Thermal Via Placement in 3D ICs

Brent Goplen
Department of Electrical and Computer Engineering
University of Minnesota
200 Union ST SE, Minneapolis, MN 55455
Phone: (612) 626-7162

bgoplen@ece.umn.edu

Sachin Sapatnekar
Department of Electrical and Computer Engineering
University of Minnesota
200 Union ST SE, Minneapolis, MN 55455
Phone: (612) 625-0025

sachin@ece.umn.edu

ABSTRACT

As thermal problems become more evident, new physical design paradigms and tools are needed to alleviate them. Incorporating thermal vias into integrated circuits (ICs) is a promising way of mitigating thermal issues by lowering the thermal resistance of the chip itself. However, thermal vias take up valuable routing space, and therefore, algorithms are needed to minimize their usage while placing them in areas where they would make the greatest impact. With the developing technology of three-dimensional integrated circuits (3D ICs), thermal problems are expected to be more prominent, and thermal vias can have a larger impact on them than in traditional 2D ICs. In this paper, thermal vias are assigned to specific areas of a 3D IC and used to adjust their effective thermal conductivities. The thermal via placement method makes iterative adjustments to these thermal conductivities in order to achieve a desired maximum temperature objective. Finite element analysis (FEA) is used in formulating the method and in calculating temperatures quickly during each iteration. As a result, the method efficiently achieves its thermal objective while minimizing the thermal via utilization.

Categories and Subject Descriptors

B.7.2 [**Integrated Circuits**]: Design Aids – *Placement and routing*; B.7.1 [**Integrated Circuits**]: Types and Design Styles – *Advanced technologies*

General Terms

Algorithms, Performance, Design, Experimentation

Keywords

3-D IC, 3-D VLSI, thermal optimization, temperature, thermal gradient, placement, routing, finite element analysis, thermal via

1. INTRODUCTION

As the technology node progresses, chip areas and wire lengths continue to increase, causing such problems as increased interconnect delays, power consumption, and temperature, all of

which can have serious implications on reliability, performance, and design effort. Three dimensional technology attempts to overcome some of these limitations by stacking multiple active layers into a monolithic structure, using special processing technologies such as silicon-on-insulator (SOI) or wafer bonding [1]. By expanding vertically rather than spreading out over a larger area, the chip space is better utilized, interconnects are decreased, and transistor packing densities are increased, leading to better performance and power efficiency [2]. Despite the advantages that 3D ICs have over 2D ICs, thermal effects are expected to be more pronounced because of higher power densities and greater thermal resistance along heat dissipation paths. With the advent of better processing technologies for 3D ICs, design tools are needed to realize their full potential and overcome thermal and efficiency issues. Current design tools used for 2D ICs can not be easily extended to 3D ICs [2] especially when taking into account thermal effects.

The idea of using thermal vias to alleviate thermal problems was first utilized in the design of packaging and printed circuit boards (PCBs). Lee *et al.* studied arrangements of thermal vias in the packaging of multichip modules (MCMs) and found that as the size of thermal via islands increased, more heat removal was achieved but less space was available for routing [3]. Li studied the relationships between design parameters and the thermal resistance of thermal via clusters in PCBs and packaging [4]. These relationships were determined by simplifying the via cluster into parallel networks using the observation that heat transfer is much more efficient vertically through the thickness than laterally from heat spreading. Pinjala *et al*. performed further thermal characterizations of thermal vias in packaging [5].

Chiang *et al*. first suggested that "dummy thermal vias" can be added to the chip substrate as additional electrically isolated vias to reduce effective thermal resistances and potential thermal problems [6]. A number of papers have addressed the potential of integrating thermal vias directly inside chips to reduce thermal problems internally [6,7,8,9]. Because of the many dielectric layers, thermal problems are greater and thermal vias can have a larger impact in 3D ICs than 2D ICs. In addition, interconnect structures can create efficient thermal conduits and greatly reduce chip temperatures.

It has become of particular interest to design efficient heat conduction paths right into a chip to eliminate localized hot spots directly. Despite all the work that has been done in evaluating

thermal vias, thermal via placement algorithms are lacking even with 2D ICs. As circuits and temperature profiles increase in complexity, efficient algorithms are needed to accurately determine the location and number of internal thermal vias. The thermal via placement method presented in this paper uses designated thermal via regions to place thermal vias and efficiently adjusts the density of thermal vias in each of these regions with minimal perturbations on routing. In the design process, thermal via placement can be applied after placement and before routing.
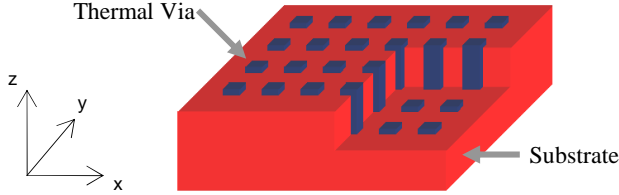
2. THERMAL VIA REGIONS

Figure 1. Uniform density of thermal vias within a region.

In order to make the placement of thermal vias more manageable, certain areas of the chip are reserved for placing thermal vias. In thermal via regions as shown in Figure 1, a uniform density of thermal vias is used, and the thermal via placement algorithm determines the density in each of these regions. In a 3D IC as shown in Figure 2, these regions are evenly placed between the rows in a standard cell design.
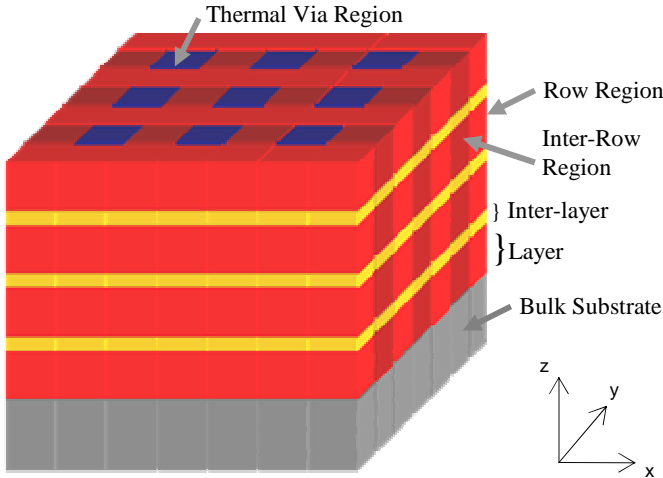

Figure 2. Thermal mesh for a 3D IC, with thermal via regions.

The thermal via regions are composed of electrically isolated vias and are oriented vertically between the rows. The density of the thermal vias determines the thermal conductivity of the region which in turn determines the thermal properties of the entire chip. They are generally obstacles to routing except for regions which require only a low density of thermal vias. Placing thermal vias in specific regions allows for predictable obstacles to routing. The density of these routing obstacles is limited in any particular area so that the design does not become unroutable. It also allows for regularity and uniformity in the entire design process.

The value of the thermal conductivity, K, in any particular direction corresponds to the density of thermal vias that are

arranged in that direction. Increasing the number of thermal vias in one direction does increase the thermal conductivity in the other directions but at an order of magnitude less. For simplicity, the interdependence can be considered to be negligible, and the K's in the x, y, and z directions can be considered to be independent to a certain extent.

Current integration technologies for producing 3D ICs result in the layers being closely stacked together and the design space being tightly compressed in the z direction. In addition, the location of the heat sink in relation to the heat sources produces a heat flux that is primarily downward in direction with very minor lateral components. Furthermore, with the thermal via regions being oriented vertically, lateral thermal vias would have little effect. As a result, lateral thermal conductivities (in the x and y directions) are generally unchanged by this method because the thermal gradients in the vertical (z) direction are almost two orders of magnitude larger than in the lateral directions. In this paper, the method will be developed for all three directions, but for the reasons outlined above, only vertically oriented thermal vias will be considered in the implementation and results.

3. TEMPERATURE CALCULATION

At steady state, heat conduction within the chip substrate can be described by the following differential equation:

$$K_x \frac{\partial^2 T}{\partial x^2} + K_y \frac{\partial^2 T}{\partial y^2} + K_z \frac{\partial^2 T}{\partial z^2} + Q(x,y,z) = 0 \qquad (1)$$

where T is the temperature, K_x, K_y, and K_z are the thermal conductivities, and Q is the heat generated per unit volume. A unique solution exists when convective, isothermal, and/or insulating boundary conditions are appropriately applied. The nature of the packaging and heat sink determines the boundary conditions. The FEA method from [10] was used for temperature calculations in these experiments. An overview of this method, as applied to 3D ICs, is presented in the remainder of this section.

3.1. FEA Background

In finite element analysis, the design space is first discretized or meshed into elements. An example of an 8-node hexahedral element is shown in Figure 3. The temperatures are calculated at discrete points (the nodes of the element), and the temperatures elsewhere within the element are interpolated using a weighted average of the temperatures at the nodes.
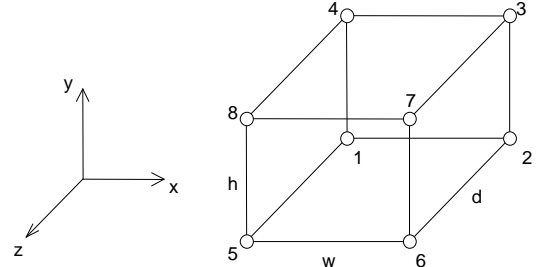

Figure 3. An 8-node hexahedral element.

For an 8-node hexahedral (rectangular prism) element, the following interpolation function is used:

$$T(x,y,z) = [N]\{t\} = \sum_{i=1}^{8} N_i t_i \qquad (2)$$

where $[N]=[N_1 \quad N_2 \quad \ldots \quad N_8]$, $\{t\}=\{t_1 \quad t_2 \quad \ldots \quad t_8\}^T$, t_i is the temperature at node i, and N_i is the shape function for node i. The shape functions are determined by the coordinates of the element's center, (x_c, y_c, z_c), the coordinates at the nodes, (x_i, y_i, z_i), the width, w, height, h, and depth, d, of the element.

$$N_i = \left(\frac{1}{2} + \frac{2(x_i - x_c)}{w^2}(x - x_c)\right)\left(\frac{1}{2} + \frac{2(y_i - y_c)}{h^2}(y - y_c)\right) \left(\frac{1}{2} + \frac{2(z_i - z_c)}{d^2}(z - z_c)\right) \quad (3)$$

From the shape functions, the thermal gradient, $\{g\}$, can be found as follows:

$$\{g\} = \left\{\frac{\partial T}{\partial x} \quad \frac{\partial T}{\partial y} \quad \frac{\partial T}{\partial z}\right\}^T = [B]\{t\} \quad (4)$$

$$\text{where } [B] = \begin{bmatrix} \frac{\partial N_1}{\partial x} & \frac{\partial N_2}{\partial x} & \cdots & \frac{\partial N_8}{\partial x} \\ \frac{\partial N_1}{\partial y} & \frac{\partial N_2}{\partial y} & \cdots & \frac{\partial N_8}{\partial y} \\ \frac{\partial N_1}{\partial z} & \frac{\partial N_2}{\partial z} & \cdots & \frac{\partial N_8}{\partial z} \end{bmatrix}.$$

Similar to circuit simulation using the modified nodal formulation, stamps are made for each element and added to the global system of equations. In FEA, these stamps are called element stiffness matrices, $[k_c]$, and can be derived as follows using the variational method for an arbitrary element type [11]:

$$[k_c] = \iiint_V [B]^T [D][B] dV \quad (5)$$

$$\text{where } [D] = \begin{bmatrix} K_x & 0 & 0 \\ 0 & K_y & 0 \\ 0 & 0 & K_z \end{bmatrix} \text{ and } K_x, K_y, \text{ and } K_z \text{ are the thermal}$$

conductivities in the x, y, and z directions.

3.2. Application of FEA

For a right prism with a width of w, a height of h, and a depth of d as shown in Figure 3, the element stiffness matrix is given in Equation (6) as an 8x8 symmetrical matrix with rows and columns corresponding to the nodes 1 through 8 [10].

$$[k_c] = \begin{bmatrix} +A & +B & +C & +D & +E & +F & +G & +H \\ +B & +A & +D & +C & +F & +E & +H & +G \\ +C & +D & +A & +B & +G & +H & +E & +F \\ +D & +C & +B & +A & +H & +G & +F & +E \\ +E & +F & +G & +H & +A & +B & +C & +D \\ +F & +E & +H & +G & +B & +A & +D & +C \\ +G & +H & +E & +F & +C & +D & +A & +B \\ +H & +G & +F & +E & +D & +C & +B & +A \end{bmatrix} \quad (6)$$

where $A = \frac{K_x hd}{9w} + \frac{K_y wd}{9h} + \frac{K_z wh}{9d}$, $B = -\frac{K_x hd}{9w} + \frac{K_y wd}{18h} + \frac{K_z wh}{18d}$

$C = -\frac{K_x hd}{18w} - \frac{K_y wd}{18h} + \frac{K_z wh}{36d}$, $D = \frac{K_x hd}{18w} - \frac{K_y wd}{9h} + \frac{K_z wh}{18d}$

$E = \frac{K_x hd}{18w} + \frac{K_y wd}{18h} - \frac{K_z wh}{9d}$, $F = -\frac{K_x hd}{18w} + \frac{K_y wd}{36h} - \frac{K_z wh}{18d}$

$G = -\frac{K_x hd}{36w} - \frac{K_y wd}{36h} - \frac{K_z wh}{36d}$, $H = \frac{K_x hd}{36w} - \frac{K_y wd}{18h} - \frac{K_z wh}{18d}$

For the entire mesh, the elements are aligned in a grid pattern with nodes being shared among at most 8 different elements. The element stiffness matrices are combined into a global stiffness matrix, $[K_{global}]$, by adding the components of the element matrices corresponding to the same node together. The global heat vector, $\{P\}$, contains power dissipated or heat generation as represented at the nodes. This is produced by distributing the heat generated by the standard cells among its closest nodes. A linear system of equations is produced, $[K_{global}]\{T\} = \{P\}$ with $\{T\}$ being a vector of all the nodal temperatures.

3.3. Isothermic Boundary Conditions

Isothermic boundary conditions are applied to the global matrix using the following procedure [11], and this results in a reduced, nonsingular system of equations. Rows and columns that correspond to fixed temperature values within the global matrix are eliminated, as are corresponding values in the power vector, and the remaining values in the right-hand side vector are modified using the fixed temperature values. For example, given the following system:

$$\begin{bmatrix} A_{11} & \vdots & A_{12} \\ \cdots & & \cdots \\ A_{21} & \vdots & A_{22} \end{bmatrix} \begin{Bmatrix} T_1 \\ \cdots \\ T_2 \end{Bmatrix} = \begin{Bmatrix} P_1 \\ \cdots \\ P_2 \end{Bmatrix} \quad (7)$$

A_{11}, A_{12}, A_{21}, and A_{22} represent arrays of elements in the global stiffness matrix. T_1 are the unknown temperatures, and T_2 are the fixed temperatures. P_1 are the known power values corresponding to the unknown values, T_1, and P_2 are the unknown power values corresponding to the known values, T_2. This system can be reduced as follows:

$$[A_{11}]\{T_1\} = \{P_1\} - [A_{12}]\{T_2\} \quad (8)$$

A_{11} is a nonsingular matrix, T_1 contain the unknowns, and the right-hand side is vector of constants so this linear system of equations can now be solved.

3.4. Meshing the 3D IC

In this method, 3D ICs are meshed into regions (elements) as shown in Figure 2. Vertically (in the z direction), the chip is separated into bulk substrate, layer, and inter-layer elements. The bulk substrate is located at the bottom attached to the heat sink. Above the bulk substrate elements are the layer and inter-layer elements. In a 3D IC, the inter-layers are composed of inter-layer vias and bonding materials that connect the layers together, and the layers contain the device and metal levels. Transistors, the primary sources of heat, are located at the bottom of each layer in the row regions. In standard cell designs, cells are placed into rows with inter-row spaces between them. These inter-row regions are necessary to accommodate interconnects between different layers. Some of the inter-row regions can serve as

thermal via regions. They can be represented with special elements in the FEA mesh with variable thermal conductivities.

4. ITERATIVE THERMAL VIA METHOD

For a given placed 3D circuit, an iterative method was developed in which, during each iteration, the thermal conductivities of certain FEA elements are modified so that thermal problems are reduced or eliminated. Thermal vias are generically added to elements to achieve the desired thermal conductivities. The goal of this method is to satisfy given thermal requirements using as few thermal vias as possible, i.e., keeping the thermal conductivities as low as possible.

4.1. Updating the Thermal Conductivities

During each iteration, the thermal conductivities of thermal via regions are modified. These thermal conductivities reflect the density of thermal vias needed to be utilized within the element. The new thermal conductivities are derived from the element FEA equations. Using Equation (6), we get:

$$[k_c]\{t\} = \{p\} \tag{9}$$

where $\{t\}$ are the nodal temperatures in the element and $\{p\}$ are the fractions of the nodal heat that transverse this particular element. Under the reasonable assumption that $\{p\}$ does not change between the old and new values, we get the following expression:

$$[k_c]_{new}\{t\}_{new} = [k_c]_{old}\{t\}_{old} \tag{10}$$

If we multiply Equation (10) by $\frac{2w}{hd}[1 \;\; -1 \;\; -1 \;\; 1 \;\; 1 \;\; -1 \;\; -1 \;\; 1]$,

$\frac{2h}{wd}[1 \;\; 1 \;\; -1 \;\; -1 \;\; 1 \;\; 1 \;\; -1 \;\; -1]$, and $\frac{2d}{wh}[1 \;\; 1 \;\; 1 \;\; 1 \;\; -1 \;\; -1 \;\; -1 \;\; -1]$, respectively, we obtain the following relations:

$$K_x^{new}\Delta t_x^{new} = K_x^{old}\Delta t_x^{old} \tag{11}$$

$$K_y^{new}\Delta t_x^{new} = K_y^{old}\Delta t_y^{old} \tag{12}$$

$$K_z^{new}\Delta t_z^{new} = K_z^{old}\Delta t_z^{old} \tag{13}$$

where $\Delta t_x^{new} = t_0^{new} - t_1^{new} - t_2^{new} + t_3^{new} + t_4^{new} - t_5^{new} - t_6^{new} + t_7^{new}$

$\Delta t_x^{old} = t_0^{old} - t_1^{old} - t_2^{old} + t_3^{old} + t_4^{old} - t_5^{old} - t_6^{old} + t_7^{old}$

$\Delta t_y^{new} = t_0^{new} + t_1^{new} - t_2^{new} - t_3^{new} + t_4^{new} + t_5^{new} - t_6^{new} - t_7^{new}$

$\Delta t_y^{old} = t_0^{old} + t_1^{old} - t_2^{old} - t_3^{old} + t_4^{old} + t_5^{old} - t_6^{old} - t_7^{old}$

$\Delta t_z^{new} = t_0^{new} + t_1^{new} + t_2^{new} + t_3^{new} - t_4^{new} - t_5^{new} - t_6^{new} - t_7^{new}$

$\Delta t_z^{old} = t_0^{old} + t_1^{old} + t_2^{old} + t_3^{old} - t_4^{old} - t_5^{old} - t_6^{old} - t_7^{old}$

The thermal gradients, $g_{new} = \{g_x^{new}, g_y^{new}, g_z^{new}\}$ and $g_{old} = \{g_x^{old}, g_y^{old}, g_z^{old}\}$, are functions of the position, (x, y, z), in the element, and at the center of the element, (x_c, y_c, z_c), they are equal to:

$$g_x^{new} = \frac{\Delta t_x^{new}}{4w}, g_x^{old} = \frac{\Delta t_x^{old}}{4w} \tag{14, 15}$$

$$g_y^{new} = \frac{\Delta t_y^{new}}{4h}, g_y^{old} = \frac{\Delta t_y^{old}}{4h} \tag{16, 17}$$

$$g_z^{new} = \frac{\Delta t_z^{new}}{4d}, g_z^{old} = \frac{\Delta t_z^{old}}{4d} \tag{18, 19}$$

g_{new} is the desired new thermal gradient, and g_{old} is the original thermal gradient of the element. The new K's can be found by combining Equations (11)-(19):

$$K_x^{new} = \frac{K_x^{old}\Delta t_x^{old}}{\Delta t_x^{new}} = \frac{K_x^{old} g_x^{old}}{g_x^{new}} \tag{20}$$

$$K_y^{new} = \frac{K_y^{old}\Delta t_y^{old}}{\Delta t_x^{new}} = \frac{K_y^{old} g_y^{old}}{g_x^{new}} \tag{21}$$

$$K_z^{new} = \frac{K_z^{old}\Delta t_z^{old}}{\Delta t_z^{new}} = \frac{K_z^{old} g_z^{old}}{g_z^{new}} \tag{22}$$

$\{g_x^{new}, g_y^{new}, g_z^{new}\}$ is chosen so that its component magnitudes are closer to some ideal thermal gradient value, g_{ideal}, than $\{g_x^{old}, g_y^{old}, g_z^{old}\}$ using the following equations:

$$\left|g_x^{new}\right| = g_{ideal}\left(\frac{\left|g_x^{old}\right|}{g_{ideal}}\right)^{\alpha} \tag{23}$$

$$\left|g_y^{new}\right| = g_{ideal}\left(\frac{\left|g_y^{old}\right|}{g_{ideal}}\right)^{\alpha} \tag{24}$$

$$\left|g_z^{new}\right| = g_{ideal}\left(\frac{\left|g_z^{old}\right|}{g_{ideal}}\right)^{\alpha} \tag{25}$$

where g_{ideal} is a nonnegative value and α is a user defined parameter between 0 and 1. If the magnitude of the old thermal gradient is below g_{ideal}, the value is increased toward g_{ideal} for the new thermal gradient. If the magnitude of the old thermal gradient is above g_{ideal}, the value is decreased toward g_{ideal}. If the magnitude of the old thermal gradient equals g_{ideal}, then the value isn't changed. This causes the K's to be decreased when the thermal gradient is below g_{ideal} and increased when the thermal gradient is above g_{ideal}. This process works to eliminate major sources of thermal impedance in areas of greatest heat transfer, but when the element is not on a critical heat sinking path, the K's are decreased to eliminate unnecessary thermal vias.

Combining Equations (20)-(25) yields the formulas used for updating the thermal conductivities during each iteration:

$$K_x^{new} = K_x^{old}\left(\frac{\left|g_x^{old}\right|}{g_{ideal}}\right)^{1-\alpha} \tag{26}$$

$$K_y^{new} = K_y^{old}\left(\frac{\left|g_y^{old}\right|}{g_{ideal}}\right)^{1-\alpha} \tag{27}$$

$$K_z^{new} = K_z^{old}\left(\frac{\left|g_z^{old}\right|}{g_{ideal}}\right)^{1-\alpha} \tag{28}$$

The ideal thermal gradients, g_{ideal}, must be chosen and adjusted specifically to satisfy some desired thermal objective. Initially it is set to the magnitude of the average thermal gradient, and in order to achieve a desired maximum temperature, it is updated during each iteration according to the following equation:

$$g_{ideal} = g_{ideal}\frac{T_{max}^{ideal}}{T_{max}} \tag{29}$$

Table 1. Thermal Conductivities of Thermal Vias Regions

	Layer			Interlayer			Chip Average		
	Thermal Conductivity		Percent Thermal Via	Thermal Conductivity		Percent Thermal Via	Thermal Conductivity		Percent Thermal Via
	Vertical	Lateral		Vertical	Lateral		Vertical	Lateral	
Minimum	1.11	2.15	0%	1.10	1.10	0%	1.11	2.06	0%
Midrange	100.33	3.21	25%	50.71	1.31	12.5%	96.13	3.05	23.9%
Maximum	199.55	5.75	50%	100.33	1.65	25%	191.14	5.40	47.9%

where T_{max}^{ideal} is the desired maximum temperature and T_{max} is the maximum temperature from the last iteration. If the maximum temperature was too low in the previous iteration, g_{ideal} is increased. Likewise, if the previous maximum temperature exceeds the desired maximum temperature, g_{ideal} is decreased to lower the thermal effects.

4.2. Thermal Via Density

As stated earlier, the thermal conductivity of a thermal via region is determined by its density of thermal vias. After thermal via placement determines the thermal conductivities, the thermal via densities can be derived. Individual thermal vias are assumed to be much smaller than the thermal via region, are arranged uniformly within the thermal via region, and change the effective thermal conductivity of the thermal via region. The percentage of thermal vias, m, (also called thermal via density) in a thermal via region is given by the following equation:

$$m = \frac{nA_{via}}{wh} \tag{30}$$

where n is the number of individual thermal vias in the region, A_{via} is the cross sectional area of each thermal via, w is the width of the region, and h is the height of the region. The relationship between the percentage of thermal vias and the effective vertical thermal conductivity is given by:

$$K_z^{eff} = mK_{via} + (1-m)K_z^{layer} \tag{31}$$

where K_{via} is the thermal conductivity of the via material and K_z^{layer} is the thermal conductivity of the region without any thermal vias. Using this equation, the percentage of thermal vias can be found for any K_z^{new} provided that $K_z^{layer} \le K_z^{new} \le K_{via}$:

$$m = \frac{K_z^{new} - K_z^{layer}}{K_{via} - K_z^{layer}} \tag{32}$$

In this implementation, only the vertical thermal conductivities will be optimized using Equation (28). Equations (26) and (27) will not be used for updating the lateral thermal conductivities because the thermal vias are only oriented vertically in these experiments. During each iteration, the new vertical thermal conductivity is used to calculate the thermal via density, m, and the lateral thermal conductivities for each thermal via region. The effective lateral thermal conductivity can be found using the percentage of thermal vias:

$$K_x^{eff} = K_y^{eff} = \left(1 - \sqrt{m}\right)K_{lateral}^{layer} + \frac{\sqrt{m}}{\frac{1-\sqrt{m}}{K_{lateral}^{layer}} + \frac{\sqrt{m}}{K_{via}}} \tag{33}$$

Figure 4 and Table 1 show the relationship between the thermal vias density and the thermal conductivities in the vertical and lateral directions for thermal via regions having vertically oriented thermal vias. In Figure 4, Equations (31) and (33) were used to plot the relationship between the percentage of thermal vias in a thermal via region and its effective thermal conductivities. As you can see, the vertically oriented vias (as shown in Figure 1) produce a much greater effect in the vertical thermal conductivity. In Table 1, the minimum, midrange, and maximum thermal conductivity values are given. In the midrange case, the average of the minimum and the maximum thermal via density values was used.
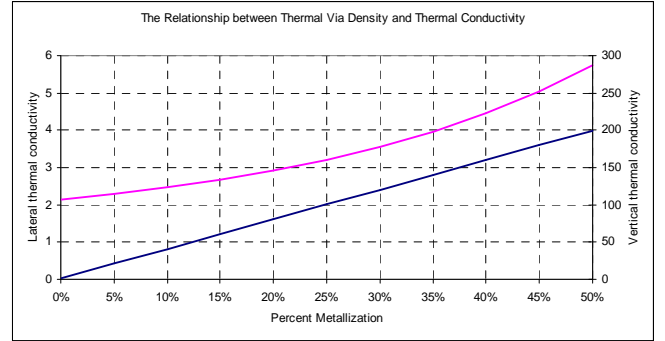

Figure 4. Percentage of thermal vias vs. thermal conductivity.

5. IMPLEMENTATION

In the thermal via placement method shown in Figure 5, the thermal gradients are used to update the thermal conductivities of the thermal via regions. The thermal gradients are compared to an ideal thermal gradient value which is modified during each iteration so that the maximum temperature converges to a given ideal maximum temperature, T_{max}^{ideal}.

```
MAX TEMP OPTIMIZATION( T_max^ideal )
{
        SET K's TO MININUM
        CALCULATE THERMAL PROFILE
        WHILE NOT CONVERGED
        {
                FOR EACH THERMAL VIA REGION
                {
                        K = K (|g|/g_ideal)^(1-α)
                }
                CALCULATE THERMAL PROFILE
                g_ideal = g_ideal T_max^ideal /T_max
        }
}
```

Figure 5. Thermal via optimization to a max. temperature.

Table 2. Relationship between Thermal Via Density and Thermal Effects

Benchmark Circuit			Thermal Via Densities of Thermal Vias Regions								
			Minimum (0%)			Midrange (23.9%)			Maximum (47.9%)		
name	cells	nets	T_{ave}	T_{max}	g_{ave}	T_{ave}	T_{max}	g_{ave}	T_{ave}	T_{max}	g_{ave}
struct	1888	1921	15.4	58.9	1.86E+05	10.9	35.0	5.67E+04	10.4	31.3	4.08E+04
biomed	6417	5743	14.6	46.0	1.72E+05	10.5	24.1	4.99E+04	10.0	20.2	3.37E+04
ibm01	12282	11754	14.2	45.1	1.71E+05	10.1	26.2	4.96E+04	9.6	22.7	3.36E+04
ibm04	26633	26451	13.5	54.0	1.51E+05	10.0	26.5	4.42E+04	9.6	21.4	2.98E+04
ibm09	51746	50679	13.8	53.0	1.56E+05	10.2	26.8	4.55E+04	9.8	21.4	3.04E+04
ibm13	81508	84297	14.6	47.3	1.84E+05	10.3	23.6	5.34E+04	9.7	19.3	3.55E+04
ibm15	158244	161580	15.1	52.8	2.01E+05	10.5	26.5	5.78E+04	9.9	20.6	3.83E+04

The thermal via placement algorithm is initialized by setting all K's to their minimum values and by calculating the temperature profile. The temperature profile is calculated using FEA, described earlier in Section 3. This gives the temperatures of the nodes and thermal gradients of the elements. In each iteration of the main loop, the thermal conductivities of the thermal via regions are modified, and the temperature profile of the chip is recalculated. For each thermal via region, the vertical thermal gradient, g, at the center of the element is calculated. Using the magnitude of the thermal gradient, a new vertical thermal conductivity is calculated using Equation (28) with α set to 0.5. If the new thermal conductivity exceeds the minimum or maximum value for that element, it is modified to the value of the bound that it exceeds. Using the new vertical thermal conductivity, the thermal via density and lateral thermal conductivities are also updated using Equations (32) and (33).

The K's are adjusted so that the maximum temperature approaches a specified value by modifying the ideal thermal gradient during each iteration based on the current maximum temperature. The algorithm terminates after the 1-norm of the change in K's is less than some small $\varepsilon > 0$. The value of T_{max}^{ideal} should be less than the maximum temperature when all the K's are minimized and greater than the maximum temperature when all the K's are maximized. In these experiments, the T_{max}^{ideal} values were obtained by using the maximum temperatures from the case where all the thermal via regions were given the midrange values as shown in Table 1. This gives a maximum temperature value that is greatly reduced from the case where no thermal vias were used and provides a comparison to the case where all the thermal via regions are given the same via density. In practice, it would be useful to use the maximum allowable operating temperature for T_{max}^{ideal}.

6. RESULTS

The program was written in C++ and run on an Intel Pentium 4 2.8 GHz machine with Linux. The conjugate gradient solver with ILU factorization preconditioning from the LASPack package [12] was used in our program to solve the FEA systems of equations. The thermal via placement method was tested using benchmark circuits from the MCNC suite [13] and the IBM-PLACE benchmarks [14].

The bulk substrate thickness was set to 500μm, the layer thicknesses were set to 5.7μm, and the interlayer thicknesses were set to 0.7μm. Four layers were used, and the chip size was fixed at 2cm x 2cm with the cell sizes adjusted accordingly. Thermal

via regions were even distributed between the rows and given 10% of the total chip area. The thermal conductivity of the silicon in the bulk substrate was set to 150W/mC, and the thermal vias were assumed to be copper with a thermal conductivity of 398W/mC. The thermal conductivities used in the layer and inter-layer elements are shown in Table 1. Thermal via regions had variable thermal conductivities ranging from the minimum to maximum values given in Table 1. All other elements used the thermal conductivities corresponding to the lateral midrange values.

A random power distribution was used with 90% of the cells having power densities ranging from 0 to 2 x 10^6 W/m^2 and 10% of the cells having power densities ranging from 2 x 10^6 to 4 x 10^6 W/m^2. The bottom of the chip was made isothermic with the ambient temperature to represent the heat sink, and the top and sides of the chip were made insulated in order to simulate the low heat sinking properties of the packaging. The ambient temperature was set to 0°C for convenience, but the temperatures can be translated by the amount of any other ambient as desired.

In the first set of experiments, thermal via densities were increased from their minimum to maximum values, and the change in the temperatures and thermal gradients was observed as shown in Table 2. In this table, T_{ave} is the average temperature, T_{max} is the maximum temperature, and g_{ave} is the average thermal gradient. The thermal via regions were given the same minimum, midrange, and maximum thermal conductivity values from Table 1. The minimum values correspond to the case where no thermal vias are used. The maximum values correspond to maximum thermal via usage. In the midrange case, thermal via regions were given thermal via densities that were the average of the minimum and maximum values. In Table 2, we can seen that as the thermal via densities of the thermal via regions were increased, the temperatures and thermal gradients decreased. The temperatures and thermal gradients in the minimum and maximum cases define the bounds on the thermal values that can be obtained from adjusting the thermal via densities. At the midrange with an average thermal via density of 23.9% in the thermal via regions, the maximum temperatures were 47.1% lower and the average temperatures were 28.3% lower than in the case where no thermal vias were used.

In Table 3, thermal via placements were obtained using the maximum temperatures from the midrange case (in Table 2) as the objective. In this table, K_{ave} is the average vertical thermal conductivity of the thermal via regions, and m is the average density of thermal vias in the thermal via regions. The thermal via placement method was very accurate in achieved its maximum

temperature objective. Maximum temperatures, T_{max}, were obtained within an average of 0.05% of their desired ideal maximum temperatures. On average, thermal via placement used only a thermal via density of 11.9% in the thermal via regions in order to obtain these maximum temperatures. This means that 50.3% fewer thermal vias were needed with thermal via placement than in the midrange case to obtain the same maximum temperatures. These maximum temperatures were 47.1% lower than in the case where no thermal vias were used. In these experiments, thermal via regions were assigned to only 10% of the chip area, and thermal vias required only 11.9% of this area. In all, thermal vias occupy only 1.19% of the total chip area, and it is expected that routability would be minimally affected from this small amount of blockages.

Table 3. Thermal Via Placement

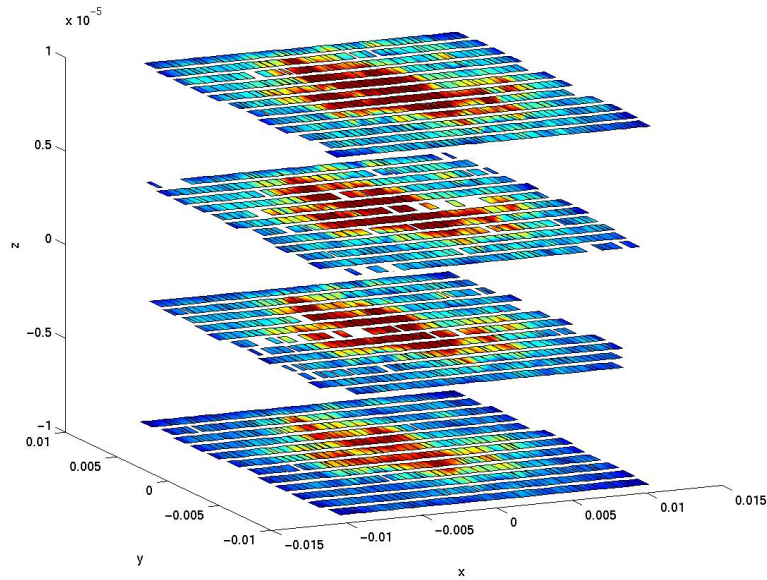| Bench-mark Circuits	Max. Temp. Optimization to T_{max} of Midrange					
	K_{ave}	m	T_{ave}	T_{max}	g_{ave}	Run Time (sec)
struct	34.9	8.5%	11.4	35.0	7.97E+04	3.9
biomed	50.1	9.2%	10.9	24.1	6.91E+04	18.2
ibm01	57.1	14.1%	10.1	26.2	5.58E+04	19.1
ibm04	51.6	12.7%	10.1	26.5	5.16E+04	43.1
ibm09	51.1	12.6%	10.3	26.8	5.40E+04	61.5
ibm13	59.8	14.8%	10.3	23.6	5.85E+04	134.0
ibm15	45.9	11.3%	10.8	26.5	7.44E+04	191.5

Figure 6. The thermal profile of struct before thermal via placement.
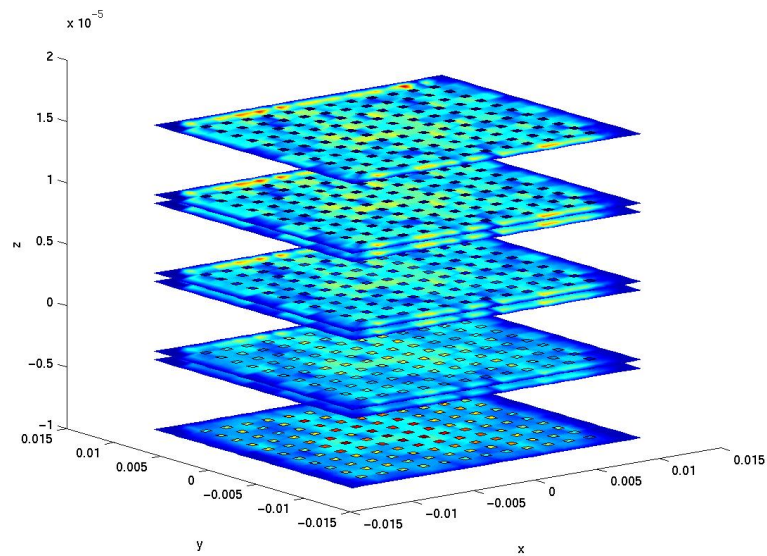

Figure 7. The thermal profile of struct after thermal via placement with thermal vias regions superimposed.

Temperature profiles before and after thermal via placement are shown in Figures 6 and 7 for the struct benchmark. In Figure 6, the heat sink is located at the bottom of the chip, and temperature contours are superimposed on the standard cells. Red indicates areas of high temperature, and blue represents areas of low temperature. As you can see, the temperatures increase as you go away from the heat sink toward the upper layers and toward the middle of the chip. After thermal via placement as shown in Figure 7, the temperatures are greatly reduced.

In Figure 7, the thermal via regions are superimposed on the resulting temperature profile after thermal via placement is performed. Thermal via regions are represented as small squares arranged in a grid pattern. The percentage of thermal vias in the thermal via regions is represented by its color. A red square represents a thermal via region with maximum thermal vias utilized. A blue square represents a thermal via region with minimum thermal vias utilized. The colored contours around the thermal via regions represent the temperatures. The heat sink is located at the bottom, and the thermal via regions were of greatest strength at the bottom of the chip where the thermal gradients are the highest and the most impact can be made in reducing thermal problems. However, at the top of the chip where the temperatures are the highest, the thermal vias are minimally used. In these areas, the thermal gradients are quite low and little impact can be made there.

7. CONCLUSION

An efficient thermal via placement method was presented that attempts to overcome the thermal issues produced in the design of 3D ICs. The resulting placements have lower temperatures and thermal gradients with the use of thermal vias kept to a minimum. In these experiments, thermal via placement uses 50.3% fewer thermal vias to obtain the same 47.1% reduction in the maximum temperatures that were obtained using the simpler method where all regions were blindly given the same midrange via densities. The iterative method modifies the thermal conductivities of thermal via regions in order to satisfy a maximum temperature requirement. Each thermal conductivity corresponds to a particular percentage of thermal vias in the thermal via region. There is a tradeoff between reducing thermal effects and the percentage of thermal vias used as seen in Table 2. There is also a tradeoff between routing space and thermal via space which results in a tradeoff between thermal problems and routability. An important observation is that placing thermal vias in areas of high temperature such as on the upper most layer has little impact in reducing thermal problems. This algorithm places thermal vias where they will have the most impact using the thermal gradient as a guide. High temperatures can only be reduced by alleviating the high thermal gradients leading up to them. The thermal resistance of these heat conduction paths is reducing by lowering the thermal conductivities of elements along it.

8. REFERENCES

[1] M. Chan and P. K. Ko, "Development of a Viable 3D Integrated Circuit Technology," *Science in China*, vol. 44, no. 4, pp. 241-248, August 2001.

[2] K. Banerjee, S. J. Souri, P. Kapur, and K. C. Saraswat, "3-D ICs: A Novel Chip Design for Improving Deep Submicrometer Interconnect Performance and Systems-on-Chip Integration," in *Proc. of the IEEE*, vol. 89, no. 5, pp. 602-633, May 2001.

[3] S. Lee, T.F. Lemczyk, and M. M. Yovanovich, "Analysis of Thermal Vias in High Density Interconnect Technology," in *8th IEEE Semi-Therm Symposium*, pp. 55-61, Feb. 1992.

[4] R.S. Li, "Optimization of Thermal Via Design Parameters Based on an Analytical Thermal Resistance Model," in *Thermal and Thermomechanical Phenomena in Electronic Systems*, pp. 475-480, 1998.

[5] D. Pinjala, M.K. Iyer, Chow Seng Guan, and I.J. Rasiah, "Thermal characterization of vias using compact models," in *Proc. of the Electronics Packaging Technology Conf.*, pp. 144-147, Dec. 2000.

[6] 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," in *IEEE Int. Electron Devices Meeting Tech. Digest*, pp. 261-264, 2000.

[7] T-Y. Chiang, K. Banerjee, and K. C. Saraswat, "Compact Modeling and SPICE-Based Simulation for Electrothermal Analysis of Multilevel ULSI Interconnects," in *Proc. of the Int. Conf. on Comput.-Aided Des.*, pp. 165-172, Nov. 2001.

[8] A. Rahman and R. Reif, "Thermal Analysis of Three-Dimensional (3-D) Integrated Circuits (ICs)," in *Proc. of the Interconnect Technology Conf.*, pp. 157–159, June 2001.

[9] T-Y. Chiang, S.J. Souri, Chi On Chui, and K.C. Saraswat, "Thermal Analysis of Heterogeneous 3D ICs with Various Integration Scenarios," in *IEEE Int. Electron Devices Meeting Tech. Digest*, pp. 681-684, Dec. 2001.

[10] B. Goplen and S. S. Sapatnekar, "Efficient Thermal Placement of Standard Cells in 3D ICs using a Force Directed Approach," in *Proc. of the Int. Conf. on Comput.-Aided Des.*, pp. 86 - 89, Nov. 2003.

[11] D. L. Logan, *A First Course in the Finite Element Method*, 3rd ed., Brooks/Cole Pub. Co., 2002.

[12] www.tu-dresden.de/mwism/skalicky/laspack/laspack.html

[13] www.cbl.ncsu.edu/pub/Benchmark_dirs/LayoutSynth92

[14] http://er.cs.ucla.edu/benchmarks/ibm-place/