



International Journal of Numerical Methods for Heat & Fluid Flow

Thermal simulations in support of multi-scale co-design of energy efficient information technology systems

Yogendra Joshi Banafsheh Barabadi Rajat Ghosh Zhimin Wan He Xiao Sudhakar Yalamanchili Satish Kumar

Article information:

To cite this document:

Yogendra Joshi Banafsheh Barabadi Rajat Ghosh Zhimin Wan He Xiao Sudhakar Yalamanchili Satish Kumar, (2015), "Thermal simulations in support of multi-scale co-design of energy efficient information technology systems", International Journal of Numerical Methods for Heat & Fluid Flow, Vol. 25 Iss 6 pp. 1385 - 1403

Permanent link to this document:

<http://dx.doi.org/10.1108/HFF-08-2014-0242>

Downloaded on: 30 August 2015, At: 14:51 (PT)

References: this document contains references to 25 other documents.

To copy this document: permissions@emeraldinsight.com

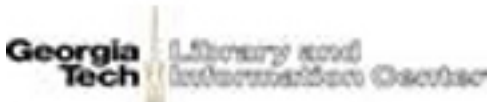
The fulltext of this document has been downloaded 14 times since 2015*

Users who downloaded this article also downloaded:

Abas Abdoli, George S. Dulikravich, Chandrajit L Bajaj, David F Stowe, Salik M Jahania, (2015), "Human heart preservation analyses using convective cooling", International Journal of Numerical Methods for Heat & Fluid Flow, Vol. 25 Iss 6 pp. 1426-1443 <http://dx.doi.org/10.1108/HFF-08-2014-0251>

Diego C. Knupp, Carolina Palma Naveira-Cotta, Adrian Renfer, Manish K. Tiwari, Renato M Cotta, Dimos Poulikakos, (2015), "Analysis of conjugated heat transfer in micro-heat exchangers via integral transforms and non-intrusive optical techniques", International Journal of Numerical Methods for Heat & Fluid Flow, Vol. 25 Iss 6 pp. 1444-1462 <http://dx.doi.org/10.1108/HFF-08-2014-0259>

Carlo Nonino, Stefano Savino, Stefano Del Giudice, (2015), "FEM for the 3-D analysis of conjugate conduction-convection heat transfer in cross-flow micro heat exchangers", International Journal of Numerical Methods for Heat & Fluid Flow, Vol. 25 Iss 6 pp. 1322-1339 <http://dx.doi.org/10.1108/HFF-07-2014-0232>



Access to this document was granted through an Emerald subscription provided by emerald-srm:168551 []

For Authors

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service information about how to choose which publication to write for and submission guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com

Emerald is a global publisher linking research and practice to the benefit of society. The company manages a portfolio of more than 290 journals and over 2,350 books and book series volumes, as well as providing an extensive range of online products and additional customer resources and services.

Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

*Related content and download information correct at time of download.

Thermal simulations in support of multi-scale co-design of energy efficient information technology systems

Energy
efficient IT
systems

1385

Yogendra Joshi

*Woodruff School of Mechanical Engineering, Georgia Institute of Technology,
Atlanta, Georgia, USA*

Banafsheh Barabadi

MIT, Cambridge, Massachusetts, USA

Rajat Ghosh

Mechanical Engineering, Georgia Tech, Atlanta, Georgia, USA

Zhimin Wan and He Xiao

*G.W. Woodruff School of Mechanical Engineering,
Georgia Institute of Technology, Atlanta, Georgia, USA*

Sudhakar Yalamanchili

*Electrical and Computer Engineering, Georgia Tech, Atlanta,
Georgia, USA, and*

Satish Kumar

Mechanical Engineering, Georgia Tech, Atlanta, Georgia, USA

Received 3 August 2014
Revised 2 January 2015
Accepted 6 January 2015

Abstract

Purpose – Information technology (IT) systems are already ubiquitous, and their future growth is expected to drive the global economy for the next several decades. However, energy consumption by these systems is growing rapidly, and their sustained growth requires curbing the energy consumption, and the associated heat removal requirements. Currently, 20-50 percent of the incoming electrical power is used to meet the cooling demands of IT facilities. Careful co-optimization of electrical power and thermal management is essential for reducing energy consumption requirements of IT equipment. Such modeling based co-optimization is complicated by the presence of several decades of spatial and temporal scales. The purpose of this paper is to review recent approaches for handling these challenges.

Design/methodology/approach – In this paper, the authors illustrate the challenges and possible modeling approaches by considering three examples. The multi-scale modeling of chip level transient heating using a combination of Progressive Zoom-in, and proper orthogonal decomposition (POD) is an effective approach for chip level electrical/thermal co-design for mitigation of reliability concerns, such as Joule heating driven electromigration. In the second example, the authors will illustrate the optimal microfluidic thermal management of hot spots, and large background heat fluxes associated with future high-performance microprocessors. In the third example, data center facility level energy usage reduction through a transient measurements based POD modeling framework will be illustrated.

Findings – Through modeling based electrical/thermal co-design, dramatic savings in energy usage for cooling are possible.

Originality/value – The multi-scale nature of the thermal modeling of IT systems is an important challenge. This paper reviews some of the approaches employed to meet this challenge.

Keywords POD, Co-design, Multi-scale, Thermal modeling, Microfluidic cooling, Transient heating

Paper type General review



Nomenclature

(c_1, c_2, c_3)	empirical constants	λ_i	eigenvalue corresponding to
C_p	specific heat	POD modes	
E_m	cumulative correlation energy	θ	normalized time
g	normalized temporal resolution	<i>Subscripts</i>	
h	heat transfer coefficient, W/m ² K	<i>amb</i>	ambient
h	normalized spatial resolution	<i>f</i>	fluid
K	thermal conductivity, W/mK	<i>cond</i>	conduction
l	size of input space	<i>conv</i>	convection
\dot{m}	mass flow rate, kg/s	<i>gen</i>	generation
\dot{q}	heat transfer rate, W	<i>Acronyms</i>	
Q	total chip power, W	BT	bismaleimide triazine
t	time, s	PCB	printed circuit board
T	temperature, K	FE	pipeline frontend and L1 instruction cache
T_0	ambient temperature, K	SC	out-of-order scheduler
T_{bottom}	temperature at the bottom surface of package, K	Int	integer unit
T_s	temperature at the bottom surface of package, K	FPU	floating-point unit
z	height, mm	DL1	L1 data cache

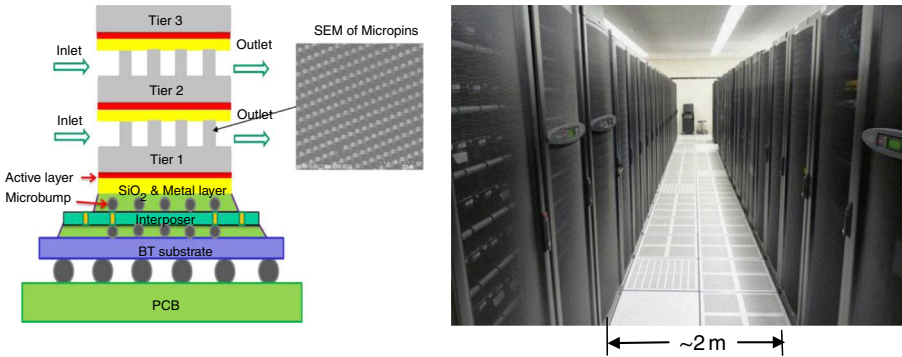
1. Introduction

Thermal management of information technology (IT) systems involves over ten decades of length scales, from on-chip interconnects (currently ~22 nm) to large data center facilities (~50-100 m in length), and a multitude of time scales. Thermal modeling at various scales is essential during design phase to insure performance and reliability. An emerging focus area is co-design, involving the consideration of two or more aspects of the overall design, with the objective of reducing overall energy consumption. Energy consumption in IT facilities and equipment is a growing concern, which is driving this interest. Figure 1 shows two dissimilar length scales, where the thermal analysis for co-design is illustrated in this paper.

On the left is a chip package. The current baseline is a single-tier package with air cooling. Temperature computations within such packages with multi-scale resolution

Figure 1.

A cm scale 3D stacked chip package with microfluidic cooling (left), and data center (right) illustrate the multi-scale spatial nature of thermal modeling needs in IT systems



Note: Both are considered in this paper to illustrate co-design approaches

are needed under both steady state and transient operation, for assessing reliability. These are explored in Sections 2-4. The packaged two-tier chip stack incorporating microfluidic cooling, shown in Figure 1 illustrates future packaging direction, to achieve the continuing goals of faster and more functional systems. In Section 5, this configuration is analyzed in an electrical/thermal co-design framework to optimize cooling. The image on the right in Figure 1 illustrates a data center aisle, with perforated tiles for cooling air delivery. Efficient computations of inlet air temperatures at servers allow for controlling cooling hardware to optimize energy consumption. A data-driven modeling framework for achieving this is presented in Sections 6-8.

2. Chip/package level thermal modeling for integrated circuit (IC) design and reliability analysis

Key factors impacting IC chip performance and reliability are the knowledge of thermal and electrical performance. As the size of on-chip interconnects reduces below the electron mean free path (~ 40 nm in Cu at room temperature), electron transport becomes dominated by scattering at the metal-dielectric interface, and at grain boundaries. This scattering can decrease the electrical and thermal conductivity to less than half that of the bulk value (Fuchs, 1938; Sondheimer, 1952; Ziman and Levy, 1961; Namba, 1970; Soffer, 1967; Gurram *et al.*, 2004) making Joule heating a major concern in IC industry, requiring a comprehensive predictive framework.

It is important to note that on-chip interconnects are not isolated entities. They are arranged as a stack of metal lines, vias, and dielectric layers, as a part of a chip that is ultimately enclosed in a package. The majority of heat is generated in the transistors and metallic interconnects, with characteristic length scale of tens of nm ($\sim 10^{-8}$ m). This heat is then conducted through the passive/active components, substrate, and heat sink, prior to being convectively removed by the ambient environment. At this level, the characteristic length scale is on the order of tens of cm (10^{-1} m). Therefore, the effect of the entire structure should be accounted for, to acquire an accurate and realistic thermal simulation of interconnects. In addition, transient thermal events in the interconnect structure can vary from tens of s (10^2 s) at the package level to μ s (10^{-6} s) at the interconnect level. Thus, the heat diffusion problem in interconnects is a multi-scale problem, both in length and time scales.

The traditional finite difference (FD) and finite element (FE) methods require a large number of computational nodes for such problems. As a result, computational times are long, even for a unit cell (micro-model), and hence unfeasible for parametric studies. Analytical methods, less complex in geometry, are not able to capture the important features of the thermal problem. Additionally, previous models have been primarily concerned with the steady-state Joule heating in interconnects, while the pulsed nature of the electric currents necessitates studying the effects of transient heat conduction in the interconnects as well. This confirms the need for the development of high-fidelity transient multi-scale thermal models that can address the aforementioned issues.

Barabadi *et al.* (2012) developed a computationally efficient hybrid multi-scale transient thermal modeling methodology. It incorporates two different multi-scale modeling approaches: “Progressive Zoom-in” (e.g. Tang and Joshi (2005)), and “Proper Orthogonal Decomposition (POD)” (Lumley, 1967). This integration resulted in a model that has the capability of handling several decades of length scales, from tens of mm at “package” level to several nm at “interconnects”, corresponding to various thermal events. This ability also applies for time scales from s to μ s, at significantly lower

computational cost, without compromising the desired accuracy. Hybrid scheme also provides the ability to rapidly predict thermal responses under different power input patterns, based on only a few representative detailed simulations (Barabadi *et al.*, 2015). It is demonstrated that utilizing the proposed model, the computational time is reduced by at least two orders of magnitude at every step of modeling.

To further describe the hybrid scheme, a chip and its function blocks, in a Flip Chip Ball Grid Array (FCBGA) package are considered. The hybrid scheme utilizes a combination of Progressive Zoom-in, and POD as an effective approach for chip level electrical/thermal co-design for mitigation of reliability concerns, such as Joule heating driven electromigration. Random transient power distributions were assumed for the chip and its individual function blocks to demonstrate the capability of POD method in predicting different thermal scenarios. To validate this methodology, the results were compared with a FE model developed in COMSOL[®]. The results of the model developed through hybrid scheme were in good agreement with the corresponding FE models.

3. Hybrid scheme for multi-scale thermal modeling

As mentioned earlier, hybrid scheme combines the implementation of POD technique and Progressive Zoom-in approach.

3.1 POD method

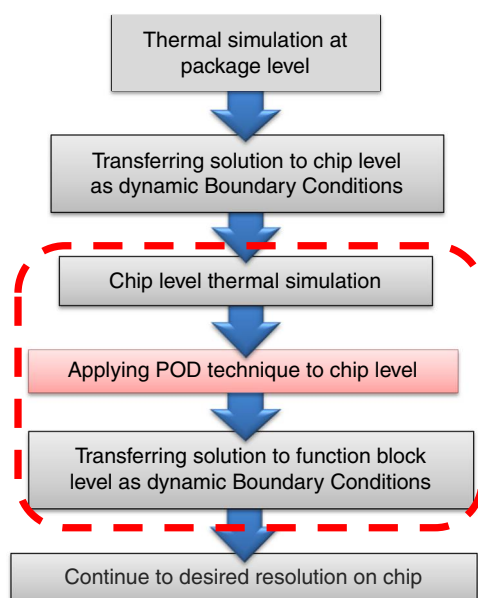
POD is a powerful method of data analysis that offers low-dimensional but accurate descriptions of a high-dimensional system. A comprehensive summary for applications of POD is given by Holmes *et al.* (1998). A detailed procedure to generate a two dimensional (2D) POD-based reduced order model is provided in Barabadi *et al.* (2011). The primary steps to generate a POD-based reduced order model are: generating the observation matrix; calculating basis functions (POD modes); calculating POD coefficients; and generating the POD temperature field (Barabadi *et al.*, 2015). Combining POD with the Progressive Zoom-in approach greatly improves the computational efficiency.

3.2 Progressive Zoom-in approach

Progressive Zoom-in method is a top-down approach to model transient thermal problems. It integrates package, chip, sub-chip level analyzes, acquiring the advantages of each (e.g. Tang and Joshi, 2005). It has the capability of covering several orders of magnitude in length and time scale (Barabadi *et al.*, 2011).

The overall hybrid approach is demonstrated in the flowchart provided in Figure 2. The section marked via dashed lines represents the area used in this paper to illustrate the proposed multi-scale transient thermal modeling approach. The scheme is outlined in Figure 2, and is further explained through a case study in the following section.

At first, the entire package, including the surrounding components such as the mold compound, solder bumps, substrate, etc. are numerically modeled and the transient thermal solution is obtained. At this level, the chip is modeled as a solid block with uniform and effective material and thermal properties and no internal details. The transient thermal solution is then transferred into the chip in the form of dynamically changing boundary conditions. The next step is to develop the thermal model of the chip using the transferred boundary conditions. At this step, the internal function blocks of the chip with their nonhomogeneous material and thermal properties are included in the modeling. The transient thermal solution at the chip level is then



Source: Reprinted with permission Barabadi *et al.* (2012)

Figure 2.
Flowchart of the
hybrid scheme for
multi-scale thermal
modeling

determined by assigning the block specific power profiles. Afterwards, by following the procedure explained in Section 3.1, the POD model is developed at the chip level. The POD model has the ability to predict the transient temperature profile in the chip for different types of power sources and power maps at a much lower computational cost. If further spatial resolution is desired, this approach can be utilized by repeating the previous steps until the desired resolution is achieved.

4. Case study and results

The focus of this study is to develop a transient thermal model at the chip and its function block level, with a spatial resolution of $1\ \mu\text{m}$. The modeling steps are described as follows.

4.1 Thermal simulation at the package level

In order to have a realistic thermal solution at the chip level, we start from modeling the entire FCBGA package (shown in Figure 3(a)), including the surrounding mold, underfill, solder bumps, substrate, etc., using COMSOL software. At this level, the chip is treated as a solid block with effective material and thermal properties, without considering internal details. The details of the package level modeling including the thermal and material property calculations are provided in Barabadi *et al.* (2012, 2015). This package is used for low power portable systems where compact form factor prohibits the use of heat sinks and forced cooling. Natural convection boundary condition with a heat transfer coefficient of $h = 15\ \text{W/m}^2\text{K}$ (typical range for air; Tang and Joshi, 2005) was applied to the top surface and a constant temperature boundary condition is assigned to the bottom surface.

HFF
25,6

A full FE model is developed for the entire package consisting of 75,919 elements, of which 343 are for the chip. The grid size is chosen after performing mesh independence analysis. Total power to the chip is $Q = 3 \sin 2\pi t + 3$ (W), which is applied for 1 s. The FE results for the spatial distribution of temperature in the package and chip after 1 s is displayed in Figure 4(a) and (b), respectively.

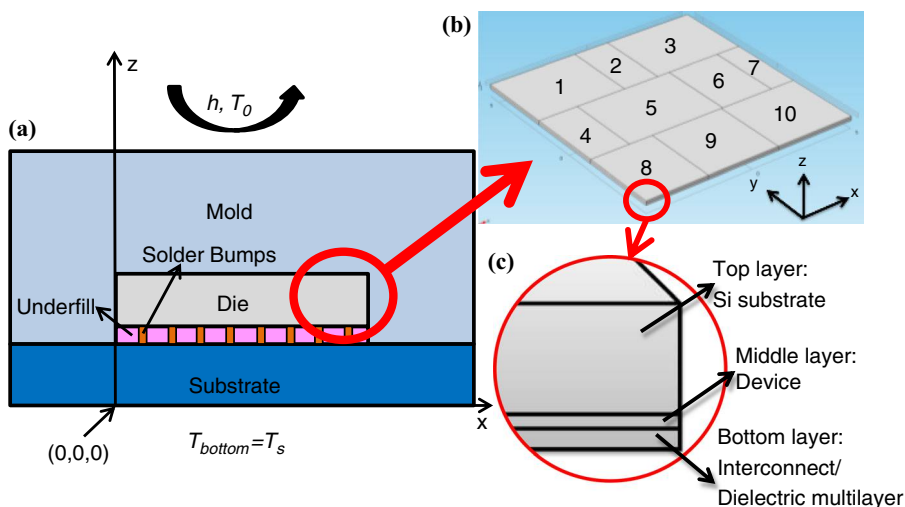
1390

4.2 Transferring the solution from package level to the chip level

Once the transient thermal solution is obtained at the package level for 1 s, the solution is then dynamically transferred into the chip every 0.1 s as boundary conditions. On the top surface temperature, and on the bottom surface heat flux are applied to the chip. Due to the high-aspect ratio of the chip, the side walls are considered adiabatic. Afterwards, the thermal solution is then linearly interpolated on each surface with much higher spatial resolution (268,033 elements to model the chip at this level vs 343 elements to model the chip at the package level).

Figure 3.

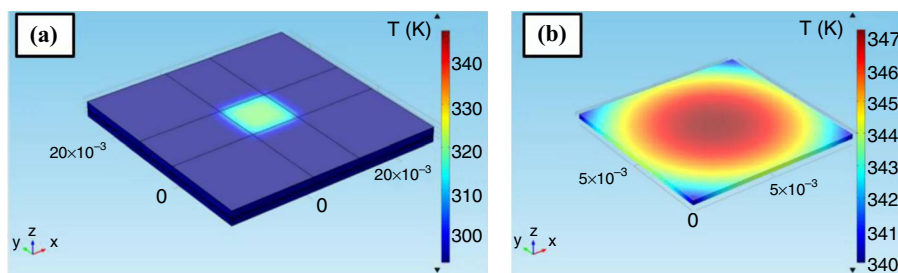
(a) Schematic of a simplified FCBGA for package level modeling, (b) zoomed-in outline of the chip with artificial block, and (c) zoomed-in schematic of the layers of the die used in chip level modeling



Source: Reprinted with permission Barabadi *et al.* (2012)

Figure 4.

Spatial distribution of temperature (K) extracted from FE method after 1 μ s for (a) FCBGA package and (b) chip



Source: Reprinted with permission Barabadi *et al.* (2012)

4.3 Chip level thermal simulation

At this level, chip is no longer assumed to be a solid block. It is now divided into ten subdomains called function blocks which symbolize various components on the chip (shown in Figure 3(b)). Each block has three layers, displayed in Figure 3(c). These layers are: top Si layer with the thickness of 0.249 mm; middle layer which is a 5 μm -thick device layer, and Interconnect/Dielectric multilayer at the bottom with the thickness of 16.72 μm . The third layer consists of 21 sub-layers including ten metal layers. Details on the calculation of thermal and material properties of the function blocks can be found in Barabadi *et al.* (2012, 2015). In order to have a realistic and accurate thermal simulation, a detailed dynamic power map of the function blocks are required which is discussed in the following section.

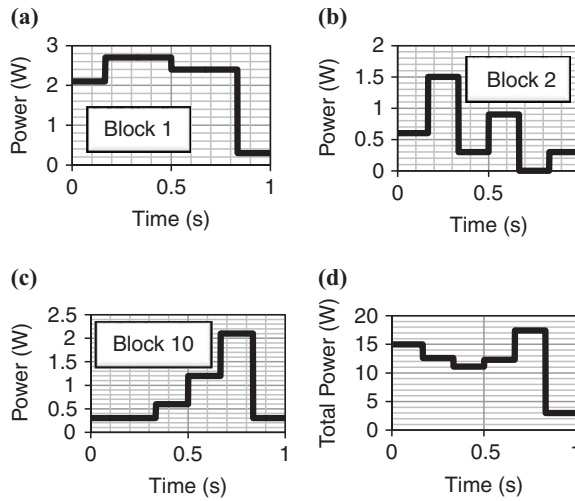
4.4 Applying POD technique to chip level

After determining the temperature distribution at package level and transferring it to the chip level as boundary conditions, a POD model is developed for chip level analysis using the procedure discussed in Section 3.1. Total of six observations based on the transient temperature solution were taken in the first 0.5 s using the package level FE thermal solution that is transferred on to the chip level model setup. These observations correspond to the temperature solutions obtained at different time instants. As previously demonstrated by Barabadi *et al.* (2011) for any linear system, POD method has the ability to predict transient thermal solution regardless of the temporal or spatial dependence of the applied heat source. In other words, the POD solutions of these transient thermal scenarios are independent of the initial observations. Essentially, for any linear system, once the solution to a sample case of chip total power is obtained, there is no need to generate new observations, or full field FE simulations. This feature provides the ability to predict temperature distributions for arbitrary heat inputs, by using a smaller sample set of applied heat sources and power maps, which results in considerably decreased simulation time.

To further verify this feature of POD, a randomly generated transient power distribution, shown in Figure 5(d), is used in the POD thermal model development. As can be implied from Figure 5(d) the new power profile differs from the chip power in duration, magnitude, and temporal dependency. Furthermore, since at this level of simulation the chip is no longer treated as a solid block and is divided into subdomains, the dynamic power grid also needs to be assigned to individual function blocks. Therefore, ten random power profiles with minimum and maximum values of 0 and 3 W were generated between zero and 1 s and assigned to each block in such a way that their sum will equal the total chip power as shown in Figure 5(d). The power sources for blocks 1, 2, and 10 are presented in Figure 5(a)-(c) as representatives.

After assigning the function blocks power profiles, the POD modes are calculated. In order to have a reliable but fast model, only five POD modes are utilized. This is selected such that the cumulative correlation energy, E_m (Bizon *et al.*, 2008) is larger than 99.0 percent. Because the POD modes are 3D, for better visualization, 2D contour plots of the first six POD modes at the interface of device and interconnect/dielectric layers are depicted in Figure 6. In this figure, POD modes are normalized with respect to the total sum of the generated modes. The POD coefficients, b_i , were then calculated as functions of time using the method of Galerkin Projection (Barabadi *et al.*, 2015).

After calculating the POD modes and coefficients, the transient thermal solution can be obtained. Figure 7 shows the spatial distribution of temperature after 0.95 s based on



Source: Adapted from Barabadi *et al.* (2015)

Figure 5. Dynamic randomly generated power profile for function blocks 1, 2, 3 (a-c) and randomly generated total chip power (d)

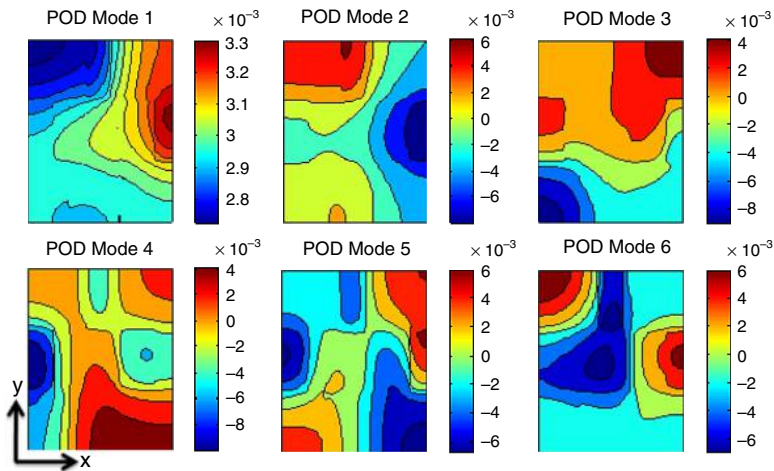
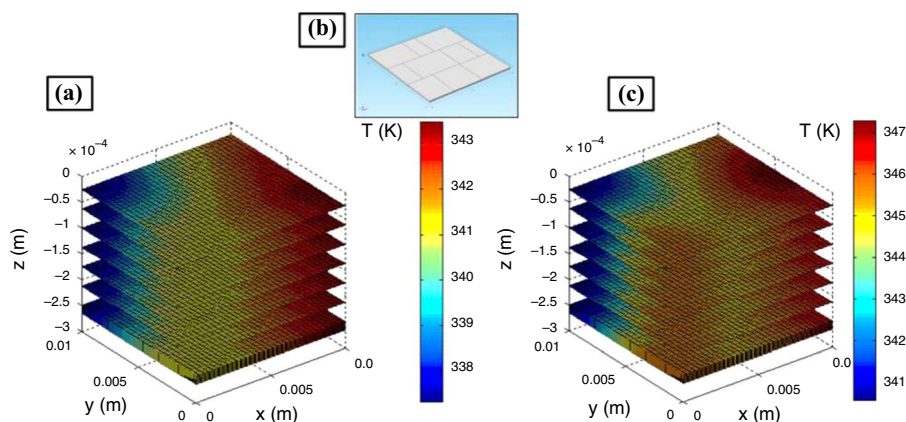


Figure 6. 2D contour plots of the first six POD modes at the interface of device and interconnect/dielectric layers

POD method (Figure 7(a)) and full field FE model (Figure 7(c)) inside the chip and its function blocks (Figure 7(b)). The temporal distribution of temperature based on POD model and FE analysis at four different points inside the function blocks are also provided in Figure 8. The mean absolute error between the POD and FE model is 8 percent over the entire space and time domain. It must be noted that the computational time required for POD model is approximately 20 s, which is two orders of magnitude less than that of FE results (~27 min).

4.5 Transferring the solution from chip level to the function block level

If higher spatial resolution on the function blocks are required, the procedure in Section 3.2 can be repeated and transient thermal solution can be transferred to any of



Notes: Extracted from the POD model (a) and FE simulation (c). The domain is sliced horizontally along the XY plane. The outline of chip with the function blocks are shown in (b)

Figure 7.
Spatial distribution
of temperature (K)
at 0.95 s

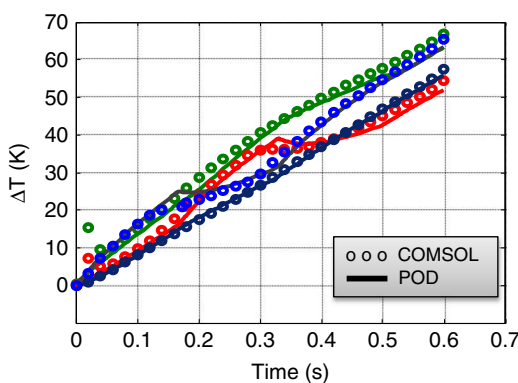


Figure 8.
Comparison of
temporal dependence
of temperature rise
between FE
(markers) and POD
(solid lines) models
at four different
points inside the die

the function blocks in terms of boundary conditions. In general, procedures 3.1-3.4 can be repeatedly applied, until the desired resolution on the chip is achieved.

5. Thermal modeling for co-design of microfluidically cooled 3D ICs

Wire interconnections become longer in 2D topology, with increasing number of transistors, which can increase the latency in the wire and reduce the expected benefit from integration (Meindl, 2003); 3D ICs as the next generation IC technology achieve higher integration with the technology of through-silicon vias. The vertical integration can reduce the global wire length by as much as 50 percent (Khan *et al.*, 2009). The wire limited clock frequency could be increased nearly fourfold (Bamal *et al.*, 2006). However, the increased power density due to the stack results in great thermal challenges. If the heat cannot be removed effectively, the temperature rise will degrade the electrical performance, and adversely impact reliability. One important drawback of high

temperature in ICs is the non-linear increase in leakage power (Roy *et al.*, 2003). The leakage power cannot be used by the transistors, and further exacerbates the thermal management challenge. Inter-tier microfluidic cooling has been proposed as a possible solution to the thermal problem of 3D ICs due to its high performance and scalability (Bakir and Meindl, 2008).

In order to design single phase microfluidic cooling, accurate thermal modeling is needed. However, computational fluid dynamics/heat transfer (CFD/HT) based detailed numerical analysis methods, such as FE or finite volume, need very fine mesh, and are time-consuming, and thus not suitable for optimization of complex 3D architectures. Compact or reduced thermal modeling is typically used to improve computational efficiency. A fast and accurate thermal circuit model for 3D ICs with an integrated microchannel heat sink was proposed (Mizunuma *et al.*, 2009). This model achieved more than $3,300\times$ speedup with less than 5 percent error, in comparison with a commercial numerical finite volume simulation tool. However, it was only applicable to steady state conditions and numerical pre-simulations were needed for every simulation executed, which further increased the calculation time.

A 3D-ICE is a transient compact thermal model, which takes into account the effects of the inter-tier cooling through a microchannel heat sink (Sridhar *et al.*, 2010). Only one node was used for each channel block, which reduced the problem size greatly. Empirical correlations for heat transfer coefficient were used, so no detailed convective heat transfer analysis was needed. However, the empirical correlations utilized were for fully developed flow and did not capture the developing flow regime.

Another technique to model the integrated microchannel cooling of 3D ICs was based on modeling the channel layer as a porous medium (Sridhar *et al.*, 2010b). A similar porous medium model as in (Sridhar *et al.*, 2010b) was developed to analyze the thermal characteristics of 3D ICs with integrated microgaps. One limitation of the previous thermal resistance method (Sridhar *et al.*, 2010a) was that the grid size must be the same as the channel pitch. Using the porous medium model removed this limitation by homogenizing the channel layers. As such, it provided freedom to increase the grid size, resulting in faster simulations.

The main problems of the existing thermal modeling of 3D ICs incorporating microfluidic cooling are: the power map is either assumed uniform or not realistic. When a specific application runs on a chip, the power should be non-uniform; the coupling effect between power and thermal is not included. Increased temperature could lead to higher consumption of power, due to increased leakage power. The leakage power raises the temperature further. So it is important to include the temperature dependent leakage power in the thermal model.

Figure 9 shows the schematic of physical model of 3D ICs with inter-tier microfluidic cooling (Wan *et al.*, 2013). The two-tiers stack with processor and L2 cache memory is placed on the substrate through interposer. Each tier consists of a SiO_2 and metal layer, an active layer, and a silicon base layer. SiO_2 is used for insulation and bonding. Active layer is a thin layer in which most of the heat is generated. Staggered micro pin fin array used for cooling is fabricated between the two tiers by Deep Reactive Ion Etching. Signal vias are embedded in the pin fins to provide communication between the two tiers. The stack is electrically connected to the outside through the Si interposer. The chip size is $8.4\text{ mm} \times 8.4\text{ mm}$. Thickness of the silicon base and is $100\text{ }\mu\text{m}$. Thickness of SiO_2 and metal is $10\text{ }\mu\text{m}$. The active layer is assumed zero thickness for simplicity.

The pin fin diameter is $100\text{ }\mu\text{m}$, height $300\text{ }\mu\text{m}$, and pitches in both directions are $200\text{ }\mu\text{m}$. De-ionized (DI) water with inlet temperature 20°C is used for cooling.

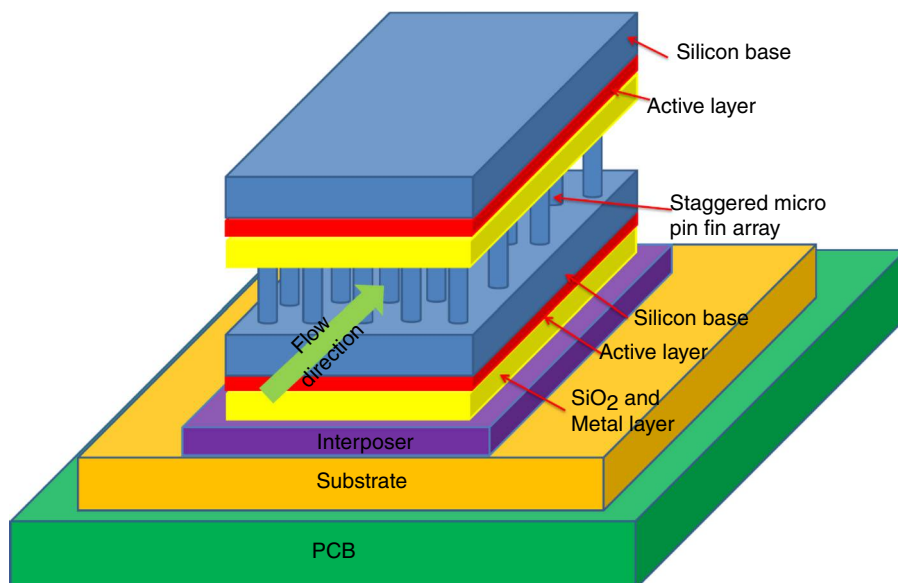
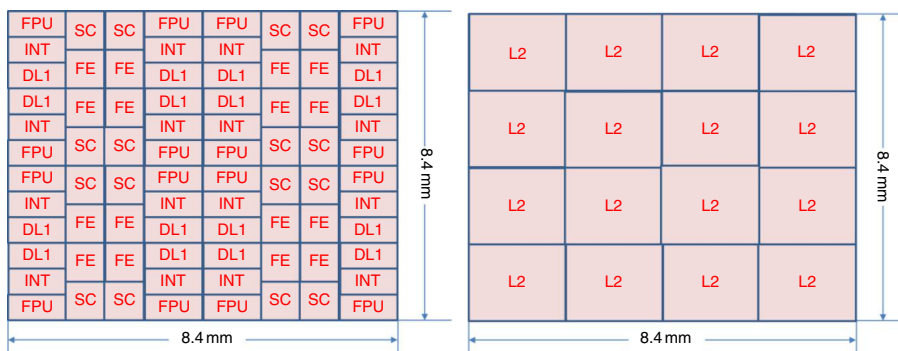


Figure 9.
Schematic of
physical model of
3D ICs with inter-tier
microfluidic cooling

A coupled power-thermal mode is used to model the 3D ICs in Figure 9. Figure 10 shows the floorplans for processor and L2 cache. The processor is a $\times 86$ 16-cores processor resemble the Intel Nehalem structure, a typical Out-of-Order microarchitecture including: FE (pipeline frontend and L1 instruction cache), SC (Out-of-Order scheduler), INT (integer unit), FPU (float-point unit) and DL1 (L1 data cache). Each core has a private L1 data cache of 128 KB and a shared L2 cache. The L2 cache layer consists of 16S RAM banks each with size of 2 MB.

A cycle-level timing model is used to execute compiled 32-bit $\times 86$ binaries. Simulations are run for 500 M clock cycles to warm up the processor state and reach a “region of interest” where the computational characteristics are representative in the benchmark program. The power at each block in the processor floor plan is sampled



Notes: (a) Processor; (b) L2 cache

Source: Wan *et al.* (2013)

Figure 10.
Floor plans

every 10 μs to produce a power trace. Table I shows the power distribution for each module. The total power of processor is 165.5 W with maximum heat flux 301.3 W/cm². The total power of L2 cache is 49.6 W. The power distribution will in input to the thermal model.

First, the 3D stack is discretized into multiple control volumes. Each of the control volumes is around one pin and consists of the fluid and solid domain. The temperatures of the two active layers and the fluid within each control volume are assumed uniform and stored in three 2D matrix. Figure 11 shows the energy flow within each control volume. Each tier has in-plane conduction from four directions. Also there are conduction between the two tiers, conduction between the stack and PCB, convection between the stack and coolant.

The thermal model is based on finite volume approximation of the energy equations for both the solid and fluid within the control volumes:

$$\text{Solid: } \dot{q}_{gen} + \dot{q}_{cond} + \dot{q}_{conv} = 0 \quad (1)$$

$$\text{Fluid: } \dot{m}C_p(T_{f,in} - T_{f,out}) + \dot{q}_{conv} = 0 \quad (2)$$

where \dot{q}_{gen} is power dissipation. It consists of two parts: dynamic power from the power simulation, and static power, which depends on temperature, \dot{q}_{cond} , \dot{q}_{conv} are the heat conduction and heat convection rate terms, \dot{m} is the mass flow rate into the control volume. C_p is the heat capacity of the fluid. A system of energy equations could be built for the 3D ICs. The conduction and convection heat rate can be determined by the temperatures of two tiers and fluid. So the system of energy equations could be transferred to a system of equations with only temperatures as variables. After iterative calculations, the temperature within each control volume can be obtained.

Figure 12 shows the temperature distribution of processor and L2 cache when the pumping power is 0.02 W and processor is located at the bottom of the stack.

Table I.

Power distribution
for each module

	FPU (W)	INT (W)	DL1 (W)	SC (W)	FE (W)	L2 (W)
Core 9	1.16	1.52	2.68	1.73	0.96	3.1
Other cores	1.69	2.06	2.81	2.31	1.63	3.1

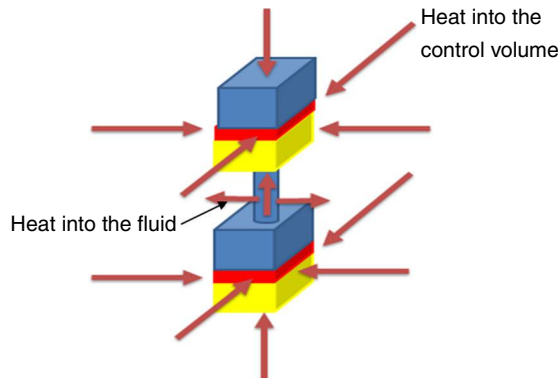


Figure 11.

Energy flow within
one control volume

The temperature of processor is non-uniform due to the non-uniform power distribution. The maximum temperature is near the exit of the fluid due to the bulk fluid temperature. Although the power distribution of L2 cache is uniform, its temperature distribution is non-uniform due to the cross-tier heat conduction. Table II shows the maximum temperature and leakage power under various pumping power. When the pumping power is 0.005 W, the maximum temperature is 123.1°C and the leakage power is 26.7 W. If the air cooled heat sink instead of inter-tier microfluidic cooling is used for cooling, the temperature and leakage power would be much higher. When the pumping power is increased to 0.12 W, the maximum temperature is reduced to 65.7°C and leakage power is reduced to 7.3 W. This proves the inter-tier microfluidic cooling could achieve significant performance enhancement.

6. Data-driven modeling of air temperatures near data center racks

POD, introduced in Section 3.1, can pave the way for an efficient decision support system to optimize the energy consumption of a data center. The most attractive feature of POD is its computational efficiency and flexibility. Currently, data center cooling hardware is controlled based on air temperature measurements. Sampling rate is a critical cost item for such transient temperature data acquisition. High sampling rate means larger sampling cost and vice versa. To solve this problem, we conducted a case study with the hypothesis that a POD-based data-driven algorithm can improve the temporal granularity of the temperature data, acquired using selected sensors. Data center air temperatures are a function of time and rack heat load. Depending upon the sampling intervals of time and rack heat load, air temperature interrogation point can lie on one of the nine possible interrogation spaces, as shown in Figure 13. Four of

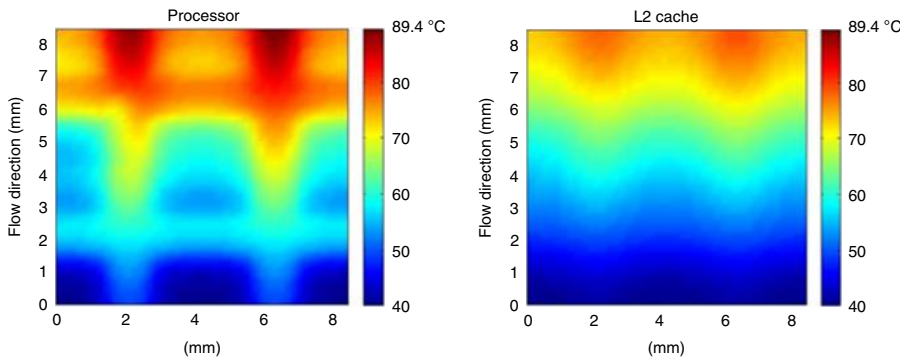


Figure 12.
Temperature
distribution of
processor and
L2 cache, pumping
power 0.02 W

Pumping power (W)	$T_{\max} (^{\circ}\text{C})$	Leakage power (W)
0.005	123.1	26.7
0.01	110.5	17.2
0.015	101.6	12.1
0.02	89.4	8.9
0.04	80.1	7.6
0.08	70.9	7.4
0.12	65.7	7.3

Table II.
Maximum
temperature and
leakage power under
various pumping
power

HFF
25,6

1398

these nine zones, namely Zones A-D convey most physical significance. Zone A represents normal mode of operation, Zone B represents failure mode, Zone C represents critical mode, and Zone D represents retrofit mode. The modeling fidelity is documented in Figure 14. It shows the maximum errors for Zones A-D are equal to 1, 9, 12.5, and 12.5 percent, respectively, similar to Ghosh and Joshi (2014b) for slightly different conditions.

7. Prognosis-based reliability modeling of data center air temperatures

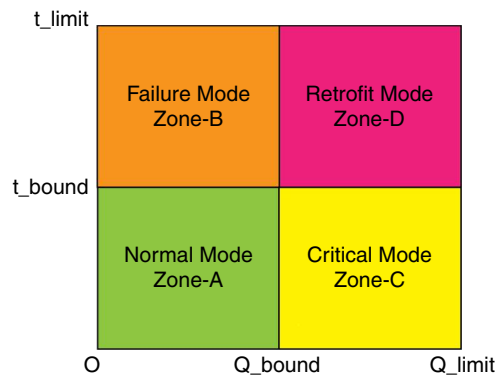
Temperature prognosis, or future prediction, at critical locations, especially at server inlets, is potentially beneficial for the reliability monitoring of data centers. POD facilitates an efficient temperature prediction tool; however, the modeling accuracy is a critical concern for the prognosis. Therefore, error conditioning is an important requirement. The error – defined as the differences between data and corresponding predictions – can be computed using a data-driven approach or by a semi-analytical formulation. The semi-analytical formulation is more cost-effective because it requires minimal hardware deployment. The analytical error is given by:

$$E = c_1(g+h) + \left[\frac{l \exp(\theta)}{1-\theta} (c_2 + c_3 g) \sum_{i=k+1}^n \lambda_i \right]^{\frac{1}{2}} \quad (3)$$

where g is the normalized temporal resolution, h is the normalized spatial resolution, l is the size of input space, θ is the normalized time, λ_i is the eigenvalue corresponding to POD modes, k is the number of retained POD modes, c_1 , c_2 , c_3 are arbitrary constants.

Figure 15 shows the numerical matching procedure to determine the empirical constants (c_1 , c_2 , c_3). The redline with open triangular markers shows how the data-driven prediction errors–deviations between data and predictions – vary with time. The predictions are computed using input data generated by measuring transient response of air temperature ensuring sudden switch-on of the test computer room air conditioning unit. The temperature data are collected in [s interval at 10 s sampling period. Figure 15 shows the extrapolation horizon starting from 201 s. The extrapolation window lasts till 224 s, because the error becomes higher than an arbitrarily-specified absolute error bound of 2°C. The empirical constants (c_1 , c_2 , c_3) are determined by the conjugate gradient method. This matching shown in Figure 15 is conducted for

Figure 13.
Two dimensional
parametric space,
comprising heat load
and time, for data
center air
temperature
prediction



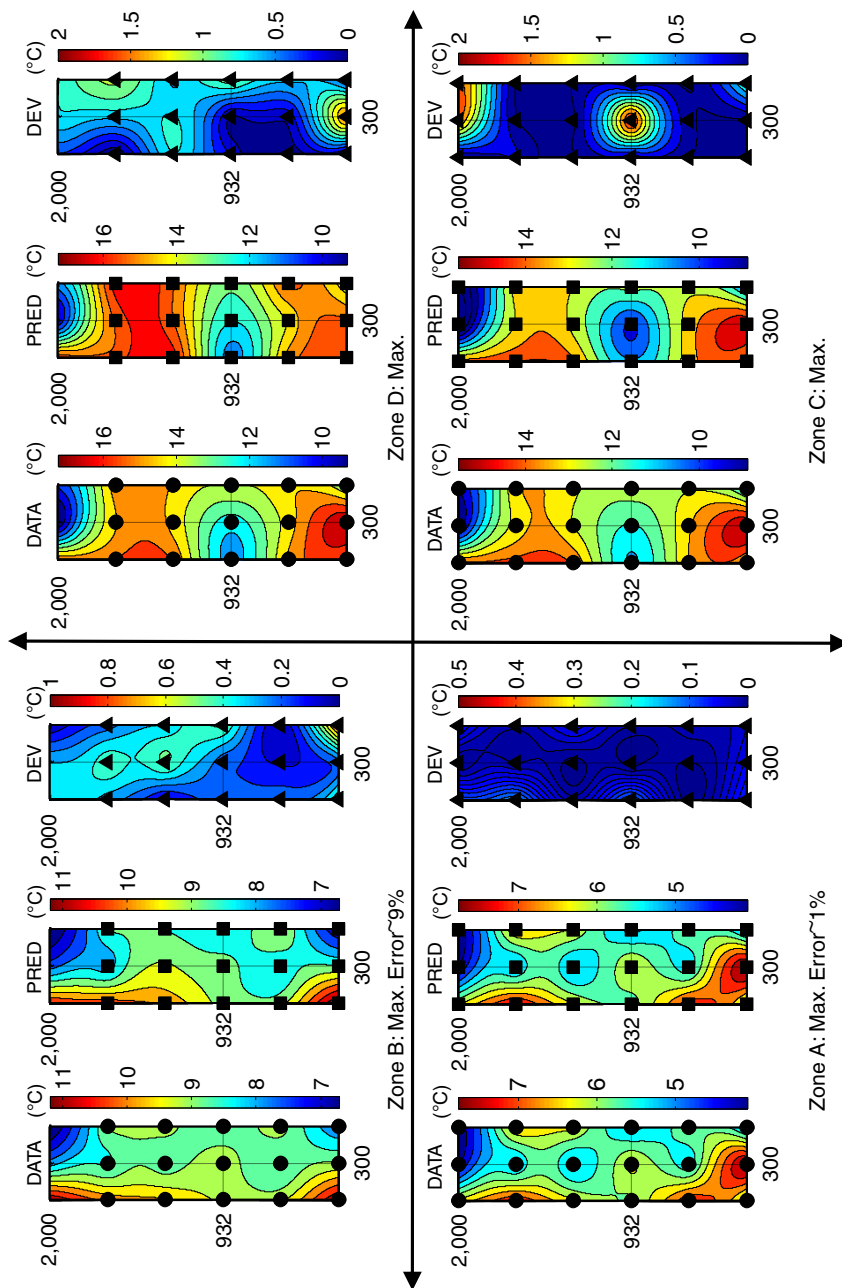
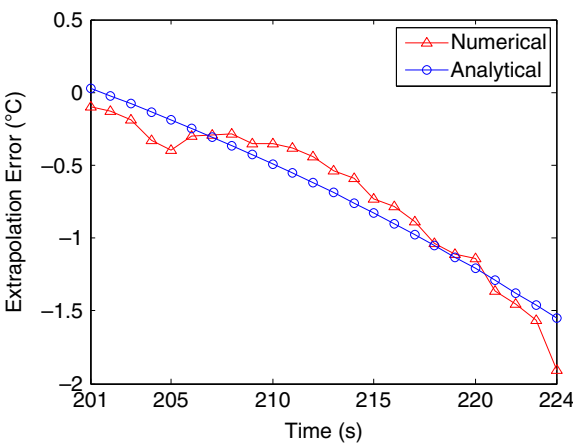


Figure 14.
Fidelity of POD-
based prediction
framework

HFF
25,6

1400

Figure 15.
Specification of a priori error by the conjugate gradient method



Notes: For this particular data set, $c_1=3.1$, $c_2=-0.02$, $c_3=-0.02$

the average air temperature. Similar procedure can be performed for all sensor points. Unless the data center condition is drastically changed, the calibration constants once determined can be used repeatedly over time.

The importance of the error conditioning is evident from Figure 16, comprising the contour plots generated from the air temperature data, predictions, and deviations

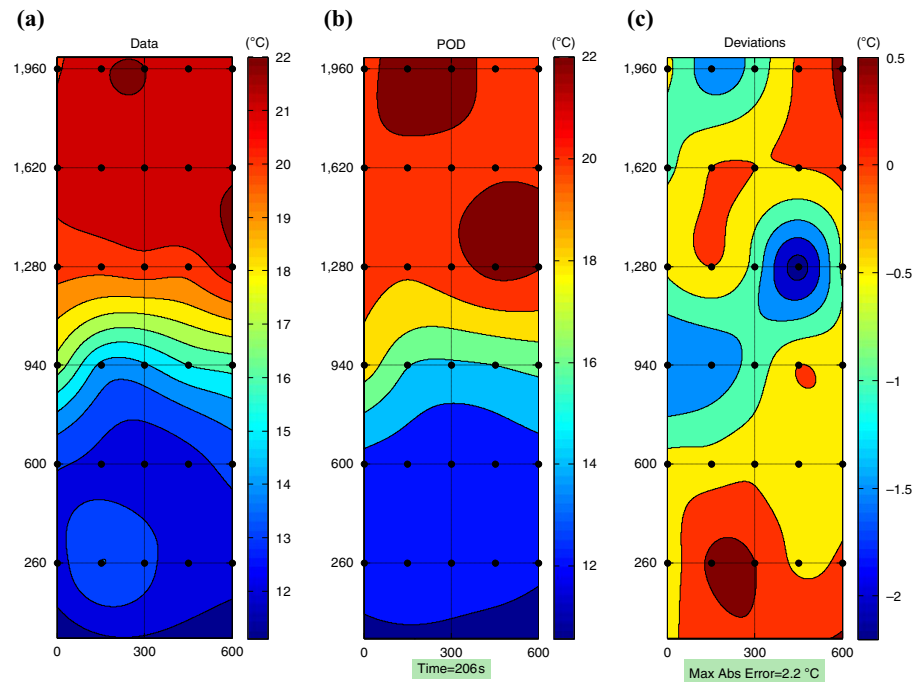


Figure 16.
Fidelity of POD-based temperature prognosis at 206 s at the test rack inlet

at the test server inlet at 206 s. Figure 16(c) shows the deviations are considerably large ranging from -2.0°C to 0.5°C , i.e. more than 10 percent of the maximum inlet air temperature. These are consistent with Ghosh and Joshi (2013).

8. Sensor fusion for reducing temperature monitoring cost

Measurement-based thermal control is of paramount importance to optimize energy consumption in data centers. Therefore, cost-effective data center management demands optimizing sensor number for monitoring air temperatures at critical locations. The optimization problem pertains to minimizing sensor numbers, while the error remains below a certain pre-assigned limit.

$$\min (\text{sensor number})$$

such that, deviation < error limit.

The effectiveness of POD for reducing the number of sensors is assessed using the temperature data collected at the test rack exhaust located near the top of the test rack as shown in Figure 17. The input air temperature data are collected by 13 sensors at the locations marked by black filled circles. These are used to predict air temperatures at the locations indicated by open circles. The prediction accuracy of the POD-based framework shows a maximum relative deviation of 2.2 percent. The high predictive accuracy at new locations indicates the usefulness of POD in reducing the number of monitoring sensors. In this particular case, the required sensor number has been reduced to 13 from 21, a 38 percent savings in number of sensors. The detailed case study can be obtained in Ghosh and Joshi (2014a).

9. Conclusion

Three examples are presented to illustrate the emerging thermal modeling needs to enable co-design of multi-scale IT systems, with the objective of energy efficient operation, while meeting the performance and reliability targets. The computationally

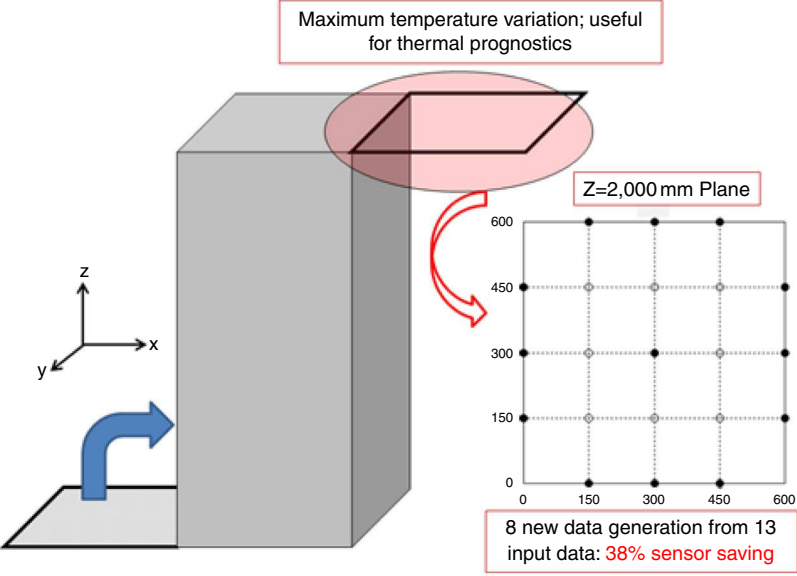


Figure 17.
Test plane for the
sensor reduction

intensive nature of full scale simulations, makes these infeasible over the several decades of length and time scales encountered. Instead, reduced order thermal modeling techniques are employed, in order to couple the thermal analyses with the overall co-design framework. Two such techniques illustrated are based on control volume energy balances, and POD. For the control volume based approach, combined conduction and convection modeling of enhanced microfluidic cooling surfaces in 3D chip stacking allows for a three-orders of magnitude speedup in computation, compared to a fully resolved finite volume simulation, while keeping prediction errors with 5 percent, thus making electrical/thermal co-design possible. For the POD technique, both computational observations, as well as a data-driven modeling framework are presented. The POD technique is found to provide a two order of magnitude faster computation, compared to a full-field simulation, with prediction errors within 10 percent. This enables their effective use in multi-scale simulations using a Progressive Zoom-in approach. For the data-driven framework, a nearly 40 percent reduction in sensors needed for creating air temperature maps for sensing based control of cooling hardware is achieved for server racks in data centers. These techniques are potentially applicable to a broad range of multi-scale microsystems.

References

- Bakir, M.S. and Meindl, J.D. (Eds) (2008), *Integrated Interconnect Technologies for 3D Nanoelectronic Systems*, Artech House, Norwood.
- Bamal, M., Stucchi, M., Verhulst, A.S., Van Hove, M., Cartuyvels, R., Beyer, G. and Maex, K. (2006), "Performance comparison of interconnect technology and architecture options for deep submicron technology nodes", *2006 International Interconnect Technology Conference, San Francisco, CA*, pp. 202-204.
- Barabadi, B., Joshi, Y. and Kumar, S. (2011), "Prediction of transient thermal behavior of planar interconnect architecture using proper orthogonal decomposition method", *ASME 2011 Pacific Rim Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Systems*, American Society of Mechanical Engineers, pp. 213-224.
- Barabadi, B., Kumar, S., Sukharev, V. and Joshi, Y.K. (2015), "Multiscale transient thermal analysis of microelectronics", *Journal of Electronic Packaging*, Vol. 137 No. 3, pp. 031002-031002-8.
- Barabadi, B., Kumar, S., Sukharev, V. and Joshi, Y. (2012), "Multi-scale transient thermal analysis of microelectronics", *ASME 2012 International Mechanical Engineering Congress and Exposition*, American Society of Mechanical Engineers, pp. 767-774.
- Bizon, K., Continillo, G., Russo, L. and Smula, J. (2008), "On POD reduced models of tubular reactor with periodic regimes", *Computers & Chemical Engineering*, Vol. 32 No. 6, pp. 1305-1315.
- Fuchs, K. (1938), "The conductivity of thin metallic films according to the electron theory of metals", *Mathematical Proceedings of the Cambridge Philosophical Society*, Vol. 34 No. 1, pp. 100-108.
- Ghosh, R. and Joshi, Y. (2013), "Error estimation in POD-based dynamic reduced-order thermal modeling of data centers", *International Journal of Heat and Mass Transfer*, Vol. 57 No. 2, pp. 698-707.
- Ghosh, R. and Joshi, Y. (2014a), "Proper orthogonal decomposition-based modeling framework for improving spatial resolution of measured temperature data", *IEEE Transactions on Components, Packaging and Manufacturing Technology*, Vol. 4 No. 5, pp. 848-858.
- Ghosh, R. and Joshi, Y. (2014b), "Rapid temperature predictions in data centers using multi-parameter proper orthogonal decomposition", *Numerical Heat Transfer, Part A: Applications*, Vol. 66 No. 1, pp. 41-63.

- Gurram, S.P., Joshi, Y., King, W.P. and Ramakrishna, K. (2004), "Numerical simulation of electron transport through constriction in a metallic thin film", *Electron Device Letters, IEEE*, Vol. 25 No. 10, pp. 696-698.
- Holmes, P., Lumley, J.L. and Berkooz, G. (1998), *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*, Cambridge University Press, Cambridge, CA.
- Khan, N.H., Alam, S.M. and Hassoun, S. (2009), "System-level comparison of power delivery design for 2D and 3D ICS", *Proceedings 3DIC 2009, San Francisco, CA*, pp. 1-7.
- Lumley, J.L. (1967), "The structure of inhomogeneous turbulent flows", *Atmospheric Turbulence and Radio Wave Propagation*, pp. 166-178.
- Meindl, J.D. (2003), "Beyond Moore's law: the interconnect era", *Computing in Science & Engineering*, Vol. 5 No. 1, pp. 20-24.
- Mizunuma, H., Yang, C.L. and Lu, Y.C. (2009), "Thermal management of 3D-ICs with integrated microchannel cooling", *Proceedings of ICCAD*, Vol. 2009, pp. 256-263.
- Namba, Y. (1970), "Resistivity and temperature coefficient of thin metal films with rough surface", *Japanese Journal of Applied Physics*, Vol. 9 No. 11, pp. 1326-1329.
- Roy, K., Mukhopadhyay, S., Mahmoodi-Meimand, H. (2003), "Leakage current mechanisms and leakage reduction techniques in deep-submicrometer CMOS circuits", *Proceedings of IEEE*, Vol. 91, pp. 305-327.
- Soffer, S.B. (1967), "Statistical model for the size effect in electrical conduction", *Journal of Applied Physics*, Vol. 38 No. 4, pp. 1710-1715.
- Sondheimer, E. (1952), "The mean free path of electrons in metals", *Advances in Physics*, Vol. 1 No. 1, pp. 1-42.
- Sridhar, A., Vincenzi, A., Ruggiero, M., Brunswiler, T. and Atienza, D. (2010a), "3D-ICE: fast compact transient thermal modeling for 3D ICs with inter-tier liquid cooling", *Proceedings of 2010 IEEE/ACM International Conf. Computer Aided Design*, pp. 463-470.
- Sridhar, A., Vincenzi, A., Ruggiero, M., Brunswiler, T. and Atienza, D. (2010b), "Compact transient thermal model for 3D ICs with liquid cooling via enhanced heat transfer cavity", *Proceedings of 2010 16th International Workshop on Thermal Investigations of ICs and Systems*, pp. 1-6.
- Tang, L. and Joshi, Y. (2005), "A multi-grid based multi-scale thermal analysis approach for combined mixed convection, conduction, and radiation due to discrete heating", *Journal of Heat Transfer*, Vol. 127 No. 1, pp. 18-26.
- Wan, Z., Xiao, H., Joshi, Y. and Yalamanchili, S. (2013), "Electrical/thermal co-design of multi-core architectures and microfluidic cooling for 3D stacked ICs", *Proceedings of 2013 19th International Workshop on Thermal Investigations of ICs and Systems*, pp. 237-242.
- Ziman, J.M. and Levy, P.W. (1961), "Electrons and phonons", *Physics Today*, Vol. 14 No. 11, pp. 64-66.

Corresponding author

Professor Yogendra Joshi can be contacted at: yogendra.joshi@me.gatech.edu

For instructions on how to order reprints of this article, please visit our website:

www.emeraldgroupublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com