3D Transient Thermal Solver using Non-conformal Domain Decomposition Approach

Jianyong Xie, Madhavan Swaminathan
Interconnect and Packaging Center (IPC), School of Electrical and Computer Engineering
Georgia Institute of Technology, Atlanta, GA 30332
Email: jianyong.xie@gatech.edu, madhavan@ece.gatech.edu

Abstract—3D integration becomes promising to be able to continue the system integration trend due to short TSV interconnection used for stacked dies. This paper proposes an efficient transient thermal modeling method using non-conformal domain decomposition approach for 3D stacked ICs and systems. To alleviate the problem arising from the feature scale difference between stacked dies as well as package and PCB, the 3D system is divided into many subdomains. Each subdomain (die, package or PCB) can be meshed independently using different gridding based on its feature size and therefore the required meshing cells are greatly reduced compared to conventional method such as finite element or finite volume method. The heat flow continuity between subdomains is captured using the introduced interface basis functions. In addition, the proposed compact micro-fluidic model based on finite volume method is proved to be compatible with the finite element model for solid medium based on introduced forced convection boundary and energy conservation. The experimental results show the proposed method offers up to 5x unknown reduction and 91x speed-up compared to conventional finite element method.

Keywords - Domain decomposition, transient, fluidic cooling, multiscale

I. INTRODUCTION

Compared to the conventional 2D integration approach, 3D integration becomes promising to be able to continuously increase integration density and improve system performance using advanced interconnect schemes such as Through Silicon Vias (TSVs) [1-2]. With the TSV fabrication process maturing and IC continuing to scale towards 22 nm node and beyond, the power density of 3D integrated systems is expected to increase dramatically, leading to the point where thermal effect on system performance such as thermal induced stress, electro-migration becomes important for chip and package design. To mitigate the thermal problem, for efficient cooling of high performance 3D stacked ICs, thermal management approaches such as micro-fluidic cooling become attractive [3-5]. For 3D system, which comprises of 3D stacked dies, package and PCB board, the thermal coupling between stacked chips can cause sharp temperature increase of the stacked chips as well as create package hot spots. Therefore, thermal modeling approaches, which are able to simulate 3D system including stacked ICs, package and PCB efficiently, are required.

For 3D stacked ICs, each die usually has its own architecture and functional blocks based on design. Therefore, each die usually requires different meshing grid based on its feature size as compared to other dies in thermal modeling. Modeling and gridding the 3D ICs or 3D system as a whole using conventional finite element or finite difference method (FDM) can result in millions of unknowns due to meshing grid propagates from one region to another. The great number of unknowns can affect the thermal modeling efficiency, which is an important aspect for 3D chip designers. To overcome the effect of increasing problem size on simulation speed, thermal modeling approach which can take into account the special structure of 3D system is needed, which is the subject of this paper.

This paper proposes an efficient 3D thermal modeling approach using non-conformal domain decomposition for simulating 3D stacked ICs or systems. Instead of modeling the 3D ICs or system using conventional finite element method (FEM) [6] or finite volume method (FVM) [7], which uses conformal meshing grids, the proposed method allows dividing the complex 3D system into many subdomains with non-matching grid at the interface. Each subdomain can be meshed independently based on its feature size. As a result, the total required meshing cells are greatly reduced and simulation efficiency is improved compared to thermal modeling approach using conformal FEM. The major contributions of this paper are listed as follows:

a) This paper first proposes a new 3D transient thermal modeling method based on non-conformal domain decomposition. It enables efficient transient thermal simulation of complex 3D system with non-conformal meshing grids on domain interfaces. The system unknowns and computational time are greatly reduced compared to conventional FEM. The accuracy of the proposed method has been validated by comparing with results from conformal FEM and measurements (the temperature error is less than 5%).

b) The proposed modeling method provides a modeling approach which can combine the FEM model for solid chip and compact FVM model for microfluidic cooling.
into a hybrid model based on energy conservation rule and introduces forced convection boundary condition. The accuracy of the model is validated against FEM and measurements. (chip temperature error is less than 5% and channel outlet temperature error is less than 2%).

c) The proposed method allows modeling 3D system including stacked dies as well as IC package and PCB as a whole. The inclusion of package and PCB provides better boundary condition for 3D stacked ICs as compared to using equivalent thermal resistance to represent the thermal convection effect on package and PCB. In addition, the temperature profile on IC package and PCB can also be obtained. Thus, it can be used for thermal induced stress and electro-migration analysis on package level, which is also an important aspect for 3D system design [8].

d) Last, the proposed method, which is based domain decomposition, provides flexibility to use different meshing size for different domains. This feature enables certain subdomains to use finer mesh grids, which will not affect the gridding of other domains among a 3D system.

The organization of the rest of the paper is as follows. In Section II, the previous work on the thermal modeling of IC chip and package including micro-fluidic cooling is discussed. Section III presents cell based thermal modeling of solid and FVM based fluid model for micro-channel. The approach for combination of the FEM model of solid and finite volume model of fluid is also introduced. Section IV discusses the 3D system thermal modeling using non-conformal domain decomposition method (DDM). Section V presents the implementation details for the proposed method. Section VI shows experimental examples and discussion. Finally, the conclusion is summarized in Section VII.

II. PREVIOUS WORK

In the past, considerable amount of approaches and efforts have been devoted for both steady state and transient thermal modeling of IC chip and package [6, 7, 9-15]. Among the methods for thermal modeling with conventional heat sink cooling, [6, 9] are based on finite element method and the approaches in [10, 11] are based on finite difference method. The finite volume based thermal modeling method has been proposed in [7, 12]. In addition, the joule heating (self-heating) effect was considered in [7] and [11]. To alleviate the effect of increasing problem size on computational cost, the thermal modeling using 3D geometrical multigrid approach has been proposed for IC chip thermal simulation in [13] and chip-package simulation in [14]. For transient thermal modeling of IC chip, a 3D solver based on alternating direction implicit (ADI) method has been proposed in [15] for obtaining transient response efficiently. To capture the microfluidic cooling effects, several approaches have been proposed in [7, 16-20] for steady state thermal analysis. For transient analysis, a compact transient thermal modeling approach has been proposed for 3D stacked ICs with inter-tier liquid cooling in [21].

For the above thermal modeling approaches, the methods have focused on modeling of IC chip, package and 3D stacked ICs, which are modeled as a single domain. Conformal meshing grids are used for meshing the system as a whole. This causes the fine meshing grids to propagate through the 3D multiscale structure. As a result, for modeling of multiscale 3D system including stacked ICs, package and printed circuit board (as shown in Fig. 1), the problem size can increase dramatically leading to long simulation time. The non-conformal DDM [22], which allows the dividing of a large complex problem into many subdomains that are smaller and easier to handle, becomes tractive. To tackle problems arising from multiscale 3D stacked ICs or 3D system, this paper proposes a transient thermal modeling method using domain decomposition and compact model for micro-fluidic cooling. It is important to note that this paper focuses on system level thermal modeling and thus modeling the internal fine details of chip is not the focus.

III. THERMAL MODELING OF SOLID AND FLUID

3.1 Cell Based Thermal Modeling of Solid

For thermal modeling of 3D problem consists of solid medium, the governing heat equation is described as:

$$\rho c_p \frac{\partial T(r,t)}{\partial t} - \nabla \cdot [k(r,T) \nabla T(r,t)] = P(r,t)$$  \hspace{1cm} (1)

where, $k(r,T)$ and $T(r,t)$ represent the thermal conductivity and temperature distribution, respectively. $P(r,t)$ denotes the total heat excitation which arises from the chip power map. To simplify the problem, the thermal conductivity is considered without temperature variation as $k(r)$. For modeling of silicon chip consisting of inhomogeneous material stacking (Fig. 2a) due to BEOL process, the cell based finite element modeling
can be applied. In this paper, the 3D problem is meshed using 3D non-uniform 8-node hexahedral elements (Fig. 2b) with trilinear basis functions [29]. The method for handling material inhomogeneity is shown in Appendix A. For modeling of realistic problem, the convection boundary condition
\[ k \frac{\partial T}{\partial n} \bigg|_{\text{wall}} = -h_c(T - T_a) \]  
needs to be considered at the periphery of chip and package to mimic the air cooling effect. \( T_a \) and \( h_c \) represent the ambient temperature and air convection coefficient, respectively.

Following the FEM modeling process [29], by multiplying testing function \( N \) at both sides of equation (1) and integrating over the volume, after using the divergence theorem, the new form of the heat equation can be obtained as

\[ \int_{\Omega} \rho c_p N \frac{\partial T}{\partial t} dV + \int_{\Omega} k \nabla N \cdot \nabla T \, dV - \int_{\partial\Omega} k N \frac{\partial T}{\partial n} \, dS = \int_{\Omega} N P \, dV \]  

For simplicity, the same basis function can also be used as testing function \( N \). The temperature can be expressed as a linear combination of basis functions:

\[ T = \sum_{i=1}^{n} T_i N_i \]  

As a result, with total of \( n \) meshed cells, the system equation with the applied convection boundary condition can be written as

\[ [C][\dot{T}] + [K][T] = f \]  

where,

\[ C = \sum_{i=1}^{n} K_c^{(e)}, \quad K = \sum_{i=1}^{n} \left( K_D^{(e)} + K_g^{(e)} \right), \quad f = \sum_{i=1}^{n} \left( f_D^{(e)} + f_g^{(e)} \right) \]

and

\[ K_c^{(e)} = \int_{V_e} \rho_c N_i N_j \, dV, \quad K_D^{(e)} = \int_{\partial V_e} k N_i N_j \, dS \]

\[ K_g^{(e)} = \int_{\partial V_e} h_c N_i N_j \, dS, \quad f_D^{(e)} = \int_{V_e} N_i P \, dV \]

\[ f_g^{(e)} = \int_{\partial V_e} N_i h_c T_a \, dS \]

In equation (5-6), \( K_c^{(e)} \) represents the elementary capacitance matrix. \( K_D^{(e)} \) and \( K_g^{(e)} \) represent the elementary stiffness matrix for each element (cell) due to heat conduction and heat convection. \( f_D^{(e)} \) and \( f_g^{(e)} \) represent the heat excitation and temperature gradient due to convection, respectively. For thermal modeling with homogenous Neumann boundary condition (natural boundary condition) [24], we can simply let \( h_c \) and \( K_g^{(e)} \) equal to zero.

### 3.2 Compact Thermal Modeling of Fluid

Because of the large number of micro-channels used for cooling of 3D ICs, to obtain fast temperature estimation at early design stage, compact thermal modeling of the fluid is required to overcome the simulation inefficient using detailed computational fluid dynamic (CFD) modeling [21]. For coolant flow in micro-channels of IC chips, the Reynolds number is usually less than 2300 and the flow is laminar [25]. Since the longitudinal direction of the channel is much larger than the lateral direction, the micro-fluidic flow can be treated as fully developed laminar flow. This property allows the developing of the compact fluid model. The governing equation for thermal analysis of micro-fluidic cooling is expressed as:

\[ \rho c_p \left( \frac{\partial T(r,t)}{\partial t} + \mathbf{v} \cdot \nabla T(r,t) \right) = \nabla \cdot (k_f \nabla T(r,t)) + P_f(r,t) \]  

Compared to equation (1) for solid medium, it has an extra term due to the coolant flow. Here, \( \mathbf{v} \) represents the velocity distribution of the fluid. \( k_f \) is the thermal conductivity of the fluid [26]. For heat transfer in coolant flow, it consists of heat conduction due to finite thermal conductivity and heat transportation due to flow velocity. In equation (7), except the second term on the left-hand side related to flow velocity, other terms can be modeled as solid. Since the coolant has certain flow direction, special consideration is required.

![Fig. 3. Discretization of micro-channel into cells. (Only bottom half part of the micro-channel is shown on the left figure)](image)

For fluidic cooling as shown in Fig. 3, since the micro-channel cross-sectional dimension is much smaller than its length, the flow velocity along the longitudinal direction is much larger than in the lateral direction. Thus, it can be assumed that the coolant only flows in the longitudinal direction and flow velocity therefore is constant. The average flow velocity \( \mathbf{v} \) along \( y \) direction (Fig. 3) has been used for simulating the fluid flow in this paper. As a result, equation (7) can be transformed as:

\[ \rho c_p \left( \frac{\partial T(r,t)}{\partial t} + \mathbf{v} \cdot \nabla T(r,t) \right) = \nabla \cdot (k_f \nabla T(r,t)) + P_f(r,t) \]  

By integrating equation (8) over the dashed finite volume cell (Fig. 3) and applying the divergence theorem, equation (8) can get rid of the 2nd order derivative and becomes

\[ \int_{V_{cell}} \rho c_p \left( \frac{\partial T(r,t)}{\partial t} \right) dV + \int_{S} \rho c_p \mathbf{v} \cdot \nabla T \cdot \mathbf{n} dS = \int_{V_{cell}} k_f \nabla T \cdot \mathbf{n} dS + \int_{S} P_f dV \]  

where, \( S \) is the surface of dashed cell (Fig. 3). It is important to note that for the second term on the left hand side, since the central finite difference scheme can generate instability in certain cases [7, 30], the backward difference approximation is used. By applying the finite difference approximation to equation (9) and applying the convection boundary condition (assuming convection coefficient of \( h_c \) at 4 sides of the
channel), the finite volume scheme for fluid flow at node \((i, j, k)\) can be expressed as:
\[
\begin{align*}
T_{i,j,k} - T_{i-1,j,k} &= \frac{W}{2} \frac{1}{h_i \cdot HL} \left( \frac{T_{i,j,k} - T_{i+1,j,k}}{W / 2} \right) + \frac{1}{h_i \cdot HL} \left( \frac{T_{i,j,k} - T_{i,j,k-1}}{W / 2} \right) + \frac{1}{h_i \cdot HL} \left( \frac{T_{i,j,k} - T_{i,j,k+1}}{W / 2} \right) + 1 \\
T_{i,j,k} - T_{i-1,j,k} &= \frac{2}{h_i \cdot HL} \left( \frac{T_{i,j,k} - T_{i+1,j,k}}{W / 2} \right) + \frac{1}{h_i \cdot HL} \left( \frac{T_{i,j,k} - T_{i,j,k-1}}{W / 2} \right) + \frac{1}{h_i \cdot HL} \left( \frac{T_{i,j,k} - T_{i,j,k+1}}{W / 2} \right) + 1 \\
T_{i,j,k} - T_{i-1,j,k} &= \frac{1}{h_i \cdot HL} \left( \frac{T_{i,j,k} - T_{i+1,j,k}}{W / 2} \right) + \frac{1}{h_i \cdot HL} \left( \frac{T_{i,j,k} - T_{i,j,k-1}}{W / 2} \right) + \frac{1}{h_i \cdot HL} \left( \frac{T_{i,j,k} - T_{i,j,k+1}}{W / 2} \right) + 1
\end{align*}
\]
\[
\sigma_{\nu} m (T_{i,j,k} - T_{i-1,j,k}) + \sigma_{\nu} c V_c \frac{\partial T}{\partial t} = P_f V_c
\]
where, \(V_c = WHL\) is the cell volume and \(m = \nu \cdot WH\) is the volumetric flow rate. Based on scheme derived in (10), the equivalent circuit representation of the fluidic cell is shown in Fig. 4b. Since the solid part can be modeled using FEM and the fluidic cooling is modeled using FVM, the combination of these two models is required. As shown in Fig. 4a, the connection between the FEM model of solid and FVM model of fluid is formed using the forced convection boundary condition. The forced convection is indicated using arrows in Fig. 4a. Since the convection boundary captures the heat transfer from chip to micro-channel effectively, the combination of these two models becomes feasible by following the energy conservation rule.

The convection strength, namely the average convection coefficient \(h\), at the 4 sides of the micro-channel can be obtained analytically from the Nusselt number as:
\[
h_s = \frac{Nu \cdot k}{D_h}
\]
where, \(D_h\) is the hydraulic diameter of the micro-channel [27]. The effect of this boundary condition is important, since eliminating it can cause incorrect chip temperatures [7, 28]. For fully developed laminar flow inside rectangular micro-channels with constant heat flux, the Nusselt number can be expressed as [25]:
\[
Nu = 8.235 \left( 1 - \frac{2.0421}{\alpha} + \frac{3.0853}{\alpha^2} - \frac{2.4765}{\alpha^3} + \frac{1.0578}{\alpha^4} - \frac{0.1861}{\alpha^5} \right)
\]
where, \(\alpha = W / H\) is the aspect ratio of the rectangular channel. After using the superposition rule for stamping the finite element solid model and FVM based fluid model based on global node number, the final stiffness matrix and system equation can be obtained as equation (13).

![Fig. 4. (a) Forced convection boundary between solid and fluid, (b) equivalent circuit model of fluidic cell.](image)

IV. 3D SYSTEM MODELING USING DOMAIN DECOMPOSITION

3D system including 3D stacked ICs, package, PCB and bump layers is shown as in Fig. 1. Since each chip has its own functional blocks, it therefore requires different meshing grid as compared to other chips. In addition, due to the feature scale difference between chip, package and PCB regions, millions of meshing cells are required using conformal FEM or FVM based meshing grids. To alleviate this problem, the problem can be divided into separate subdomains including separated chip domains, package and PCB domains, as shown in Fig. 5a. Each domain can be meshed independently using 3D non-uniform grids. As a result, the meshing grids from one chip domain do not interfere with grids from other domains and therefore the required meshing cells can be greatly reduced. For simplicity, the domain decomposition method is explained with 2D rectangular grids (Fig. 5a).

For transient thermal modeling, at each time step, the heat transfers out of one domain equals the heat flows into another domain through the common interface. At the interface, the continuity of heat transfer needs to be ensured to capture the coupling between separated domains. For two adjacent subdomains with a common interface (Fig. 5b), by assuming \(\lambda^{(i)} = k \partial T^{(i)} / \partial n_i \) \( (i = 1, 2)\), we have the relationship of \(-\lambda^{(1)} = \lambda^{(2)} = \lambda\) [22]. Then the weak continuity of heat flow across the interface can be established. In this section, the thermal modeling based non-conformal DDM is discussed for transient analysis.

By introducing \(\lambda\) (the Lagrange multiplier) for each interface, the coupling between domains can be captured using the coupling matrix (Fig. 5b). For simplicity, we can assume the system only has two separated domains. The final system equation using DDM can be obtained as:
\[
\begin{bmatrix}
C_1 & K_1 & B_1^T \\
C_2 & K_2 & -B_2^T \\
0 & B_1 & -B_2
\end{bmatrix}
\begin{bmatrix}
\dot{T} \\
K_1 T + B_1^T f \\
K_2 T - B_2^T f
\end{bmatrix} = f
\]
The detailed derivation of equation (14) from equation (3) is listed in Appendix B. For system with total \( N \) subdomains, the generalized system equation can be derived and obtained easily using the superposition rule based on node order. By using the backward time difference approximation

\[
\dot{T} = \frac{T^{(n+1)} - T^{(n)}}{\Delta t}
\]  (15)

Eq (14) can be converted to a linear equation as:

\[
AT^{(n+1)} = \begin{bmatrix}
A_1 & B_1^T & T_1^{(n+1)} \\
A_2 & -B_2^T & T_2^{(n+1)} \\
B_1 & -B_2 & 0
\end{bmatrix}
\begin{bmatrix}
f_1^{(n)} \\
f_2^{(n)} \\
0
\end{bmatrix} = f^{(n)}
\]  (16)

where,

\[
A_i = \frac{C_i}{\Delta t} + K_i, f_i^{(n)} = \frac{C_i}{\Delta t} T_i^{(n)} + f_i \ (i=1, 2)
\]

Here, the superscripts \((n+1)\) and \( n \) represent the time steps. The \( C_i \) and \( K_i \) represent the capacitance and impedance matrices for \( i \)-th domain. \( T_i^{(n+1)} \) represents the temperature vector of \( i \)-th domain. \( f_i^{(n)} \) denotes the heat excitation in \( i \)-th domain calculated at time step \( n \). The numerical scheme based on Crank–Nicolson (CN) method [30], which has second order accuracy in time, can also be obtained by using \((T^{(n+1)} + T^{(n)})/2\) to approximate the term \( T \) in Eq (14).

However, since the CN scheme has time step limitation and can result in temperature oscillation for large time steps [15], the scheme of Eq. (16), which is based on backward Euler method, is used in our simulation.

For simulating 3D stacked ICs using DDM, it is important to note that the connection between chip domains is through the transitional subdomain of bump layer (Fig. 5a). Since the domain of bump layer can be meshed using much coarser grids as compared to the chip, the required interface basis functions can be greatly reduced and thus \( B \) matrix has small dimension. To obtain the stiffness matrix for each domain, the associated boundary conditions need to be used for the corresponding subdomains. In addition, the homogeneous Neumann boundary condition needs to be assigned at the common interface.

V. TOOL IMPLEMENTATION AND FEATURES

A thermal simulation tool based on the proposed modeling method was implemented using Matlab. The input of the tool includes: a) 3D geometric stack-up information and material index, b) material information including heat capacity, thermal conductivity, etc. It accepts inhomogeneous materials with different thermal conductivity in \( x, y \) and \( z \) directions, c) time step and time dependent power excitation, which allows to define IC blockwise heat excitation, d) domains and interface parameters including domain volume, domain number and interface area, e) user defined meshing parameters for each domain including maximum mesh size in \( x, y \) and \( z \) directions and mesh refinement level. With higher mesh refinement level, finer mesh grids can be generated, f) boundary conditions including constant temperature and convection boundaries. It allows specifying natural air or forced convection boundary condition at the periphery of ICs (except the inlets and outlets of micro-channels), package and PCB surfaces. This tool accepts structures with inhomogeneous material stack-up. The method for handling material inhomogeneity is shown in Appendix A.

This tool can generate 3D non-uniform 8-node hexahedral cells for each domain including micro-fluidic cells based on the input file. With the generated hexahedral cells, the global node order is assigned for both solid medium and fluidic cells. The global stiffness matrix \( K \) and capacitance matrix \( C \) are formed using the superposition rule [23, 29] and stored as sparse matrices. For 3D problem, the interface becomes a surface. Since the interface surface can have several thousands of nodes, the 4-point Gaussian quadrature for rectangular element is used to calculate the coupling matrix \( B \) effectively.

To solve the linear equation \( AT^{(n+1)} = f^{(n)} \), the sparse matrix solver in Matlab is used for both conventional FEM and the proposed method for fair comparison. A PC with a 3.19 GHz Xeon(TM) CPU and 3.0 GB memory is used for simulation.

VI. EXAMPLES AND DISCUSSION

6.1 Model Verification Example (Test Set 1)

To evaluate the accuracy of the proposed method, an experimental example with both conventional heat sink and fluidic cooling is simulated, as shown in Fig. 6. The test vehicle in Fig. 6b consists of a silicon chip with fluidic cooling using micro-channels. The chip size is 1cm \( \times \) 1cm and the uniform power consumption is 45 W. A total number of 51 micro-channels are distributed uniformly on the chip as described in [3]. The cross-sectional dimension of each micro-channel is 0.1 mm \( \times \) 0.2 mm. The input water temperature at the inlets of micro-channels is set to be 22 \( ^\circ \)C as in measurement. Because of the scale difference between chip and package, the examples in Fig. 6a and Fig. 6b are both divided into two domains including chip domain and package domain. The detailed material properties and geometrical information are listed in Table I. Since the measurement [3] was carried out at natural convection condition, the convection coefficient of \( 5 \) \( W/(m^2\cdot K) \) is applied to both the top and bottom surfaces of the package in our simulation.

![Fig. 6. Test set 1 with (a) heat sink cooling and (b) micro-fluidic cooling.](image)
This test vehicle with heat sink cooling and fluidic cooling has been simulated using the proposed method and conventional FEM solver. The non-conformal and conformal rectangular meshing grids have been used for the proposed method and FEM solver, respectively. Note that the fluidic model is combined into the FEM solver. The comparison of simulated chip temperature and channel outlet temperature is shown in Fig. 7b. As seen from Fig. 7, the results from the proposed model using DDM agree very well with results from the conventional FEM solver. The maximum temperature difference is less than 0.2 and 0.5 Degree for heat sink and fluidic cooling, respectively. For fluidic cooling, the measured steady state chip temperature and outlet temperature are 40.8 and 32.2 Celsius in [3]. The differences between the simulated converged chip, channel outlet temperatures and measurements [3] are 1.7 °C and 0.4 °C, respectively. The relative errors are about 4.2% and 1.2% with respect to the measurements, respectively. The unknowns (including unknowns for interface) and simulation time using the proposed method and FEM are shown in Table II. It shows that the proposed method can reduce unknowns about 2-4 times. Due to the reduced unknowns, the simulation time speed up can reach 35x for simulation with fluidic cooling.

![Graph](image1)

Fig. 7. Temperature waveform comparison using proposed method and conventional FEM with (a) heat sink cooling, (b) fluidic cooling.

<table>
<thead>
<tr>
<th>Test set 1</th>
<th>Test set 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of layers</strong></td>
<td>4 (die: 10 mm x 10 mm)</td>
</tr>
<tr>
<td></td>
<td>4 (package: 4 cm x 4 cm)</td>
</tr>
<tr>
<td><strong>Channel width * height * length</strong></td>
<td>0.1 mm x 0.2 mm x 10 mm</td>
</tr>
<tr>
<td><strong>Channel pitch</strong></td>
<td>196 micron</td>
</tr>
<tr>
<td><strong>Bottom and top silicon height</strong></td>
<td>100 micron, 0 micron</td>
</tr>
<tr>
<td><strong>Fluid flow rate</strong></td>
<td>65 mL/min</td>
</tr>
<tr>
<td><strong>Pyrex glass heat capacity, thermal conductivity</strong></td>
<td>820 J/Kg-K, 1.1 W/m-K</td>
</tr>
<tr>
<td><strong>TIM heat capacity, thermal conductivity</strong></td>
<td>610 J/Kg-K, 1.6 W/m-K</td>
</tr>
<tr>
<td><strong>Heat sink boundary temperature</strong></td>
<td>25 Celsius</td>
</tr>
<tr>
<td><strong>Number of layers</strong></td>
<td>4 (die: 10 mm x 10 mm), 4 (package: 3 cm x 3 cm)</td>
</tr>
<tr>
<td></td>
<td>10 (board: 10 cm x 10 cm)</td>
</tr>
<tr>
<td><strong>Channel width * height * length</strong></td>
<td>0.1 mm x 0.2 mm x 10 mm</td>
</tr>
<tr>
<td><strong>Channel pitch</strong></td>
<td>196 micron</td>
</tr>
<tr>
<td><strong>Top and bottom silicon height</strong></td>
<td>100 micron, 0 micron</td>
</tr>
<tr>
<td><strong>Fluid flow rate (per chip)</strong></td>
<td>65 mL/min</td>
</tr>
</tbody>
</table>

**Common Parameters**

| **Thermal conductivity of fluid, silicon, BEOL** | 0.6, 110, 2 (W/m-K) | 4187, 700, 520 (J/Kg-K) |
| **Heat capacity of fluid, silicon, BEOL layer** | 4187, 700, 520 (J/Kg-K) | 610 J/Kg-K |
| **Ambient temperature** | 25 Celsius | 25 Celsius |

**Table II.** PROBLEM SIZE AND SIMULATION TIME COMPARISON

<table>
<thead>
<tr>
<th>Test</th>
<th>DDM Solver</th>
<th>FEM Solver</th>
<th>Time Speed-up</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Size (K)</strong></td>
<td><strong>Time (s)</strong></td>
<td><strong>Size (K)</strong></td>
<td><strong>Time (s)</strong></td>
</tr>
<tr>
<td>Test</td>
<td>Fluidic cooling</td>
<td>50.4</td>
<td>5.59</td>
</tr>
<tr>
<td>Set1</td>
<td>Heat sink</td>
<td>36.9</td>
<td>6.84</td>
</tr>
<tr>
<td>Test</td>
<td>3D system</td>
<td>53.1</td>
<td>3.41</td>
</tr>
<tr>
<td>Set2</td>
<td>3D ICs only</td>
<td>33.2</td>
<td>2.68</td>
</tr>
</tbody>
</table>

6.2 3D Stacking Example (Test Set 2)

A 3D system with inter-tier micro-fluidic cooling is also simulated using the proposed method. The 3D system includes three stacked chips, 4-layer package and 10-layer PCB, as shown in Fig. 8a. Each chip has 20 micro-channels. The layer stack-up for the stacked chip and microchannel is shown in Fig. 8b. The geometry and material parameters are summarized in Table I. Air convection with heat transfer coefficient of 10 W/(m²·K) is applied to the top surface of the package and both sides of the PCB. 3 chips are supplied with same water flow rate. Non-uniform heat dissipation is used for chip 2, as shown in Fig. 8c. Uniform power consumption of 40 W/cm² is used for chip1 and chip3. This test set is divided into 7 domains including 3 chip domains, 2 domains for micro-bump layer, one package domain and one PCB domain using DDM. The three chips are meshed independently using different mesh size. This example is simulated for 2.2 s. In the first second, the 3 chips operate with uniform power density of 40 W/cm². From 1.0-2.0 s, the 3 mm wide middle region of chip2 (Fig. 8c) is switched between 70 W/cm² and 40 W/cm² periodically. The simulated temperature using the proposed method and conventional FEM is shown in Fig. 9. It shows the simulated results using the proposed method agree very well with the FEM. The maximum temperature difference is less than 0.5 degree and the temperature error is about 1%. From Fig. 9, it is also observed that due to the convection on package and PCB, the bottom chip temperature is about 3 Celsius lower than the top chip. To show the capability for simulating only 3D ICs, the example (Fig. 8a) is also simulated without IC package and PCB. The required unknowns and simulation time comparison are shown in Table II. As seen from Table II, the proposed method can reduce unknowns about 2.3-4.9 times for simulating 3D ICs and 3D system. Due to the reduced unknowns, a simulation speed up to 91x can be obtained, indicating the efficiency of the proposed method. The side view (in yz plane) of temperature distribution of 3D ICs at t = 1.9s is shown in Fig. 10 with non-conformal gridding and hotspot illustrated.
In this paper, we proposed an efficient 3D transient thermal modeling method for simulating 3D systems using non-conformal domain decomposition approach. The proposed modeling method can capture the transient behavior of 3D ICs with microfluidic cooling. Detailed analysis of non-conformal domain decomposition for 3D system and 3D stacked IC oriented thermal modeling is presented. The method for compact modeling of microfluidic cooling is discussed. The accuracy and simulation efficiency of the proposed method have been validated against conventional FEM solver as well as measurements. Based on the experimental results, the proposed method shows the maximum temperature error of 5%. Due to the non-conformal domain decomposition approach used, the proposed method can reduce the meshed cells up to 5x, which can result in up to 91x speed-up as compared to the conventional FEM solver, which enables fast temperature estimation of 3D system or 3D ICs in their early design stage.

REFERENCES

Appendix

A. Handling Inhomogeneous Material

3D system can consist of inhomogeneous materials, as shown in Fig. 2a. To handle the inhomogeneity, cell based FEM formulation needs to be used [29]. With non-uniform meshing, each hexahedral cell only contains a single material with 8-nodes, as shown in Fig. 2b. Due to non-uniform meshing, the cell length is different in x, y and z directions. The elementary stiffness matrix for one cell due to heat conduction can be formed, described by:

\[
K_D^{(c)} = k_x \frac{\Delta y \Delta z}{2} K_x^{(c)} + k_y \frac{\Delta x \Delta z}{2} K_y^{(c)} + k_z \frac{\Delta x \Delta y}{2} K_z^{(c)}
\] (17)

where, \(K_x^{(c)}\), \(K_y^{(c)}\) and \(K_z^{(c)}\) represent the \(8 \times 8\) standard elementary stiffness matrix in x, y and z directions with normalized cell thermal conductivity \(k_x = k_y = k_z = 1\) and \(\Delta x = \Delta y = \Delta z = 2\) [29]. Based on the relationship between elementary cell node number and its global node order, the global stiffness matrix can be formed using the superposition rule, described as:

\[
K_D = \sum_{c=1}^{n} K_D^{(c)}
\] (18)

B. Formulation using Non-conformal DDM

Assuming the system is divided into two subdomains, the two subdomains own a common interface, as shown in Fig. 11. By assuming \(\lambda^{(i)} = k \partial T^{(i)} / \partial n_i\) \((i = 1,2)\), we have the relationship of \(\lambda^{(1)} = \lambda^{(2)} = \lambda\) [22]. Then the weak discontinuity across the interface can be established. Based on equation (3) and introduced Lagrange multiplier \(\lambda\), the following equations for domains and interface can be derived as

\[
\begin{aligned}
\int_{\Omega_1} k \nabla N_1 \cdot \nabla T_1 \, dV &- \int_{\Gamma_1} k N_1 \frac{\partial T_1}{\partial n} \, dt + \int_{\Gamma_{\text{int}}} \lambda N_1 \, dt + \\
\int_{\Omega_2} \rho c_p N_2 \frac{\partial T_2}{\partial t} \, dV &- \int_{\Gamma_1} N_1 P_1 \, dV \\
\int_{\Omega_1} k \nabla N_1 \cdot \nabla T_2 \, dV &- \int_{\Gamma_1} k N_2 \frac{\partial T_2}{\partial n} \, dt - \int_{\Gamma_{\text{int}}} \lambda N_2 \, dt \\
\int_{\Omega_2} \rho c_p N_2 \frac{\partial T_2}{\partial t} \, dV &- \int_{\Gamma_1} N_2 P_2 \, dV \\
\int_{\Gamma_{\text{int}}} (T_1 - T_2) \psi \, dt &- 0
\end{aligned}
\] (19)

where, \(N_1\), \(N_2\) and \(\psi\) represent the basis functions for domain 1, domain 2 and Lagrange multiplier space, respectively [22, 31]. With temperature \(T\) being expressed as a linear combination of basis functions and \(\lambda = \sum_{i=1}^{n} b_i \psi_i\), the system equation for the problem with two subdomains (Fig. 11) can be written as equation (14) as

\[
\begin{bmatrix}
C_1 & C_2 \\
C_2 & 0
\end{bmatrix} \dot{T} + \begin{bmatrix}
K_1 & B_1^T \\
B_2 & -B_2^T
\end{bmatrix} T = \begin{bmatrix}
f_1 \\
f_2
\end{bmatrix} = f
\] (20)

In the above equations, \(K_k\), \(B_k\) and \(C_k\) represent the conduction impedance matrix, coupling matrix and capacitance matrix for \(k\)-th domain, respectively. \(f_k\) represents the excitation vector for \(k\)-th domain. For Lagrange multiplier of the interface, its basis function can be constructed base on the interface grids from either side. To reduce the number of unknowns for the interface, it can be constructed based on a coarser meshing grid. To ensure the coupling matrix \(B\) for the interface has full rank, the basis function for the interface cannot be randomly selected. For 2D problem with 4-node (bilinear) elements (Fig. 11), since the interface becomes a line, the interface basis function (Fig. 12) can be constructed based on linear shape functions, expressed as

\[
\psi_i = \begin{cases}
\varphi_1 + \varphi_2 & (i = 1) \\
\varphi_{i+1} & (1 < i < n - 2) \\
\varphi_{n-1} + \varphi_n & (i = n - 2)
\end{cases}
\] (21)

where, \(\varphi_i\) is the linear shape function associated with node \(i\).

For 3D problems, the interface becomes a surface connecting two subdomains. For 2D interface with \(N_x \times N_y\) nodes, the interface basis function can be obtained based on 2D bilinear shape functions. For easier representation, it can be described based on 1D basis function in two directions as:

\[
\psi_{ij} = \psi_i \psi_j, \quad (1 \leq ix \leq N_x - 2, 1 \leq iy \leq N_y - 2)
\] (22)