CPU Overclocking: A Performance Assessment of Air, Cold Plates, and Two-Phase Immersion Cooling

Bharath Ramakrishnan, Husam Alissa, Ioannis Manousakis, Robert Lankston, Ricardo Bianchini, *Fellow, IEEE*, Washington Kim, Rich Baca, Pulkit A. Misra, Inigo Goiri, Majid Jalili, Ashish Raniwala, Brijesh Warrier, Mark Monroe, Christian Belady, Mark Shaw, *Member, IEEE*, and Marcus Fontoura, *Senior Member, IEEE*

***Abstract*— Computing capacity has always been on the upward climb due to the constant technological improvements in semiconductor manufacturing and packaging industry. This growth in computing capability is usually accompanied by a steep rise in heat flux density associated with the electronic component (CPUs or GPUs, for example). High-performance computing (HPC) data centers often employ several of these high-performance devices for crunching their enormous artificial intelligence (AI) or scientific computing workloads. State-of-the-art air cooling technologies throttle after a certain heat flux level and would require bigger heat sinks driving enormous airflow through it, which may not be desirable from a data center operation standpoint. Hence, a lookout for advanced thermal management techniques is quite imperative. In this article, a commercially available intel overclockable CPU i9-9900k was exercised with high performance workloads while employing and evaluating three different cooling technologies namely air-cooled, cold plates, and two-phase (2P) immersion. The high heat carrying capacity associated with cold plates and 2P immersion techniques outperformed the air-cooled solution by constantly yielding higher CPU clock rates up to 41% utilizing cold plates and 51% employing 2P immersion. The performance of cold plates and 2P immersion was evaluated at different functional points (different coolant operating temperatures and 2P coolant types for example) to understand how the technologies would respond in a legacy or modern data center operation. It was again observed that both 2P immersion technology and cold plates provided the least thermal resistance (as low as 0.247 °C/W) path to heat transfer and therefore provided higher computational performance and efficiency compared to air cooling. Further improvement in performance is not limited by cooling but by packaging constraints. Advanced packaging techniques coupled with minimizing the interfacial resistance can reap enhanced performance and energy efficiency gains using these modes of cooling. This article demonstrates the potential increase in virtual machine (VM) performance that can be attained using cold plates or 2P immersion. Authors recommend adoption of liquid cooling in early chip design and architecture phases. Finally, the article provides a generic overclocking methodology that the industry can use to qualify postturbo server class chip performance.**

***Index Terms*— Air cooling, cold plates, data center thermal management, high-performance computing (HPC), overclocking, two-phase (2P) immersion.**

Manuscript received January 31, 2021; revised May 29, 2021; accepted July 26, 2021. Date of publication August 19, 2021; date of current version October 28, 2021. Recommended for publication by Associate Editor M. Iyengar upon evaluation of reviewers' comments. *(Corresponding author: Bharath Ramakrishnan.)*

Bharath Ramakrishnan, Husam Alissa, Washington Kim, Mark Monroe, and Christian Belady are with Microsoft CO+I, Redmond, WA 98052 USA (e-mail: bhramak@microsoft.com).

Ioannis Manousakis, Robert Lankston, Rich Baca, Ashish Raniwala, Brijesh Warrier, Mark Shaw, and Marcus Fontoura are with Microsoft Azure, Redmond, WA 98052 USA.

Ricardo Bianchini, Pulkit A. Misra, Inigo Goiri, and Majid Jalili are with Microsoft Research, Redmond, WA 98052 USA.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TCPMT.2021.3106026.

Digital Object Identifier 10.1109/TCPMT.2021.3106026

NOMENCLATURE

AI	Artificial intelligence.
ML	Machine learning.
VM	Virtual machine.
HPC	High-performance computing.
CPU	Central processing unit.
GPU	Graphical processing unit.
NPU	Neural processing unit.
TPU	Tensor processing unit.
RAM	Random access memory.
TDP	Thermal design power (W).
SKU	Stock keeping unit.
TIM	Thermal interface material.
IHS	Integral heat spreader.
BIOS	Basic input–output system.
DUT	Device under test.
CDU	Coolant distribution unit.
R_{th}	Thermal resistance (°C/W).
R_{TIM}	TIM thermal resistance (°C/W).
R_{IHS}	IHS/CASE thermal resistance (°C/W).
R_{hs}	Heat sink thermal resistance (°C/W).
T_a	Ambient (coolant) temperature (°C).
T_j	Junction temperature (°C).
Q_{cpu}	CPU core power/input power to CPU (W).
t_{TIM}	Thickness of TIM (mm).
t_{IHS}	Thickness of IHS (mm).
k_{TIM}	Thermal conductivity of TIM (W/m°C).
k_{IHS}	Thermal conductivity of IHS (W/m°C).
h_{HS}	Heat-transfer coefficient of heat sink (W/cm^2 °C).
A	Area available to heat transfer (cm^2).

I. INTRODUCTION

WE ARE running up against the limits of all the compute scaling laws. Moore's scaling (1965) of transistor' density has slowed down due to lithography limitations [1].

Similarly, and almost a decade ago, the breakdown of Dennard scaling (1974) happened as transistors leakage power continued to increase. Today, the number of cores as well as their performance are continuously increased by chip manufacturers for better monetization. Advanced packaging technologies and multiple chip modules (MCMs), and high bandwidth memory (HBM) on chip are being incorporated. AI/GPU/NPU/TPU workloads have also been pushing toward higher performance/bandwidth, and lower latency chip architectures. These developments have led to significant increases in chip thermal design power (TDP, Watts) and heat flux (W/cm^2). This can also be accompanied by stricter thermal requirement of lower case temperature for the chip. As rightly predicted by Bar-Cohen [1], this increased power dissipation trend will continue requiring new cooling and packaging technologies such as 2.5-D/3-D packaging to dissipate the thermal load and deliver stable, reliable, and efficient higher chip performance.

For some historical perspective: The microprocessor power dissipation plateaued at around 100 W/chip in 2005 and similarly the clock frequency stabilized at around 3 GHz in 2010 suggesting a single thread performance "stall." Even though as the chip core counts increase, because of the inability to remove the dissipated heat thereby has led to "dark" or unused Silicon where the increase in core/transistor count are only moderately utilized. In simpler words, "As chip transistor counts increase, the inability to remove heat is limiting the number of transistors that can be concurrently active." Studying these trends, Fuller and Millet [2] rightly concluded that the growth in computing performance would be limited by thermal and power requirements, hence new cooling and packaging technologies are clearly needed. As more and more transistors are provided on a single chip, the losses associated with communication and propagating the signal off the chip are greatly reduced. This is especially good from a HPC sector perspective.

Super computers for example: Japanese K computer [3], US' Jaguar [4], and IBM' 775 [5] being large systems dissipated 5–10 MW from space occupying 1000 m^3. These large heat-carrying capacities are attributed to the computing system being liquid cooled. Strikingly, the computational efficiency associated with the K computer was reported to be 1 petaflops/W and it can be credited to the increased transistor count aided by liquid cooling using water or refrigerants. The Cray-2's [6] innovative cooling scheme involves immersed dense stacks of PCB in engineered dielectric fluid. Hence, from some of these renowned examples, it is evident that the high-performance systems embrace liquid cooling technologies for a variety of technoeconomic reasons.

We list the thermophysical properties of air, water and 2P immersion fluid in Table I. For 2P immersion we use a dielectric fluid. In particular, we use a low boiling point dielectric fluid, HFE 7000 [7] from 3M. At atmospheric pressure, the saturation temperature of HFE 7000 is 34 °C. One of the most significant differences between these fluids would be their density. Referring to Table I, the density of air at 34 °C is 1.14 kg/m^3, while the density of water is 994 kg/m^3 and the density of the dielectric fluid HFE 7000 is 1400 kg/m^3. Another important property would be the specific heat of the fluid. The specific heat of air is 1.01 kJ/kg. °K, the specific heat of water is 4.18 kJ/kg. °K and the specific heat of HFE 7000 is 1.30 kJ/kg. °K. During two-phase (2P) immersion, the dielectric fluid changes phase at the intended temperature (isothermal process), hence the specific or sensible heat is no longer a useful parameter rather the latent heat of vaporization, which is relatively quite high equaling 142 kJ/kg. Even after multiplying the specific heat of air and water by reasonable high temperature rises of 11 °C and 6 °C, the volumetric heat-transfer capacity of the dielectric fluid is quite large compared to air (15 743:1) and water (8:1). Thus, 2P dielectric fluids can carry more heat.

TABLE I

HEAT-TRANSFER AND COP COMPARISON BETWEEN AIR, WATER, AND HFE 7000 FOR A TYPICAL DATACENTER OPERATION

Cooling Medium/ Properties	Unit	Air	Water	HFE 7000
Density	**(kg/m^3)**	**1.14**	**994.10**	**1400.00**
Specific Heat	**(kJ/kg. °K)**	**1.01**	**4.18**	**1.30**
Latent Heat	**(kJ/kg)**	**--**	**--**	**142.00**
Volumetric Heat Capacity	**(kJ/m^3 °K)**	**1.15**	**4153.35**	**198800**
Typical heat Rise	**(°K)**	**11.00**	**6.00**	**--**
Volumetric Heat Transfer (HT) Content	**(kJ/m^3)**	**1.63**	**24920.10**	**198800**
Volumetric HT Ratio to Air			**1973.43**	**15743.07**
Volumetric HT Ratio to Water				**7.98**
Flow Rate	**(m^3/s)**	**1.667E-05**	**1.67E-05**	**1.67E-05**
Heat Transfer/Unit Time	**(J/s)**	**0.21**	**415.33**	**3313.33**
Typical System Pressure Drop	**(kPa)**	**0.623**	**100**	**150**
Pressure Drop Ratio to Water			***160***	***240***
Pumping Power	**(J/s)**	**0.0103833**	**1.666667**	**2.5**
COP	**--**	**20.26**	**249.20**	**1325.33**
COP Ratio to Air			***12.29***	***65.39***
COP Ratio to Water				***5.32***

From a pumping power or energy transport perspective, the power required to pump a given volume of water or dielectric fluid is much higher than for a given volume of air plainly due to the high density of liquid coolants. Referring to Table I, representative fan pressures of 0.62 kPa for air cooling and conservative pump pressures of 100 and 150 kPa for water and HFE 7000, respectively, were assumed in pumping power estimation. In this example, the pressure needed to drive/pump a given volume of liquid coolant is greater than air by a ratio of between 160:1 (Water) and 240:1 (HFE 7000). Dividing the volumetric heat capacity of these coolants, namely, air, water, and HFE 7000 by the pumping power (expressed as equivalent heat), coefficient of performance (COP) of these different coolants are obtained. The COP of air transport is about 20, the COP of water transport is about 250 and the COP of the dielectric fluid HFE 7000 is about 1325. In this example, hence the dielectric fluid is the most efficient and air is the least efficient. It should be noted that, every project will have a specific temperature rise, fan and pump pressure expectations. And as shown in Table I, the differences in energy transport can be dramatic. This example was extended to other 2P dielectric coolants like FC 3284 [8] (At atmospheric pressure, the saturation temperature of FC 3284 is 50 °C) and similar performance/energy boosts were observed.

With this information as background, let us look at a particular problem of cooling down an overclockable processor using air, water (cold plates), HFE 7000 and FC3284 (boiler plates) in Sections II–V in detail. In the subsequent sections, let us talk about the overclockable CPUs first, followed by the list of hardware used to run these tests, then followed by the test set up and the performance comparison in the end.

II. Overclocking

Internet offers many definitions to overclocking. Intel [9] defines "overclocking" any unlocked electronic component [processor, RAM, or motherboard (MB)] to custom tune the performance of server or personal computer. The power, voltage, frequency or clock rates, core, memory settings, and other key system values of the unlocked electronic component can be adjusted for extracting maximum compute performance.

A. Advantages of Overclocking

Overclocking results in speeding up of components' performance or throughput; thus an enhanced gameplay for example. Overclocking in other computing scenarios can help with processor intensive tasks such as image rendering and transcoding. In certain other circumstances, overclocking is deemed a way to extend the usable life on the equipment. At a data center or cloud-level operation, overclocking can be beneficial by means of "oversubscription of servers" and "virtual machine (VM) auto scaling" [10].

B. Implications of Overclocking

Overclocking an electronic component at high-performance levels by increasing the clock rate or frequency increases the heat generated. The TDP of the component usually increases linearly with the clock frequency rate. There is a maximum frequency threshold up to which the frequency can be increased comfortably without compromising the stability. After which the voltage of the component needs to be increased to maintain stable clock frequencies. Thus, increasing the voltage increases the heat produced further (All electronic components produce heat by the movement of current through the electronic circuit. For example, in a linear circuit, heat produced is proportional to the square of the voltage). This increased heat generation within the device needs to be properly addressed for the longevity, reliability, and proper functioning of the component.

A wide variety of thermal solutions ranging from using liquid Nitrogen or Helium, or other advanced liquid or air-cooling solutions readily support this overclocking feature and can let individual PC enthusiasts' clock their electronic devices at higher frequency levels. But not all the thermal solutions are practically feasible or adoptable at scale when it gets to data centers. It gets increasingly difficult for air cooling to support high heat flux density components after a certain heat flux level. Air cooling would require bigger heat sinks and would expect enormous air flow through it to achieve desired device junction temperatures. This is undesirable at the data center level where we tend to stack many of these devices on top of other. On the other hand, cold plates which uses water and 2P immersion which uses low boiling point dielectric fluids have high heat-transfer coefficient associated with them compared to air-cooled heat sinks and can offer reduced **s**ize, **w**eight, and pumping **p**ower (SWaP). In addition to this, cold plates and 2P immersion technologies are becoming more mature and scalable to rack/row level at the data center level compared to other nascent cooling techniques like 2.5-D/3-D embedded cooling or microfluidics. In this article, three thermal solutions namely air cooling, cold plates [11], and 2P immersion cooling [12], [13] were used to support overclocking of the processor.

TABLE II
EXAMPLES OF HIGH-END DESKTOP UNLOCKED CPUs

Model	Cores/ Threads	Base/Boost (GHz)	TDP (W)
Intel i9 9900k	8/16	3.6/5.0	95
AMD TR 3970X	32/64	3.7/4.5	280
Intel W- 3175X	28/56	3.1/4.8	255
Intel Xeon W 3265	24/48	2.7/4.6	205

III. Hardware Used and Benchmarking

A wide variety of CPUs from Intel and AMD have commercially available unlocked versions which let the user go to higher frequency levels for extracting maximum performance.

Some of the high-end desktop CPUs are provided in Table II. As shown, the CPUs can be exercised to 5 GHz or above when overclocked carefully. The associated TDP at base frequency can be as high as 280 W. During overclocking, they can go to even higher power levels than the associated TDP rating. Every CPU has an associated maximum junction temperature T_j (°C) that needs to be kept under control during operation. The image of the CPU tested in this article (intel i9-9900k) and their associated dimensions/specifications along with a schematic is shown in Fig. 1. This SKU was chosen because, from a thermal standpoint, the associated die-level heat flux is much higher and challenging to cool compared to other CPUs listed. It was observed that, for this CPU, 90% of the heat was generating out of the 50% of the die area (which are primarily occupied by the cores). Thus, the heat flux associated with cores could go above 220 W/cm^2 under peak load conditions. Going into a future, the thermal challenge lies in heat flux (W/cm$_{\wedge 2}$) rather than TDP (W). Devices with chiplets should produce manageable flux values even when the

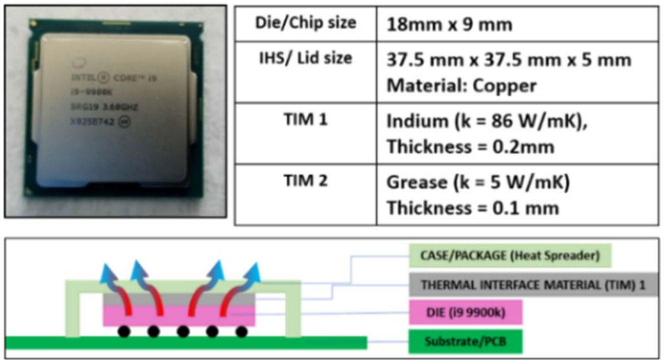


Fig. 1. Intel i9 9900k specs and schematic.

TDP increases. It is noted that some of the newer high-power GPUs are lower in heat flux density when compared to lower TDP CPUs but with higher heat flux like the one discussed in this article.

The schematic of the i9 package in Fig. 1 exposes the various layers inside the package starting from die to thermal interface material (TIM) 1 to the package case/integral heat spreader (IHS). From a mechanical standpoint, IHS acts as a cover and provides safeguard to the die from shock and vibrations. From a thermal perspective, IHS let the heat spread to a wider area thus making it easier for the heat sink to remove heat from top. Strictly talking, these layers are in fact layers of resistance to heat dissipation from the die to the case. The thermal conductivity, sectional area, and thickness of different layers (as provided in Fig. 1) influence the conductive thermal resistance to heat-transfer R_{cond}(°C/W). Typically, the heat sinks are attached on top of the Case/IHS with another layer of TIM 2 typically grease or could be another layer of Indium too.

A. Benchmarking Methodology

Before getting into the overclocking process, it should be noted again that altering the clock frequency or voltage could void the CPU manufacturer's warranty and may result in reduced stability, security, performance, and life of the device. Thus, proper consideration was taken during every step of overclocking process so as not to incur any damage. An ASUS Z390-F Gaming MB was used with an appropriate power supply and RAM configurations to power the CPU.

Overclocking is done by carefully tuning the knob for CPU clock frequency or CPU core voltage in the basic input–output settings (BIOS) while well supported by a cooling solution. Tests were initially run at base frequency to provide a baseline. Next, the computer was rebooted, and the BIOS was entered to access the CPU, DRAM, and Cache specifications. These settings were maintained uniformly across different overclocking tests using different cooling platforms. The core ratio limit and CPU core voltage were varied and recorded for each test in the BIOS settings to achieve stability. HWinfo64 was used to record the MB data containing clock frequency, core power, temperature at a sampling rate of 1 Hz.

B. Computing Workloads

Three stress testing tools Prime95, Intel XTU, and Cinebench were used to exercise intensive workloads to the CPU at the selected BIOS settings. The loads could be simulating a graphically intensive or mathematically-intensive application. Prime 95 offered the maximum possible workload compared to XTU or Cinebench and thus the CPU package power readings associated with Prime95 tests were consistently higher as reported later in this article.

The following steps were followed to overclock the processor in a systematic way and to get consistent test results.

1) Initially the CPU clock frequency was manually changed in the BIOS in the increments on 0.1 GHz. Run the desired workloads.
2) If the stress tests fail due to being thermally throttled, reduce the core voltage by 0.01 V and see if the system stabilizes. If the system does not stabilize and gets thermally throttled again, reboot the system, and reduce the core frequency by 0.1 GHz in the BIOS.
3) If the system fails immediately with a blue screen of death (BSOD), reboot the system, and increase the core voltage by 0.01 V.
4) If the workloads execute successfully, reboot the system, and increase the core frequency by 0.1 GHz and repeat the procedure in the previous two steps. Continue ramping up the frequency until the point where the system destabilizes.
5) Make sure that the desired operating parameters like inlet coolant temperature or flow rate are stable before recording test data. Obtain all the readings after the system reaches a steady state. As the stability criteria, run XTU and Prime95 workloads for 5 min, and loop the Cinebench workloads three consecutive times with no errors or application crashes or BSOD.

IV. Test Setup and Parameters Tested

Three different heat sinks, namely, air-cooled, cold plates, and 2P immersion boiler plates were used while overclocking the CPUs. At a data center level, the server chassis equipped with different cooling configurations would be looking as shown in the Fig. 2(a). But for the actual tests discussed in this article, data were obtained using a gaming MB populated with DIMMs and equipped with a power supply. The picture of an air-cooled i9 9900k in a MB is as shown in Fig. 2(b).

Parameters that were tested during this study are captured in Table III. The following section describes the different test setups, instruments, schematics, and the chosen test parameters as well as the reasoning behind them. The device under test (DUT) in the following set of figures refers to the MB having CPUs with the different heat sink configurations.

A. Air Cooling

A commercially available air-cooled heat sink (model no: intel XTS100H) as shown in Fig. 3 was used to overclock the CPUs. The chosen heat sink may not be the best-in-class air cooling solution but could very well be one of the top heat

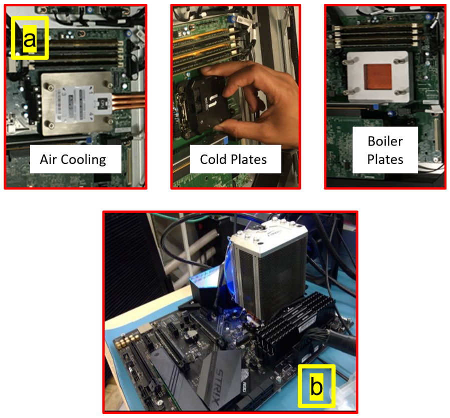


Fig. 2. (a) Picture of server chassis equipped with different heat sinks to realize different cooling/heat-transfer capabilities. (b) Picture of an air-cooled setup showing the MB populated with DIMMs and the heat sink.

TABLE III
PARAMETERS TESTED IN THIS STUDY FOR THREE DIFFERENT COOLING CONFIGURATIONS

Parameters	Air Cooling	Cold Plates	2P Immersion
Fluid Type	Air	Water:PG	HFE7000, FC 3284
Inlet or Saturation Temp (˚C)	34˚C	24,34,39, 47,50˚C	34˚C, 50˚C
Inlet Flow Rate	158 cfm/kW	1.5-5 lpm/kW	--
TIM 1	Indium	Indium	Indium
TIM 2	Grease	Grease	Indium
Stress Tools	Prime95, Intel XTU, Cinebench		

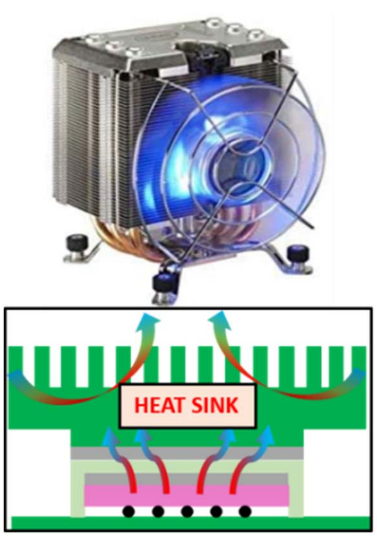


Fig. 3. Picture of an air-cooled heat sink for Intel i9 9900k along with a sketch exposing the layers of package resistance.

sink contenders in terms of both performance and cooling. The heat sink was attached on top of IHS with a layer of thermal grease (TIM 2). The thermal conductivity of the grease used

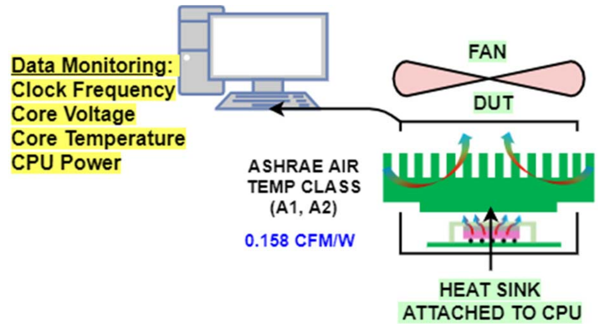


Fig. 4. Schematic of air-cooled setup.

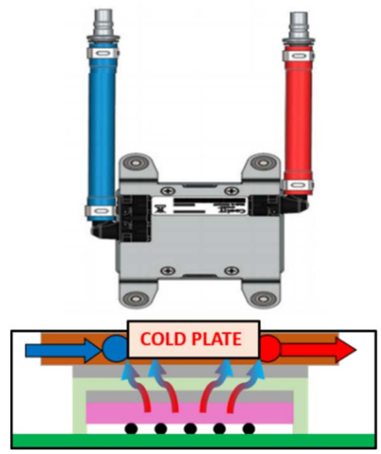


Fig. 5. Picture of cold plate used for overclocking along with a schematic exposing the layers of package resistance.

is 8.5 W/mK. Proper care was taken while attaching the heat sink. The TIM pattern indicating good thermal contact was observed at the interface while removing/swapping the heat sinks for subsequent tests.

The fan speed was controlled using Corsair Link V3 software. The fan speeds were constantly maintained to deliver 0.158 CFM/W during all stages of testing. DUT containing the CPUs were ducted and the temperature of the inlet air was maintained at 34 °C (equivalent to ASHRAE air temperature class: A2 [14]).

Fig. 4 shows the schematic of a simple air-cooled set up. The chosen flow rate and the inlet air temperature are a representative of a typical data center operation. The CPU was overclocked carefully, and the core frequency, temperature, and voltage data were monitored using HWinfo64. Steady state was maintained before recording the data. The data from air-cooled heat sink set the benchmark for the subsequent set of performance tests on the CPU using cold plates or 2P immersion boiler plates.

B. Cold Plates

A commercially available cold plate (model no: CoolIT R4 passive loop) as shown in Fig. 5 was used to overclock the CPUs using water:polypropylene glycol (PG) as the heat-transfer fluid. The cold plate contains copper microchannels in a split flow arrangement to maximize heat transfer with

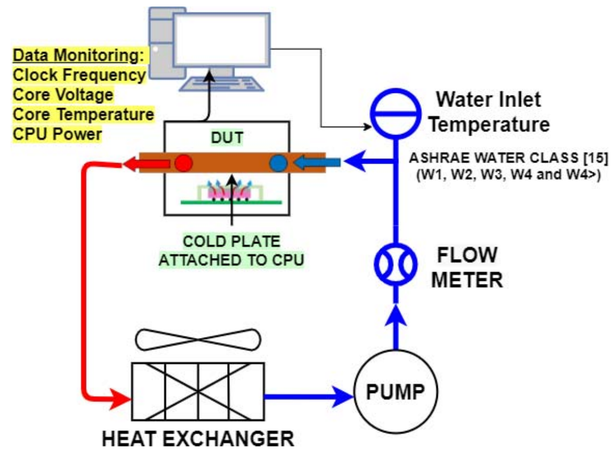


Fig. 6. Schematic of liquid cooled cold plate setup.

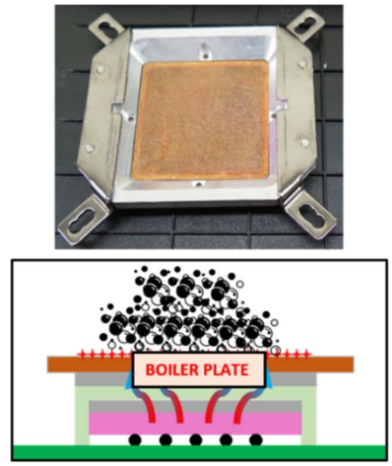


Fig. 7. Picture of a 2P immersion boiler plate used for testing along with a schematic exposing the layers of package resistance.

minimal penalty to the pumping power. The heat sink was attached on top of IHS with a layer of thermal grease (TIM 2). The thermal conductivity of the grease used is 8.5 W/mK. Proper care was taken while attaching the cold plates to maintain optimal thermal contact.

A coolant distribution unit (CDU) (model no: CoolIT AHx2 CDU) was used to propel fluids through the cold plate. A valve was used to precisely adjust the flow rate of the coolant entering the cold plate. A flowmeter was installed in the flow loop to measure the flow rate of coolant getting into the cold plate. Different inlet fluid temperatures, respectively, 24 °C, 34 °C, 39 °C, 47 °C, and 50 °C were able to be achieved using a ceramic heater attached to AHx2 which heats up the coolant. The different coolant temperature tested are representative of different ASHRAE water class server temperatures [15] $W2$, $W3$, $W4$, and $>W4$ with each has different set of implications like, performance or possible heat reuse for example. The hot coolant coming out of the cold plate was cooled by an air-liquid heat exchanger located in the CDU. While running test using cold plates, for the results discussed through this article, the air inlet temperature was maintained at about 34 °C around the DUT to cool down other server components like DIMMs and storage. CPU was overclocked carefully, and the core frequency, temperature, and voltage data were monitored using HWinfo64. Steady state was maintained before recording the data.

C. 2P Immersion Cooling

2P immersion heat sinks also known as *Boiler Plates* were used to overclock the CPU in an immersion system. The boiler plate as shown in Fig. 7 has boiler enhancement coating (BEC) applied on top which act as active nucleation sites. BEC dictates the amount of vapor bubble that can form, grow, and later detach from the boiling surface. Three different boiler plates namely sintered mesh copper (from Cooler Master), sintered powder copper (from Boyd Corporation) and copper pinfins with BEC (from 3M) were used while overclocking the CPUs. It was observed that the performance of pinfin boiler plates outdid the other two boiler plates during testing by yielding minimal junction temperature for the same operating CPU power. Hence, in the subsequent tests as reported in the article, pinfin boiler plates were used while stress testing the processor. The pinfin boiler plate was soldered on top of the IHS using Indium with values of thermal conductivity and thickness as 86 W/mK and 0.2 mm, respectively. It should be carefully noted that the higher thermal conductivity from Indium (TIM 2) will help in bringing down the overall thermal resistance compared to using solder paste or grease as TIM2 like what been used in the case of air cooling and cold plates in this article. The influence of TIM 2 on the overall performance will be stressed again in the results section.

DUT was immersed in a test tank filled with dielectric fluid HFE 7000 (and then later using FC 3284). The tank has a small recirculating pump which circulates the fluid through a filter to remove any contaminants. Tank also has a desiccant box which absorbs any moisture which could form during service and operation. The tank has a copper condenser located on top. The dielectric vapor bubbles rising from the heat generating component (overclockable CPUs in this test) hit the copper condenser located on top and the condensed liquid drips back into the tank. Water at room temperature was used as the condenser fluid with a dry cooler to reject the heat picked up from the tank. The CPU was overclocked carefully, and the core frequency, temperature and voltage data were monitored using HWinfo64. Steady state was maintained before recording the data.

V. Performance Comparison

As mentioned in Sections III and IV, the i9 9900k processor was carefully overclocked by tweaking its core frequency and core voltage. Sections V-A–V-C talk about the stable CPU frequencies and voltage that were achieved while overclocking using different coolant mediums and its corresponding impact on the CPU junction temperature and CPU power draw. A deep dive analysis of thermal resistance wraps up the section followed by recommendation to packaging industry and researchers/developers on how the performance could be maximized even further. Before getting into the results, the reader should bear in mind that the tested CPU was

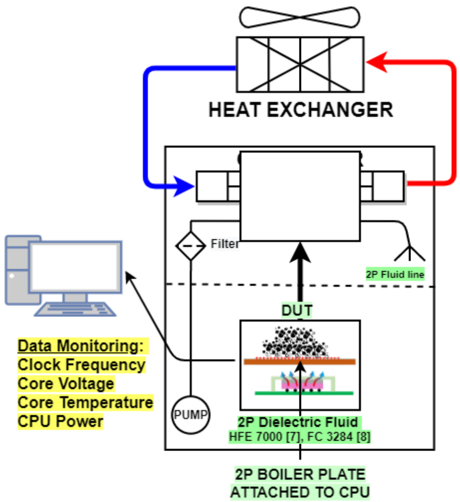


Fig. 8. Schematic of 2P immersion cooled setup.

purchased online, so there could be differences at the silicon level which could have an influence on the results. While it should also be carefully noted that the chosen or tested heat sinks may not represent the best-in-class solution as there could be plenty of room for improvement as revealed by the results later.

A. CPU Frequency Versus Voltage

Initially, tests were carried out to benchmark the performance of the CPU at its base frequency (GHz) using air-cooled heat sink. Later, tests were carried out at the intended test conditions captured in Table III. The primary axis in Fig. 9 shows the maximum stable core frequencies that were achieved. X-axis in Fig. 9 shows the coolant medium temperatures that were maintained across different test conditions. The secondary axis in Fig. 9 shows the corresponding core voltages that were maintained to achieve stable operating frequencies. Using air cooling at an inlet temperature of 34 °C, maximum stable frequencies of 3.3, 4, and 3.6 GHz were attained using Prime95, XTU, and Cinebench workloads, respectively. It should be noted from the figure that both cold plates and 2P immersion cooling technologies consistently produced more stable higher frequencies than what air cooling could offer, credits to the high heat-carrying capacity associated with those technologies. 2P immersion cooling using HFE 7000 (having a boiling point of 34 °C at 1 atm) yielded the maximum possible frequencies while subjecting the CPU to different workloads. Compared to air cooling, the relative core frequencies using 2P immersion cooling soared by 51%, 32%, and 44% while subjecting the CPUs to Prime95, XTU, and Cinebench workloads, respectively. In case of the cold plates, though the solution produced higher stable frequencies compared to air cooling, the values start to drop as the coolant temperature was increased up from 24 °C to 50 °C. It is also interesting to observe a striking similarity between the air-cooled test at 34 °C and cold plates tests at 50 °C which provides the best user case for possible waste heat reuse using cold plates.

Achieving stable frequencies comes at a price of rise in the core voltage too as seen in Fig. 9. The rise in CPU core voltage translates to a rise in the CPU core power as captured in the upcoming sections.

> *Takeaway 1:* Increase in CPU frequency is equivalent to improvement in CPU performance which could be realized as more VMs, auto-scaling and over subscription for example.

B. Junction Temperature Versus CPU Power

TDP and junction temperature of the CPU while operating at the base frequency is about 95 W and 95 °C. As the CPU was carefully overclocked, the associated power with CPU starts to rise as you can notice in Fig. 10 while the CPU still maintains a maximum allowable junction temperature of about 95 °C to ensure safe operation of the device. When the silicon junction temperature limit is exceeded, the device starts to throttle by jumping to a lower frequency rating. The primary axis in Fig. 10 shows the junction temperature for different workload conditions using varying cooling configurations. The secondary axis in Fig. 10 shows the measured CPU core power for achieving the stable frequencies. It is very clear from the Fig. 10 that both cold plates and 2P immersion technologies were comfortably dissipating high-power levels than what was feasible using air cooling. The cold plates operating at 34 °C yielded a rise in CPU power by 38%, 42%, and 48% compared to air cooling while running Prime95, XTU, and Cinebench workloads, respectively. Likewise, 2P immersion cooling using HFE 7000 yielded an increase in CPU power of 77%, 30%, and 46% using Prime95, XTU, and Cinebench workloads, respectively. A closer apples-to-apples comparison between the three different coolant configurations operating at 34 °C inlet temperature as indicated in the chart (marked with orange borders) reveal that 2P immersion yielded high-power levels while maintaining the least junction temperature. This result is especially important because reduced junction temperature means improved CPU reliability/life. This is due to the low boiling point of HFE 7000 and partially attributed to the 2P heat-transfer mechanism as well. While it is also important to note that moving to a higher boiling point dielectric fluid like FC3284 would result in a higher junction temperature while delivering similar or degraded CPU performance as shown by the lone bar in Fig. 10.

The majority of conclusions from the article on 2P immersion pertain to using HFE 7000. Section V-C tries to explain why and how 2P immersion using HFE 7000 was able to deliver better performance by estimating thermal resistance and their associated heat-transfer coefficients.

> *Takeaway 2:* Two phase immersion using HFE 7000 with its low boiling point and by innate nature of high heat carrying capacity helps in dissipating larger heat levels with the least junction temperature. This means higher performance while at the same time improved lifetime or reliability. Lower junction temperature would also imply lower leakage current.

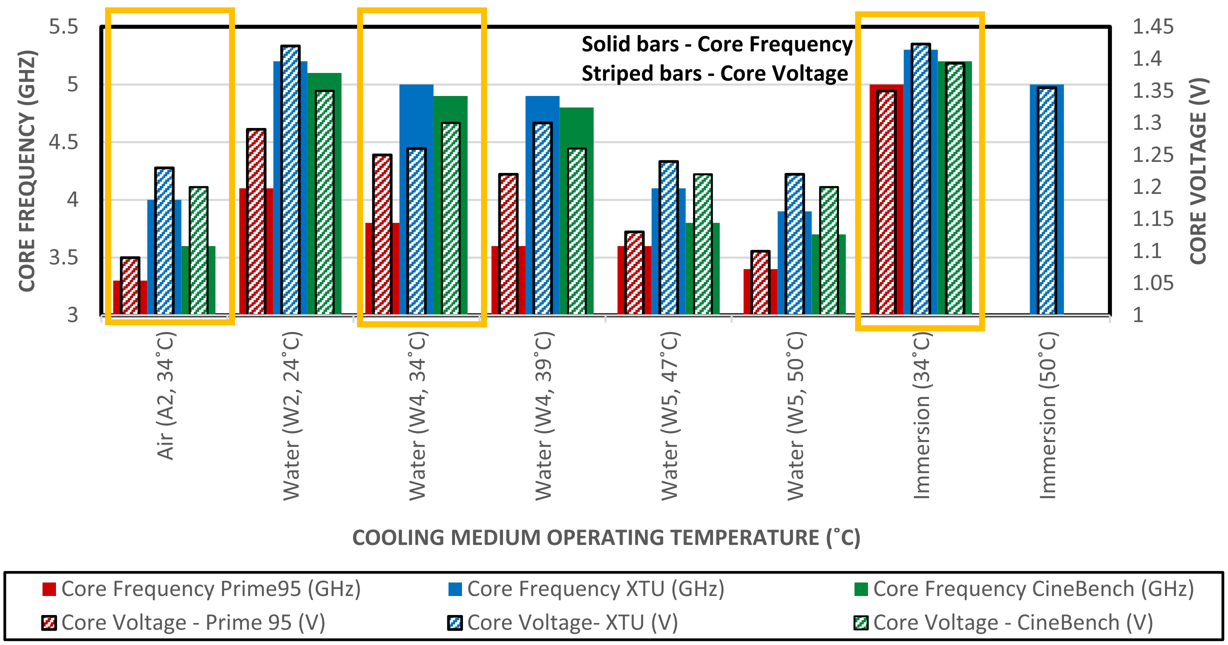


Fig. 9. CPU core frequency (GHz) (solid bars in the chart) versus CPU core voltage (V) (striped bars in the chart) for all the cooling configuration while stressing the CPU with different workloads, respectively; Prime95, XTU, and Cinebench.

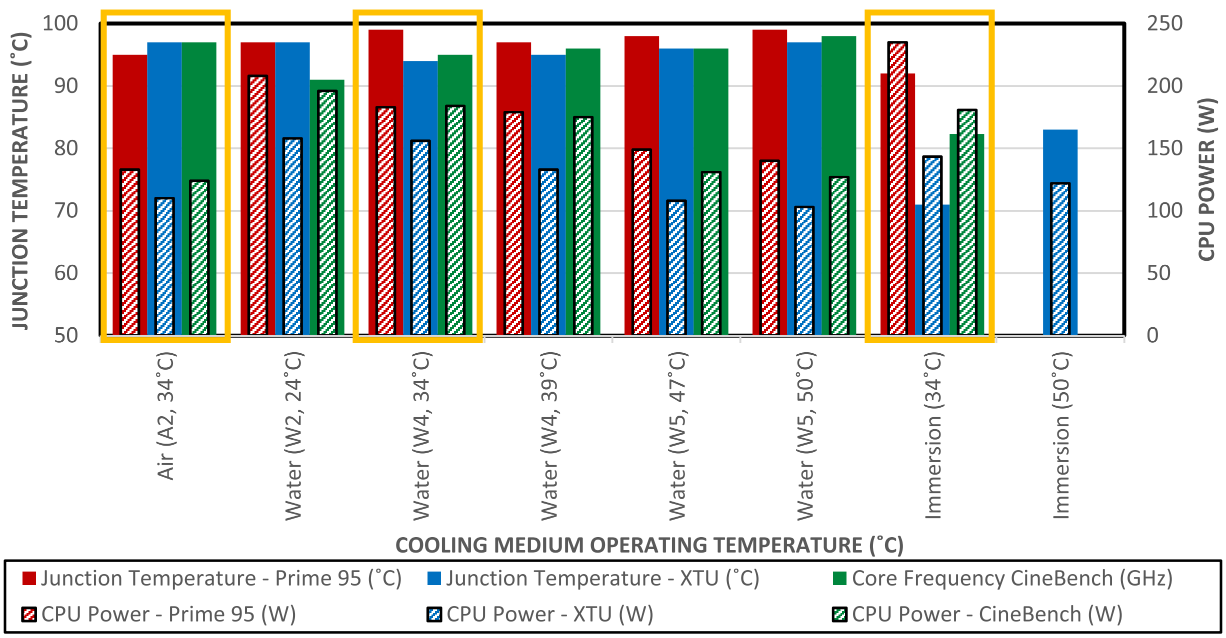


Fig. 10. CPU Junction temperature (solid bars in the chart) versus CPU power (striped bars in the chart) for all the cooling configuration while stressing the CPU with different workloads, respectively; Prime95, XTU, and Cinebench.

C. Thermal Resistance

$$R_{\text{th}} = \frac{(T_j - T_a)}{Q_{\text{cpu}}}(^\circ\text{C/W}). \tag{1}$$

Thermal resistance as given in (1), R_{th} is defined as the ratio of difference between Junction (T_j) and Ambient temperature (T_a) to the power dissipated from the CPU (Q_{cpu}). Estimating the thermal resistance is an important property for an electronic thermal/mechanical engineer as it captures how much the object or material (CPU in our case) resist the heat flow for the given temperature difference. Fig. 11 shows the estimated thermal resistance for different cooling configurations when subjected to variable workloads, namely, prime95, XTU and Cinebench, respectively. The thermal resistance value drops for cold plates and 2P immersion indicating reduced hindrance to heat transfer from the CPU to the ambient fluid. This is attributed to the efficiency in the heat-transfer medium compared to air cooling. Air cooling at an inlet temperature of

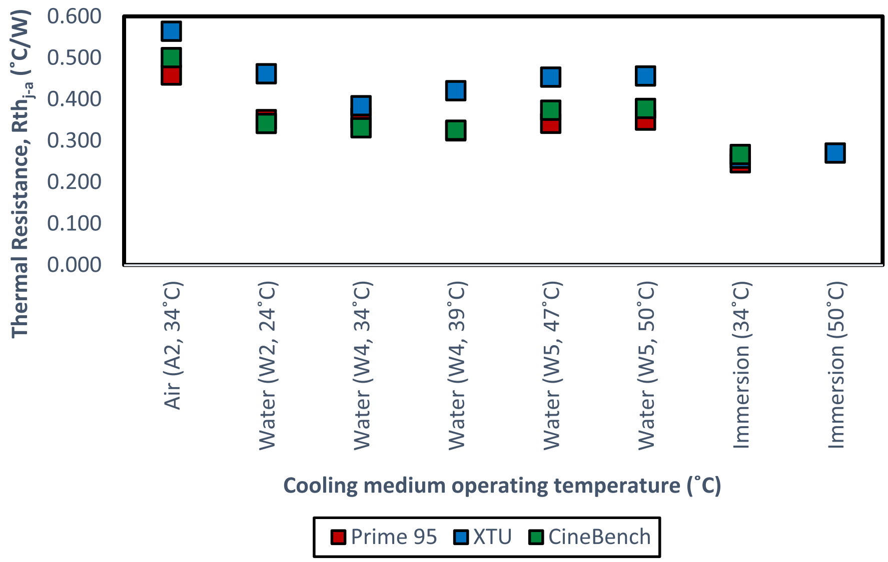


Fig. 11. Measured thermal resistance (°C/W) for all the cooling configuration while stressing the CPU with different workloads, respectively; Prime95, XTU, and Cinebench.

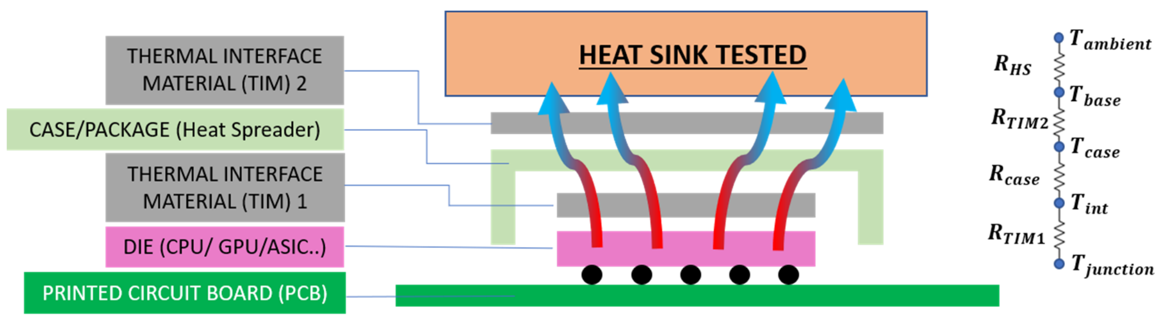


Fig. 12. Schematic of the thermal resistance stack in a package exposing layers of physical impedance going all the way from die to the heat sink on top.

34 °C resulted in maximum thermal resistance of 0.56 °C/W. Cold plates resulted in thermal resistance values in the range of 0.35–0.45 °C/W between different inlet coolant temperature and workload configuration. 2P immersion resulted in the least thermal resistance of about 0.25 °C/W and hence was able to deliver both maximum CPU power and reduced junction temperature. It was also important to realize that the thermal resistance did not change with respect to changes in workloads indicating a more stable CPU operation compared to cold plates and air. This narrower spread in the thermal resistance using 2P cooling could also be ascribed to its ability to carry high heat loads with lower junction temperature. It should again be carefully noted in the case of 2P immersion scenario that the TIM2 resistance coming from Indium is an order of magnitude lower than the TIM2 resistance coming from thermal grease used in the case of air cooled or cold plate tests reported in this article. Please refer to Fig. 13 for identifying the differences in thermal resistance and hence the estimated heat-transfer coefficient.

A deeper dive into the thermal resistance is needed to isolate the different elements that contribute or hinder heat transfer originating from die to the ambient fluid. As shown in the schematic in Fig. 12, there are physical layers originating from the die all the way to the heat sinks on top. The properties of different layers which includes material thickness, size, and thermal conductivity are provided in Fig. 1. Thus, the total thermal resistance is a combination of conduction resistance from the TIM1, TIM2, and IHS, in addition to convection resistance from the heat sink. Thus

$$R_{\text{th}} = R_{\text{TIM1}} + R_{\text{IHS}} + R_{\text{TIM2}} + R_{\text{hs}}. \quad (2)$$

With the knowledge of known material properties and the measured CPU junction temperatures and CPU power, the resistance from different layers were estimated according to

$$R_{\text{TIM}} = \frac{(t_{\text{TIM}})}{k_{\text{TIM}} \cdot A} \quad (3)$$

$$R_{\text{IHS}} = \frac{(t_{\text{IHS}})}{k_{\text{IHS}} \cdot A} \quad (4)$$

$$R_{\text{hs}} = \frac{1}{h_{\text{hs}} \cdot A}. \quad (5)$$

D. Thermal Resistance Breakdown

The primary axis in Fig. 13 shows the thermal resistances of individual layers contributing to the total resistance. Prime

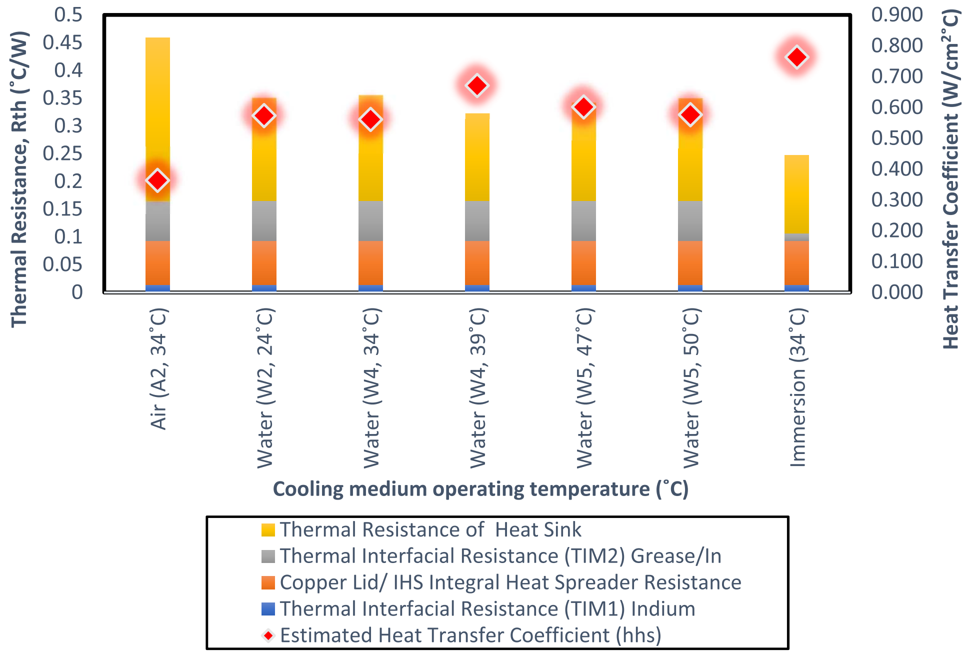


Fig. 13. Thermal resistance breakdown (°C/W) and heat-transfer coefficient estimation (W/cm^2 °C) for all the cooling configuration while stressing the CPU using Prime95.

95 workload was chosen to plot this chart as the associated heat dissipation is higher than XTU, or Cinebench workloads. It could be observed that the TIM1 and IHS resistances were not influenced by the cooling configuration. It should be noted that the TIM2 resistance associated with 2P immersion cooling is lower as Indium was used instead of thermal grease. The heat sink thermal resistance as indicated by the yellow bars in Fig. 13 shows that cold plates and 2P boiler plates offered lower thermal resistance than air cooling. The heat sink thermal resistance is not entirely convection resistance but includes a fair contribution of conduction, caloric, and spreading resistance within the heat sink. Typically, it is a rule of thumb to consider the spreading resistance to be one-third of the total resistance. The heat-transfer coefficient in the secondary axis in Fig. 13 was estimated according to (5) and the heat-transfer coefficient associated with cold plates and 2P immersion are much higher than what air cooling could offer. The heat-transfer coefficient value of 0.76 W/cm^2 °C associated with 2P immersion cooling corroborated well with [16]. Again, for apples-to-apples comparison at 34 °C cooling medium temperature, cold plates and boiler plates offered 55% and 110% rise in heat-transfer coefficient, respectively. It should be carefully noted that using a high conductive TIM2, the total thermal resistance of the cold plates can be reduced or in other words the heat-transfer coefficient of cold plates can be enhanced further. Similarly, using FC 3284 having a saturation temperature of 50 °C, a heat-transfer coefficient value of 0.65 W/cm^2 °C was achieved while running XTU workloads. The vigorous bubble sweeping motion along with the latent heat of vaporization associated with 2P immersion cooling results in such a rise in heat-transfer efficiency.

Takeaway 3: Cold plates and two phase immersion have higher heat transfer coefficient when compared to air. This results in a lower thermal resistance to heat transfer. Meaning it can dissipate more heat with minimum rise in junction temperature.

E. Recommendations Into the Future

All the above outcomes suggest that both 2P immersion and cold plates result in greater CPU performance compared to that of air cooling. Having said that

1) The CPU performance in air cooled environment can be enhanced further by increasing the heat sink fin density or/and by increasing the airflow flow through the fins which may or may not come with a penalty to the pumping power.
2) In the case of *boiler plates*, further efficiency gains can be obtained by further optimizing the boiler plate geometry and design. As pointed out in [17] research should focus in developing BEC coatings which aim toward improving the critical heat flux (CHF) of high-density components and also aims at reducing the onset of nucleate boiling (ONB) from not-so-hot components or in other words having a minimal superheat. While at the CPU package level, high thermal conducting TIMs which can sustain large thermal cycles can be developed in reducing the interfacial resistance. Even rigorous approach of applying the BEC directly on top of IHS/Case can yield even lower thermal resistances by eradicating layers of TIM and metallic tiers.

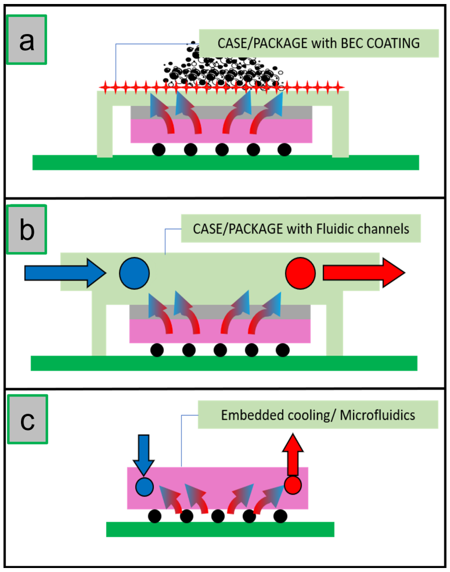


Fig. 14. Schematic of modified package designs to gain more thermal advantage using liquid cooling. (a) IHS/case with BEC coating. (b) Package provided with microfluidic channels. (c) Embedded cooling with die containing etched channels.

The schematic of such a package will be as shown in Fig. 14(a).

3) In the case of *cold plates*, further performance gains can be achieved by increasing the flow rate, increasing the fin density or by adopting jet impingement without compromising the pumping power expenditure. Other avenues could be looking at 2P coolants instead of water. Even rigorous approach would be to eliminate a layer of TIM2 and have the fluid flow through the package lid as shown in Fig. 14(b). Packaging and semiconductor industries ought to bring these modifications either at the manufacturing or supply chain level to combat the decline of Moore's law.
4) Printing on *top of Silicon* or as shown in Fig. 14(c), *embedded cooling* of 2.5-D/3-D chip architectures using dielectrics would be the ultimatum.

VI. Conclusion

Historically, with the advancements in packaging and semiconductor manufacturing industry, chip performance hence chip power has constantly been on the rise. Air cooling of high-density processors is becoming increasingly inefficient and expensive. Therefore, liquid cooling using cold plates and 2P immersion using boiler plates were used to observe the performance of a commercially available overclockable CPU. It was observed that both 2P immersion and cold plates yielded the maximum CPU performance compared to air cooling by constantly operating at stable higher frequencies above 5 GHz. At a data center level, this rise in stable frequency can be realized and monetized in many ways like server virtualization or oversubscription. The junction temperatures associated with 2P immersion is lower than what was observed using cold plates and air thus enabling lower leakage current and higher device lifetime.

In terms of pure thermal performance, at the maximum workload accomplished, the heat-transfer coefficient associated with 2P immersion using HFE 7000 along with an enhanced TIM2 of Indium was estimated to be 110% greater than what was achieved using air cooling and 50% greater than cold plates. This increased heat-transfer coefficient and the lower saturation temperature of HFE 7000 along with an enhanced TIM2 results in the lowest junction to ambient thermal resistance of 0.25 °C/W which indirectly translates to higher CPU performance in commercial terms. For example, the CPU performance using cold plates can also be improved substantially using an enhanced TIM2 such as Indium.

Below is a summary of observations and remarks regarding the distinct cooling technologies and chip performances based on the current study.

1) Liquid cooling enables higher performance on an unlocked overclockable desktop part and hence the server parts when unlocked can also yield greater performance along with further efficiency gains at the system level.
2) Both cold plates and immersion are important technologies that enable next generation chips and performance. Embracing such technologies and others by chip manufacturers, data center operators, and service providers will accelerate their adoption, optimization and reduce associated cost.
3) The efficacy of the aircooled solution could be increased by increasing airflow, heat sink fin density, enhanced heat pipes and/or vapor chambers (VCs). All of that comes at the expense of size, power, cost, and engineering overhead.
4) The efficacy of the cold plate solution could be increased by increasing the water flow, fin density, impingement flow scheme, cold plate integration into IHS, or with an enhanced TIM.
5) The efficacy of the 2P immersion solution could be increased by better BEC metallurgies and structure, better spreading with VCs, printing BEC directly on the IHS or all the way on top of the flip-chip.
6) Typical thermal resistances of cold plates and 2P immersion are comparable for well-optimized systems.
7) Each technology has its own unique characteristics. Cold plates are a quick enabler for high-power chips for hybrid cooled systems, while 2P immersion could enable disruptive 3-D packaging and architecture.
8) Testing with actual chip-silicon with HPC like workloads can reveal thermal challenges such as high localized heat flux at the cores tile. Such findings are difficult to ascertain using TTVs. There is a margin of error when running tests using off the shelf CPU parts. This is mainly governed by the allowed silicon quality variation by the chip supplier.
9) Liquid cooling can help chip manufacturers to bring overclocking capabilities to server class parts. Frequency boosts beyond turbo can be enabled in reliable fashion.
10) The thermal challenge lies in heat flux (W/cm^2) rather than TDP (W). Going to a future that uses chiplets

should produce manageable flux values even when the TDP increases. Note that some of the newer high-power GPUs are lower in flux when compared to lower TDP CPUs but with higher heat flux.

11) Creating an ecosystem for liquid cooling with a mature supply chain is essential for this technology to scale.

Acknowledgment

This article was prepared in the memory of late Dr. Avaram Bar Cohen who was a pioneer of electronics and data center thermal management. This article elucidates one of his favorite topics "2P heat transfer" for high-power electronics. The authors would like to dedicate this article to him for nurturing so many students/professionals and for contributing so much to the industry and academia.

The authors would like to sincerely extend many thanks to Austin Agar and Cam Turner (from CoolIT Systems) and Phil Tuma and Thomas Crandall (from 3M). They also like to acknowledge Microsoft Azure HPC team.

References

[1] A. Bar-Cohen, "Gen 3 'embedded' cooling: Key enabler for energy efficient data centers," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 7, no. 8, pp. 1206–1211, Aug. 2017.

[2] S. H. Fuller and L. I. Millett, *The Future of Computing Performance: Game Over or Next Level?*. Washington, DC, USA: National Academy Press, 2011.

[3] *Japanese Supercomputer: K Computer—Fujitsu Global*, Fujitsu, Japan, 2017.

[4] *US Supercomputer: Oak Ridge 'Jaguar' Supercomputer is World's Fastest | ORNL*, CRAY, USA, 2009.

[5] M. J. Ellsworth, G. F. Goth, R. J. Zoodsma, A. Arvelo, L. A. Campbell, and W. J. Anderl, "An overview of the IBM power 775 supercomputer water cooling system," *J. Electron. Packag.*, vol. 134, no. 2, Jun. 2012, Art. no. 020906.

[6] *Smaller and Faster: The Cray-2 and 3—CHM Revolution (Computer-history.Org)*, CRAY, USA, 1985.

[7] *3M Novec 7000 Engineered Fluid | 3M United States*.

[8] *3M Fluorinert Electronic Liquid FC-3284 | 3M United States*.

[9] *Overclock Your CPU With Unlocked Intel Processors*, Intel, USA.

[10] M. Jalili *et al.*, "Cost-efficient overclocking in immersion-cooled datacenters," in *Proc. ACM/IEEE 48th Annu. Int. Symp. Comput. Archit. (ISCA)*, Jun. 2021, pp. 623–636, doi: 10.1109/ISCA52012.2021.00055.

[11] B. Ramakrishnan, Y. Hadad, S. Alkharabsheh, P. R. Chiarot, and B. Sammakia, "Thermal analysis of cold plate for direct liquid cooling of high performance servers," *J. Electron. Packag.*, vol. 141, no. 4, Dec. 2019, Art. no. 041005.

[12] H. Bostanci, M. A. Shareef, and P. E. Tuma, "Immersion cooled ARM-based computer clusters towards low-cost high-performance computing," in *Proc. 19th IEEE Intersociety Conf. Thermal Thermomechanical Phenomena Electron. Syst. (ITherm)*, Orlando, FL, USA, Jul. 2020, pp. 450–456.

[13] P. E. Tuma, "Design considerations relating to non-thermal aspects of passive 2-phase immersion cooling," in *Proc. 27th Annu. IEEE Semiconductor Thermal Meas. Manage. Symp.*, San Jose, CA, USA, Mar. 2011, pp. 1–9.

[14] (2016). *ASHRAE Air-Cooled Server Requirements*. [Online]. Available: https://ASHRAE_TC0909_Power_White_Paper_22_June_2016_REVISED.pdf

[15] (2019). *ASHRAE Water Cooled Server Requirements*. [Online]. Available: https://tc0909.ashraetcs.org/documents/ASHRAE_TC0909_Water_Cooled_Servers_15_Oct_2019.pdf

[16] I. Mudawar, "Assessment of high-heat-flux thermal management schemes," *IEEE Trans. Compon. Packag. Technol.*, vol. 24, no. 2, pp. 122–141, Jun. 2001.

[17] S. V. Garimella, T. Persoons, J. A. Weibel, and V. Gektin, "Electronics thermal management in information and communications technologies: Challenges and future directions," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 7, no. 8, pp. 1191–1205, Aug. 2017, doi: 10.1109/TCPMT.2016.2603600.

Bharath Ramakrishnan received the master's degree in mechanical engineering (ME) from CIT, Coimbatore, India in 2010, the M.S. degree in ME from Auburn University, Auburn, AL, USA, in 2014, and the Ph.D. degree in ME from Binghamton University, Binghamton, NY, USA, in 2019.

He is currently a Mechanical Engineer working with the Microsoft CO&I's Advanced Development Team, Redmond, WA, USA. His research background is on data center thermal management with an emphasis on all forms of high-density liquid cooling techniques including two phase immersion and microchannel cooling. His recent research interests are into exploring advanced cooling for next generation IC packages at a microlevel and into developing data center waste heat recovery and direct air capture systems at a macrolevel.

Husam Alissa is a Principal Engineer and Technical Lead with the Microsoft's Data Center Advanced Development Team, Redmond, WA, USA. His focus areas include systems (chip-server-data center), cooling (air, direct to chip, immersion, and cryogenics), performance, architecture, reliability, efficiency, sustainability, and TCO, with more than 50 publications in these fields.

Mr. Alissa is a member of ASHRAE TC9.9, ASME, OpenCompute, and iMasons. His work has received many recognitions including NewYork State Assembly Early Career Achievement, the ASME InterPACK Outstanding Paper Award, the IEEE TCPMT Electronics Packaging Society Best Paper Award, and the S3IP Distinguished Doctorate Dissertation Award.

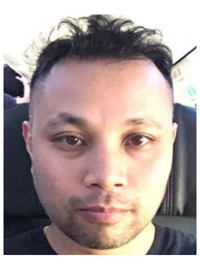
Ioannis Manousakis received the B.Sc. and M.Sc. degrees in computer science from the University of Crete, Crete, Greece, both in 2010, and the Ph.D. degree in computer science from Rutgers University, New Brunswick, NJ, USA, in 2018.

He is currently a Principal Software Engineer at Microsoft Azure, Redmond, WA, USA. He is currently working on increasing the efficiency of Azure by co-optimizing the datacenter, server, and software architectures. His research interests include cost-efficient datacenter and systems architectures, and risk-aware computer systems.

Robert Lankston is a Principal Mechanical Engineering Manager with the Microsoft's Cloud Hardware Engineering Team, Redmond, WA, USA, with a focus on thermal design of servers. At Microsoft, he contributed to the thermal design and architecture of the Olympus infrastructure as well as many of the server platforms that integrate into it.

Ricardo Bianchini (Fellow, IEEE) was a Professor of computer science at Rutgers University, New Brunswick, NJ, USA. He is currently a Distinguished Engineer at Microsoft, where he leads efforts to improve the efficiency of the company's online services and datacenters. He also manages the Systems Research Group at Microsoft Research, Redmond, WA, USA. He has authored nine award articles.

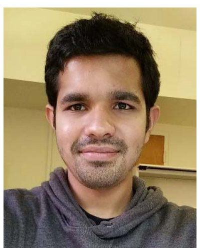
Mr. Bianchini received the CAREER Award from the National Science Foundation. He has given several conference keynote talks and served on numerous program committees, including as Program Co-Chair of ASPLOS, EuroSys, and ICDCS. He is an ACM Fellow.



Washington Kim was working on signal integrity testing and IoT. He is currently a Design Verification Engineer at Microsoft in CO&I's Advanced Development, Redmond, WA, USA. His background was in hardware testing. He is currently working on different types of liquid cooling technologies to find more efficient ways for cooling servers and data centers.

Rich Baca received the Ph.D. degree from the University of Illinois Urbana-Champaign, Champaign, IL, USA, in 1996.

He is currently a Principal Optics Hardware Engineer at Microsoft, Redmond, WA, USA, working on low-cost server optics and optical connectivity for the data center. Mostly recently, he has been involved with two-phase immersion cooling along with optical solutions for efficiently transmitting optical signals across the tank interface. He has over 30 years of experience working with electromagnetics/laser simulation, microwave photonics and MMIC design, high-speed digital optical design working at various companies such as Texas Instruments, Raytheon, Finisar, Tellabs, CommScope, and other startups along the way. His research interests are in optics for the data center including novel server optics solutions.



Pulkit A. Misra received the Ph.D. degree from Duke University, Durham, NC, USA, in 2019.

He is currently a Research Engineer at Microsoft Research, Redmond, WA, USA. His current research is on improving the power efficiency of datacenters.

Inigo Goiri received the Ph.D. degree from the University Politecnica de Catalunya (UPC), Barcelona, Spain, in 2010.

He is currently a Research Software Developer at Microsoft Research, Redmond, WA, USA. His current research focuses on the efficiency of large-scale distributed systems.

Majid Jalili received the B.Sc. degree from the Shahid Bahonar University of Kerman, Kerman, Iran, in 2010, and the M.Sc. degree from the Sharif University of Technology, Tehran, Iran, in 2013. He is currently pursuing the Ph.D. degree with the LPH Group, University of Texas at Austin, Austin, TX, USA.

His current research interests include memory systems, multicore, and datacenters.

Ashish Raniwala received the B.Tech. degree in computer science and engineering from IIT Delhi, Delhi, India, in 1998, and the Ph.D. degree in computer science from Stony Brook University, Stony Brook, NY, USA, in 2009. His Ph.D. dissertation focused on enterprise-class multihop wireless backbone networks and his work won IEEE Infocom Test of Time Award.

He worked on inference and training infrastructure for Azure Machine Learning, and on Bing's Engineering Infrastructure. He worked at Google on Flume that is now the de facto interface for writing data processing pipelines at Google and formed the basis of GCP's DataFlow offering. He is currently a Partner Architect with Microsoft's Azure organization, Redmond, WA, USA, where he is leading the two-phase liquid immersion cooling program.

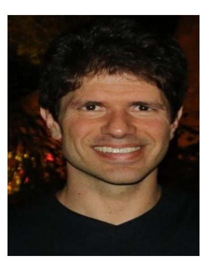
Brijesh Warrier focused on leveraging software and ML to drive efficient use of resources in the datacenter, driving toward a fully software-defined datacenter. He has built and delivered large-scale innovative solutions around predictive hardware failure and mitigation, system for safely oversubscribing data center power, and machine learning for systems solution. He is currently a Partner Engineering Manager at Microsoft, Redmond, WA, USA, working in Azure core.

Mark Monroe received the M.Sc. degree in mechanical engineering from Virginia Tech, Blacksburg, VA, USA, in 1982.

He is currently a Principal Engineer with the Microsoft's Datacenter Advanced Development Group, Redmond, WA, USA, working on future systems for energy, water conservation, and automation. He is a Six Sigma Master Black Belt, with a focus on change acceleration in corporate settings. He led industry consortium including Infrastructure Masons and The Green Grid, and served on the board of directors for the Center for ReSource Conservation, Boulder, CO, USA.

Christian Belady is a seasoned veteran in managing Datacenter/Infrastructure Development at global scale. He has developed one of the largest data center footprints in the world. He is currently a Distinguished Engineer and the Vice President of Cloud Infrastructure Strategy and Architecture for Microsoft's Cloud Infrastructure Organization, Redmond, WA, USA. In recent years, he has been responsible for driving the strategy and delivery of server and facility development worldwide, including research, engineering, construction, and operations for Microsoft's data center portfolio. He has authored dozens of articles, is frequently quoted in the press, and is a featured speaker on data center trends in the industry. Since the late 1990s, his publications have focused on data centers and the industry's need for engineering efficient computing environments.

Mark Shaw (Member, IEEE) is a Partner Architect, a Manager, and a Developer of server hardware for Microsoft's Cloud + Enterprise Division, Redmond, WA, USA. He and his team are working with the industry on developing the next set of server hardware.



Marcus Fontoura (Senior Member, IEEE) worked on the production infrastructure for Bing and in several Bing Ads projects. He worked as Staff Research Scientist and Researcher in Google, Yahoo, and the IBM Almaden Research Center. He is currently a Technical Fellow at Microsoft, Redmond, WA, USA, where he works as a Chief Architect for Azure compute.

Mr. Fontoura is an ACM Distinguished Member. He has more than 25 issued patents (and many others filed) and more than 50 published articles. He has been serving in several program committees over the years, including SIGIR, WWW, WSDM, KDD, and CIKM. You can find more about him at his personal website Fontoura.org.