

Experimental Characterization of Cold Plates used in Cooling Multi Chip Server Modules (MCM)

Bharath Ramakrishnan¹, Yaser Hadad¹, Sami Alkharabsheh¹, Paul R. Chiarot¹

Kanad Ghose¹, Bahgat Sammakia¹, Vadim Gektin², Wang Chao³

¹State University of New York at Binghamton, Binghamton, NY, 13902, USA

²Futurewei Technologies, Santa Clara, CA, 95050, USA

³Huawei Technologies, Shenzhen, Guangdong, PRC, 518129

Email: bramakr1@binghamton.edu

ABSTRACT

Miniaturization of microelectronic components comes at a price of high heat flux density. By adopting liquid cooling, the rising demand of high heat flux devices can be met while the reliability of the microelectronic devices can also be improved to a greater extent. Liquid cooled cold plates are largely replacing air based heat sinks for electronics in data center applications, thanks to its large heat carrying capacity. A bench level study was carried out to characterize the thermohydraulic performance of two microchannel cold plates which uses warm DI water for cooling Multi Chip Server Modules (MCM). A laboratory built mock package housing mock dies and a heat spreader was employed while assessing the thermal performance of two different cold plate designs at varying coolant flow rate and temperature. The case temperature measured at the heat spreader for varying flow rates and input power were essential in identifying the convective resistance. The flow performance was evaluated by measuring the pressure drop across cold plate module at varying flow rates. Cold plate with the enhanced microchannel design yielded better results compared to a traditional parallel microchannel design. The study conducted at higher coolant temperatures yielded lower pressure drop values with no apparent change in the thermal behavior using different cold plates. The tests conducted after reversing the flow direction in microchannels provide an insight at the effect of neighboring dies on each other and reveal the importance of package specific cold plate designs for top performance. The experimental results were validated using a numerical model which are further optimized for improved geometric designs.

KEY WORDS: Data center thermal management, direct liquid cooling, warm water cooling, microchannels, cold plate, specific thermal resistance, heat transfer coefficient

NOMENCLATURE

TIM	Thermal Interface Material
MCM	Multi-Chip Module
IHS	Integrated Heat Spreader
SKX	Skylake EX Purley Chip
FPGA	Field Programmable Grid Array Chip
lpm	liters per minute
ΔT	Temperature difference (°C)
T_{in}	Inlet temperature (°C)
T_{out}	Outlet temperature (°C)
Q_w	Heat picked by the coolant
$\dot{m}$	Mass flow rate of the coolant (kg/s)
C_p	Specific heat of the coolant (kJ/kgK)
R_{th-DIE}	Die Specific thermal resistance (°C.cm ² /W)

R_{cp}	Cold Plate thermal resistance (°C.cm ² /W)
R_{cond}	Conductive resistance (°C.cm ² /W)
R_{calo}	Caloric resistance (°C.cm ² /W)
R_{conv}	Convective resistance (°C.cm ² /W)
R_{TIM}	TIM resistance (°C.cm ² /W)
T_{case}	Case temperature (°C)

Greek symbols

ρ	mass density (kg/m ³)
μ	dynamic viscosity (kg/ms)

Subscripts

f	liquid state
s	solid state

INTRODUCTION

A wider introduction to the application of current work can be provided by asking a set of questions with answer to first question often marks the genesis to next. The cold plates developed for characterization in this study are typically deployed in servers connected to a rack level cooling system in data centers. *What are data centers?* Data center are a large group of networked servers used by organizations for the remote storage, processing and distribution of large amounts of data. Even the cloud service providers, an off-premise form of computing that stores data on the internet use data center to house cloud services and its resources. Over the last decade, it is astonishing to see how every one of us was/is linked to data center use in one way or the other be it social, professional, health, economic or scientific reasons. The dependence on data centers are more likely to grow in the near future as well for the above-mentioned reasons. *Why data center thermal management is important?* Data centers play an inevitable role in global economy or power usage. In 2014, data center consumed an estimated 70 billion kWh in U.S. alone which is equivalent to powering up 6.4 million average American homes [1]. Based on the current trends and aided by the advent of hyperscale data centers and server virtualization, data centers are projected to consume a staggering 73 billion kWh in 2020. While cooling was predicted to utilize 30-40% of the total power consumed by data centers [2], the average power usage effectiveness (PUE) value of 1.8 indicates poor efficiency in the cooling system [1]. Thus, it is incumbent on thermal engineers to develop efficient methods in order to minimize the expenditure towards cooling while not disregarding the diminishing global power resources. *Why liquid cooling in data centers?* Processor speeds and cooling needs go hand in hand. High processor density is a result of the following a) increase in core count, b) having multiple chips inside a single package (MCM) and c) reduction in die

size. Increase in the server package density with a desire to provide faster computing and connectivity translates to a rise in server rack density when the high-performance servers are brought together in a rack as like in a typical data center environment. While air cooling is effective in the range 10-15 kW per rack they are still far off from being energy efficient as the thermal resistance from the chip to the incoming air is very high and delivering cold air through the racks using fans involve moving parts and are often expensive. Liquid cooling with its broad range of classification on the other hand is capable of removing 25kW per rack easily and have potential to go further above 100kW, thanks to the heat carrying capacity associated with liquids. It was reported that for a water-cooled system, the facility power consumption reduces by 45% compared to an air based system [2]. In another work [3], it was shown that, energy savings of 25% was achieved by adopting a chiller less facility with warm water and recirculated air-cooled servers. This translates to about 90% reduction in cooling energy usage. The excess heat picked from the chips can be used in several ways like storage, space heating or facility water preheating. Thus, liquid cooling help in boosting the performance per wattage while limiting the total energy budget. Many advantages of liquid cooling include saving valuable space and providing a less noisy working environment. *Why cold plates?* A rack level system with direct liquid cooling (DLC) uses warm water to carry the heat away from the chips using cold plates. The low thermal resistance allows the cold plates to operate at a higher coolant temperature in which a water side economizer replaces the chilled water system thus enabling a chiller less data center operation. Latest commercial efforts have reinstated the benefits of cooling the server modules using direct liquid cooling (DLC) technology, as they not only improve the energy efficiency but also the reliability of the system. A recent study [4] on the characterization of a rack level system employing DLC, reveal that, the cold plates, hoses and the quick disconnects are the major individual components which can be optimized for improving the overall performance of system. An opportunity is taken in the current work to optimize the performance of a cold plate employed in a rack level system. A properly designed cold plate thus will not only have high heat transfer performance but also have minimal pressure drop across them as higher pressure drop results in a penalty to the pumping power. Energy savings therefore begin at the chip level.

Forced liquid cooling of chips is not a new technique in electronics thermal management. Following the pioneering works of Tuckerman and Pease, there blossoms a plethora of research [5-7] which uses single phase coolant typically forced through microchannels to dissipate the heat from chip. With the dawn of packages with multiple chips (MCM), cold plates of next generation are expected to address multiple hot spots at the same time. Package specific cold plate designs are needed to facilitate efficient heat removal from the chips. The thermohydraulic performance can be augmented in several ways like using an enhanced fin design or by breaking the boundary layer through flow constriction or incorporation of mixing devices to enable better mixing between the coolant and channel walls [8]. Cold plates discussed in the present

article uses one such augmentation technique, by means of having channel constrictions within the cold plate module in order to improve the thermohydraulic efficiency.

Microchannel cold plates tested in this study are designed to cool a MCM package having two dies. Copper block heaters having the size of actual dies integrated with a common copper heat spreader was used as a mock package. Two cold plates with different microchannel designs were developed and characterized using the mock package for a set of operational parameters like coolant temperature, flow rate and input power to the dies. The package case temperature across the heat spreader and the pressure drop across the cold plates were reported for a variety of input parameters. The results obtained using the cold plate with enhanced microchannel design was compared to the cold plate design having conventional parallel microchannels. The study not only show improved performance using an enhanced microchannel design but also signifies the importance of package specific cold plate designs. Study further shed light into the effect of neighboring dies on each other while highlighting the significance of flow direction in the microchannels relative to the package.

EXPERIMENTAL FACILITY

Mock Thermal Package

As discussed in the literature, future generation processors will have multiple chips inside a single package such that the cold plates of next generation will have to address multiple hot spots. In a typical package, the chips/dies are covered by a metal case or cap or lid to protect the chip against any thermo-mechanical strain and to increase the reliability of the silicon device. The case is typically soldered to the dies using a thermally conducting material like Indium. The microchannel cold plates are then installed on top of the case with thermal interface material in between. The microchannel cold plates developed in the present study aims to address the cooling problem of one such package with two dies under a common case. Thermal Load Board (TLB) or Thermal Test Vehicle (TTV) are usually used in industries by thermal management designers to mimic a package. The major downsides of using TTVs or TLBs are a) they are not cost efficient and typically takes around 5-7 weeks time to develop a PCB with essential packages attached to it. b) Thermal Test Vehicles (TTV) are not readily available in the package style as the actual component it is trying to simulate. The mock package used in the current study have copper block heaters mimicking; a SKX die and a FPGA die with an effective heat transfer area at the top surface of 6.34 cm² and 3.12 cm² respectively. Figure 1 shows the mock package which houses two copper block heaters mimicking two processor dies. The mock dies will be referred to as Die 1 and Die 2 in the sections to follow.

The copper blocks are equipped with cartridge heater and the input power to dies can be varied by changing the power to the cartridge heaters using DC power supply capable of delivering 2kW. RTDs are installed across the length of the copper block heaters to measure the temperature gradient in order to calculate the input power to respective dies. The copper block heaters are thoroughly insulated on all the sides

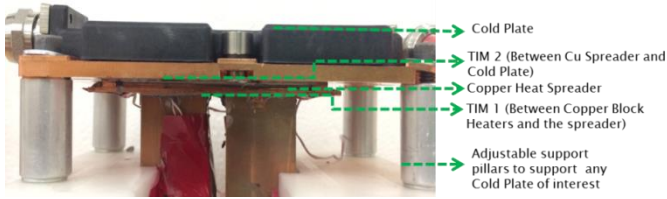


Fig. 1 Microchannel cold plates installed on top of the mock package housing two dies; Die 1 and Die 2.

leaving out the top surface such that heat flows in 1-D from the dies to the microchannel cold plates with minimal loss to the surroundings. Building a mock package this way has its many advantages like a) they are cost-effective, easily manufacturable, and the time taken to put together the fixture is quicker than for TTV or TLB and b) the researcher will have complete control over the input power to the dies with an added ability to test at power ratings higher than what would be achievable using a TTV.

Integrated Heat Spreader (IHS).

A copper case or lid or otherwise known as an Integrated Heat Spreader (IHS) is installed on top of the copper block heaters with thermal interface material in between them. The fluctuating power to the dies results in an on-chip temperature gradient. IHS reduces the on-chip thermal strain by spreading the heat from the dies to a wider area [9]. Identifying the case temperatures are of prime importance as they set the elementary design criteria for a cold plate. The local heat transfer coefficient from the channels reported in the article are also based on the measured local case temperatures. The IHS in the current package is 2.93mm tall and houses five T-type thermocouples as shown in the figure 2 to represent local case temperatures. Five Grooves of size 0.02"x 0.03" were cut to contain T-type thermocouples and the tip of the groove was made 0.035" deeper to hold the thermocouple bead. T-type thermocouples used are calibrated using a precision oven. The thermocouples are named in the alphabetical order 'A to E'. Thermocouples A and E are located on the IHS at the coolant inlet and the outlet side relative to the cold plate flow direction. Thermocouples B and D are installed on locations corresponding to die 1 and die 2 to represent the local case temperature or in other words hot spot temperatures. Thermocouple C is located between the B and D.

Microchannel Cold Plates

Cold plates developed in the study were designed and fabricated to address two hot spots on a single package. The cold plates labeled as A and B in the figure 3 are the two different microchannel designs explored in this article. The cold plates would be referred to as model A and model B in the discussions to follow. Model A design is the conventional cold plate design with parallel microchannels throughout the length of dies. Results from the parallel microchannels are essential as they often define the benchmark for novel cold plate designs optimized for performance. The detailed flow and thermal characterization of Model A were reported in an earlier article [10] from the authors. Model B is an improvised version of model A having die specific channels. The central idea behind the model B design is that, the performance of the

microchannels can be augmented by introducing flow constrictions to enable good mixing between bulk fluids and channel wall [8].

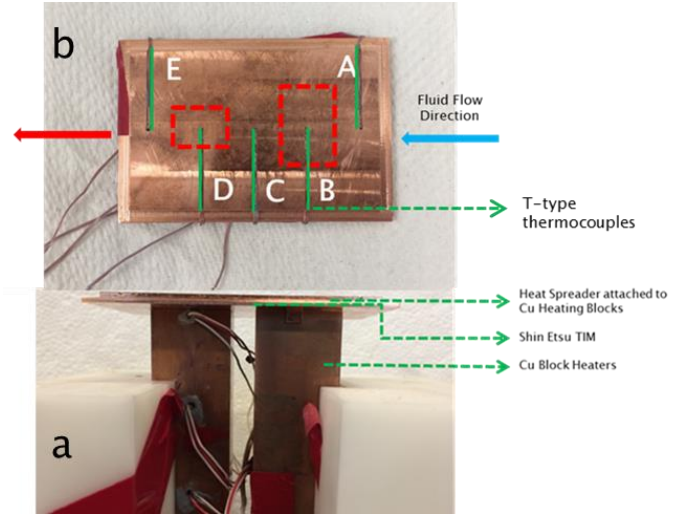


Fig. 2 a) Arrangement of IHS over the copper block heaters b) Top view of the IHS relative to the position of the dies with T-type thermocouples to represent location specific case temperatures

In case of Model B, the concentrated channels on top of die 2 results in increased coolant velocity (Reynolds number) by means of pushing the same amount of fluid in lesser number of channels. This translates to an increase in local Nusselt number or in other terms a reduction in the local case temperature. Thus, model B was expected to perform better than model A especially in addressing the hot spot corresponding to die 2 as the thermal boundary layer develops again with respect to channels above die 2. The observed results indicated the same. At the same time, by using model B, the case temperature corresponding to that of die 1 also drops because of what was happening downstream. The results will be explored further in the forthcoming sections. The physical dimensions are same for both the cold plates measuring 108mm x 72mm x 11.5mm. The copper base part contains the microchannels, while the plastic top part houses the inlet, exit manifold and a rubber seal to direct the coolant through the microchannels and not by-pass them.

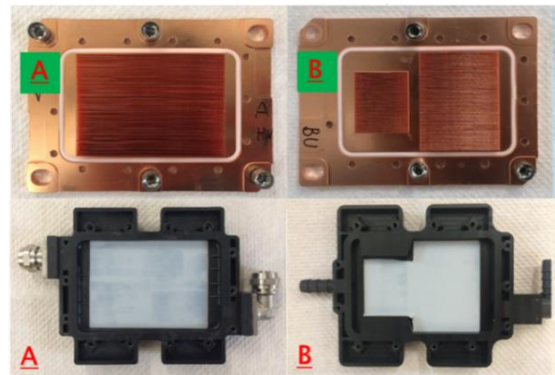


Fig. 3 Microchannel cold plate designs Model A (parallel channels) and Model B (concentrated channels)

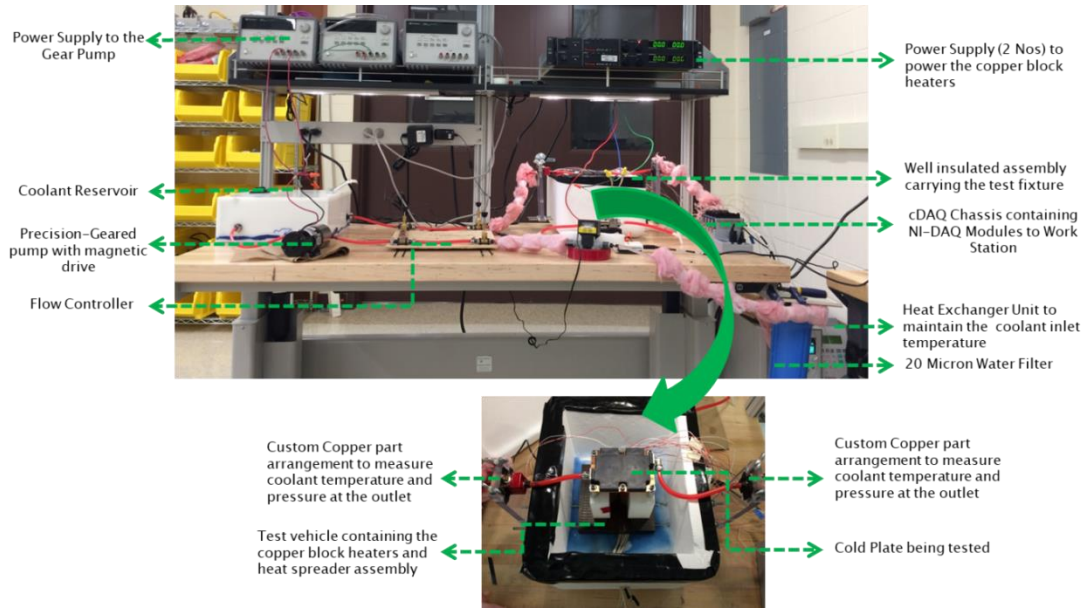


Fig. 4 Bench top experimental setup (Arrowed image: arrangement of cold plate on the mock package for characterization)

Bench-Top Experimental Setup

A bench top setup as shown in the figure 4 was assembled to characterize the microchannel cold plates. DI water from the coolant reservoir was driven by a magnetic head gear pump through the coolant flow loop. The flow rate was varied precisely by controlling the power supply to the gear pump. A flow meter capable of measuring from 0.3 to 3lpm was installed to measure the volumetric flow rate of coolant. A micro filter was coupled to the flow loop to filter out any particulate matter from the coolant before entering the microchannels. The temperature and the pressure of the coolant were measured using RTDs and pressure transducer installed at the inlet and the exit of the cold plate. The hoses along with the mock package fixture were insulated using fiber wool to minimize heat loss to the surroundings. Two DC power supplies were employed to vary the input power to dies. RTDs installed across the length of the copper block were used to measure the input power to the dies. Data Acquisition System (DAS) from NI instruments were programmed using LabVIEW VI to read and write pressure, coolant, input power and case temperature data to the excel files.

The details like brand and capability of various devices like gear pump, DC power supplies, micron filter, heater/chiller unit and detailed information about flow meter, temperature and pressure sensing instruments, their associated uncertainties and the accompanying data acquisition devices were discussed at length in an earlier article from the authors which also discusses the machining tolerance maintained while building the package along with the details of assembly procedure followed to mount the cold plates. The particulars are not provided in the present article to avoid repetition.

EXPERIMENTAL RESULTS AND DISCUSSION

Two set of experiments; a) Flow resistance experiments and b) Thermal resistance experiments were conducted to characterize the performance of cold plates at various parameters like coolant temperature, flow rate and input power to the dies. Flow resistance tests were carried out to capture the pressure drop across the cold plate which includes the pressure drop across the manifolds and that of microchannels. While, thermal resistance experiments were performed to calculate the local heat transfer coefficient by measuring the case temperature at varying input parameters.

Flow Resistance Experiments

The pressure drop across the cold module was measured using a pressure transducer capable of measuring in the range 0.2-2 psi (1 psi = 6894.76 Pa) with a measurement accuracy of $\pm 0.25\%$ RSS. The measured value accounts for the pressure drop from microchannels and that of inlet and exit manifolds. The error bars plotted in figure 5 indicate that the tests can be repeated with reasonable levels of accuracy.

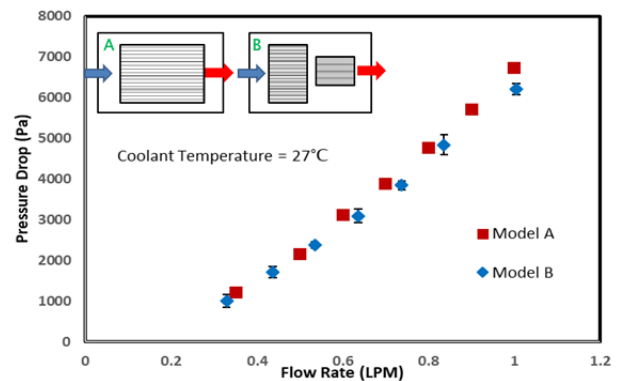


Fig. 5 PQ test comparing two different cold plate designs for an inlet coolant temperature of 27°C

The flow rate was varied from 0.35 to 1 lpm by adjusting the gear pump while maintaining the inlet coolant temperature at a constant value of 27°C. Data were continuously monitored using LabVIEW VI and were recorded once the system reaches a steady-state. Considering constant pressure drop across the cold plate inlet and exit manifolds for both the cold plate designs, the difference in the pressure drop measurements was apparently due to the length of the channels associated with cold plate modules. At the maximum flow rate tested, model B experiences a pressure drop value of 6.2 kPa, 7.5% lower than model A's pressure drop value of 6.7 kPa, thanks to the shorter microchannel lengths associated with model B. In other words, larger contact surface area results in larger pressure drop using model A compared to that of model B.

In an actual rack level system employing DCLC, coolant would be entering the cold plates at a much higher temperature thus eliminating the need for a chiller in a data center environment. At the same time, the excess heat captured from the electronics can be reused in several ways like space heating, or facility water preheating. Tests were conducted to simulate the actual operating temperature by adjusting the external heater/chiller unit to raise the coolant temperature value to 55°C. It was interesting to observe that the pressure drop value decreased by 18% and 20% as the coolant temperature was increased from 27°C to 55°C when using model A and model B respectively. The drop-in viscosity with rise in coolant temperature facilitates smoother flow through channels thus leading to a lower pressure drop value. These results at higher coolant temperature are especially good as they not only show the reduction in pumping power required to drive the coolant through the cold plates but also emphasizing the advantages of adopting 'warm water cooling'.

Thermal Resistance Experiments

Tests were conducted to measure the case or package temperature at varying parameters like coolant flow rate, coolant temperature and input power to the dies. The measurement of case temperatures is necessary in identifying the die specific thermal resistance and the local heat transfer coefficient associated with the channels. Plentitude of tests were performed at varying levels of power to die 1 and die 2 and for flow rates ranging from 0.35 to 1 lpm at a constant inlet coolant temperature of 29°C. Data were recorded using LabVIEW VI at each steady state point after changing the input variable. Input power was calculated using voltage (V) and current (I) data obtained from power supply while the output power was calculated using $Q = mC_p(T_{out} - T_{in})$. Obtained results indicate a heat loss of less than 5% within the setup. But in an actual scenario, there would be significant amount of heat loss from the dies to the board and other components in the form of conduction or convection. Thus, the case temperature values reported in this article are higher than what would be expected in reality.

The flow direction was maintained in such a way that die 1 receives the coolant first. The flow direction is referred to as 'A-E' (based on thermocouple locations relative to the flow)

in the sections to follow. Figure 6 shows the case temperature distribution in the IHS at different thermocouple locations for two different cold plate designs. The results were obtained for the case when a maximum power of 300W was supplied to the dies (Die 1: 210W, Die 2: 90W) and at a maximum flow rate of 1 lpm. The error bars plotted in case of model A results ensures repeatability. Model B with the ingenious microchannel design outperformed model A by having reduced case temperature values especially at thermocouple location corresponding to that of die 2 and die 1.

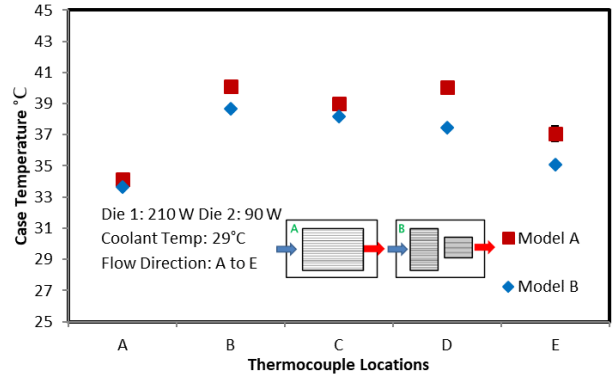


Fig. 6 Case temperature distribution comparison using two different cold plate designs

Results showing the case temperature in figure 6 could be understood better with the help of a schematic. Figure 7 shows the heat transfer path from the dies to the channels for two different cold plate designs. In case of model B which are having reduced number of channels or flow constrictions, the coolant velocity increases in channels corresponding to die 2 area by means of pushing the same amount of coolant in relatively lesser number of channels. The breakage in the channel not only enable good mixing of the fluid, but also lead to the formation of a developing boundary layer in the channels corresponding to die 2, thus resulting in an increase in local heat transfer coefficient or in other words, reduction in local case temperature. At the same time, local case temperature corresponding to that die 1 is also lower when employed model B. This is because of the conduction or diffusion from the die 2 choosing a low resistance path when using model A thus heating up the die 1 area as indicated in figure 7.

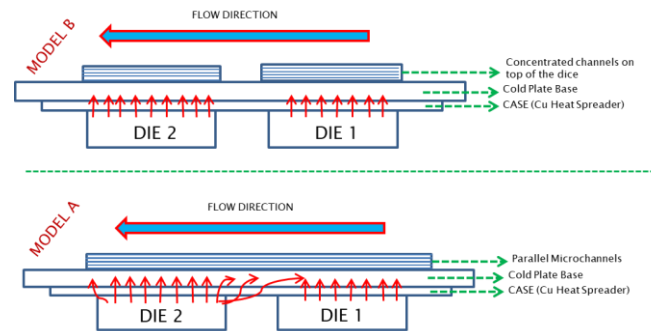


Fig. 7 Schematic of the heat transfer paths from the dies to the channels for two different cold plate designs

Model B performs better than that of model A, a closer look at the specific thermal resistance values can help the researcher realize the underlying physics. The specific thermal resistance value defined in equation 1 and 2 was estimated based on the measured local case temperatures i.e. at thermocouple locations B and D (please refer to figure 2).

$$R_{th-DIE\ 1} = \frac{T_{case,B} - T_{in}}{Q_{DIE\ 1}} \quad (1)$$

$$R_{th-DIE\ 2} = \frac{T_{case,D} - T_{in}}{Q_{DIE\ 2}} \quad (2)$$

where $T_{case,B}$ and $T_{case,D}$ represent case temperature with respect to die 1 and die 2 at thermocouple locations B and D or in other words hot spot temperatures. T_{in} represent the inlet coolant temperature. Q_{die} represents heat flux corresponding to the respective die. The controlled test setup with minimal heat loss facilitates the definition of both specific thermal resistance values based on T_{in} . Uncertainty in the measurements was calculated following [11]. The uncertainty in the specific thermal resistance measurement was estimated to be $\pm 8\%$.

Specific thermal resistance corresponding to die 1 and die 2 locations was calculated for two different cold plate designs as shown in the figure. 8. In order to plot these results, tests were conducted at different flow rates, and at different input power rating to the dies, the power to the die 1 was varied in the order of 50-100-150-210W while the power to die 2 was maintained constantly at 90W in all the tests. Thus, a maximum total power of 300 W was delivered to the dies. At every flow rate tested, the specific thermal resistance $R_{th-DIE\ 1}$ and $R_{th-DIE\ 2}$ did not vary with input power to the dies. At the same time, the specific thermal resistance value drops with rise in coolant flow rate. The specific thermal resistance approaches a saturation value as the coolant flow rate is increased. Any further increase in flow rate would be a penalty in terms of pumping power. The results in figure 8 shows that the specific thermal resistance with respect to both the die 1 and die 2 did not change when using model A which have parallel microchannels through the length of dies. The specific thermal resistance values corresponding to die 1 and die 2 did vary when using model B.

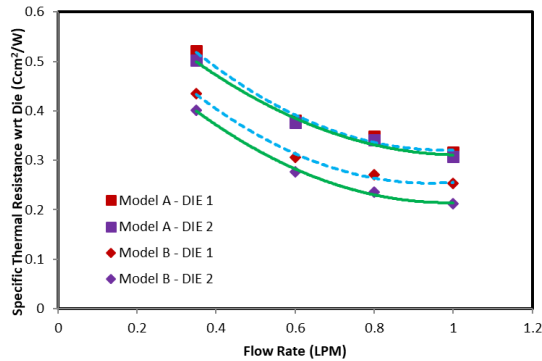


Fig. 8 Specific thermal resistance corresponding to die 1 and die 2 at different coolant flow rates using two different cold plate designs

While the differences when using model B are majorly due to the varying length (size) of die specific channels and partly

due to good fluid mixing at the channel interval as explained using figure 7. A resistance network based analysis is always a useful tool to understand the heat captured at or transferred between different layers. The specific thermal resistance defined in equation 1 and 2 encompasses resistance from TIM and that of the cold plate. The cold plate thermal resistance contains three individual resistances namely; 1. Conduction/bulk resistance at the copper heat sink base 2. Caloric resistance due to the flowing coolant and 3. Convective resistance from the channels to the incoming coolant.

$$R_{th-DIE} = R_{ColdPlate} + R_{TIM} \quad (3)$$

$$R_{ColdPlate} = R_{cond} + R_{calo} + R_{conv} \quad (4)$$

The percentage contribution from the conduction and caloric resistances to the total cold plate resistance is usually very small compared to that of the influence from the convective resistance. Assuming negligible spreading resistance and considering constant conduction and caloric resistance for the both cases when model A and model B were employed, it is innocuous to state that the local heat transfer coefficient is inversely proportional to the measured specific thermal resistance minus TIM resistance. Tight machining tolerance maintained while manufacturing the mock package and a properly followed assembly procedure facilitates minimal TIM thickness between the mating or contact surfaces. The thickness of the TIM was considered to be 100 microns with a thermal conductivity value of $6 \frac{W}{mK}$ which is confirmed by the numerical model as well. As shown in figure 9, the local convective heat transfer coefficient increases with coolant flow rate.

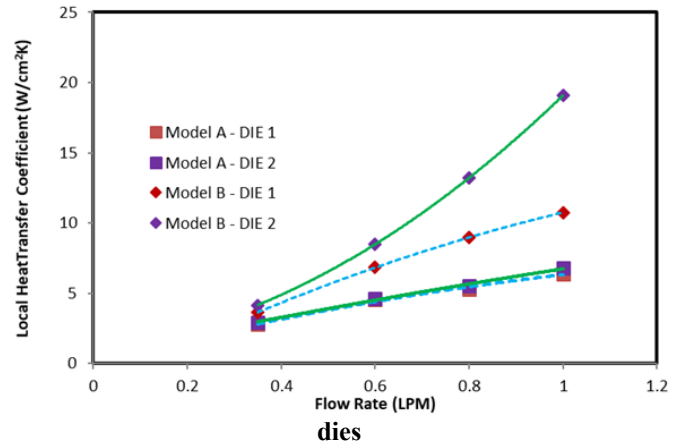


Fig. 9 Local heat transfer coefficient corresponding to die 1 and die 2 or hot spot temperatures at different coolant flow rates using two different cold plate designs

The local heat transfer coefficient calculated when using model A did not change with respect to die locations. This behavior was expected with conventional parallel microchannels. At a flow rate of 1lpm, the local heat transfer coefficient at the hot spots from die 1 and die 2 increases by 68% and 181% respectively when using model B compared to that of model A. The results obtained when using model B

clearly shows an improvement in the local heat transfer coefficient corresponding to the hot spots, thanks to the subtle tweaks to its microchannel design which facilitates good fluid mixing and leads to the formation of a developing boundary layer in channels with respect to die 2 thus enabling better heat transfer performance.

The designed cold plates with a total height of 11.5mm are deployed in low profile footprints like blade servers or IU servers or custom chassis which typically have two cold plates connected in series as shown in figure 10. Tests were extended using both the cold plate modules; model A and B to simulate a real-time situation when the coolant comes in at 59°C from the supply manifold to the cold plates. Assuming all the heat from the dies to be picked by the first cold plate, the inlet temperature to the second cold plate was estimated to be 63°C for a coolant flow rate of 1lpm. Experiments were conducted at two different inlet coolant temperatures; 59°C and 63°C with the total power to first and second package maintained at 300 W (210+90W) and 225W (165+70W) respectively. The observed case temperature value rises linearly with respect to coolant temperature. Thus, the calculated specific thermal resistance values remain constant with rise in coolant temperature, or in other words the convective heat transfer efficiency did not change with coolant temperature. But recall that, the pumping power decreases with rise in coolant temperature due to the drop in coolant viscosity.

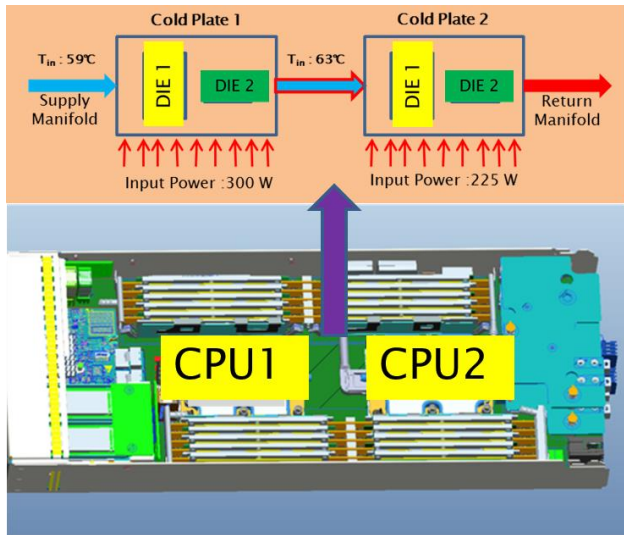


Fig 10. Tests conducted to address the effect of coolant temperature (simulating real time situation)

Tests were conducted after reversing the flow direction in the cold plates to observe their effects on the thermal performance. This was done by merely rotating the mock package fixture by 180°. The arrangement was such that the die 2 receives the coolant first. The flow direction is referred to as 'E-A' in the results and discussions. The results shown in figure 11 indicate that the case temperature with respect to FPGA die drops significantly by about 8% when the flow direction is reversed using both the cold plate models. Results clearly show that it is advantageous to let the die 2 receive the coolant first. While changing the flow direction when using model B do not have a significant effect on the case

temperature with respect to die 1, reversal in flow direction has a detrimental effect when using model A as the high powered die 1 would be at the rear end of the microchannels resulting a natural rise in local temperature. A quick model calculation performed for when the flow direction is 'A-E' assuming all the heat from die 1 to be picked by the incoming coolant, reveal that the higher case temperature with respect to die 2 is because of the heated coolant leaving the die 1 area and not due to conduction or spreading. This condition also helps in defining the specific thermal resistance in the way it is. These results are especially good as they not only show the effect of neighboring dies on each other but also let the board developers decide how closely the dies can be located to each other such that interconnect lengths can be curtailed creating shorter power paths. Also interconnects length plays a crucial part in data transmission and retrieval time. Identifying die locations on the board at the design stage can facilitate hot spot management when retrofitting the heat sinks at a later stage.

Comparison of different results obtained using two different cold plate modules model A and model B reveal that model B proved to be a better design both in terms of pumping power and effective heat transfer performance. These results are specifically good as they imply the importance of package specific cold plate designs used in the servers deployed in a rack level system. Consider the following example, for a high-performance package (in HPCs) having a size similar to the tested package but with a higher power rating, the thermal design requirement for the package can be met using the existing model A (nevertheless with a harsh penalty to the pumping power) but the performance would still be far off from being energy efficient and thus an abundant room for improvement and optimization blossoms as a result.

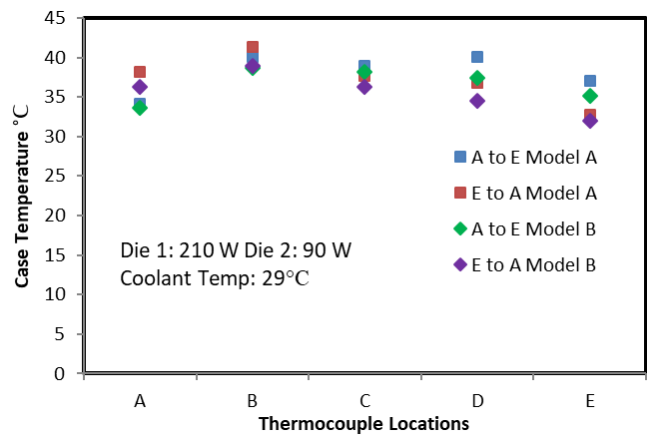


Fig. 11 Tests conducted to show the effect of coolant flow direction using two different cold modules.

NUMERICAL MODEL

An all comprehensive numerical model was established for model A to validate the experimental results and further to explore package specific microchannel designs and to develop an optimized model in the end. For the numerical simulation

dies of microchannel cold plates, the following are the set of basic assumptions applied on to the governing equations.

1. The flow is three dimensional, steady, laminar and incompressible.
2. The effects of gravity are negligible.
3. The thermo-physical properties like conductivity, specific heat, viscosity of the coolant and solid phases are constant.
4. Viscous heating and radiation heat transfer are omitted.
5. Coolant distribution is uniform through the channels.

The governing equation for conjugate (conduction-convection) heat transfer after applying the above assumptions turn out to be;

Continuity equation (liquid phase)

$$\nabla \cdot \vec{V} = 0 \quad (5)$$

Equation of motion (liquid phase)

$$\rho_f (\vec{V} \cdot \nabla) \vec{V} = -\nabla P + \mu_f \nabla^2 \vec{V} \quad (6)$$

Energy equation (liquid phase)

$$\rho_f c_{p,f} (\vec{V} \cdot \nabla) T_f = k_f \nabla^2 T_f \quad (7)$$

Energy equation (solid phase)

$$k_s \nabla^2 T_s = 0 \quad (8)$$

Only one half of the computational domain was considered as there existed a geometric symmetry. Figure 12 shows the schematic of the computational domain with associated boundary conditions. DI water as coolant enters the microchannels seeing the die 1 first and flows over the die 2 specific area before exiting. A constant pressure drop boundary condition was applied at the exit of the channels. A symmetry boundary condition was used at the front plane of the computational domain (where the heat sink is divided into symmetric parts). The key dimensions of model A design are; fin height: 3mm, channel width: 0.2mm and fin thickness: 0.2mm. The base plate thickness is taken to be 3.5mm. Constant heat flux boundary conditions were applied to the dies at the bottom.

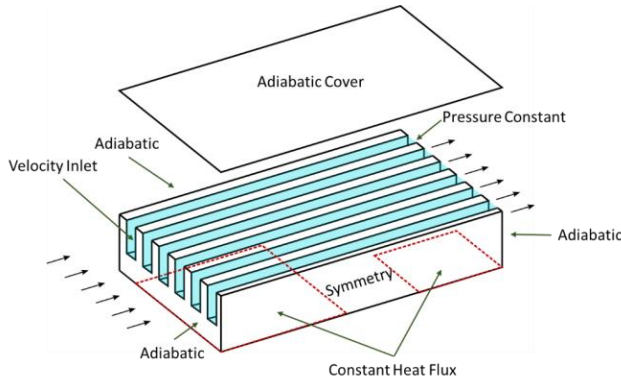


Fig. 12 Computational domain considered for the numerical model and the associated boundary conditions

The governing equations were solved numerically using the finite volume SIMPLEC algorithm with a pressure based solver. The equation of motion and energy were discretized using a second order upwind scheme. A structured grid with hexahedral cells was employed for the whole domain which

includes both the solid and liquid sections. A grid sensitivity analysis was also performed based on the flow and thermal resistances. The results shown in the figure 13 were simulated using model A for the case when a total power of 300 W was delivered to the dies; 210 W to die 1 and 90W to die 2 respectively. The coolant flow rate and temperature was maintained at 1lpm and 59°C. The thermophysical properties of different layers were obtained from the literature, except for the TIM where the conductivity value of $6 \frac{W}{mK}$ was provided by the manufacturer. The TIM thickness was assumed to be 100 microns in the simulations performed. The numerically simulated case temperatures with respect to die 1 and die 2 agree well with the experimental findings with less than 1% difference.

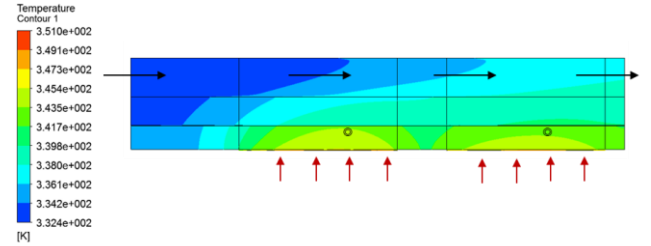


Fig. 13 Numerically simulated case temperatures using model A when a total power of 300 W was delivered to the dies and when the coolant temperature was at 59°C

A very detailed numerical analysis was performed for different parameters like coolant temperature, flow rate, chip power and heat sink material. The numerical study also included exploring different microchannel designs (like model B reported in the article) for efficient flow and thermal performance. A thorough statistical based optimization study was carried out in the end. The comprehensive study of the cold plate designs and results from optimization scheme are reported in an exclusive article dedicated for numerical observations from the authors. Moving forward, a quick validation was required to benchmark the experimental results and the detailed numerical analysis are out of scope in the current article.

CONCLUSION

Cold plates having enhanced microchannel designs were characterized using a mock package with two chips capable of delivering 300W and above. Tests results show that the cold plate model B having the enhanced microchannel design, performed better than cold plate Model A having conventional parallel microchannels design. The local heat transfer coefficient at the die hot spots increases by 68% and 181% respectively using enhanced microchannel design. The study not only shows improved performance using an enhanced microchannel design but also signifies the importance of package specific cold plate designs. Study further shed light into the effect of neighboring dies on each other while highlighting the significance of flow direction in the microchannels relative to the package. The intricacies to address hot spots will be even harder in packages with more than two chips. While high coolant temperature facilitates

lesser pumping power, they do not have an effect on the thermal behavior. The experimental results are in good agreement with the numerical model developed while the numerical simulation suggests more room for improvement in terms of innovative geometric designs. Optimized cold plates mounted in server modules when grouped to rack level arrangement yield better system efficiency. Thus, 'Energy savings at the cold plate level (component level) translates to savings at the server level and ultimately at the rack level (system level)'.

Acknowledgements

This work is supported by NSF IUCRC Award NO. IIP-1134867. The author would like to thank Pat Crowley and Ron from Progressive Tool Co, Endwell, NY for their guidance and assistance in putting the mock package together. The author would like to express gratitude towards Steven Schon (Quantacool), Mark Seymour (Futurefacilities) and Russ Tipton (Vertiv Co) for their valuable inputs during the work.

References

- [1] Shehabi, Arman, et al. "United States data center energy usage report." (2016)
- [2] Ellsworth MJ, Jr., Iyengar MK. Energy Efficiency Analyses and Comparison of Air and Water Cooled High Performance Servers. ASME. International Electronic Packaging Technical Conference and Exhibition, *ASME 2009 InterPACK Conference, Volume 2* ():907-914.
- [3] M. P. David et al., "Experimental characterization of an energy efficient chiller-less data center test facility with warm water cooled servers," 2012 28th Annual IEEE Semiconductor Thermal Measurement and Management Symposium (SEMI-THERM), San Jose, CA, 2012, pp. 232-237.
- [4] S. Alkharabsheh, B. Ramakrishnan and B. Sammakia, "Pressure drop analysis of direct liquid cooled (DLC) rack," *2017 16th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm)*, Orlando, FL, 2017, pp. 815-823.
- [5] D. B. Tuckerman and R. F. W. Pease, "High-performance heat sinking for VLSI," in *IEEE Electron Device Letters*, vol. 2, no. 5, pp. 126-129, May 1981
- [6] R. W. Knight, D. J. Hall, J. S. Goodling and R. C. Jaeger, "Heat sink optimization with application to microchannels," in *IEEE Transactions on Components Hybrids, and Manufacturing Technology*, vol. 15, no. 5, pp. 832-842, Oct 1992.
- [7] D. Copeland, "Optimization of parallel plate heatsinks for forced convection," Sixteenth Annual IEEE Semiconductor Thermal Measurement and Management Symposium (Cat. No.00CH37068), San Jose, CA, 2000, pp. 266-272
- [8] Kandlikar SG, Grande WJ. Evaluation of Single Phase Flow in Microchannels for High Flux Chip Cooling: Thermohydraulic Performance Enhancement and Fabrication Technology. ASME. International Conference on Nanochannels, Microchannels, and Minichannels, ASME 2nd International Conference on Microchannels and Minichannels ():67-76.
- [9] Wei J. Challenges in Package Cooling of High Performance Servers. ASME. International Electronic Packaging Technical Conference and Exhibition, ASME 2007 InterPACK Conference, Volume 2 ():501-507
- [10] Ramakrishnan, B., Alkharabsheh, S., Hadad, Y., Chiarot, P., Ghose, K., Sammakia, B., Gektin, V., Chao, W., 2018, "Experimental Investigation of Direct Liquid Cooling of Two-Die Package," 34th Annual Thermal Measurement, Modeling, and Management Symposium, SemiTherm, San Jose, CA, March 19-23, 2018 (Article accepted. Yet to be published at the time of current article submission)
- [11] Thomas G Beckwith, Roy D Marangoni, and John H Leinhard. Mechanical measurements. Pearson Prentice Hall, 2007.