6-8 October 2010, Barcelona, Spain

# Compact Transient Thermal Model for 3D ICs with Liquid Cooling via Enhanced Heat Transfer Cavity Geometries

Arvind Sridhar
EPFL, Lausanne
arvind.sridhar@epfl.ch

Alessandro Vincenzi
EPFL, Lausanne
alessandro.vincenzi@epfl.ch

Martino Ruggiero
EPFL, Lausanne
martino.ruggiero@epfl.ch

Thomas Brunschwiler
IBM, Zurich
tbr@zurich.ibm.com

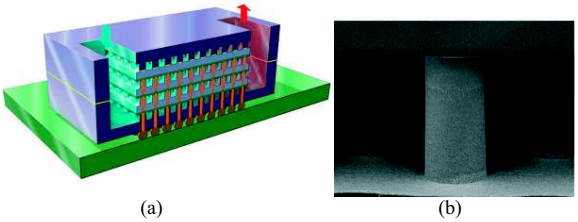
David Atienza
EPFL, Lausanne
david.atienza@epfl.ch

***Abstract*- The advent of 3D stacked ICs with accumulating heat fluxes stresses thermal reliability and is responsible for temperature driven performance deterioration of the electronic systems Hot spots with power densities typically rising up to 250 W/cm$^2$ are not acceptable, with the result of limited performance improvement in next generation high-performance microprocessor stacks. Unfortunately traditional back-side cooling only scales with the chip stack footprint, but not with the number of tiers. Direct heat removal from the IC dies via inter-tier liquid cooling is a promising solution to address this problem. In this regard, a thermal-aware design of a 3D IC with liquid cooling for optimal electronic performance and reliability requires fast modeling and simulation of the liquid cooling during the early stages of the design. In this paper, we propose a novel compact transient thermal modeling (CTTM) scheme for liquid cooling in 3D ICs via microchannels and enhanced heat transfer cavity geometries such as pin-fin structures. The model is compatible with the existing thermal-CAD tools for ICs and offers significant speed-up over commercial computational fluid dynamics simulators (13478x for pin-fin geometry with 1.1% error in temperature). In addition, the model is highly flexible and it provides a generic framework in which heat transfer coefficient data from numerical simulations or existing correlations can be incorporated depending upon the speed/accuracy needs of the designer. We have also studied the effects of using different techniques for the estimation of heat transfer coefficients on the accuracy of the model. This study highlights the need to consider developing flow conditions to accurately model the temperature field in the chip stack. The use of correlation data from fully developed flows only results in temperature error as high as 9 K (about 41%) near the inlet.**

## I. INTRODUCTION

With the fast approaching end of classical CMOS scaling [1], vertical integration of IC dies using through silicon vias (TSVs) is envisioned to be the most viable solution to continue the performance increase in the electronics industry. Three dimensional stacked integrated circuits (3D ICs) offer massive bandwidth improvements while reducing effective chip footprint. They are extremely attractive to overcome the barriers in interconnect scaling, and thus, open up opportunities to continue the CMOS performance trends (Moore's law) for the next decade. However, this means an increase in power dissipation per unit area of the chip stack and results in higher chip temperatures and thermal stress, hence, limiting the performance and reliability of the chip-stack [2].

Conventional back-side heat removal strategies such as air cooled heat sinks and microchannel cold-plates only scale with the die size and are insufficient to cool 3D ICs with hotspot heat fluxes up to 250W/cm$^2$ [3]. On the contrary, inter-tier forced convective liquid cooling, as shown in Fig. 1(a), scales with the number of dies and it is compatible with area-array TSVs [4]. Hence, it is capable of removing heat from multi-processor 3D ICs. Recently, the heat transfer performance of enhanced cavity geometries, such as pin-fins, have been studied and shown to be much more effective due higher permeability and reduced convective thermal resistance compared to conventional microchannel cooling [3-5].

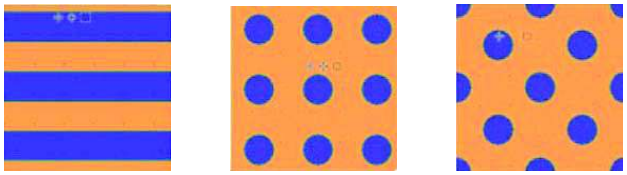


(a) (b)

**Figure 1: (a) 3D stacked IC with inter-tier liquid cooling (b) SEM micrograph showing the fluid cavity with a pin-fin structure.**

A magnified image of the area-array TSVs compatible pin-fin heat transfer geometry is shown in Fig. 1(b). The use of such complex cooling cavities necessitates the development of a modeling platform to enable "cooling aware" design of ICs during the early design-stages. Accurate thermal models are needed for predicting the costs of operating the liquid cooling (pumping power), determining the overall energy budget and performing run-time thermal management. However, it is impractical to accurately analyze thermal effects and model temperature distribution of a system in full detail. Detailed numerical analysis methods, such as finite-element methods (FEM), are a time-consuming process not suitable for design-time and run-time thermal management of very complex 3D architectures. Multi-scale modeling concepts are typically used to improve computational efficiency. We propose a compact thermal model to provide the tradeoff between accuracy and speed, giving precise temperature predictions with the use of effective parameters such as convective thermal resistance derived from sub-domain modeling or correlations.

In this paper, a generalized compact transient thermal model (CTTM) for 3D ICs with liquid cooling via complex heat transfer geometries (HTGs) is presented. This model is based on the finite-difference approximation imposed on the governing equations of energy transport for forced convection [6,7]. Also, this model simplifies the computation of temperatures in the channel layers by homogenizing them into a *porous medium* [8]. The proposed model is flexible and provides a general framework for thermal simulation, in which, convective heat transfer models of any kind can be incorporated depending upon the speed/accuracy needs of the designer. The proposed model is simple to implement and is as accurate as the convective heat transfer model incorporated.



(a) (b) (c)

**Figure 2: Forced convective heat transfer geometries (*dark*: pin or fin)- (a) microchannel (b) pin-fin inline (c) pin-fin staggered.**

In addition, the proposed model has been used to study the impact of heat transfer coefficients obtained with different methods on the accuracy of the temperature estimation. As a result, the effect of thermally developing boundary layers is found to be important. The use of correlations considering only fully-developed conditions result in conservative temperature estimation and could be insufficient for some cases.

The rest of the paper is organized as follows. Section II provides the background for compact modeling and reviews the recently advanced 4 resistor model based-CTTM for microchannels [8]. Section III describes the proposed CTTM based on the porous media approach for both microchannel and pin-fin HTGs. The experimental results are detailed in Section IV and finally, the conclusions are presented in Section V.

## II. Related Work

Compact modeling of the thermal behavior of an IC is a very useful tool in the early stages of IC design, when the final design layout and thermal management strategies are not yet known [9]. Compact modeling enables the quick simulation of a large number of test cases and is a practical alternative to fine-grained finite element CFD simulators which take hours to simulate even straightforward problems. Compact thermal modeling is performed by invoking the well-known analogy between heat flow and electrical conduction in solids [6,7]. A given 3D IC is discretized into resistive-capacitive "thermal cells" to create an equivalent RC circuit based on this analogy. Until now, in the context of IC design, compact modeling has been limited to the simulation of only solid structures.

Numerous empirical correlation and numerical simulation-based methods have been proposed in the literature for the simulation of ICs with liquid cooling[10-12]. However, all these methods are meant for only steady state simulations and do not exhibit the unique advantages of compact modeling.

A 4 resistor model-based CTTM (4RM-based CTTM) was recently developed for microchannels in [13]. This model was built by discretizing the point form of energy conservation equation for heat transfer in fluids using 1) finite-difference to represent the conduction part of the equation and 2) central-differencing to represent the convective part. That is, given the partial differential equation,

$$C_v \frac{dT}{dt} - \nabla \cdot (k \nabla T) + C_v \vec{u} \cdot \nabla T = \dot{q}, \qquad (1)$$

it is converted into an ordinary differential equation (for coolant flowing in the $y$ direction in Cartesian coordinates) as follows:

$$\begin{aligned} C_v \Delta x \Delta y \Delta z \frac{dT}{dt} - k_{xx} \frac{T_{i+1,j,k} - 2T_{i,j,k} + T_{i-1,j,k}}{\Delta x^2} \\ - k_{zz} \frac{T_{i,j,k+1} - 2T_{i,j,k} + T_{i,j,k-1}}{\Delta z^2} \\ + C_v u_{avg,y} \Delta A_y (T_{S2} - T_{S1}) \\ = \dot{q} \Delta x \Delta y \Delta z. \end{aligned} \qquad (2)$$

Here, $C_v$ is the volumetric specific heat of the coolant, $k$ is the thermal conductivity in various directions, $u_{avg,y}$ is the average coolant velocity, $T$ is the temperature, ($\Delta x, \Delta y, \Delta z$) are the discretization lengths and $\Delta A_y = \Delta x \cdot \Delta y$ If the microchannels are discretized in such a way that the entire cross section of the channel forms one face of a thermal cell (Fig. 3), then the first term in the above equation is translated into a capacitance term in the equivalent RC circuit. The second and the third terms are translated into the 4 convective resistance terms from the fluid to the four walls of the channel, which can be either calculated using empirical correlations or numerical presimulations. The fourth term is translated, using central differencing, into two "voltage controlled current source terms"- one representing convective heat transport from the previous cell to the current cell along the channel and the other representing the heat transport away from the current cell to the next cell.

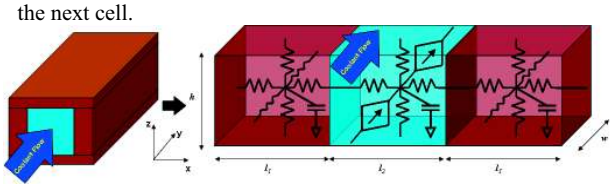



**Figure 3: Modeling a three dimensional microchannel subdomain (left) as a three dimensional 4RM-based CTTM thermal cell (right)**

In [13], the heat transfer coefficient from silicon walls into the fluid from all the 4 sides was calculated using the formula:

$$h_{f,vertical} = h_{f,side} = \frac{k_{coolant} \cdot Nu}{d_h}, \qquad (3)$$

where $k_{coolant}$ is the thermal conductivity of the coolant, $d_h$ is the hydraulic diameter of the channel and $Nu$ is the Nusselt number calculated using the correlations for fully developed flows provided by London and Shah [14] by considering imposed axial heat flux and radial isothermal conditions (given microchannels with low aspect ratio fins).

The above representation for the coolants, combined with the conventional transient compact modeling for solids [9], produces the 4RM-based CTTM for 3D ICs with Microchannel liquid cooling. This model was demonstrated to be flexible and accurate, given the availability of accurate estimation of convective resistances for Microchannels. Also, this model was shown to be significantly faster than commercial CFD simulators [13].

One disadvantage of the 4RM-based CTTM is that the user is forced to use the channel width as the discretization length along the $x$-direction (i.e. transverse to the flow direction, $\Delta x$), since each fluid cell must encompass the entire cross section of the channel. Given the typical microchannel width of 50-100μm in most inter-layer cooling HTGs, this results in a significantly finer mesh for the thermal grid than what is necessary for the accuracy/resolution purposes of an electronic designer.

In this work, the porous media based 2 resistor model (2RM) advanced in [5,8] is incorporated in the CTTM and its scope has been extended to include enhanced heat transfer geometries (HTGs) such as pin-fins (Fig. 2). Using the porous media based CTTM also allows the designer greater freedom to increase the discretization size, resulting in smaller problem sizes and faster simulations.

## III. Proposed 2-Resistor Model-based CTTM

The porous media approach [5,8] homogenizes the cavity layer into a porous medium, where the heat is transferred from the dies to the coolant via only 2 convective thermal resistances- one in the top and the other in the bottom. The heat transfer coefficients for convection and conduction are calculated based on the relative fraction of the volume of the cavity occupied by the fluid - called the *porosity*. Hence, the three dimensional heat transport from the solid domain to the fluid domain is reduced to a two dimensional circuit.

### A. 2RM-based CTTM for Microchannels

The 2RM-based CTTM for microchannels is illustrated in Fig. 4. The basic idea here is to translate the convective heat transfer from silicon to the coolant via two directions (vertical from top and bottom walls, and horizontal from the two side walls) into a single direction (vertical), by projecting the heat transfer through the side walls on to the top and bottom surfaces, that is,

$$h_{efff,porous} = \frac{\iint h \cdot dA_{wetted}}{A_{projected}}. \quad (4)$$

Here, $A_{wetted}$ represents the actual area that is wetted by the coolant, and $A_{projected}$ is the final area of projection of the heat transfer in the model. For example, if the vertical and the side heat transfer coefficients for the microchannel are equal (as in (3)), then the effective porous media heat transfer coefficient for the top and the bottom wall are given by:

$$h_{eff,porous} = h_f \cdot \frac{(w_c + t_c)}{p_c}, \quad (5)$$

where, $w_c$ is the width of the microchannel, $t_c$ is the height of the cavity and $p_c$ is the pitch of the channels. In Fig. 4, $R_{cond}$ represents the conductive resistance between the top wall and the bottom wall via the silicon walls separating the microchannels. $R_{downstream}$ represents the conductive resistance of these walls along the channel direction. With the discretization lengths of $l$, $w$, and $h$, these parameters are given by:

$$\begin{aligned}
g_{top/bottom} &= \frac{1}{R_{conv}} = h_{eff,porous} \cdot (l \cdot w),\\
J_{conv} &= C_v \cdot u_{avg,y} \cdot (\Delta A_y / 2) T_{i,j\pm1,k} \cdot \varepsilon,\\
c_{coolant} &= C_{v,coolant} \cdot (l \cdot w \cdot h) \cdot \varepsilon,\\
g_{cond} &= \frac{1}{R_{cond}} = k_{Si} \cdot \frac{l \cdot w}{(h/2)} \cdot (1-\varepsilon),\\
g_{downstream} &= \frac{1}{R_{downstream}} = k_{Si} \cdot \frac{l \cdot h}{(w/2)} \cdot (1-\varepsilon),\\
c_{Si} &= C_{v,Si} \cdot (l \cdot w \cdot h) \cdot (1-\varepsilon),
\end{aligned} \quad (6)$$

where $\varepsilon = w_c / p_c$, the porosity of the microchannel cavity [8]. With this homogenization of the cavity layer, the user is free to use any discretization for the CTTM resulting in a reduced problem size, and in turn, reduced CPU time of simulation.

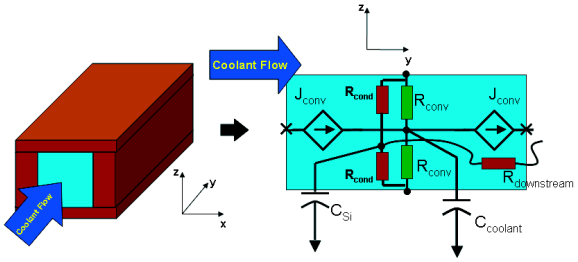



**Figure 4: Modeling a three dimensional microchannel subdomain (left) as a two dimensional 2RM-based CTTM thermal cell (right)**

### B. 2RM-based CTTM for Pin-Fins

A similar 2RM-based CTTM for pin-fins is illustrated in Fig. 5. Here, a pin-fin staggered HTG is considered. The 3-dimensional velocity (vectors) and temperature (color) profile of a unit pin-fin staggered subdomain is shown along with its conversion into a 2-dimensional CTTM. The model parameters are computed similarly as for the case of microchannels given in (6). The porosity for the pin-fin HTGs, with pin-fin diameter $d$, is given by

$$\varepsilon_{pin-fin} = \left(1 - \frac{\pi d^2}{4} \cdot \lambda\right), \quad (7)$$

where λ is the pin density (number of pins per unit area) in the cavity. The effective heat transfer coefficient for the pin-fin HTG is obtained from the correlations presented in [8]. That is,

$$\begin{aligned}
h_{pin-fin,inline} &= \left[\left\{25.27 \cdot (v_{darcy} + 1.35)^{0.64}\right\} + 1.533\right]^{-1} \times 10^6,\\
h_{pin-fin,stag} &= \left[\left\{25.27 \cdot (v_{darcy} + 1.35)^{1.52}\right\} + 1.533\right]^{-1} \times 10^6,
\end{aligned} \quad (8)$$

where, $v_{darcy}$ is the darcy velocity of the coolant in **m/s**. Note that there is no $R_{downstream}$ component in the pin-fin CTTM because there is no solid structure transporting heat via conduction along the flow direction in pin-fins.

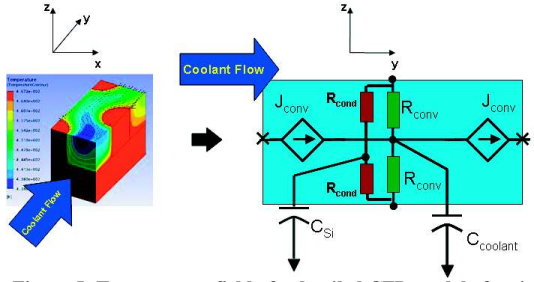



**Figure 5: Temperature field of a detailed CFD model of a pin-fin staggered unit-cell (left) and the corresponding two dimensional 2RM-based CTTM thermal cell (right)**

## IV. NUMERICAL EXAMPLES

In this section, the accuracy and the performance of the proposed 2RM-based CTTM will be studied. In addition, the effects of thermally developing boundary layers on the accuracy of the proposed model, when correlations from fully developed flows are used, are examined in detail. Finally, a performance comparison is made between the 2RM-based CTTM and a 4RM-based CTTM for an actual 3D-IC stack with realistic floorplan and power dissipation. The CTTMs were written and simulated in Matlab 7.9. The system matrices were built and solved using Matlab's internal sparse matrix routines. All the simulations described in this section were run on a Intel Core2Duo 3.16GHz machine with 2 GB RAM and running Windows XP OS.

### A. Performance of 2RM-based CTTM for Microchannels

The structure considered in **Test Stack 1** consists of 3 active dies and four microchannel cavities adjacent to them. **Test Stack 1** has a footprint of 10mm X 10mm. A small slice of the IC, showing the composition of the stack and the cross section of the channels is shown in Fig. 6 (a). The material and properties used in all our experiments henceforth are tabulated in Table1 [5].

**Table 1: Geometrical and material properties of Test Stacks**

| | |
|---|---|
| Number of layers- **Test Stack 1** | 15 (3 active dies) |
| Number of layer- **Test Stack 2** | 5 (2 active dies) |
| Cavity height- $t_C$ | 100µm |
| Si slab height- $t_S$ | 50µm |
| Heater height- $t_h$ | 2µm |
| Back-end of line (BEOL) height- $t_B$ | 12µm |
| Top Si layer height- $t_T$ (**Test Stack 1**) | 100µm |
| Channel width, pitch- $w_C$, $p$ (**Test Stack 1**) | 50 µm , 100 µm |
| Pin diameter, pitch- $d_{pin}$, $p_{pin}$ (**Test Stack 2**) | 50 µm , 100 µm |
| Silicon thermal conductivity, heat capacity | 130W/m.K, 702J/kg.K |
| BEOL thermal conductivity, heat capacity | 2.25W/m.K, 517J/kg.K |
| Fluid thermal conductivity, heat capacity | 0.6069W/m.K,4181J/kg.K |
| Fluid density, viscosity | 998kg/m3, 0.889mPa.s |
| Average darcy velocity (**Test Stack 1**) | 0.81m/s |
| Average darcy velocity (**Test Stack 2**) | 1.1067m/s |
| $R_{conv}$, $R_{cond}$, $R_{downstream}$ (**Test Stack 1**) | 17.6µΩ $m^2$,1.5µΩ $m^2$,153Ω m/m |
| $R_{conv}$, $R_{cond}$ (**Test Stack 2**) | 45µΩ $m^2$, 3.9µΩ $m^2$ |

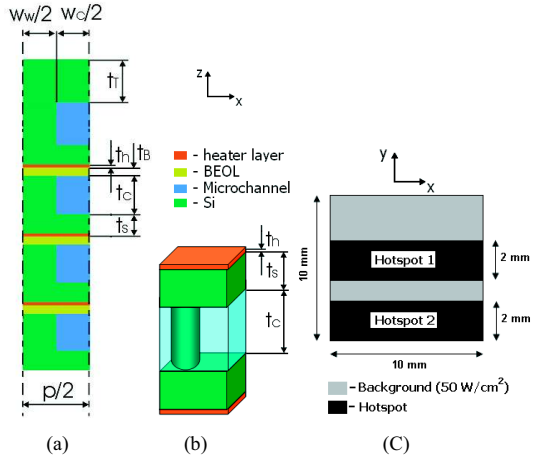



**Figure 6: (a) Test Stack 1 with microchannel HTG and (b) Test Stact 2 with pin-fin inline HTG; (c) Power dissipation map for Test Stack 1**

To validate the accuracy and performance of the proposed model, we compared the results from the model with a commercial computational fluid dynamics (CFD) simulation tool, i.e., Ansys CFX [15]. In all our experiments, the fluid flow through the channel was assumed to be hydrodynamically fully developed. Therefore, periodic hydrodynamic boundary conditions with a fixed pressure gradient ($10^5$Pa/cm) were imposed, to derive the velocity field of the coolant. In a subsequent transient run, the velocity field was used as initial condition and the energy equation with an imposed power step was computed.

In order to minimize the model complexity, only a 50µm wide stack with symmetry boundary conditions was taken into account in all our experiments using the CFD model. In the case of **Test Stack1**, this resulted in a 50µmX10mm long computation domain with only half of the microchannel and microchannel wall.

The resulting computational domain contained an unstructured hexahedral mesh with 176k nodes. Invariant heat fluxes transversal to the flow direction (i.e. along *x* direction if the microchannels are laid out in the *y* direction) were imposed, resulting in a uniform temperature along the x direction. For the purpose of analogous comparison, a structure similar in size and composition to the one used in CFX was simulated using the proposed CTTM in all the experiments in subsections *IV-A* to *IV-C*. The corresponding 2RM-based model for **Test Stack 1**, with a thermal cell size of 50µmX100µm, contained 8.1k nodes.

Test Stack 1 was simulated for a time period of 0.7 seconds. For the first 0.1 seconds all the active dies were given a uniform background heat flux of 50W/cm$^2$. Between $t$=0.1s and $t$=0.2s, Hotspot 1- a 2mm wide region in the center of the second die (as shown in Fig. 7) was switched alternatively between the powers 250W/cm$^2$ and 50W/cm$^2$. And then between $t$=0.3s and $t$=0.6s, Hotspot 2- a similar 2mm wide region near the inlet of the microchannels (Fig. 6 (c)) was switched alternatively between the powers 250W/cm$^2$ and 50W/cm$^2$ , and between 450W/cm$^2$ and 50W/cm$^2$. Through out the simulation, the temperature at the center of the 2$^{nd}$ die was monitored. Such a switching activity was chosen to simulate an actual non-uniform IC power dissipation during the course of its operation and its effect on the temperature near it and at some location downstream along the microchannel. The CPU times for simulation and the resulting speed-up is shown in Table 2.

The resulting temperature waveforms from the proposed 2RM-based CTTM and CFX are plotted in Fig. 7 (a). It can be seen that while the temperature fluctuations due to Hotspot 2 is captured fairly accurately (maximum error= 0.43$^o$C, or about 2.7% with respect to the corresponding maximum deviation of temperature from the initial state), the error is slightly higher for the temperature fluctuations caused by Hotspot 1 (maximum error=0.7$^o$C or about 3.4%). Reasons for this varied behavior will be discussed in subsection *VI-C*. Note that these errors arise only for the steady state temperature, while the transients are captured well, as shown by the normalized temperature rise near t=0.1s (temperatures normalized between 0 and 1) in Fig. 7(b) (error in $e^{-1}$ time constant=0.1ms).

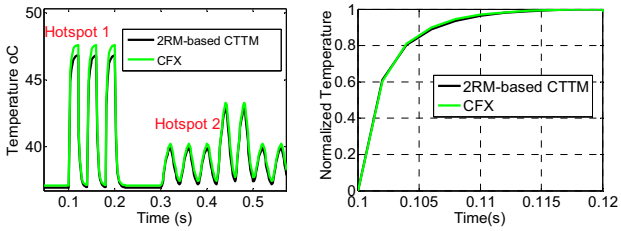



**Figure 7: (a) Absolute and (b) Normalized Temperature waveforms for Test Stack 1 at the chip center location on die 2.**

**Table 2: CPU times for simulation and speed-ups**

| Problem | CPU Time (hh:mm:ss) | | Speed-up |
|---|---|---|---|
| | CFX | 2RM CTTM | |
| **Test Stack 1** | 1:40:53 | 3.29 sec | **1857x** |
| **Test Stack 2** | 4:15:29 | 1.14 sec | **13478x** |

### *B. Performance of 2RM-based CTTM for Pin-Fins (inline)*

We now consider **Test Stack 2**- a symmetric 3D IC with 2 active dies and Pin-Fin inline HTG as shown in Fig. 6 (b). Inline pins with a diameter of 50µm and pitch of 100µm were modeled inside the cavity. Due to the increased complexity and the limited memory of 2GB the CFD model could only handle a 50µmX2mm slice of a 2mmx2mm chip stack. In addition, exploiting the symmetrical structure of the stack, only half the height of the stack was simulated with symmetric vertical boundary conditions in CFX. However, the entire height was simulated using the proposed 2RM-based CTTM. The resulting CFX model contained 720k nodes while the proposed model, with a thermal cell size of 50µmX100µm, contained only 960 nodes.

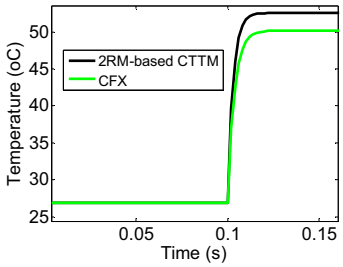



**Figure 8: Temperature waveforms for Test Stack 2 at the center of the die.**

In this experiment, **Test Stack 2** was simulated for 0.2 seconds. Until $t$=0.1s, no heat input was applied to the dies. At $t$=0.1s, a uniform 125W/cm$^2$ background heat flux was switched on both dies. The CPU time for the simulation and the resulting speed-up is shown in Table 2. The temperature waveform measured at the center of the die, from the proposed 2RM-based CTTM and CFX, are plotted in Fig.8. Here, we can identify a signification difference between the steady state temperatures from CFX and the proposed model (maximum error= 2.4$^o$C, or about 10%). However, as in the case of microchannels, the transients were captured accurately (error in $e^{-1}$ time constant=0.6ms).

### *C. Effects of Thermally Developing flow on the accuracy of the proposed model*

To examine inaccuracies in the steady state temperature estimation by the 2RM-based CTTM in the two experiments described above, the steady state thermal maps for both cases are studied against the corresponding results from CFX.

First, we consider **Test Stack 1** with uniform heat flux in all the three dies. In steady state, the area averaged wall temperatures and heat flux densities from solid to fluid at different points along the cavities are measured from both CFX and the 2RM-based CTTM. From the heat flux densities and corresponding fluid temperatures, the heat transfer coefficients (HTCs) are calculated for these points. The resulting wall temperatures and HTCs are plotted (for the second cavity alone) in Fig. 9 (a)-(b). As can be seen, there is relatively good agreement between CFX and the proposed model except for the region near the inlet. This is because the $R_{conv}$ used in the CTTM simulations are calculated using (3) and (5) and are therefore constant throughout the cavity, derived from correlations for fully developed flows from London and Shah [14]. However, the HTC is higher near the inlet due to the locally thin thermal boundary layer (Fig. 9 (b)) and approaches the fully developed value as we move along the cavity.

Next, the $R_{conv}$ measured from CFX (for all cavities) were substituted in the proposed model and the resulting steady state wall temperatures and HTCs are plotted in Fig. 9(c)-(d). As can be seen, the results now show considerable agreement.

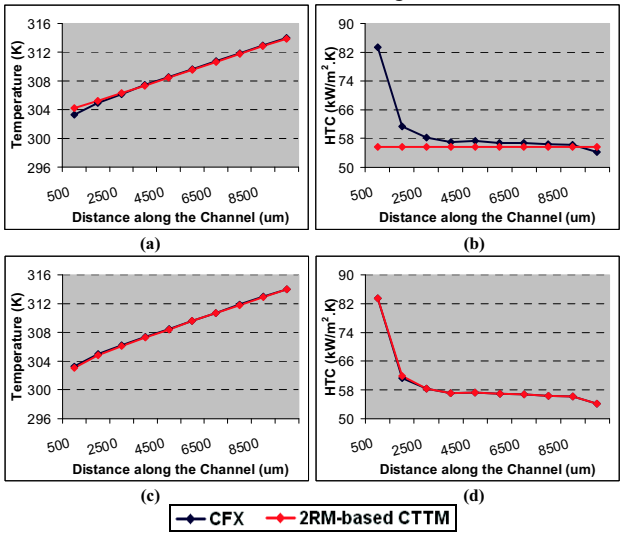



**Figure 9: Test Stack 1 in steady state without hotspot: (a) Wall temperatures and (b) HTC in case of constant $R_{conv}$ from (3) for the CTTM; (c) Wall temperatures and (d) HTC in case of locally modified $R_{conv}$ for the CTTM, as derived from CFX measurements.**

Next, the same experiment was repeated for **Test Stack 1**, but now with a 250W/cm2 hotspot at the center of the second die as shown in Fig. 6. The corresponding steady state measurements are plotted in Fig. 10(a)-(b). Here, in addition to the deviation near the inlet, we can see another deviation of the wall temperatures at the center of the cavity. This can be explained by the fact that on entering a region with a different heat flux from the walls, the thermal boundary layer again develops, resulting in yet another shift in the HTC (Fig. 10 (b)). Once again, after substituting for the $R_{conv}$ measured from CFX in the proposed model, the measurements show agreement in the steady state (Fig. 10(c)-(d)). Notice that in all cases, although the difference in the HTCs between CFX and CTTM is large (about 48%), it does not linearly translate to a large difference in wall temperatures (about 12%). This is because the thermal gradient along the channel walls does not depend upon the convection into the coolant alone, but is also a function of conduction along and around the channel walls, and the rise in the coolant temperature.

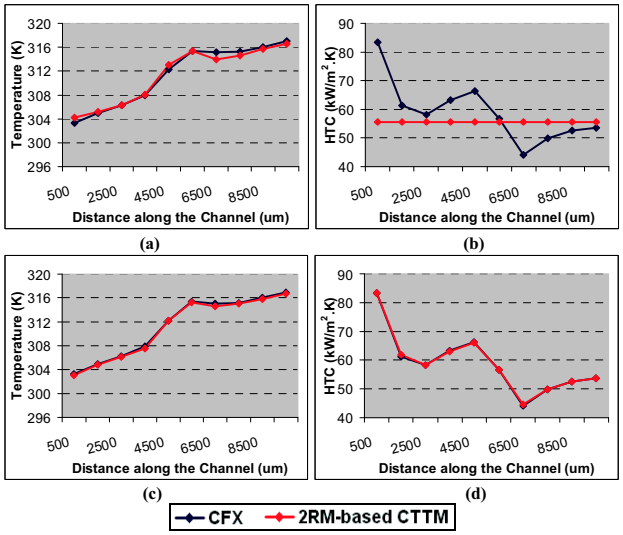



**Figure 10: Test Stack 1 in steady state with hotspot: (a) Wall temperatures and (b) HTC in case of constant $R_{conv}$ from (3) for the CTTM; (c) Wall temperatures and (d) HTC in case of locally modified $R_{conv}$ for the CTTM, as derived from CFX measurements.**

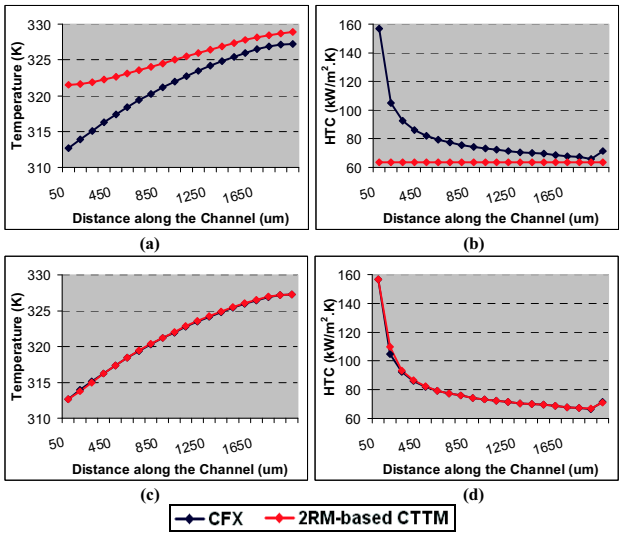



**Figure 11: Test Stack 2 in steady state: (a) Wall temperatures and (b) HTC in case of constant $R_{conv}$ from (8) for the CTTM; (c) Wall temperatures and (d) HTC in case of locally modified $R_{conv}$ for the CTTM, as derived from CFX measurements**

Finally, this experiment is repeated for **Test Stack 2**, excited with a uniform background heat flux as in subsection *IV-B*. The resulting wall temperatures and HTCs are plotted in Fig. 11 (a)-(b). It can be seen from these plots that the effects of thermally developing coolant is even more pronounced, with the inlet wall temperatures from the proposed 2RM-based CTTM showing a deviation of about 9K (about 40%). Again this because the $R_{conv}$ used in the proposed model was calculated using (8) derived from the correlations for fully developed flows from Brunschwiler et. al. [8]. In addition, the heat spreading due to conduction is not as significant for pin-fins as it is for microchannels. This results in highly uneven heating of the pin-fin cavity along its walls resulting in a much larger deviation in the measured wall temperatures.

### D. 2RM-based CTTM vs 4RM-based CTTM

A performance and accuracy comparison is drawn between the proposed 2RM-based CTTM and the 4RM-based CTTM advanced in [13] to highlight the advantages of the new approach.

For this, a 2-die **Test Stack 3** was constructed using the floorplan architecture studied in [16], with a single microchannel cavity in between a core die and a memory die. The cavity dimensions used are the same as that used for **Test Stack 1**. **Test Stack 3** was simulated using both the models and the resulting CPU times and the steady state temperature maps were recorded. The thermal cell size for the 4RM-based CTTM was 50µmX100µm (since the channel width and the channel wall width are 50µm) resulting in a problem size of 110500 nodes and simulation time of 402.17 seconds (marked using a red line in Fig. 13). The 2RM-based CTTM was simulated using various discretization cases- (i)100µm x 100µm, (ii)125µm x 125µm, (iii) 200µm x 200µm and (iv) 500µm x 500µm, resulting in various problem sizes.

The temperature map for the core layer calculated using the 2RM-based CTTM (Case (i)) is shown in Fig. 12. The CPU times and error percentages for all the cases are plotted against 2RM-based CTTM problem sizes in Fig. 13. The CPU times are plotted in a log scale. The error percentages were calculated as:

$$\text{error}\,\% = \frac{\max\left(T_{2RM} - T_{4RM}\right)}{\Delta T_{max,4RM}} \times 100, \qquad (8)$$

where $\Delta T_{max,4RM}$ is the maximum deviation of the temperatures calculated using the 4RM-based CTTM from the initial state.

As can be seen from Fig. 13, the 2RM-based CTTM allows for far coarser discretization of the problem domain improving the simulation speed, while retaining the accuracy of the 4RM-CTTM (For Case (iii), the CPU simulation time is 9.7 seconds and the error % 5.5%). This experiment demonstrates the advantages of the new approach.

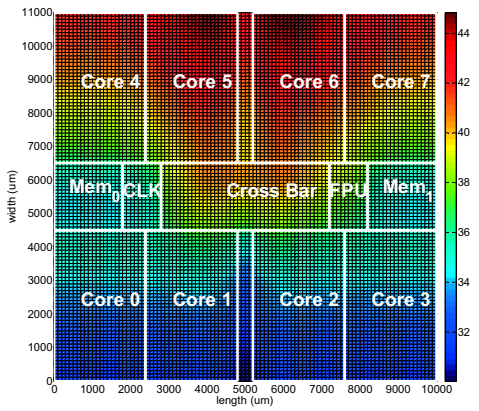



**Figure 12: Temperature map of the core layer of Test Stack 3 using 2RM-based CTTM (Case (i)).**

## V. CONCLUSIONS

In this paper we have presented a new approach for compact and transient thermal modeling of 3D ICs with inter-tier liquid cooling. The model, 2RM-based CTTM, is highly generic and flexible, and can incorporate heat transfer data from any source and produce fast thermal analyses of 3D ICs. The proposed model is as accurate as the heat transfer coefficients incorporated and offers significant speed-ups compared to commercially available CFD simulators. The accuracy, speed and flexibility of the proposed model have been demonstrated for both microchannels and pin-fin heat transfer geometries. In the studies performed using the proposed model, the effect of thermally developing flows was found to be significant in determining the final accuracy of the temperature estimation in some cases. Finally, the proposed model is shown to have a computational edge over the previous 4RM-based CTTM.

ACKNOWLEDGMENTS

This research has been partially funded by the Nano-Tera RTD project CMOSAIC (ref.123618), which is financed by the Swiss Confederation and scientifically evaluated by SNSF, and the PRO3D project, financed by the European Community 7th Framework Programme (ref.FP7-ICT-248776).

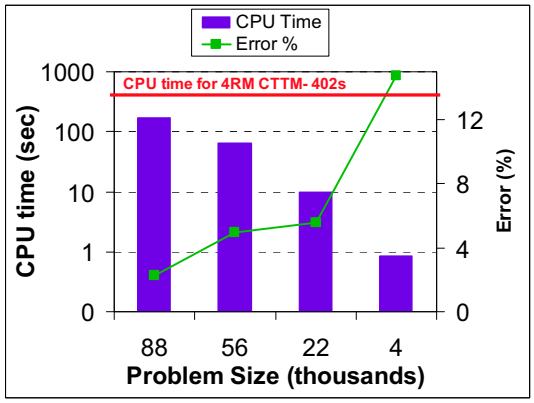



**Figure 13: CPU time comparison and error incurred by the 2RM-based CTTM w.r.t. 4RM-CTTM for various discretizations.**

REFERENCES

[1] T. Skotnicki et. al., "The end of CMOS scaling: toward the introduction of new materials and structural changes to improve MOSFET performance," *IEEE Circuits and Devices Magazine*, vol.21, no.1, pp. 16- 26, Feb 2005.

[2] F. Li et. al., "Design and management of 3D chip multiprocessors using network-in-memory," in Proc. ISCA, pp. 130–141, 2006.

[3] T. Brunschwiler et. al., "Forced convective interlayer cooling in vertically integrated packages," in Proc. ITHERM, 2008.

[4] T. Brunschwiler et. al., "Interlayer cooling potential in vertical integrated packages," Microsystem Technologies: MNSISPS, vol. 15, no. 1, pp. 57–74, 2009.

[5] T. Brunschwiler et. al., "Heat-removal performance scaling of interlayer cooled chip stacks", Proc. ITHERM 2010, pp. 1-12, June 2010.

[6] J. Lienhard-IV and J. Lienhard-V, *A heat transfer textbook*, Cambridge, Massachusetts: Phlogiston Press, 2006.

[7] F. Incropera, D. Dewitt, T. Bergman and A. Lavine, *Fundamentals of heat and mass transfer*, New York: John Wiley and Sons, 2007.

[8] T. Brunschwiler et. al., "Validation of the Porous-Medium approach to model interlayer-cooled 3D-chip stacks", Proc. 3DIC, 2009.

[9] W. Huang et. al., "HotSpot: A compact thermal modeling methodology for early-stage VLSI design," *IEEE Trans. VLSI Sys.*, vol. 14, pp. 501–513, May 2006.

[10] D. Tuckerman and R. Pease, "High-performance heat sinking for VLSI", *IEEE ED Letters*, vol.2, no. 5, pp. 126-129, 1981.

[11] J. Koo et. al., "Integrated microchannel cooling for 3-dimensional electronic circuit architectures", *ASME Journal HT*, vol. 127, pp. 49-58, 2005.

[12] H. Mizunuma et. al. "Thermal modeling for 3D-ICs with integrated microchannel cooling", Proc. ICCAD 2009, pp. 256-263, Nov 2009.

[13] A. Sridhar et. al., "3D-ICE: Fast Compact Transient Thermal Modeling for 3D ICs with inter-tier Liquid Cooling", in Proc. ICCAD, 2010.

[14] R. Shah and A. London, *Laminar flow forced convection in ducts*, New York: Academic Press, 1978.

[15] http://www.ansys.com/products/fluid-dynamics/cfx/.

[16] A. Coskun et. al., "Energy-efficient variable-flow liquid cooling in 3D Stacked architectures", Proc. DATE 2010, pp. 111-116, 2010.