A Multi-Scale Random-Walk Thermal-Analysis Methodology for Complex IC-Interconnect Systems*

R.B. Iverson, Y.L. Le Coz
Center of Integrated Electronics, CII-6225
Rensselaer Polytechnic Institute
Troy, NY 12180 USA

B. Kleveland, S.S. Wong
Center for Integrated Systems
Stanford University
Stanford, CA 94305 USA

Abstract—**We have developed and demonstrated a 3D multi-scale thermal-analysis methodology for multiple and stacked-chip configurations. This approach employs a global-local problem-domain discretization in conjunction with the floating RW (Random-Walk) method. Emphasis has been placed on capturing complex thermal effects due to interconnect layers. We have analyzed a hypothetical stacked-chip geometry derived from a Stanford interconnect test chip. 2D gdsII layout data was automatically processed and converted into a 3D problem domain. On a 400MHz Apple PowerBook G3™, execution time was about three minutes per temperature data point. Temperature at each evaluated point was computed with 1000 RWs, yielding a 1-σ statistical error of about 5%. Based on heuristic formulas that we have deduced, a local window of ±20μm relative to the RW start point achieved a reasonable global-local discretization error.**

I. Introduction

We have developed a 3D multi-scale thermal analysis methodology for multiple and stacked-chip configurations. Emphasis has been placed on capturing effects arising from the complex conductor-dielectric interconnect structures found within modern, high-performance digital ICs. We employ a novel global-local floating RW (Random-Walk) method to solve Poisson's equation at any particular point of interest:

$$\nabla \cdot [-\kappa(\mathbf{r}) \nabla T] = G(\mathbf{r}). \qquad (1)$$

T is the 3D temperature field, G is the volumetric heat-generation rate (W/cm^3), and κ is the thermal conductivity (W/cm-K). The source function G represents electrical power dissipation within transistors and interconnects. We model $\kappa(\mathbf{r})$ as a piecewise-constant function.

Our approach is to use a global-local multi-scale procedure in conjunction with the floating RW method. We present here a comprehensive methodology that depends solely on the RW method. The RW method has previously proven its numerical efficiency in the area of IC-interconnect capacitance extraction.[1] It has also been successfully used with finite-element and variational methods in the global-local thermo-mechanical analysis of multi-chip modules.[2,3]

In brief, the RW method requires no numerical meshing, efficiently calculates point-to-point field values, and is eminently suitable for parallel computation.

II. Multiscale Analysis

A. Random-Walk Algorithm

The commercial RW Laplace solver *QuickCap API*™ was used as the computational kernel for the "Laplace" part of the problem ($G = 0$). Each step of a walk passing through a region of power generation was handled externally from *QuickCap* by adding a temperature contribution ΔT to the RW temperature estimate. RWs terminate at a user-defined Dirichlet boundary (ambient or package). The boundary temperature is added to the sum of ΔTs along the walk, yielding an estimate for T at the origin of the walk. As more RW estimates for a given start point are averaged together, the accuracy improves.

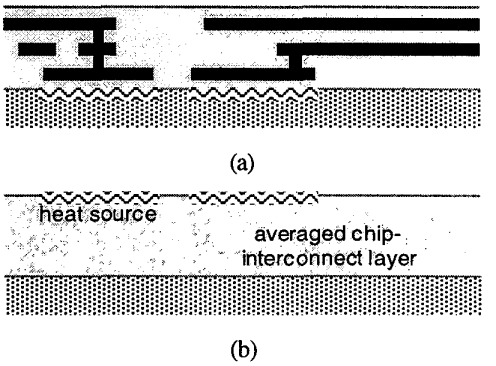

Fig. 1. Detailed and averaged Poisson environments experienced by a random walk. (a) DPE: Interconnect embedded in dielectric with underlying planar heat-generation sources (wavy lines). (b) APE: Global geometry in which the interconnect layers are replaced by a single effective thermal conductor. Power sources are nudged upward to facilitate RW execution.

* This work has been supported by the Defense Advanced Research Projects Agency (DARPA), the New York State Science and Technology Foundation (NYSSTF), and the SRC Microelectronics Advanced Research Corporation (MARCO).

0-7803-6279-9/00/$10.00 © 2000 IEEE

At the beginning of each RW, the computational environment includes all interconnect structures. We call this local description a Detailed Poisson Environment (DPE). A user-defined local window bounds the origin by $\pm W$. When the walk exits the local window, the DPE is replaced with an averaged global description for the remainder of that walk. We call this global description the Averaged Poisson Environment (APE). DPE and APE heat-generation sources were modeled as planes. Figures 1 depict the DPE and APE. The APE is user defined. Also, to facilitate RW evaluation, APE heat-generation planes were moved slightly (7μm), to the top of the chip from the top of the underlying *Si* substrate material.

We have analytically calculated temperature at the center of a materially uniform cube bisected by a uniform planar heat source. Cube boundaries were held at $T = 0$. We thus obtained a numerical prefactor for calculating temperature contribution due to a random step from the center of a cube to its surface:

$$\Delta T = 0.2819 gL / \kappa , \qquad (2)$$

where g is the heat generation rate in the plane (W/cm^2), κ is a constant thermal conductivity (W/cm-K), and $2L$ is the cube edge length (cm). The numerical prefactor in formula (2) was verified by analysis of a large planar heat source midway between two large planar Dirichlet boundaries. The observed discrepancy between the RW and exact analytical results was less than 1%.

B. DPE Window Size

As mentioned earlier, a user-defined window size W establishes the point in a RW at which the DPE is replaced with the APE. To help guide the selection of W, we have developed approximate, heuristic formulas to determine when a walk is likely to leave a given interconnect wire and "wander" into an adjacent one. To find the distance beyond which a RW has a high probability of escaping an interconnect, we consider the quasi-1D dependence $T(z)$ along the axis of a highly conductive cylindrical wire, centered (co-axially) within a cylinder of temperature $T = 0$. The distance in the wire where T decays by one-half is the distance where a RW has a 50% chance of escaping the wire. We have

$$T(z) = T_0 \exp(\pm z/\alpha), \qquad (3)$$

for which $\alpha^2 = \rho_0^2 \kappa_0 \ln(\rho_1/\rho_0)/2\kappa_1$.

In Eq. (3), (κ_0,ρ_0) and (κ_1,ρ_1) are the thermal conductivity and radius of the wire and surrounding cylinder, respectively. Given, for example, a 0.25μm *Cu*-interconnect layout, we can use $\rho_0 = 0.125\mu m$ and $\kappa_0 = 4$W/cm-K to represent a *Cu* wire; and $\rho_1 = 0.375\mu m$ and $\kappa_1 = 0.014$W/cm-K to represent the SiO_2 dielectric between wires. These values yield an average escape distance of $\alpha \approx 1.5\mu m$. A window size $W = 10\mu m$ would be seem quite justified in this case: 99.9% of the walks would have escaped to a neighboring wire in 10μm. We have also obtained a similar conclusion using a heuristic formula for a planar interconnect geometry.

III. Computational Results

A. RW Thermal Solver

We have created a RW thermal solver that processes a structure file containing material parameters and generates the necessary DPE and APE. Interactive user input is then used to generate thermal solution data.

The structure file defines the problem bounds, the boundary temperature, the DPE, and the APE. DPE specification involves references to secondary input files containing chip-technology data and 2D gdsII layout data.

Materials in the structure file are defined by blocks (x_0, y_0, z_0 to x_1, y_1, z_1) of specified thermal conductivities. Materials in the chip-technology file are defined by layers (z_0 to z_1) of specified thermal conductivities. For interconnects, a gdsII layer ID is specified. For background dielectric layers, no layer ID is specified.

Power layers are defined in the chip technology file by planes of specified power generation rates (W). Although gdsII layer IDs can be specified to allow nonuniform power generation, this feature was not tested in this study.

In the DPE, materials and power planes defined in the chip files are superimposed on materials defined in the structure file. In the APE, the chip data is ignored, except for the power planes, which are moved to the "top" of each chip. (The top is the maximum chip z before performing any spatial transformation such as flipping.)

B. Test Chip

In digital circuits, the existence of several metal layers without a reference ground plane gives rise to uncertain current return paths for high-speed clocks and buses. Parasitic interconnect inductance can cause undesired clock skew or ringing. A Stanford test chip with random interconnect lines, a power-distribution grid, and ring-oscillator circuits was previously designed to characterize the inductance of realistic on-chip interconnects.[4]

TABLE I
TEST-CHIP SPECIFICATIONS FOR EACH OF TWO CHIPS.

Specification
5 metal interconnect layers
7.8M geometric-element primitives
0.5μm interconnect feature scale
10mm × 5mm chip size
3.2W/cm^2 power consumption*

*Power consumption was taken uniform within a planar layer in each chip.

Our thermal-analysis methodology was used to analyze a 3D stacked IC-chip geometry derived by scaling the Stanford test-chip layout *act15top*. Scaled test-chip specifications are listed in Table I. The test chip was duplicated and inverted to construct the 3D structure shown in Fig. 2. The bottom chip rests on 200μm of SiO_2. Both chips have 200μm substrates and are separated from each other by 200μm of SiO_2. *Cu* blocks at left and right edges provide a thermal path between the two chips. Table II lists thermal-conductivity data.

TABLE II
THERMAL CONDUCTIVITY κ OF MATERIALS USED IN THE STANFORD STACKED TEST-CHIP EXAMPLE.

Material	κ (W/cm-K)
copper	4.0
silicon	1.0
silicon dioxide	0.014
air	0.0003

C. Temperature Profile

RW results for temperature deviation from ambient (T = 0) across each chip centerline (left-to-right in Fig. 2), at the *Si* substrate-interconnect interfaces, are plotted in Fig. 3. Each temperature point was generated from 1000 RWs with a 1-σ statistical error of about 5%. The DPE window size W corresponded to ±20μm relative to the RW start point. W is about seven times the mean wire-escape distance. Less than 0.1% of the walks that reach the edge of the DPE window would remain in a single wire. On a 400MHz Apple PowerBook G3™, execution time was about three minutes per data point, or two hours overall.

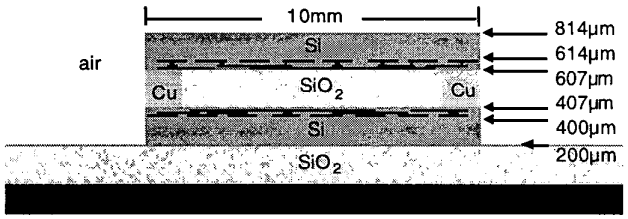

Fig. 2. Cross-sectional view of the stacked-chip structure. Interconnect layers are within 7μm of the chip surfaces. Chips are separated by a 200μm interfacial *Cu*/SiO_2 layer. They lie on a 200μm SiO_2 layer situated on a 10cm×10cm substrate (in black, at bottom) where T = 0. The computational domain extends up to z = 10cm and out to x = ±10cm, and y = ±10cm, where we apply an ambient boundary condition T = 0.

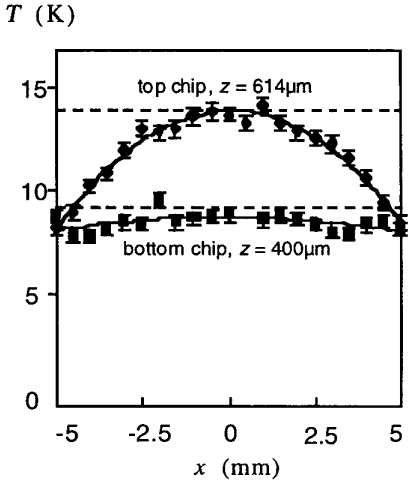

Fig. 3. Stacked test-chip results. 1-σ error bars are shown for each temperature point. Each point corresponds to 1000 RWs. The solid lines are 2nd-order polynomial fits to the data. Near the chip centers x = 0mm, ambient temperature deviation approaches analytical 1D solutions (dashed lines). Heat flow through *Cu* at the chip edges (x = ±5mm) causes an expected reduction in top-chip temperature.

IV. CONCLUSION

We have developed and demonstrated a 3D multi-scale thermal-analysis methodology for multiple and stacked-chip configurations. This approach employs a global-local problem-domain discretization in conjunction with the floating RW method. In the future, we will test the procedure for nonuniform power generation. We will also apply our methodology to realistic structures.

REFERENCES

[1] Y.L. Le Coz, H.J. Greub, and R.B. Iverson, *Solid-St. Electron.*, **42**, pp.58–68, 1998.

[2] T.-L. Sham, Y.L. Le Coz, *et al.*, *ASME Int'l. Electron. Packaging Conf.*, pp. 551–62, **4-2**, Sept. 1993.

[3] M.S. Shephard, Y.L. Le Coz, R.B. Iverson, *et al.*, *Int'l. Symposium on Highly Advanced Computing*, pp. 123–34, Aug. 1994.

[4] B. Kleveland, X. Qi, L. Madden, R. W. Dutton, S. S. Wong, *IEDM*, 1999, pp. 901–2.