Open access • Proceedings Article • DOI:10.1109/3DIC.2009.5306530

Validation of the porous-medium approach to model interlayer-cooled 3D-chip stacks

Thomas Brunschwiler, Stephan Paredes, Ute Drechsler, Bruno Michel ...+4 more authors

Institutions: IBM, École Polytechnique Fédérale de Lausanne

Published on: 30 Oct 2009

Topics: Heat flux, Heat transfer, Thermal resistance, Stack (abstract data type) and Coolant

VALIDATION OF THE POROUS-MEDIUM APPROACH TO MODEL INTERLAYER-COOLED 3D-CHIP STACKS

T. Brunschwiler, S. Paredes, U. Drechsler, and B. Michel,
IBM Research GmbH, Zurich Research Laboratory, 8803 Rüschlikon, Switzerland,
tbr@zurich.ibm.com, +41 44 724 86 81,

W. Cesar, G. Töral, Y. Temiz, and Y. Leblebici,
Swiss Federal Institute of Technology in Lausanne (EPFL), Microelectronic Systems Laboratory,
1015 Lausanne, Switzerland.

ABSTRACT

Interlayer cooling is the only heat removal concept which scales with the number of active tiers in a vertically integrated chip stack. In this work, we numerically and experimentally characterize the performance of a three tier chip stack with a footprint of 1cm^2. The implementation of 100µm pitch area array interconnect compatible heat transfer structures results in a maximal junction temperature increase of 54.7K at 1bar pressure drop with water as coolant for 250W/cm^2 hot-spot and 50W/cm^2 background heat flux. The total power removed was 390W which corresponds to a 3.9kW/cm^3 volumetric heat flow.

An efficient multi-scale modeling approach is proposed to predict the temperature response in the complete chip stack. The experimental validation confirmed an accuracy of +/- 10%. Detailed sub-domain modeling with parameter extraction is the base for the system level porous-media calculations with thermal field-coupling between solid – fluid and solid – solid interfaces.

Furthermore, the strength and weakness of microchannel and pin fin heat transfer geometries in 2-port and 4-port fluid architectures is identified. Microchannels efficiently mitigate hot spots by distributing the dissipated heat to multiple cavities due to their low porosity. Pin fins with improved permeability and convective heat dissipation are advantageous at small power map contrast and aligned hot spots on the different tiers.

Large stacks of 4cm^2 can be cooled sufficiently by the 4-port fluid delivery architecture. The flow rate is improved four times compared to the 2-port fluid manifold. The non-uniformity of the flow in case of the 4-port demands a more careful floor-planning with hot spots placed in the chip stack corners. This is especially true in case of communicating heat transfer geometries such as pin fin structures with zero fluid velocity in the stack center. This large velocity contrast can be reduced by the implementation of non-communicating microchannels.

KEYWORDS: Interlayer cooling, microchannel, pin fin, cross-flow, multi-scale modeling, porous-media, field-coupling, forced convective single-phase heat transfer, vertically integrated packages, 3D chip stacks.

1. INTRODUCTION

Thermal management in high-performance chip packages is one of the major challenges in vertical integration according to the ITRS roadmap [1]. First products will adopt traditional back-side heat removal. This scheme scales with the die size, but not with the number of stacked tiers. In multi-tier packages, both heat flux and thermal resistance from junction to the coolant accumulate. This constrains the electrical design to a single logic layer with subsequent memory dies or two logic tiers with non-aligned hot-spots [2] demanding different floor-plans for each logic layer, reducing the economy of scale of these products.

To exploit the full potential of 3D integration, scalable heat-removal concepts are necessary. In forced convective interlayer cooling, the coolant is pumped between the active layers and removes the heat right at the source. This concept scales with the number of tiers in the stack (Figure 1). Former studies defined heat transfer coefficients and friction factors in single-fluid-cavity experiments for various heat transfer structures at single and double side uniform power dissipation, respectively [3], [4]. Pin fin in-line geometries perform best at pressure drop boundary conditions: Despite the very low volumetric flow rate of 300 mL/min for a 1cm^2 cavity uniform heat fluxes up to 180 W/cm^2 at 100 µm interconnect pitch can be removed.

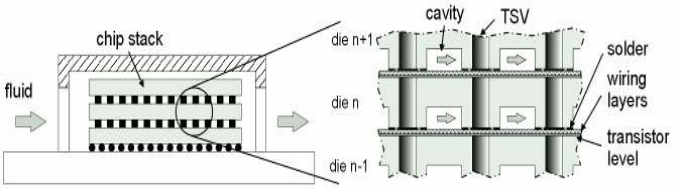

Figure 1: Cross-section trough a chip stack with through-silicon vias (*TSVs*) embedded in a fluid containment. The detailed view demonstrates the implementation of the fluid cavity utilizing a solder sealing concept to prevent fluid contact with the electrical interconnects.

As pointed out: the main limiting factor in interlayer cooling is the low coolant flow rate due to the small hydraulic diameters constrained by the interconnect pitch and the through-silicon via (*TSV*) aspect ratio. Compared to back-side cold plates [5] the flow rate is 10 fold reduced. Therefore, the fluid temperature increase from inlet to outlet dominates the thermal budget.

To enhance the cavity volumetric flow rate a 4-port fluid delivery architecture utilizing the complete periphery of the chip stack was proposed [6]. Compared to the 2-port configuration the fluid velocity in the corners is drastically enhanced due to the short fluid path and results in efficient hot-spot heat removal in the corners (Figure 2).

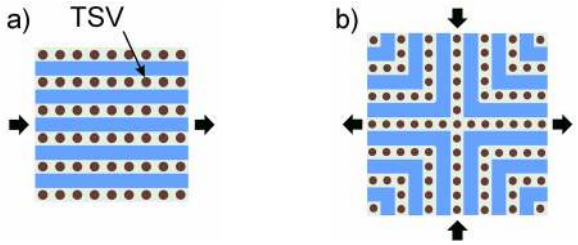

Figure 2: Top view of a) 2-port and b) 4-port microchannel fluid delivery architecture compatible with area-array interconnects.

To demonstrate interlayer cooling performance on a complete chip stack Takahashi and Chen performed a conjugate heat and mass transfer model considering peripheral interconnects only [7],[8]. To efficiently predict the junction temperature in a single cavity test section considering symmetric but non-uniform power dissipation from two sides the porous-media approach was proposed [6]. The fluid flow in the cavity is modeled as a two-dimensional problem considering velocity and direction dependent permeability. This effective media model reduces the number of nodes in the computational domain by several orders of magnitude, since detailed geometries and fluid boundary layers do not have to be resolved. The heat conduction in the solid is modeled in three dimensions capturing also heat spreading effects. The solid – fluid temperature field-coupling was accomplished by a predefined heat transfer coefficient. The study does not consider interconnect mediated tier-to-tier heat flow and is therefore only valid for single cavities with symmetric power dissipation from both cavity sides. Up to now the experimental validation of these concepts on an interlayer cooled, multi-tier, and multi-cavity chip stack with an area array interconnect compatible heat transfer geometry has not been performed.

The goal of this study is to combine the proposed heat transfer building blocks in a chip stack thermal demonstrator to discuss their performance considering uniform and hot-spot dominant power maps. Furthermore, the existing porous-media concept will be extended for the use in multi-cavity devices at non-symmetric power dissipation including heat flow through interconnects. Finally, the model accuracy is validated by experimental temperature readings. As a projection, we demonstrate interlayer cooling performance for a realistic chip stack with thinned dies and current wiring layers.

2. MULTY-CAVITY STACK DESIGN

All thermal demonstrator chip stacks include three power dissipating tiers and four heat removing fluid cavities (Figure 6). To reduce the process complexity, a pyramid chip stack configuration was realized with lateral electrical I/Os utilizing wire-bonds instead of TSVs (Figure 3a). This also allows the integration of the fluid in- and outlets into the chip stack. The fluid cavity spans a quadratic area of 1cm^2 and is populated either with microchannel (CH) or pin fin in-line (PF) heat transfer geometries which are area array TSV compatible (Figure 3b). The nominal channel and pin dimensions are listed in Table 1. Fluid in- and outlets with an aperture of 1.5mm are arranged in 2 and 4-port configuration. The later represents a single quadrant of a 4cm^2 chip stack. Due to symmetries this is sufficient to predict the total heat transfer performance (Figure 4) (compare with Figure 2b).

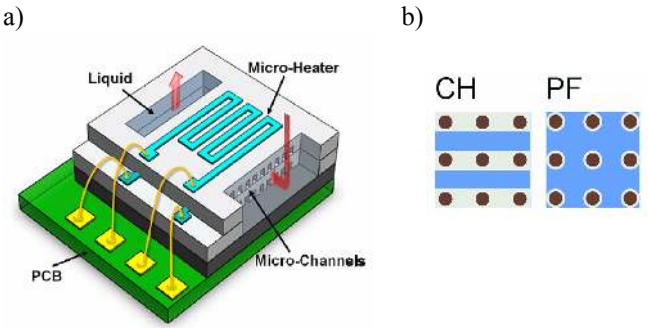

Figure 3: a) Sketch of the interlayer-cooled thermal demonstrator with the pyramid chip stack design and resulting lateral I/O. b) Top view to cavity showing area array interconnect compatible heat transfer structures, namely microchannels (CH) and pin fins (PF) with TSVs (brown).

Table 1: Test vehicle specification

Test vehicle	In-/outlet	Heat transfer geometry
2-port CH	2-port	Channel
2-port PF	2-port	pin fin in-line
4-port CH	4-port	Channel
4-port PF	4-port	pin fin in-line

Parameters	Values
Heat transfer area	10x10 mm^2
Heat transfer geometry:	
- channel / pin pitch	100 µm
- cavity height	100 µm
- channel wall width / pin diameter	50 µm
Hot-spot area per cavity	4 times 10 mm^2

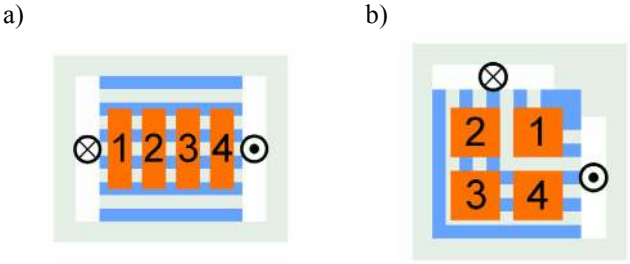

Figure 4: Top view of a) 2-port and b) 4-port microchannel configuration. The 4-port test vehicle represents only one quadrant of a 4cm2 cooled chip stack. Evenly distributed hot-spot areas (orange) are marked with identification numbers.

Power can be dissipated independently in four hot spot heaters per tier on an area of 10mm^2 each (2x5mm^2 2-port / 3.33x3.33mm^2 4-port). This results in a <40% heat transfer area coverage. The heaters are distributed equidistant with a spacing of 0.42mm for the 2-port and 0.92mm for the 4-port respectively. A meander design is used to meet resistance specification of 30Ω. The heater wire is divided into five parallel strips to reduce current crowding in the meander bends resulting in a high heat flux uniformity. The hot spot temperature (T_{HS}) is recorded with a four-point resistive measurement of the resistive temperature probe (*RTD*) located along the heater symmetry line (Figure 5).

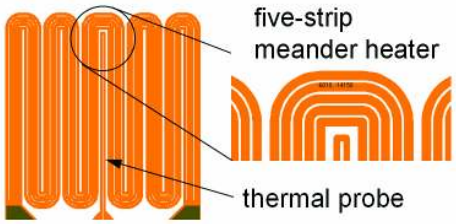

Figure 5: Five-strip hot spot heater design with integrated resistive thermal probe. Resistor metallization (orange), electrical leads (brown).

3. PYRAMID CHIP STACK

The test vehicle fabrication sequence started with wafer level metal deposition onto 525µm thick silicon substrates covered with 200nm SiO_2 wet-oxide dielectrics. Aluminum (*Al*) strips with a thickness of 250nm for the heaters and sensors followed by an additional 400nm of *Al* acting as electrical leads and wire-bond pads are sputter deposited and patterned with lift-off technique. Atomic layer deposition (*ALD*) was used to cover the metal layers with a pinhole-free, 200nm thick Alumina (Al_2O_3) layer, to prevent hydrolysis in the water. The dielectric on bond pads was removed by buffered hydrofluoric acid. A 4µm thick polyimide layer (HD3003, DuPont) was then spin-coated and structured in a oxygen plasma reactor with a positive photoresist mask. Cavities and ports were fabricated into the silicon die by double-side deep reactive ion etching. After a first electrical inspection the known-good-dies were singulated by wafer dicing. The alignment of the five silicon dies representing the chip stack was done with a brass stencil. This complete assembly was placed into a membrane oven. The polyimide bond was performed at 350°C and in a 1mbar vacuum under an applied load of 7bar on the stack top surface through the oven membrane (Figure 6 and 7a and b). Alignment accuracy was better than 10µm, which is sufficient for the demonstrator. A leak test with water at 2bar over pressure proofed bond line quality.

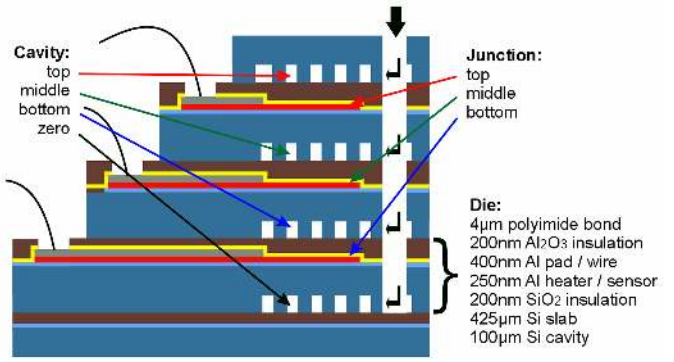

Figure 6: Thermal demonstrator cross-section showing layer and stacking sequence. The nomenclature of the cavities and junctions is given in words and colors used in the subsequent graphs.

The stack was then glued to the printed circuit board using a mechanical compliant silicon adhesive (Sylgard 577, Dow Corning) to minimize thermo-mechanical stress. Wedge – wedge wire bonding with 25µm thick *Al*-wires was performed to support a maximal current load of 0.1A (Figure 7c). To protect the wires a UV curable epoxy (Norland 65, Optical Adhesives) was used as globe top. Finally a PMMA manifold with fluid connections was attached to the stack with a underfill epoxy (EpoTek 302-3M, Epoxy Technologies) at a defined gap of 50µm forming the capillary (Figure 7d).

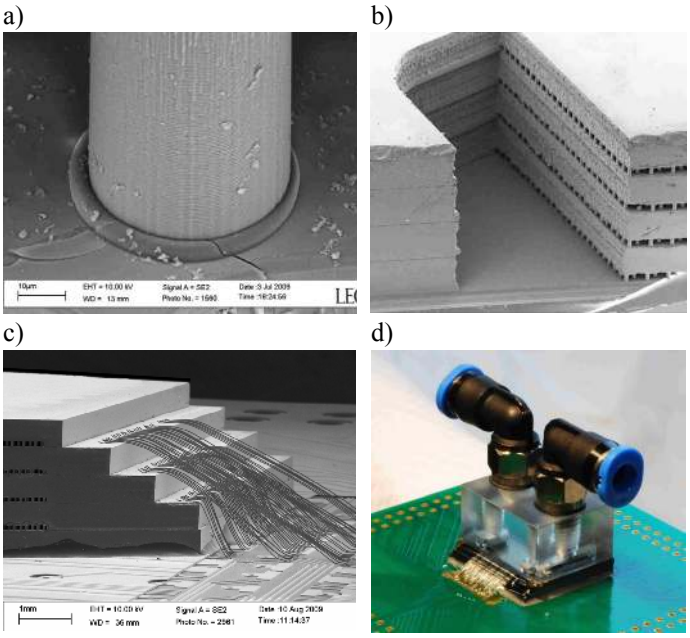

Figure 7: Scanning electron microscope and photographic close-ups of the pyramid chip stack: a) pin fin bond showing polyimide meniscus formation, b) cross-section through fluid port and cavities, c) view at stack to board wire-bond I/Os, d) complete test vehicle mounted on the printed circuit board with fluid manifold and connection.

Compared to realistic chip stacks the test vehicle silicon slab thickness is 425µm instead of 50µm to reduce wafer handling complexity. This will enhance the heat spreading capability in each layer. A realistic slab-thickness results from a maximal TSV height of typically 150µm minus the cavity depth. Furthermore, we used a compliant polyimide layer for leak-tight bonding. This layer represents a thermal impedance of 20 K*mm^2/W and emulates the wiring levels of a real processor die with a typical thermal resistance of 7 K*mm^2/W. The offset of 13 K*mm^2/W needs to be considered in further discussion of the thermal performance. The bars in Figure 8 demonstrate the significance of these process-induced adjustments.

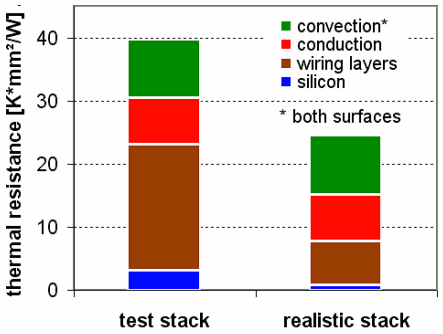

Figure 8: Thermal resistance value comparison of layers present in the thermal test stack compared to a realistic product chip stack. Most significant deviation can be noticed in the wiring layers.

The thermo-fluidic characterization of the test vehicles was performed on a single-phase fluid-loop with water as coolant, temperature controlled through a secondary chiller loop (ProLine RP855, Lauda). The primary loop is equipped with a magnetically coupled gear pump (Fluidotech), a 10µm particle filter, a Coriolis-flow mass flow meter (MFS 3000-S03) with an accuracy of 0.3 %, a differential pressure sensor (PD23-V-2, Omega, accuracy 0.1 %), and T-type thermocouples measuring the in- and outlet fluid temperature. The hot spots are powered by multi-purpose DC power supplies. The dissipated power and the hot spot temperature is measured with a Keithley 2701 multimeter and a Keithley 7700 multiplexer card. The data acquisition was performed through a LabView platform.

4. POROUS-MEDIA APPROACH

It is possible to predict the junction temperature in the chip stack with conjugate heat and mass transfer modeling. The fluid thermal and hydrodynamic boundary layers are most important in case of convective heat transfer. Typically, 20'000 nodes with five degrees of freedom are needed to resolve the boundary layers in one pin fin unit cell. Multiplied with the number of pins per layer (in this case 10'000) and the amount of cavities the model complexity of the fluid only is 0.8 billion nodes. This detailed approach is computationally very demanding and can only be solved on a high performance cluster system with a slow response time.

Multi-scale modeling helps to reduce the complexity by orders of magnitudes. The cavity can be represented as a two dimensional porous-media if the length-scale of interest is multiples of the heat transfer unit-cell dimension. An effective permeability (κ) accounts for the viscous-dissipation in the specific cavity. In computational fluid dynamics (*CFD*) the permeability can be implemented as a negative momentum source term added to the Navier-Stokes equation. From this the pressure and velocity field can be derived. Additionally, the heat flux from solid – fluid is defined by temperature field-coupling considering a velocity dependent thermal resistance on each cavity side. To account for solid – solid heat conduction through the pin or channel walls a fill factor depend conductive thermal resistance is applied between the adjacent tiers. With this approach it is possible to solve the velocity and temperature field of the complete chip stack in case of periodically arranged heat transfer unit-cells in individual domains with a single desktop computer within minutes, including also heat spreading in the solid (Figure 9).

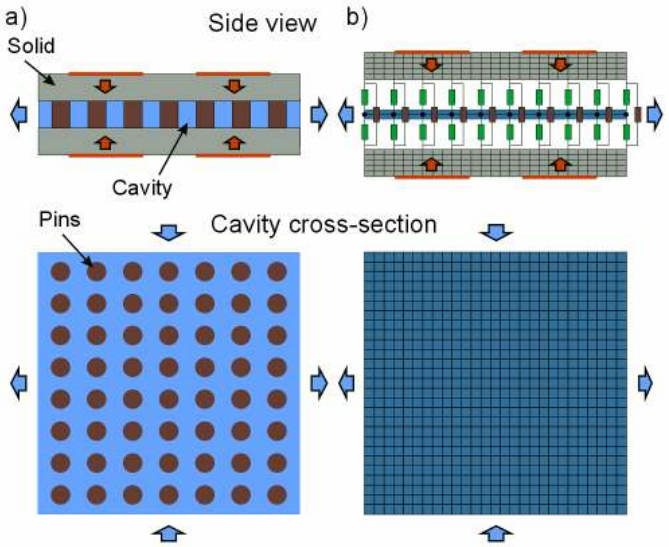

Figure 9: a) Detailed conjugate heat and mass transfer problem with large complexity due to thousands of pins per cavity. b) Complexity reduction by porous-media approximation of fluid cavity utilizing field-coupling to transfer heat from solid – fluid (green resistors) and solid – solid (brown resistors).

To derive the effective model parameters, detailed heat and mass transfer modeling is performed in the sub-domain representing a single heat transfer unit cell at imposed periodic boundary conditions valid for a pin array with equidistant spacing [9]. From this analysis the permeability and the convective thermal resistance is extracted. (Figure 10).

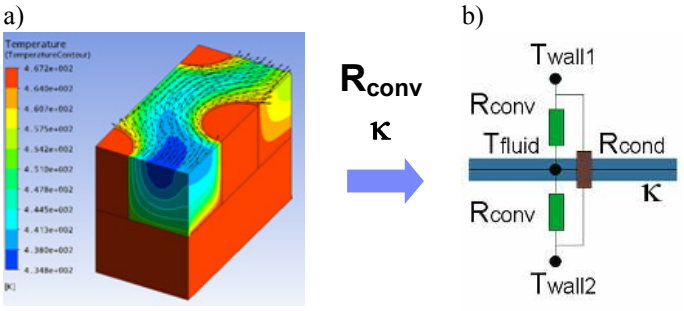

Figure 10: a) Isometric view of the temperature field with normalized velocity vectors of a pin fin staggered sub-domain. The model resolves the boundary layers accurately and therefore consists of 26'000 nodes. The convective thermal resistance (R_{conv}) and the cavity permeability (κ) is extracted from this results. b) Resistor network representing the thermal field-coupling of the 2D-porous media (T_{fluid}) with the adjacent 3D-solid walls (T_{wall}).

The permeability is defined through Darcy's law as

$$\nabla p = -\frac{\mu}{\kappa}\vec{v}_{Darcy} \qquad (1)$$

with the linear dependence of pressure gradient (∇p) on the superficial velocity, also called the Darcy velocity (v_{Darcy}), and the dynamic viscosity (μ) as material coefficient. The Darcy velocity is the average fluid velocity (v_{bulk}) in the cavity multiplied with the cavity porosity (ε):

$$\vec{v}_{Darcy} = \varepsilon \cdot \vec{v}_{bulk}\,. \qquad (2)$$

The cavity porosity is the ratio of the cavity fluid volume (V_{fluid}) to the total cavity volume including the fluid and solid part (V_{tot})

$$\varepsilon = \frac{V_{fluid}}{V_{tot}}\,. \qquad (3)$$

The projected convective thermal resistance (R_{conv}) mapping the heat transfer on a single cavity side is computed by

$$R_{conv} = \frac{\overline{T}_{wall} - \overline{T}_{fluid}}{\dot{q}_1} \qquad (4)$$

with average wall ($\overline{T}_{wall}$) respectively fluid ($\overline{T}_{fluid}$) temperature and the heat flux ($\dot{q}_1$) dissipated on one cavity wall in case of a symmetric heat flux boundary conditions.
The solid – solid (tier to tier) conductive thermal resistance is defined as

$$R_{cond} = \frac{t_{cavity}}{k_{solid} \cdot (1-\varepsilon)} \qquad (5)$$

with cavity thickness (t_{cavity}), pin or channel wall thermal conductivity (k_{solid}) and the porosity (ε).

The permeability and convective thermal resistance of a microchannel at fully developed boundary layers is independent of the Reynolds number and fluid velocity, respectively. In case of pin fins these parameters are in general velocity and direction dependent. Therefore, the parameters were extracted at different Darcy velocities for pin fin in-line and staggered orientation. The regression for each orientation is defined by parameter fitting and is considered to be the upper and lower bound. Values for other orientations are interpolated assuming a sinusoidal behavior (Table 2). The effective permeability is velocity dependent and reduced in case of staggered pin fin compared to in-line orientation. The convective thermal resistance depends in both cases on the velocity (Figure 11, 12).

Table 2: Effective model parameters for interconnect pitch of 100µm.

Microchannel (CH): test vehicle dimensions - 101µm height, 46µm wall width
$\varepsilon = 0.540$
$\kappa = 8.76E\text{-}11\ m^2$
$R_{cond} = 7.3\ K*mm^2/W$
$R_{conv} = 18.76\ K*mm^2/W$
Pin fin (PF): regressions valid for v_{darcy} 0 to 7m/s test vehicle dimensions - 103µm height, 45µm pin diameter
$\varepsilon = 0.841$
$(dp/dx)_{stag} = -9.591E6\ Pa*s^2/m^3*v_{darcy}^2 - 9.363E6\ Pa*s/m^2*v_{darcy}$
$(dp/dx)_{in\text{-}line} = -1.100E7\ Pa*s/m^2*v_{darcy}$
$(dp/dx)(\alpha, v_{darcy}) = ((dp/dx)_{in\text{-}line}+(dp/dx)_{stag})/2 - ((dp/dx)_{in\text{-}line}-(dp/dx)_{stag})/2*\cos(4*\alpha)$
$\kappa_{in\text{-}line} = 1.12E\text{-}10\ m^2$ (valid for v_{darcy} 0 to 1.3m/s)
$\kappa = -\mu / (dp/dx)(\alpha, v_{darcy}) * v_{darcy}$
$R_{cond} = 19.7\ K*mm^2/W$
$R_{conv\ stag} = 2.527E7\ K*m^2/W/((v_{darcy}+1.35m/s)/(1m/s))^{1.52}+1.533E6\ K*m^2/W)$
$R_{cond\ in\text{-}line} = 2.527E7\ K*m^2/W/((v_{darcy}+1.35m/s)/(1m/s))^{0.64}+1.533E6\ K*m^2/W)$
$R_{conv}(\alpha, v_{darcy}) = (R_{cond\ in\text{-}line} + R_{conv\ stag})/2 - (R_{cond\ in\text{-}line} - R_{conv\ stag})/2 * \cos(4*\alpha)$

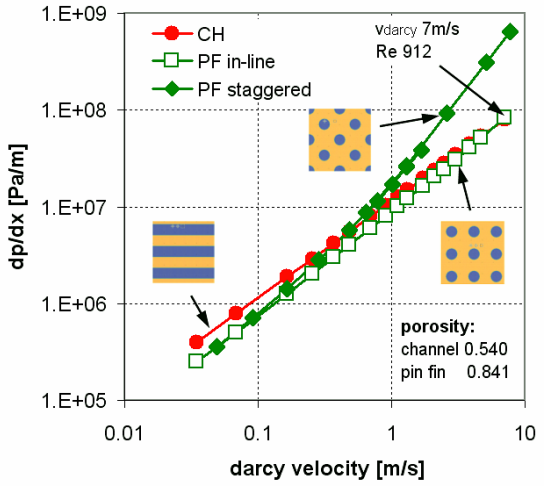

Figure 11: Velocity dependent pressure gradient for microchannel and pin fin heat transfer structures.

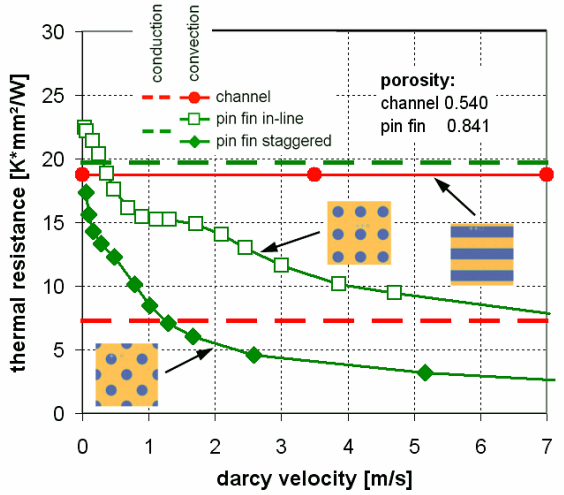

Figure 12: Velocity dependent convective and constant conductive thermal resistance values for microchannel and pin fin heat transfer structures.

It should be noticed, that the pressure gradient at low velocities of the pin fin in-line is lower than the one of the microchannel, but approaches the microchannel permeability asymptotically with increasing velocity. Periodic momentum changes of the fluid in the pin fin staggered unit-cell cause a strongly non-linear pressure gradient velocity dependency. These changes are also responsible for thin, non-developed thermal boundary layers with superior heat removal performance. In general the pin fin structure outperforms the microchannel with respect to reduced pressure drop in in-line orientation and increased heat transfer coefficients. As a result from the high porosity of the pin fin its only disadvantage is the poor solid – solid (tier to tier) coupling (R_{cond}).

On the system-level the pyramid chip stack is modeled with all three tiers represented by the silicon slab, the wiring layers and the power map imposed at the contact surface between these two materials. The four cavities are represented in a quasi two-dimensional domain, with only one node and infinite fluid heat conduction in z-direction. They are thermally field-coupled to the solid as described previously. The modeling concept was implemented on a commercially available computational fluid dynamic platform (CFX V12, ANSYS) (Figure 13).

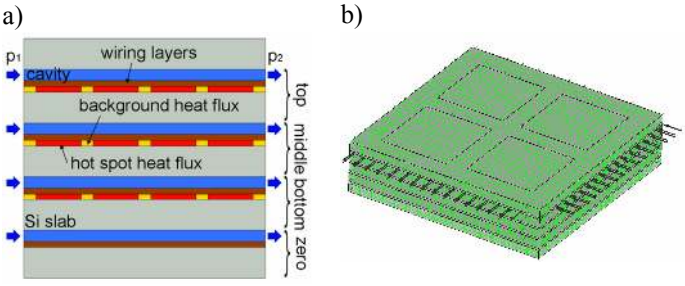

Figure 13: a) Cross-section of a 2-port model presenting all implemented layers. b) 4-port model with indicated boundary conditions created with the ANSYS CFX pre-processor. The green lines represent the mesh.

5. RESULTS AND DISCUSSION

To compare the test vehicle performance a benchmark operating point was defined at a applied pressure drop (Δp) of 1bar reasonable for server applications, fluid inlet temperatures (T_{in}) of 20°C and a hot-spot power (P_{HS}) of 12W being the upper limit of reliable operation. Temperatures for increased power dissipation can be scaled easily due to the linear nature of heat transfer problems in case of constant material properties. This is in first approximation the case for all material properties expect the fluid viscosity.

Mass transfer performance

Pressure drop measurements are presented in Figure 14. At these low Reynolds numbers (<124) the pin fin permeability is highest as predicted from sub-domain modeling. Interestingly the flow rates from the 2- and the 4-port case for a given structures nearly coincide. To compare the port architecture performance the cavity size needs be scaled to 4cm^2. Since the 4-port test vehicle only represents a single quadrant its flow rate needs to be multiplied by four. Doubling the cavity length of the 2-port reduces its flow rate by a factor of two, but doubling the cavity width increases its flow rate by a factor of two. The result is a cavity size independent flow rate at a constant length to width cavity aspect ratio. Finally, the flow rate in the 4-port cavity compared with the 2-port is four times increased at equal chip size.

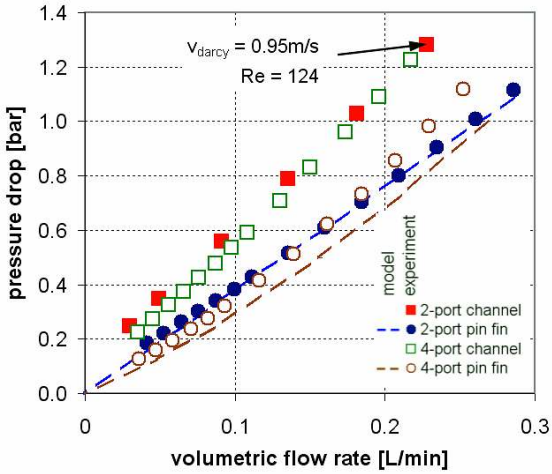

Figure 14: Comparison of pressure drop performance of all test vehicles measured. Dots representing experimental values, dashed lines porous-media model results.

The only non-linear behavior was detected for the pin fin in 4-port mode, were the fluid flow orientation from inlet to outlet is a smooth transition from in-line to staggered to in-line flow (Figure 15a). The staggered flow is responsible for the non-linearity as derived from sub-domain modeling. The numerical results for the 2-port PF test vehicle nicely represent the experiment. The deviation in case of the 4-port PF is -17% compared to the experiment.

The velocity at the inlet and outlet and at the diagonal position of 4-port are plotted in Figure 15b). The velocity is increasing hyperbolically at shorter fluid path from inlet to outlet and reaches 5m/s, but drops to less than 1m/s at the left end of the inlet. In the lower left corner the velocity even drops to zero. This stagnation point would also exist in a full four quadrant 4-port due to symmetry reasons. At this point hot spots are problematic and have to rely on heat spreading. The velocities of the 4-port with microchannels are in general smaller, but do not drop to zero in the central symmetry point due to fluid guiding.

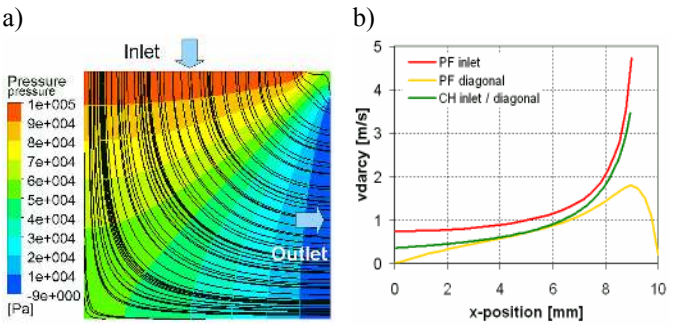

Figure 15: a) Pressure field and stream-lines of the 4-port PF device at 1bar inlet pressure. b) Superficial velocity at the in- / outlet and the diagonal of the 4-port PF and CH device at 1bar inlet pressure.

Heat transfer performance

To demonstrate the temperature response in the 2-port CH device a test case with a random power map was computed. Figure 16 presents the result at benchmark conditions (Δp=1bar, Tin=20°C, P_{HS}=12W) with hot-spots (*HS*) top HS 2,3 / middle HS 1,4 / bottom HS 2,4 active. The non-uniform junction temperature and the heat spreading are visible. The heat pick-up of the fluid can also be noticed. To identify the individual temperature gradients in the chip stack the temperature normal to the cavity plane at the center of hot spot two is plotted in figure 17. As expected, the largest gradients are caused by the poor thermal conductivity of the polyimide layer (0.2 W/(m*K)) and the convective heat transfer from the solid to the fluid.

To validate the temperature field-coupling approach the modeled hot-spot temperature defined as the average temperature along the sensor (T_{jla}) is compared with the measured hot-spot junction temperature (T_{HS}). The model estimates are conservative with a deviation ranging from zero to 21% (Figure 18). The origin of this difference is a superposition of mainly three effects. First: an estimated 3.7% of the total hot-spot power is dissipated in the lead wires. Second: a central gap in the hot-spot heater design of 200µm width serving for the thermal probe placement interrupts the uniform power dissipation. This discontinuity in heat flux locally reduces the junction temperature. Third: the polyimide thickness is considered to be 4µm. This is the case between the bonding areas were heat is dissipated through the polyimide into the fluid. However, the polyimide bond line thickness between heat transfer structure top and silicon slab is 3.2 µm thick. This results in an improved thermal coupling between the slab and the pin or channel wall. Without this parasitic effect in the experiment the estimated deviation would be +/- 10% which seems reasonable for device performance investigations and predictions. Further, the junction temperature (T_j) in the flow direction and in the center of the chip is plotted on figure 18 for each tier. Even with improved heat spreading capability due to the 425µm silicon slab thickness the hot spot contrast is still strong (remember hot spot width of 2mm).

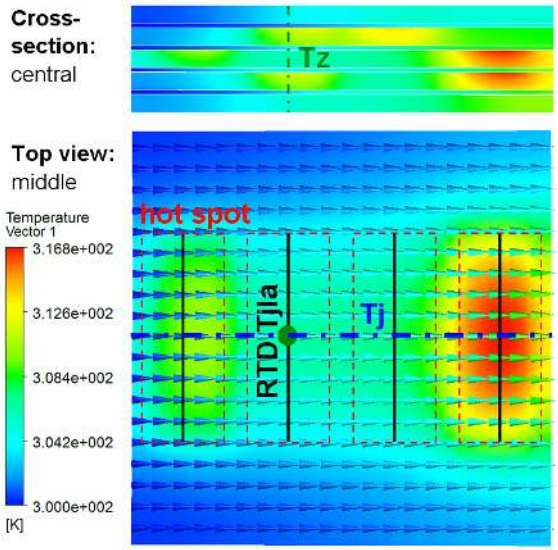

Figure 16: Central cross-section and top view to the middle junction of the random-powered test case showing the temperature map. Arrows represent the fluid flow direction with colors indicating the fluid temperature. Hot-spot areas are marked (red dashed squares) as well as the thermal probe location (black lines). Further model junction temperatures are depicted either from the center line (T_j) (blue, dash-dotted) or as the line average temperature (T_{jla}) representing the measured hot spot temperature (T_{HS}).

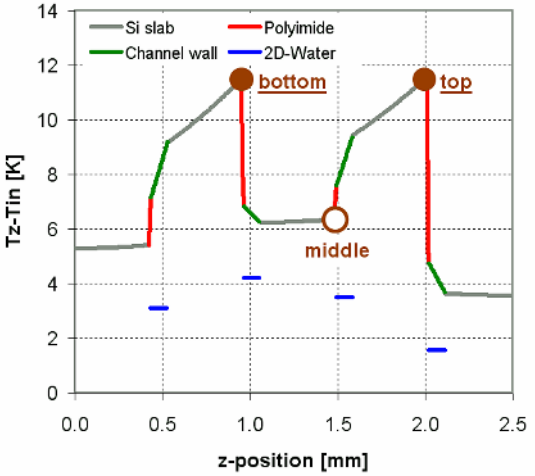

Figure 17: Temperature (T_z) through the stack compared to the fluid inlet temperature (T_{in}) at the center of hot spot 2, normal to the cavity plane. Filled brown dots indicate dissipating, empty dots represent inactive hot spots. The line colors refer to individual materials in the chip stack.

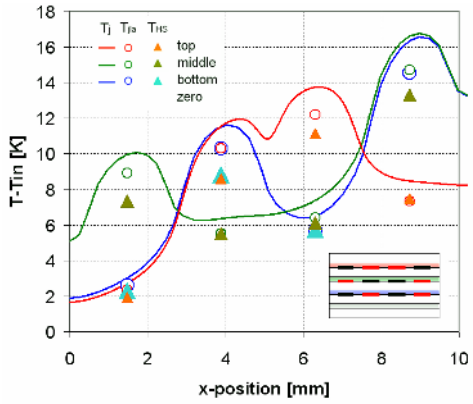

Figure 18: Measured (T_{HS}, triangles) and modeled hot-spot temperature (T_{jla}, circles) is presented together with the central junction temperature (T_j, lines) of the 2-port CH test vehicle at benchmark boundary condition and random hot spot pattern: top HS 2,3 / middle HS 1,4 / bottom HS 2,4.

To analyze the characteristics of the microchannel and the pin fin structures the 2-port test vehicles were operated at different regular hot-spot patterns and varying pressure drops. The maximal hot-spot temperatures are reported in figure 19. Despite its lower flow rate and higher convective thermal resistance the hot-spot temperature of the CH devise is equal compared to the PF in case of a single active hot-spot. The reason for this is its stronger thermal coupling between tiers caused by its low porosity. This results in efficient heat distribution between the four cavities. This experimental finding was also confirmed by the model, with a constant offset of about 20%. If HSs on the top layer are powered heat spreading becomes asymmetric. In this case the spreading benefit of the CH is limited. By activating all three HS2 or even all HSs in the stack, the power dissipation pattern is quasi periodic with minimal heat spreading to cavities of other tiers. In this mode the improved convective heat transfer and increased permeability of the PF results in lower stack temperatures.

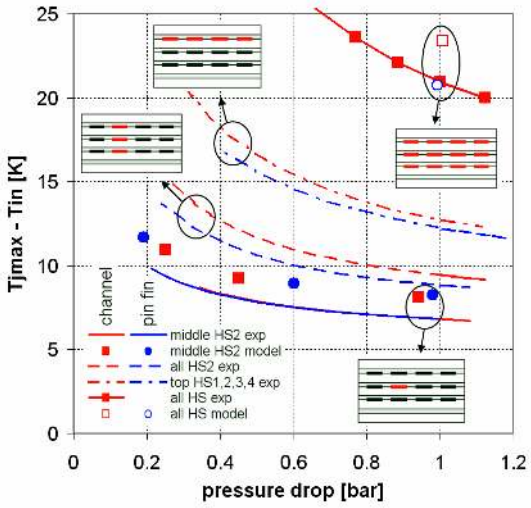

Figure 19: Experimental (T_{HS}) and modeled (T_{jla}) maximal hot-spot temperatures of a 2-port CH and PF device at different power maps versus applied pressure drop. The deviation between model and experiment is ~20%.

The strength of 4-port fluid delivery is well demonstrated at benchmark operation and four active HSs on the top tier (Figure 20). In 4-port flow only hot-spots 2, 3, 4 are thermally coupled through the fluids temperature, but not HS1. Furthermore, the coolant velocity at HS1 is highest (Figure 15b). These are the reasons for the low temperature at HS1 and HS2 in case of the 4-port. The temperature increase from HS2 to HS3 is most dominant due to the dramatic velocity drop towards the lower left corner (stagnation point). It is less pronounced for the CH device since the velocity does not drop to zero. Important to notice is the fact, that the 4-port test demonstrates the cooling performance of a 4cm^2 chip stack compared to the 1cm^2 in case of the 2-port.

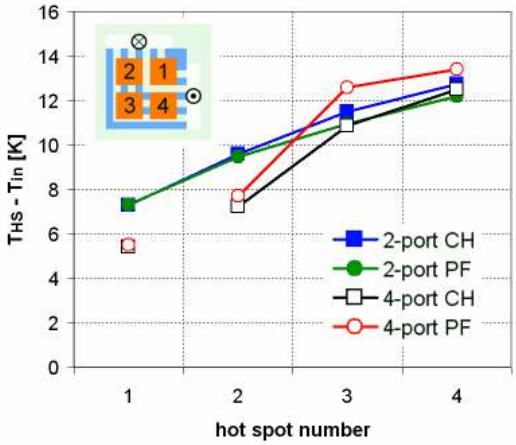

Figure 20: Experimental hot-spot temperature comparison for 2- and 4-port fluid delivery and PF / CH heat transfer structures at benchmark operation and all top HS operational.

Realistic product performance

Finally, we compare the 2-port pin fin test vehicle central junction temperature response at benchmark operation with one of a realistic chip stack product (Si slab thickness of 50µm and wiring thermal resistance of 7 K*mm^2/W) (Figure 21). For this test case the hot-spots are operated at 25W resulting in a heat flux of 250W/cm^2 which is a realistic value for high performance processors. Furthermore, a heat flux of 50W/cm^2 was imposed on the residual chip surface representing the background power dissipation of the cache area. In total 390W are dissipated on a 1cm^2 footprint corresponding to an average volumetric heat flow of 3.9kW/cm^3 if a 1mm stack height is considered. The maximal junction temperature is reached in the middle tier. Due to its central location it has to share the two adjacent cavities with the top and bottom tier. The values are well within typical temperature margins of 60K. Interestingly, the top junction has a lower temperature than the bottom tier. This is astonishing because the bottom junction is more efficiently coupled to its own fluid cavity (zero) than the upper tier which has to dissipate the heat through the lower conduction wiring levels to its top cavity. The reason is the asymmetry in heat flux. The fluid temperature in cavity zero is increasing more rapidly compared to the top cavity, indicating a heat flux crowding from the upper layers in the bottom section.

The test vehicle junction temperature maximum is comparable to the realistic product temperature even with larger silicon thickness and wiring resistance. Enhanced heat spreading in the thicker silicon slab helps to mitigate hot-spot effects and suppresses maximum junction temperatures. This compensates for the increased temperature drop across the low conductive polyimide layer. The hot-spot contrast is much more dominant in the product example. Analyzing the thermal gradient ratio induced by thermal conduction and convection compared to the fluid temperature increase from inlet to outlet indicates the significance of a high flow rate of coolant representing the heat capacity flow through the package. This will be further accentuated at smaller interconnect pitches with cavities of reduced permeability due to reduced hydraulic diameters.

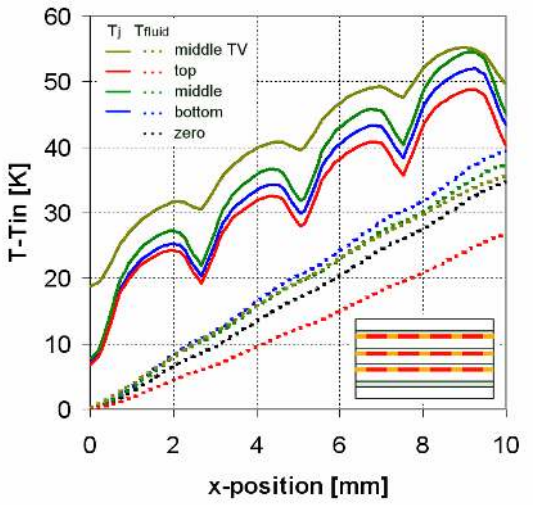

Figure 21: Thermal response in a realistic product chip stack with pin fin height of 100µm, silicon slab thickness of 50µm and wiring layer thermal resistance of 7 K*mm^2/W at benchmark operation. All hot spots dissipate 250W/cm^2, other chip area is powered with 50W/cm^2. Junction temperature of all tiers and fluid temperature in all cavities are presented. The middle junction temperature of the corresponding test vehicle 2-port PF is plotted in comparison.

6. SUMMARY AND CONCLUSION

Interlayer cooling performance was experimentally demonstrated on a pyramid chip stack with three power dissipating tiers and four heat removing cavities with area array compatible interconnect heat transfer structures. The power map was varied by activating individual sets of hot-spots. The readings were compared with the proposed multi-scale modeling approach and deviate less than +/- 10% excluding experimental parasitic effects. Effective parameters such as permeability, convective and conductive thermal resistance of the heat transfer structure are extracted from unit-cell sub-domain modeling with imposed periodic boundary conditions. To compute the chip stack temperature response this values are utilized to represent the cavity as a two-dimensional porous-media using thermal field-coupling to connect the fluid to the solid and adjacent tiers. With this approach temperatures of complex chip stacks can be computed on a single desktop machine.

Finally, we have demonstrated the potential of interlayer cooling in a realistic 1cm^2 chip stack of footprint with 250W/cm^2 hot-spot on 40% and 50W/cm^2 background heat flux on the residual chip surface. With 2-port and pin fin heat transfer structure at a 1 bar pressure drop the maximal junction temperature increase is 54.7K.

4-port fluid delivery is preferred in case of larger 4cm^2 chip stacks and hot-spot locations in the corners. The mass flow rate is four times higher than in the 2-port configuration. This is important since the largest portion of the thermal budget is consumed by the fluid temperature increase.

The performance of the tested heat transfer structures depends on the global cavity geometry and applied power maps in the package (Table 3). Heat transfer geometries with high permeability and low convective thermal resistance such as pin fins are superior in case of periodic power maps from tier to tier and for low hot-spot contrasts. For strongly localized power dissipation microchannels with a low porosity are distributing the heat more efficiently between the cavities by improved tier to tier coupling.

Table 3: Heat transfer structure characteristics

Structure	Strength
Microchannel (CH)	- tier to tier thermal coupling → hot-spot mitigation
Pin fin (PF)	- high permeability, high mass flow rate - moderate fluid temperature increase - small convective thermal resistance → uniform, periodic power maps
2-port	→ uniform heat removal → simple to design and integrate
4-port	- reduced average coolant path resulting in maximal mass flow - high velocities in stack corner → large chip stack footprint → corner hot spot performance
4-port CH	- fluid guiding improves center fluid velocity
4-port PF	- highest overall mass flow

7. NOMENCLATURE

CH	microchannel	
HS	hot-spot	
PF	pin fin	
x-port HSy	test vehicle abbreviation x: number of ports HSy: hot-spot location	
k_{solid}	thermal conductivity of solid	[W/(m*K)]
Δp	pressure drop	[bar]
∇p	pressure gradient	[Pa/m]
dp/dx	pressure gradient in x-direction	[Pa/m]
P_{HS}	hot-spot power	[W]
$\dot{q}$	cavity wall heat flux	[W/m^2]
R_{cond}	conductive thermal resistance	[K*mm^2/W]
R_{conv}	convective thermal resistance	[K*mm^2/W]
t_{cavity}	cavity thickness	[m]
T_{fluid}	fluid temperature	[K]
T_{HS}	experimental hot-spot temp.	[K]
T_{in}	fluid inlet temperature	[K]
T_j	modeled junction temperature along the center line of a tier	[K]
T_{jla}	line averaged junction temp. of modeled hot-spot → equivalent to T_{HS}	[K]
T_{wall}	cavity wall temperature	[K]
T_z	stack temperature	[K]
$\dot{V}$	volumetric flow rate	[L/min]
v_{bulk}	average fluid velocity or bulk velocity	[m/s]
v_{Darcyl}	superficial or Darcy velocity = $v_{bulk} * \varepsilon$	[m/s]
V_{fluid}	fluid volume in the cavity	[m^3]
V_{tot}	solid and fluid volume in the cavity	[m^3]
α	flow direction	[rad]
ε	porosity	[-]
κ	permeability	[m^2]
η	kinematic viscosity	[m^2/s]
μ	dynamic viscosity	[Pa*s]

Acknowledgements

We acknowledge Martin Witzig, Reto Wälchli, Werner Escher, and Richard Stutz from the IBM Research GmbH and Steffen Peters from CADFEM for their technical contribution, and Walter Riess for financial support. This project was scientifically evaluated by SNSF and financed by the Swiss Confederation through the Nano-Tera.ch program.

References

[1] International Technology Roadmap for Semiconductors (ITRS), 2008 Update – Packaging and Assembly http://www.itrs.net/Links/2008ITRS/Update/2008_Update.pdf

[2] T. Brunschwiler, H. Rothuizen, U. Kloter, H. Reichl, B. Wunderle, H. Oppermann, B. Michel, "Forced Convective Interlayer Cooling Potential in Vertically Integrated Packages", Proc. 11th "ITHERM 2008," Orlando, FL, May 2008 (IEEE, 2008), pp. 1114-1125.

[3] Y. Peles, A. Kosar, C. Mishra, C. Kuo, and B. Schneider, "Forced Convective Heat Transfer across a Pin Fin Micro Heat Sink", Int. J. Heat Mass Transfer, 2005, Vol. 48, 3615-3627.

[4] T. Brunschwiler, B. Michel, H. Rothuizen, U. Kloter, B. Wunderle, H. Oppermann, and H. Reichl, "Interlayer Cooling Potential in Vertically Integrated Packages", Microsystem Technol., Vol. 15, No. 1, pp. 57-74 (2009, published online August 21, 2008).

[5] E.G. Colgan et al., "A Practical Implementation of Silicon Microchannel Coolers for High Power Chips", IEEE Trans. Components Packaging Technol., Vol. 30, No. 2., June 2007, 218-225.

[6] T. Brunschwiler, H. Rothuizen, U. Kloter, H. Reichl, B. Wunderle, H. Oppermann, B. Michel, " Hotspot-optimized interlayer cooling in vertically integrated packages ", Proc. MRS fall meeting 2008 p.223-234, Boston, MA, December 2008.

[7] K. Takahashi et al., "Process Integration of 3D Chip Stack with Vertical Interconnection", Proc. 54th IEEE Electronic Components and Technology Conference, 2004, Vol. 1, 601-609.

[8] X. Chen, K. Toh, and J. Chai, "Direct Liquid Cooling of a Stacked Multichip Module", Proc. Electronics Packaging Technology Conference, 2002, 380-384.

[9] R. Waelchli, T. Bruschwiler, B. Michel, and D. Poulikakos, "Combined local microchannel-scale CFD modeling and global chip scale network modeling for electronics cooling design", International Journal of Heat and Mass Transfer, in-press (2009).