Very Fast Chip-level Thermal Analysis

Keiji Nakabayashi†, Tamiyo Nakabayashi‡, and Kazuo Nakajima*

†Graduate School of Information Science, Nara Institute of Science and Technology
Keihanna Science City, Nara, Japan, keiji-n@is.naist.jp
‡Graduate School of Humanities and Sciences, Nara Women’s University
Kitauoyahigashi-machi, Nara, Japan, nakaba@ics.nara-wu.ac.jp
*Dept. of Electrical and Computer Engineering, University of Maryland
College Park, MD 20742, USA, nakajima@umd.edu

Abstract
We present a new technique of VLSI chip-level thermal analysis. We extend a newly developed method of solving two dimensional Laplace equations to thermal analysis of four adjacent materials on a mother board. We implement our technique in C and compare its performance to that of a commercial CAD tool. Our experimental results show that our program runs 5.8 and 8.9 times faster while keeping smaller residuals by 5 and 1 order of magnitude, respectively.

1. Introduction
As technology advances towards the physical limits of VLSI devices, a variety of physical phenomena play much heavier roles in their performances. For example, heat generation and thermal conduction within silicon substrate and from metal layers have negligible effects on the characteristics of VLSI chips in the post-90 nm era [5]. Such thermal phenomena are also an important factor for system design on a mother board [4] [12] [16].

Among main research activities on thermal analysis [9] [15] is the modeling of thermal conduction by a Laplace equation and its solution by such methods as finite difference (FDM), finite element (FEM), and boundary element methods (BEM) [10] [12] [16]. As such a number of CAD tools for thermal analysis are commercially available. In fact, Raphael from Synopsis was developed as a Laplace equation solver and has successfully been used in many relevant areas [11]. This software uses FDM to discretize a Laplace equation and solves the resultant huge system of linear equations by Incomplete Cholesky Conjugate Gradient method (ICCG) [11].

Indirect or iterative methods such as ICCG can produce numerical solutions with less accuracy for large scale systems of linear equations faster than direct methods such as LU decomposition. Recently, however, a new efficient direct method, called Symbolic Partial Solution Method (S-PSM) was introduced in the area of computational fluid dynamics [2] [3]. This method was derived from the general linear equation solver, called Partial Solution Method (PSM) [8], and is applicable to the case of tridiagonal coefficient matrices. S-PSM was later applied in a Laplace equation solution process for thermal conduction analysis of two adjacent materials [6].

We extend this S-PSM-based Laplace equation solver and present a new chip-level thermal analysis technique. We consider a realistic case of four layer materials of different thermal conductivities on a mother board. We first describe thermal conduction in these materials by a set of Laplace equations. We then use FDM to discretize the equations and generate large scale systems of linear equations. Finally we apply S-PSM to get numerical solutions for the linear systems. We conclude the paper in Section IV.

2. The Problem
Consider a multi-layer structure as depicted in Fig. 2.1, where four layers of materials $p$, $q$, $r$, and $s$ of thermal conductivities $k_p$, $k_q$, $k_r$, and $k_s$, respectively, are stacked together. This structure may represent portion of a mother board design, where the layers are from the bottom, a VLSI chip, a thermal interface material (e.g., adhesive), a heat spreader, and a heat sink. Heat is generated from the heat source of active silicon bulk, which is located at the bottom of the VLSI chip. The heat travels through a heat transfer pass consisting of the chip die, the adhesive, the heat spreader, and the heat sink, and goes out to the ambient air. Our problem is to find temperature distribution through two dimensional steady-state thermal conduction analysis.
Fig. 2.1: Two dimensional analysis for steady-state heat conduction of four materials.

3. S-PSM-based Solution Process

We describe the concept and a computational procedure of the linear equation solver, called S-PSM as applied to Laplace equations. Given a system of linear equations derived by FDM, the S-PSM decomposes it into its subsystems and finds the values of the variables shared by each pair of adjacent subsystems. Figure 3.1 shows an overall flow of the major algebraic computations to take place at each subsystem with their relevant equations and solutions specified.

It should be noted from the figure that the S-PSM-based solution process goes through many levels of repeated operations of decomposition and merging. In the following section, the level information is attached to variable vectors and coefficient submatrices as their superscripts with parentheses such as (0) and (ε).

Fig. 3.1: An overall computation flow with relevant equation numbers of a S-PSM-based solution process.

3.1. Laplace and Finite Difference Equations

Heat diffusion is described by the Laplace equation

$$ k_u \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 $$  \hspace{1cm} (3-1)

where \( u \) is the variable associated with the material \( u \) under consideration and \( k_u \) represents its thermal conductivity. In our analysis, the bold lower case letter \( u \) represents one of the four materials \( p, q, r \) and \( s \) and the lower case letter \( u \) denotes one of their corresponding variable names \( p, q, r \) and \( s \). So we have a set of four Laplace equations to solve.

When we apply FDM to Eq. (3-1), we decompose each of the corresponding rectangle domains into a grid of \((n + 2) \times m\) points to which variables are assigned as shown inside each rectangle of Fig. 3.2. Furthermore, two sets of \((n + 2)\) variables are assigned to the boundary points on the left and right borders and two sets of \( m \) variables to those on the upper and lower borders of each rectangle in the figure. They all represent the boundary conditions to be described later.

Fig. 3.2: The arrangement of interior grid and boundary points for the four material domains.

For the sake of simplicity, we make two assumptions:

1. Each cell of the four grids is a square of the same size, that is, \( \Delta x = \Delta y \).

2. The number of X-axis grid points for each material \( u \) (\( p, q, r \), \( s \)) is a power of 2, that is, \( n_p = 2^{e_p}, n_q = 2^{e_q}, n_r = 2^{e_r}, n_s = 2^{e_s} \) \hspace{1cm} (3-2)
In the following, whenever two indices \( n \), their relevant vectors, their respective ranges are \( i = 1, 2, \ldots, n \) and \( j = 1, 2, \ldots, m \). As the thermal conductivity is the same at each grid point in each domain, the application of FDM yields

\[
k_i \left( \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{\Delta x^2} + \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{\Delta y^2} \right) = 0
\]

(3-3)

As \( \Delta x = \Delta y \), we derive

\[
u_{i+1,j} + u_{i+1,j} + u_{i,j+1} + u_{i,j+1} - 4u_{i,j} = 0
\]

(3-4)

### 3.2. Boundary Conditions

As shown in Fig. 2.1, the upper and lower boundary conditions are the ambient temperature \( T_{amb} \) of 55°C and the heat source temperature \( T_j \) of 120°C. Likewise the left and right boundary conditions are the reference temperatures \( T_{left} \) and \( T_{right} \) of 20°C each.

The left and right boundary conditions for each material \( u \) are expressed in the equations as

\[
f_u^{(0)} = (-u_{i,0} 0 0 \ldots 0 -u_{i,m+1})^T
\]

(3-5)

The top and bottom boundary conditions for the combined four layers of materials are given as

\[
v^{(0)} = (v_{0,1} v_{0,2} \ldots v_{0,n-1} v_{0,m})^T
\]

\[
w^{(0)} = (w_{0,0} v_{0,2} \ldots w_{0,n-1} w_{0,m})^T
\]

(3-6)

(3-7)

Note that the superscripts (0) for the above vectors indicate that their element values are given at the start of the S-PSM process.

We set the thermal conductivity of the boundary between adjacent materials, say \( q \) and \( r \) of thermal conductivities \( k_q \) and \( k_r \) as

\[
k_q = \frac{2k_k k_k}{k_k + k_r}
\]

(3-8)

At the boundary between each pair of materials, we use the first order approximation for heat conduction as follows. From the viewpoint of material \( q \), the following difference equation holds at its boundary with material \( p \):

\[
k_{p} q_{i+1,j} - q_{i,j} + k_{q} q_{i,j} - q_{i,j} + k_{q} q_{i,j} - q_{i,j} = 0
\]

(3-9)

Similarly, at its boundary with material \( r \), we have

\[
k_{r} q_{i+1,j} - q_{i,j} + k_{q} q_{i,j} - q_{i,j} + k_{q} q_{i,j} - q_{i,j} = 0
\]

(3-10)

Thus, using the variable vector notation

\[
u^{(0)} = (u_{1,1} u_{1,2} u_{1,3} \ldots u_{1,n-1} u_{1,m})^T
\]

(3-11)

we have the following four equations for the pair of adjacent materials, \( q \) and \( r \).

We now get the following equations for the first and second material,

\[
\begin{align*}
q_1^{(1)} &= f_1^{(1)} - A_1^{(1)} q_1^{(1)} - A_1^{(1)} q_2^{(1)} \\
q_2^{(1)} &= f_1^{(1)} - A_1^{(1)} q_2^{(1)} - A_1^{(1)} q_3^{(1)} \\
r_1^{(1)} &= f_2^{(1)} - A_2^{(1)} r_1^{(1)} - A_2^{(1)} r_2^{(1)} \\
r_2^{(1)} &= f_2^{(1)} - A_2^{(1)} r_2^{(1)} - A_2^{(1)} r_3^{(1)}
\end{align*}
\]

(3-12)

\[
\begin{align*}
k_{p} q_{i+1,j} + A_{q}^{(1)} q_{i}^{(1)} + k_{q} q_{i,j} + \frac{n}{2} r_{i,j}^{(1)} + k_{q} r_{i,j} + \frac{n}{2} q_{i,j}^{(1)} = k_{r} f_{r}^{(1)} \\
k_{r} r_{i,j} + A_{r}^{(1)} r_{i,j} + k_{r} r_{i,j} + \frac{n}{2} q_{i,j}^{(1)} + k_{r} q_{i,j} + \frac{n}{2} r_{i,j}^{(1)} = k_{p} f_{p}^{(1)}
\end{align*}
\]

(3-13)

We apply the above procedure of derivation of Eqs. (3-12), (3-17) and (3-18) to the remaining boundaries. Then we get the following matrix-vector form of Eq. (3-19) for the three boundaries:
\[
\begin{align*}
\text{3.3. System Decomposition and Partial Solutions for Each Subsystem/Material} & \\
\text{We decompose Eq. (3-19) into four subsystems of} & \\
\text{equations that correspond to the four materials.} & \\
\begin{pmatrix}
A_1^{(p)} & I & O & O \\
A_2^{(p)} & I & O & A_2^{(p)} \\
A_3^{(p)} & O & I & A_3^{(p)} \\
O & k_0 & A_4^{(p)} & A_4^{(p)}
\end{pmatrix}
\begin{pmatrix}
p_0^{(p)} \\
p_0^{(p)} \\
p_0^{(p)} \\
p_0^{(p)}
\end{pmatrix}
& = 
\begin{pmatrix}
0 \\
v_0^{(p)} \\
0 \\
v_0^{(p)}
\end{pmatrix}
& \\
& +
\begin{pmatrix}
f_0^{(p)} \\
f_0^{(p)} \\
f_0^{(p)} \\
f_0^{(p)}
\end{pmatrix}
& (3-20)
\end{align*}
\]

\[
\begin{align*}
\text{We denote the inverse matrix of each coefficient matrix of the above equations by the matrix of 16} & \\
\text{submatrices of the form} \ B^u & \\
\text{as shown below. The partial solutions for each subsystem are given as} & \\
\text{follows.} & \\
\text{(For the correspondence between the blocks of the inverse matrix above and the} & \\
\text{B matrix below, see [7].)} & \\
\begin{pmatrix}
p_0^{(p)} \\
p_0^{(q)} \\
p_0^{(u)} \\
p_0^{(v)}
\end{pmatrix}
& = 
\begin{pmatrix}
B_{01} & B_{02} & B_{03} & B_{04} \\
B_{01} & B_{02} & B_{03} & B_{04} \\
B_{01} & B_{02} & B_{03} & B_{04} \\
B_{01} & B_{02} & B_{03} & B_{04}
\end{pmatrix}
\begin{pmatrix}
v_0^{(p)} \\
0 \\
0 \\
0
\end{pmatrix}
& \\
& +
\begin{pmatrix}
f_0^{(p)} \\
f_0^{(p)} \\
f_0^{(p)} \\
f_0^{(p)}
\end{pmatrix}
& (3-24)
\end{align*}
\]

\[
\begin{align*}
\text{As both of the above coefficient matrices are of} & \\
\text{special structure, their inverse matrices are expressed as} & \\
\text{shown below with two full middle columns of} & \\
\text{submatrices of the forms} \ B^p & \\
\text{and} \ B^q, & \\
\text{respectively. The partial solutions for each pair of materials are now} & \\
\text{obtained as follows.} & \\
\text{(See [7].)} & \\
\begin{pmatrix}
p_0^{(p)} \\
p_0^{(q)} \\
p_0^{(u)} \\
p_0^{(v)}
\end{pmatrix}
& = 
\begin{pmatrix}
I & B_{01}^{p} & B_{02}^{p} & \cdots \\
0 & B_{01}^{p} & B_{02}^{p} & \cdots \\
0 & 0 & B_{01}^{p} & \cdots \\
0 & 0 & 0 & B_{01}^{p}
\end{pmatrix}
\begin{pmatrix}
p_0^{(p)} \\
p_0^{(p)} \\
p_0^{(p)} \\
p_0^{(p)}
\end{pmatrix}
+ 
\begin{pmatrix}
f_0^{(p)} \\
f_0^{(p)} \\
f_0^{(p)} \\
f_0^{(p)}
\end{pmatrix}
& (3-31)
\end{align*}
\]

\[
\begin{align*}
\text{3.5. Final Solutions} & \\
\text{Finally, for the two pairs of materials} \ p & \\
\text{and} \ q, & \\
\text{and of} \ r & \\
\text{and} \ s, \text{we extract equations with respect to variables} & \\
\text{and} \ r^{(s)} & \\
\text{from Eqs. (3-31) and (3-32). Their merging then produces the following equation.} & \\
\begin{pmatrix}
1 & \cdots \\
0 & \cdots \\
0 & \cdots \\
0 & \cdots
\end{pmatrix}
& \\
\times
\begin{pmatrix}
B_{01}^{p} & B_{02}^{p} & \cdots \\
B_{01}^{q} & B_{02}^{q} & \cdots \\
B_{01}^{u} & B_{02}^{u} & \cdots \\
B_{01}^{v} & B_{02}^{v}
\end{pmatrix}
& \\
\begin{pmatrix}
v_0^{(p)} \\
v_0^{(q)} \\
v_0^{(u)} \\
v_0^{(v)}
\end{pmatrix}
& +
\begin{pmatrix}
f_0^{(p)} \\
f_0^{(p)} \\
f_0^{(p)} \\
f_0^{(p)}
\end{pmatrix}
& (3-33)
\end{align*}
\]
where
\[
\mathbf{f}_{ij} = B_{i}^{(m)q} \mathbf{f}_{q} + B_{i}^{(n)q} \mathbf{f}_{q} + \mathbf{f}_{q}
\]

\[
\mathbf{f}_{ij} = B_{i}^{(m)q} \mathbf{f}_{q} + B_{i}^{(n)q} \mathbf{f}_{q} + \mathbf{f}_{q}
\]

We then take the inverse matrix of the coefficient matrix of Eq. (3-33) and derive the final solutions for variables \( q_{0}^{(m)} \) and \( u_{0}^{(m)} \) as follows. (See [7].)

\[
q_{0}^{(m)} = \mathcal{L}_{0}^{-1} \left( \mathcal{L}_{0} \cdot q_{0}^{(m)} \right)
\]

\[
u_{0}^{(m)} = \mathcal{L}_{0}^{-1} \left( \mathcal{L}_{0} \cdot \nu_{0}^{(m)} \right)
\]

\[
T_{j} = 120 \degree C
\]

\[
\text{room temperature of } 20 \degree C
\]

Reference [10] used the reference temperature for our analysis. This implies that our technique produces correct results.

Table I shows the CPU times required and the residuals produced by our program and Raphael. The results demonstrate that for the largest grid, our program ran 5.8 and 8.9 times faster while keeping smaller residuals by 5 and 1 order of magnitudes, than Raphael.

Table I: Processing CPU times of and residuals by our program and Raphael for difference arrangements of grid points.

| Number of grid points | S-PSM CPU(s) | Raphael CPU(s) | Raphael CPU(s) |
|-----------------------|--------------|----------------|----------------|
|                       | Residual     | EPS=1.0E-10    | EPS=1.0E-10    |
| 128                   | 0.05         | 0.26           | 0.26           |
| 256                   | 0.05         | 0.26           | 0.26           |
| 512                   | 0.05         | 0.26           | 0.26           |
| 1,024                 | 0.05         | 0.26           | 0.26           |

4. Experiments

We applied the technique described above to two dimensional steady-state heat conduction analysis at the chip level of four material layers mounted on a mother board. They are a VLSI chip, a thermal interface material (e.g., adhesive), a heat spreader, and a heat sink as shown in Fig. 2.1. The parameter values, such as the thermal conductivity and the depth and width of each material given in the figure are those used in the industrial design settings.

In our example, we assume that the active silicon bulk is the only heat source and that no Joule heat is generated from the metal layer of wire. So the temperature of the heat source, which is the value of junction temperature of \( T_{j} = 120 \degree C \), and the ambient temperature of \( T_{amb} = 55 \degree C \) are set as the boundary conditions at the bottom and top edges, respectively, of the region analyzed. These values are also used in the industrial design settings. As for the left and right boundary conditions, we set \( T_{right} = T_{left} = 20 \degree C \) as the reference (room) temperatures for our analysis. This room temperature of 20 \degree C was also used in reference [14]. Reference [10] used the reference temperature of 0 \degree C.

To evaluate the performance of our technique, we developed a C program for solving the four Laplace equations described above. Although it took more time to develop this C code than expected, the resultant program was rather simple. We ran it on an IBM Z Pro (OS Linux, CPU Xeon 3.60GHz, memory 32GB). We also ran a commercial Laplace equation solver software, called Raphael [11]. Figure 4.1 depicts the temperature distributions obtained by our program and Raphael. Both figures resemble very close to each other. This implies that our technique produces correct results.

Table I shows the CPU times required and the residuals produced by our program and Raphael. The results demonstrate that for the largest grid, our program ran 5.8 and 8.9 times faster while keeping smaller residuals by 5 and 1 order of magnitudes, than Raphael.

Table I: Processing CPU times of and residuals by our program and Raphael for difference arrangements of grid points.
5. Conclusions

We have proposed a new technique of thermal analysis for multi-layer VLSI chips. It starts with the modeling of multiple layers of materials of different heat conductivities by a set of Laplace equations. It then discretizes the equations by FDM and finally uses S-PSM to obtain numerical solutions for the resulting systems of linear equations. We applied our technique to steady-state heat conduction analysis at the chip level for a realistic case of four layer materials. The experimental results showed 5.8 and 8.9 times speed-up and simultaneous residual improvement by 5 and 1 order of magnitude, respectively, by our program over the commercial tool, called Raphael.

In our experiments, we assumed that the temperature of the heat source was given. The generation of the heat may be expressed by a Poisson equation and its solution could easily be computed by S-PSM. Extension of our work to the process of a Poisson equation and other cases, such as three dimensional and transient heat conduction analysis are under way. We can also easily extend our work to the case of complex shapes and boundary conditions of materials.

6. Acknowledgment

The authors would like to express their gratitude to Dr. T. Hiramoto for his introduction of S-PSM to them. They are also grateful to anonymous reviewers for their suggestions.

References

[1] Y. K. Cheng, P. Raha, C. C. Teng, E. Rosebaum, and S. M. Kang, “ILLIADS-T: An Electrothermal Timing Simulator for Temperature-sensitive Reliability Diagnosis of CMOS VLSI Chips,” IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems, Vol. 17, No. 8, pp. 661-681, Aug. 1998.

[2] T. Hiramoto, M. Osano, and K. Oshima “A New Efficient Solver for Huge Linear Algebraic Equations Derived from Discretization of the Fluid Dynamic Equations,” Computational Fluid Dynamics Jour., Vol. 11, No. 3, pp. 346-357, Oct. 2002.

[3] T. Hiramoto, M. Osano, and K. Oshima, “Symbolic PSM Solver for the Finite Difference Equations Derived from the Poisson Equation,” Computational Fluid Dynamics Jour., Vol. 12, No. 1, pp. 19-28, Apr. 2003.

[4] W. Huang, M. R. Stan, K. Skadron, K. Sankaranarayanan, S. Ghosh, and S. Velusamy, “Compact Thermal Modeling for Temperature Aware Design,” Proc. 41st Design Automation Conf., San Diego, CA, June 2004, pp. 878-883.

[5] S. Im, N. Srivastava, K. Banerjee, and K. E. Goodson, “Scaling Analysis of Multilevel Interconnect Temperatures for High-Performances ICs,” IEEE Trans. on Electron Devices, Vol. 52, No. 12, pp. 2710-2717, Dec. 2005.

[6] K. Nakabayashi, T. Hiramoto, M. Osano, and K. Oshima, “Mixed Interface on Symbolic PSM Solver for the Finite Difference Equations,” Proc. 5th Asia Workshop on Computational Fluid Dynamics, Xian, China, Aug. 2006.

[7] K. Nakabayashi, T. Nakabayashi, and K. Nakajima, “A New Technique of Multi-layer Thermal Analysis for VLSI Chips,” manuscript in preparation for publication.

[8] M. Osano, K. Nakajima, and M. Tanimoto, “A New Efficient Solution Method for a System of Linear Equations: Partially Solving Method (PSM),” Japan Jour. of Industrial Math., Vol. 13, No. 2 pp. 243-256, June 1996.

[9] M. Pedram, and S. Nazarain, “Thermal Modeling, Analysis, and Management in VLSI Circuits: Principles and Methods,” Proceedings of the IEEE, Vol. 94, No. 8, pp. 1487-1501, Aug. 2006.

[10] 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-A, Vol. 21, No. 3, pp. 406-411, Sept. 1998.

[11] Synopsys, Raphael Interconnect Analysis Program Tutorial, Mar. 2006.

[12] T-Y. Wang and C. C-P. Chen, “3-D Thermal-ADI: a Linear-Time Chip Level Transient Thermal Simulator,” IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems, Vol. 21, No. 12, pp. 1434-1445, Dec. 2002.

[13] T. Y. Wang and C. C-P. Chen, “Spice-compatible Thermal Simulation with Lumped Circuit Modeling for Thermal Reliability Analysis Based on Model Reduction,” Proc. 5th Int’l Symp. on Quality Electronic Design, San Jose, CA, Mar. 2004, pp. 357-362.

[14] B. Wang and P. Mazumder, “Accelerated Chip-Level Thermal Analysis Using Multilayer Green’s Function,” IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems, Vol. 26, No. 2, pp. 325-344, Feb. 2007.

[15] Z. Yu, D. Versege, R. W. Dutton, S. Nakagawa, N. Chang, S. Lin, and W. Xie, “Full Chip Thermal Simulation,” Proc. 1st Int’l Symp. on Quality Electronic Design, San Jose, CA, Mar. 2000, pp. 145-149.

[16] Y. Zhan, B. Goplen, and S. S. Sapatnekar “Electrothermal Analysis and Optimization Techniques for Nanoscale Integrated Circuits,” Proc. 2006 Conf. on Asia South Pacific Design Automation, Yokohama, Japan, Jan. 2006, pp. 219-222.