dy} \left[ -wtk \frac{dT(y)}{dy} \right] \Delta y.
\]

The net heat generation in the box, named \( p_{gen} \), is given by

\[
p_{gen} = \left[ p(y) - p^{+}_v(y) \right] w t \Delta y.
\]

The law of energy conservation requires that

\[
q_r - q_l = p_{gen}.
\]

Insert (3), (4), and (5) into (6). Then making \( \Delta y \) infinitesimal leads to the governing equation for interconnect temperatures:

\[
(k + \beta_1) \frac{d^2 T(y)}{dy^2} - \frac{T(y)}{R} = -f(y)
\]

where

\[
 f(y) = p(y) - \frac{T^a(y)}{R} - \frac{T^u(y)}{R} - \beta_2 \frac{\partial^2 T^a(y)}{\partial y^2} - \beta_3 \frac{\partial^2 T^u(y)}{\partial y^2}. \]

The heat convection boundary conditions are specified for (7):

\[
k \frac{dT(y)}{dy} \bigg|_{y=0} = \frac{T(0) - T^a(0)}{R_l}, \quad k \frac{dT(y)}{dy} \bigg|_{y=L} = \frac{T(L) - T^u(L)}{R_r}.
\]
chip can be neglected or combined into (9) by modifying the values of \( R_l, R_r, T^0(0), \) and \( T^a(L). \)

The new model (7) is more general than traditional ones by considering the ambient temperatures, thermal impedances of vias, and the effect of interconnect temperature gradients. As shown in (8), the model can handle distributed heat sources.

D. Solution by Equivalent Transmission Line Circuit

Compare (7) to the \( s \)-domain transmission line equation

\[
\frac{d^2V(y)}{dy^2} - \lambda^2 V(y) = -\lambda Z_c I_c(y)
\]

where \( V, I, \) and \( I_c \) are voltage, current, and external current source for a TL oriented in the \( y \) direction. \( \lambda \) is the TL propagation constant. \( Z_c \) is the TL characteristic impedance.

Clearly, (7) describes an equivalent TL circuit, with temperature \( T(y) \) equivalent to voltage \( V(y) \), as shown in Fig.3. In the circuit, \( \lambda = \frac{1}{\sqrt{(k+\beta_1)R_l}} \) and \( Z_c = \frac{\sqrt{(k+\beta_1)R_l}}{k} \). By (9), the two TL ends are driven by voltage sources \( T^0(0) \) and \( T^a(L) \) through resistors \( R_l \) and \( R_r \), respectively. Along the TL is an external distributed current source \( I_c(y) = \frac{k}{k+\beta_1} f(y) \).

As shown in Fig.3, interconnect temperature \( T(y) \) can be obtained by solving the voltage at the location \( y \) of the TL:

\[
T(y) = T^0(y) + T_f(y) + T_r(y).
\]

Here

\[
T_f(y) = \frac{k}{k+\beta_1} \int_0^L f(y') Z(y|y') dy'
\]

and \( Z(y|y') \) denotes the transfer impedance from the location \( y' \) to the location \( y \) of the TL:

\[
Z(y|y') = Y_l (R_l \cosh \lambda L_l + Z_c \sinh \lambda L_l) [R_r \cosh \lambda L_r + Z_c \sinh \lambda L_r]
\]

where

\[
Y_l = \frac{Z_c}{Z_c (R_l + R_r) \cosh \lambda L + (Z_c^2 + R_l R_r) \sinh \lambda L} \\
L_l = \min(y, y') , L_r = \min(L - y, L - y')
\]

The second term in (10), \( T_f(y) \), is equivalent to the voltage at the location \( y \) of the TL raised by the external distributed current source \( \frac{k}{k+\beta_1} f(y) \) when the voltage sources at the two ends of the TL are both set to zero. \( T^0(y) \) or \( T^a(y) \) is equivalent to the voltage at the location \( y \) of the TL raised by the voltage source \( T^0(0) \) or \( T^a(L) \) when both the external distributed current source and the voltage source at the opposite end of the TL are set to zero:

\[
T^a(y) = T^0(0) \frac{Z_l H_l (L - y)}{Z_l + R_l}, T_r(y) = T^a(L) \frac{Z_r H_r (y)}{Z_r + R_r}
\]

where

\[
H_l(y) = \frac{R_l f(y)}{R_l f(y) \cosh \lambda L + Z_c \sinh \lambda L} \\
H_r(y) = \frac{R_r f(y)}{R_r f(y) \cosh \lambda L + Z_c \sinh \lambda L}
\]

E. An \( O(n) \) Algorithm

In general, \( f(y) \) is given at discrete locations: \( f(y_0), \ldots, f(y_n) \), where \( 0 = y_0 < \cdots < y_n = L \). That is, \( f(y) \) is usually a piecewise-linear function or a piece-wise-smooth function. Using generic numerical integral methods to evaluate (10) for \( T(y) \) at all \( y_0, \ldots, y_n \) requires \( O(n^2) \) time. To improve the efficiency, an \( O(n) \) algorithm is introduced. From (10), \( T(y_i) \) is rewritten in the form of

\[
T(y_i) = T_f(y_i) + \alpha_L(y_i) S_L + \alpha_r(y_i) S_r + T_r(y_i)
\]

and

\[
\alpha_{\{t,r\}}(y) = Y \frac{Y_k}{k + \beta_1} (R_{\{t,r\}} \cosh \lambda y + Z_c \sinh \lambda y)
\]

\[
S_L = \sum_{j=1}^{n-1} \int_{y_j}^{y_{j+1}} f(y) g_l(y) (L - y) dy \\
S_r = \sum_{j=1}^{n-1} \int_{y_j}^{y_{j+1}} f(y) g_r(y) (L - y) dy
\]

End ComputeTempDist

Begin ComputeTempDist
1.let \( S^0_L = 0; \)
2.let \( S^0_r = \sum_{j=0}^{n-1} \int_{y_j}^{y_{j+1}} f(y) g_r(L - y) dy; \)
3. for \( i = 0 \) to \( n \)
   compute \( T^0_f(y_i), T^0_r(y_i), \) \( \alpha_L(y_i), \) and \( \alpha_r(y_i); \)
   let \( T(y_i) = T^0_f(y_i) + T^0_r(y_i) + \alpha_L(L - y_i) S_L + \alpha_r(y_i) S_r; \)
   let \( S^1_L = S^1_L + \int_{y_j}^{y_{j+1}} f(y) g_l(y) (L - y) dy; \)
   let \( S^1_r = S^1_r - \int_{y_j}^{y_{j+1}} f(y) g_r(y) (L - y) dy; \)
4.end for
End ComputeTempDist

Fig.4: Algorithm ComputeTempDist for evaluating the interconnect temperature \( T(y) \) at locations \( y_0, \ldots, y_n \).
p = 2.02 × 10^{13} \text{ W/m}^3. \text{ Two types of dielectric materials were tested: SiO}_2, \text{ with a thermal conductivity of 1.2 W/(m K), and polymer, with a thermal conductivity of 0.3 W/(m K).}

First, thermal impedances for the cross-sectional area of the interconnect array were obtained by 2-D FLUENT simulation. The thermal conductance (1/R) from the cross-sectional boundary of the middle metal line to the substrate was obtained by measuring the total heat flux out of the substrate when applying a 1 °C temperature to the boundary (Fig.5b). Then algorithm ComputeTempDist was used to compute the temperature distribution along the middle metal wire.

Fig.6 compares algorithm ComputeTempDist and FLUENT for the case that s = 0.3 μm and h = t = 0.8 μm. The thermal conductance (1/R) obtained was 3.55 × 10^{12} \text{ W/(K m^3)} when the ILD was SiO_2, and 8.875 × 10^{11} \text{ W/(K m^3)} when the ILD was polymer. The maximum temperature of the middle metal line increased from 5.685 °C to 21.863 °C when the ILD changed from SiO_2 to polymer. As shown in Fig.6b, the temperature errors of algorithm ComputeTempDist were within 2% of the 3-D FLUENT simulation results.

To observe the effect of interconnect temperature gradients, different β_1 factors were tested, while β_2 and β_3 were set to zero. Fig.7 shows the temperature results for the middle metal line with s = 0.5μm, h = 1.6μm, and k_{il} = 0.3 W/(m K). The maximum temperature errors of algorithm ComputeTempDist, when compared to FLUENT, are given in the table. Note that increasing the value of β_1 to 0.8 reduced the maximum temperature error by around 70%, although the maximum temperature error was very small when compared to the actual maximal temperature in the tested case. The above results demonstrate the accuracy of the new model and the effect of interconnect temperature gradients.

IV. CONCLUSIONS

This paper introduces a new interconnect thermal model capable of handling complicated ambient conditions as well as distributed heat sources. The model also considers the effect of interconnect temperature gradients. The paper obtains the temperature distribution from the model by solving the voltages in an equivalent TL circuit. The paper presents an O(n) algorithm to compute the interconnect temperatures. Experimental results have demonstrated the accuracy of the new model and the efficiency of the O(n) interconnect thermal analysis algorithm, by comparisons to 3-D FLUENT simulation.

V. ACKNOWLEDGMENT

This work was supported in part by NSF and AFOSR grants.

REFERENCES