

ASSESSMENT OF RELIABILITY ENHANCEMENT IN HIGH-POWER CPUs AND GPUS USING DYNAMIC DIRECT-TO-CHIP LIQUID COOLING

**Pardeep Shahi,* Amith Mathew, Akash, Satyam Saini,
Pratik Bansode, Rajesh Kasukurthy, & Dereje Agonafer**

*Mechanical and Aerospace Engineering Department, The University of Texas at
Arlington, 500 W 1st St., Arlington, Texas 76019, USA*

*Address all correspondence to: Pardeep Shahi, Mechanical and Aerospace Engineering
Department, The University of Texas at Arlington, 500 W 1st St., Arlington, Texas 76019,
USA, E-mail: pardeep.shahi@mavs.uta.edu

Original Manuscript Submitted: 3/20/2022; Final Draft Received: 5/4/2022

Thermal management of high-performance computing servers is becoming a prevalent challenge for the data-center cooling industry due to increasing power densities at the server level and data-center level. Efficient heat dissipation is also directly related to electronic package reliability. Improved cooling technologies like direct-to-chip liquid cooling can address the rising cooling demands due to the higher thermal performance of water-based coolants. A methodology to further enhance the efficiency of direct liquid cooling (DLC) is experimentally investigated using the concept of dynamic cooling. A flow control device (FCD) was developed to regulate the flow rate to four custom-made thermal test vehicles (TTV) using ceramic heaters. The TTV assembly was placed at four different levels in a standard 19-inch information technology equipment (ITE) rack in test fixtures mounted with cold plates. The flow regulation to each of the TTVs was done based on the power dissipated by each TTV. The power dissipation of each TTV was varied for various non-homogeneous power-distribution values in the entire rack. The influence of coolant inlet temperature and flow rate on the TTV temperature and rack pressure drops was analyzed. The results indicated more uniform temperatures on the TTVs and a reduction in the maximum temperature on the TTV with maximum power. The impact of the temperature uniformity on package level reliability was also analyzed by comparing the results obtained with the published literature.

KEY WORDS: *data center, liquid cooling, electronics cooling, cold plate*

1. INTRODUCTION

Because power consumption for conventional air-cooled data centers usually accounts for 30–40% of the total energy consumption of the data center (Iyengar, 2010), the most optimistic trends in the recent survey show the increasing power consumption (Shehabi et al., 2016) due to inefficiencies in the cooling systems and report an average power usage effectiveness value of around 1.8 (Stansberry and Kudritzki, 2013). Several thermal management technologies are available in the market to reduce data-center inefficiencies (Alkharabsheh et al., 2015; Chainer et al., 2017). The limitations of air-cooling technologies make direct liquid cooling (DLC) for

high-performance computing servers, especially for GPU servers, a more efficient cooling technology. The developments in liquid cooling have been, in the past, thwarted due to significant capital associated with and potential risks related to coolant leakage, restricting its wide implementation in data centers (Beaty, 2004; Ellsworth et al., 2008; Helsin, 2014). But, increasing power densities have again sparked an interest in direct-to-chip liquid cooling as compared to efficient air-cooling techniques like contained aisles, in-row cooling, rear-door heat exchangers, etc. This is due to the requirement of significant facility infrastructure upgrades for the aforementioned air-cooling technologies.

Heat dissipation (W/cm^2) capability is significantly larger for liquid cooling than comparative air-cooling techniques because of the large thermal mass and smaller fluid transport energy requirements (Patterson et al., 2016). Recent investigations point out that implementing liquid cooling can save 45% of energy as compared to hot/cold aisle containment-type air-cooled data centers (Ellsworth and Iyenger, 2009). Liquid cooling is quite different from air cooling where cold plates are used instead of heat sinks and pumps replace the fans for the circulation of the cooling fluid. There has been a great deal of interest in liquid-cooling technology since the 1990s. With the innovations in complementary metal-oxide-semiconductor (CMOS) transistor technology, the power densities again drastically reduced heat dissipation, and once again the use of air cooling became economically more viable. But, from the last few years, heat dissipation from the processing components is increasing causing air cooling to reach its cooling limits (Chu et al., 2004; Ellsworth et al., 2008; McFarlane, 2012; Schmidt, 2005), shifting the focus towards liquid cooling.

Reviewed literature suggests that for a typical data center, the average CPU utilization remains below 50% (Barroso and Hölzle, 2007) with a mean value of 36.44% with 95% confidence. A traditional liquid-cooled data center operates at constant coolant flow rates irrespective of the IT load in each server. These excessive flow rates lead to higher pumping-power consumption as well. Similarly, at the component level, there is an uneven distribution of power on each core in multi-core processors and GPUs causing a temperature gradient across the processors. The present investigation experimentally determines the increased reliability of processors due to the selective distribution of coolant at rack level and is termed rack-level dynamic cooling. This problem was addressed in the past by proposing a self-regulated flow control device (FCD) (Kasukurthy et al., 2018). In the present research, a novel FCD is used to dynamically vary the flow rates for a rack-level experimental setup.

The junction temperature of microelectronic devices is directly proportional to the device's reliability (Kumar et al., 2018; Lall et al., 1997). Microelectronics reliability and failure rates can be attributed to several failure mechanisms. These failure mechanisms can be due to interconnect failure, corrosion, electrochemical migration of metals, thermal stresses, etc. A common parameter in the acceleration of all these failure modes is a temperature gradient, temperature cycling, or rate change of temperature (Lall et al., 1995). The Arrhenius model has been typically used to predict the mean time to failure for semiconductor devices, using the relation given by (Shahi et al., 2021):

$$t_{f,\text{mean}} = t_{f,\text{mean,ref}} e^{-E_a/(kT)}, \quad (1)$$

where $t_{f,\text{mean}}$ is mean time to failure in hours, $t_{f,\text{mean,ref}}$ is time to failure at a reference temperature, k is Boltzmann's constant, T is temperature in kelvin, and E_a is activation energy (eV) of the failure mechanism.

The motivation behind the present investigation arises from the fact that there is a scope of further enhancing the reliability of processors at higher utilization using a targeted coolant

flow rate for each server depending on the operational IT load. The present work experimentally analyzes the concept of dynamic cooling at rack level using novel FCD. The lab-based experimental setup was built to mimic heat dissipation by four high-power servers in a rack using a custom-made high-power heater setup. Each of these four heater arrangements, placed at different elevations in a rack, dissipates nonuniform heat, depicting varying IT load in the servers in a typical rack. A commercially used cold plate is placed on top of these heaters connected to the coolant manifold. At each of the four levels in the rack, the coolant loop has flow sensors, pressure sensors, and thermocouples placed at the inlet and outlet of the cold plate. The outlet of each cold plate has the FCD dynamically causing a variation of the coolant flow rate based on the coolant outlet temperature.

2. EXPERIMENTAL SETUP AND METHODOLOGY

2.1 Experimental Setup

A polycarbonate fixture was designed and fabricated in the lab for these experiments as shown in Fig. 1. The test fixture consists of two plates, a bottom plate, and a top plate. The heater was fixed on the bottom plate followed by thermal interface material (TIM) and cold plate and the top plate was used to put the required pressure on the cold plate using hex nuts as shown in Fig. 2. The heater and its bottom plate which acts as insulation is denoted as a thermal test vehicle (TTV) and the entire cold-plate heater fixture assembly represents a single server. A schematic of the entire test fixture and assembly along with the instrumentation used for each TTV is shown in Fig. 3.

The heater used in these experiments to design the mock package/TTV was a Watlow ceramic heater with dimensions of 50.4×50.4 mm. These heaters were specifically chosen as they are readily manufactured with an embedded K-type thermocouple and can produce a total power of up to 2000 W. The power given to these heaters for the present study was up to approximately 480 W due to the power supply and chiller constraints. The whole fixture of the cold plate and the heater assembly was kept in place using two polycarbonate slabs and screws. A thin layer of TIM (Arctic MX-4 TIM) with thermal conductivity of 8.5 W/mK is applied between the cold plate and the heater, ensuring good thermal coupling. In the series of experiments carried out, the heater was provided with two power sources; when high-power outputs were required, it was connected to 120 V AC voltage and when regulated power outputs were needed, it was connected to an external DC power supply (make: Agilent, N5749A, 100 V/7.5 A, 750 W).

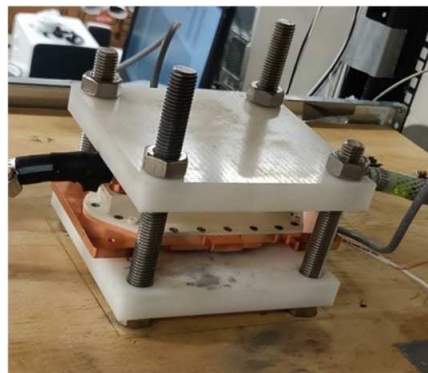


FIG. 1: Test fixture designed for housing the TTV assembly and cold plate used for the experiment

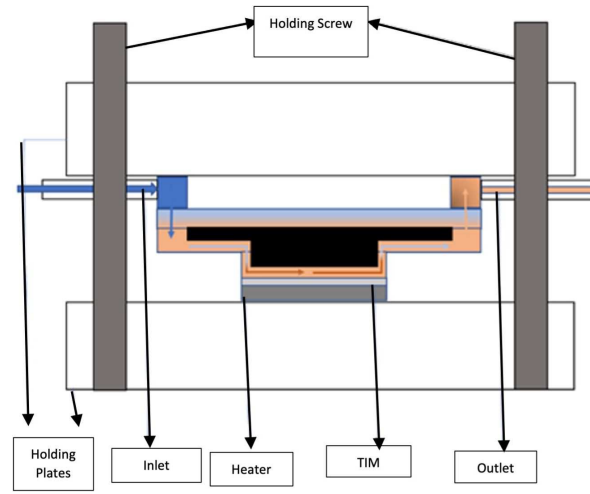


FIG. 2: Schematic of the TTV and cold-plate assembly in the fixture

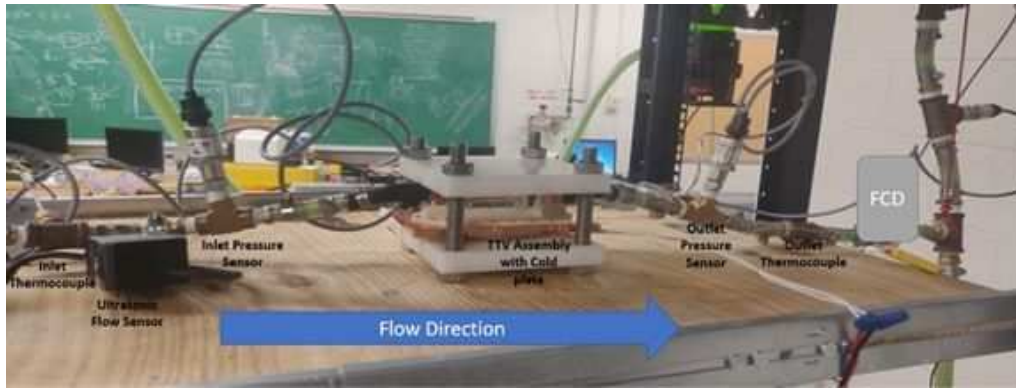


FIG. 3: Experimental arrangement showing the cold plates, test fixture, pressure and flow sensors, and FCD

Each cooling loop as shown in Fig. 3 in the rack was connected to an inlet and an outlet manifold through which the coolant entered and left the rack. The rack consisted of a flow meter (Keyence FDX-A1), inlet and outlet pressure sensors (Honeywell pressure sensors), inlet and outlet thermocouples (K-type thermocouple), and an FCD to control the flow rate within the cooling loop or server. The specifications of the above-mentioned components are listed in Table 1. The standard 42U rack used in the experiment was populated with four cooling loops with temperature and pressure sensors at the inlet and outlet manifold. The pump used was a Thermaltake Pacific DIY LCS PR22-D5, with an operating voltage of 8–12 V DC, providing 3.8 lpm to the rack. The 2 kW chiller was used to stabilize the inlet temperature of the coolant to servers at 35, 40, and 45°C. A schematic of the arrangement of the cooling loops and sensors in the rack is shown in Fig. 4. All the pressure sensors used in this setup as well as the thermocouples used were directly connected to the data acquisition unit (DAQ). A LabVIEW program was made for the collection of data in real time. A 30% inhibited ethylene glycol and deionized water solution was used as a coolant for the present investigation. This solution is non-corrosive

TABLE 1: Details of sensor accuracy, measurement range, and operating voltages of the sensors used in the experiment

Sensor	Operating voltage	Range of measurement	Accuracy
Keyence FDX-A1	20–30 V	0.02–20 L/min	+0.003 ml/min
Honeywell pressure sensors	4.75–5.25 V	0–50 psi	+0.25%
K-type thermocouple	—	0–400 °C	+0.75%

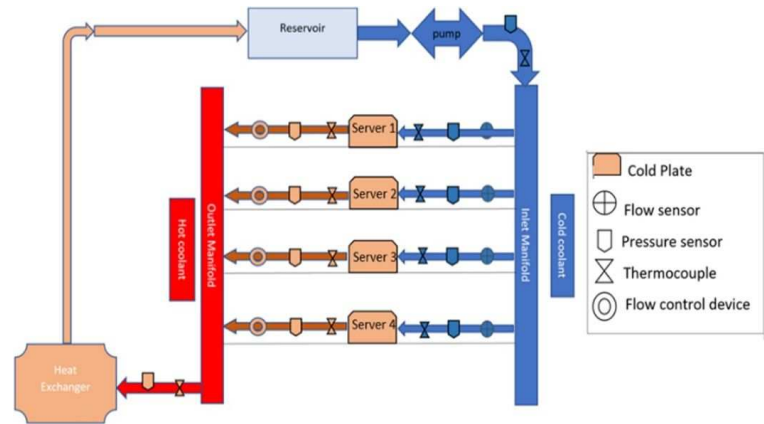


FIG. 4: Schematic of the entire experimental setup used in the study

and non-toxic with compatible materials as per ASHRAE liquid cooling guidelines and offers excellent heat-transfer properties in the desired temperature range.

2.2 FCD Design and Control

The FCD design for this study was chosen based on the expected flow rates specific to the current setup that would be sufficient to dissipate the maximum power from the heater TTVs. As shown in Fig. 5, the FCD used in this study is a micro-servo controlled, ball-valve assembly that was designed and validated for operation under a range of flow rates (Shahi et al., 2021). The V-cut ball valve of this FCD was designed in such a way that it always allows a minimum flow rate to pass through it to accommodate idle server load in real-world applications. The entire FCD assembly was 3D printed after preliminary design computational fluid dynamics (CFD)

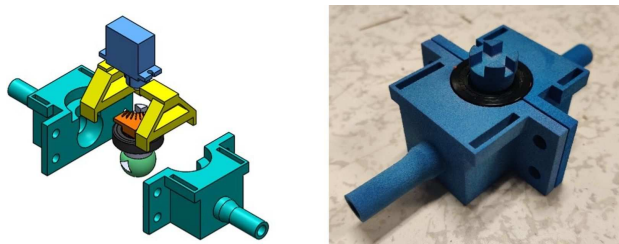


FIG. 5: Exploded view of the CAD model (left) and final 3D-printed FCD (right)

and finite element analysis (FEA) simulations for a range of flow rate range of 0–4 lpm. The V-cut ball valve inside the FCD allows a minimum of 0.2 lpm flow rate through the device cavity in the fully closed position. The flow rate through the FCD is controlled using a shaft rotated by a servo motor at pre-decided angles based on the flow and pressure drop characterization study on the FCD (Shahi et al., 2021).

2.3 Experimental Procedure

The experimental setup consists of four TTVs mimicking four servers in a rack at different heights, each built identically. The coolant is pumped from the reservoir, placed at top of the rack, to the inlet manifold from where it is distributed to each of the four servers. Once the coolant passes through the server, it reaches the outlet manifold from which it goes to the heat exchanger (chiller) and then again recirculates back to the reservoir. The experiments were conducted at three different inlet coolant temperatures of 35, 40, and 45°C with the experimental set conditions as shown in Table 2. Experiment 1 is the base case where all four servers are at 100% utilization and provisioned with a maximum flow rate. For experiment 2, server 1 was brought down to 5% server utilization and the inlet flow rate varied to 0.24 lpm, whereas all the other servers were kept at 100% utilization. 0.24 lpm is based on the minimum allowable flow rate that the FCD allows during the operation when integrated with the current experimental flow loop. The power utilization values were selected based on the minimum supply load from the power supply to all the TTVs. In experiment 3, servers 1 and 2 were brought down to 5% server utilization and flow rate of 0.24 lpm, and the other two servers were held at 100% utilization. Finally, in experiment 4, servers 1, 2, and 3 were brought down to 5% utilization and at a flow rate of 0.24, 0.24, and 0.3 lpm, respectively, whereas server 4 was at 100% utilization. In the next set of experiments the flow-rate variation was kept constant for all experiments as detailed above but the server utilization was brought to 35% instead of 5% as shown in Table 3. These set experiments were performed to see the maximum power that can be given to the TTVs with lower utilization without increasing their temperature beyond their temperature as in the base case.

3. RESULTS AND DISCUSSION

As discussed in the previous section, the experiments were performed in two sets: the first, where the power utilization of the servers was varied from a maximum of 100% to an ideal power of

TABLE 2: Experimental conditions for servers at maximum (100%) and ideal power utilization

Server no.	Experiment 1 (base case)		Experiment 2		Experiment 3		Experiment 4	
	Flow rate (lpm)	Power (Watts)/(%) Utilization	Flow rate (lpm)	Power (Watts)/(%) Utilization	Flow rate (lpm)	Power (Watts)/(%) Utilization	Flow rates (lpm)	Power (Watts)/(%) Utilization
Server 1	1	479 (100)	0.24	20.75 (5)	0.24	20.75 (5)	0.24	20.69 (5)
Server 2	0.95	490 (100)	—	490 (100)	0.24	20.75 (5)	0.24	20.69 (5)
Server 3	0.91	480 (100)	—	480 (100)	—	480 (100)	0.3	20.69 (5)
Server 4	0.79	476 (100)	—	476 (100)	—	476 (100)	—	476 (100)

TABLE 3: Experimental conditions for servers at maximum (100%) and 35% power utilization

Server no.	Experiment 1 (base case)		Experiment 2 (one server idle)		Experiment 3 (two servers idle)		Experiment 4 (three servers idle)	
	Flow rate (lpm)	Power (Watts)/(%) Utilization	Flow rate (lpm)	Power (Watts)/(%) Utilization	Flow rate (lpm)	Power (Watts)/(%) Utilization	Flow rates (lpm)	Power (Watts)/(%) Utilization
Server 1	1	479 (100)	0.24	165 (33)	0.24	165 (33)	0.25	165 (33)
Server 2	0.95	490 (100)	—	490 (100)	0.24	165 (33)	0.25	165 (33)
Server 3	0.91	480 (100)	—	480 (100)	—	480 (100)	0.32	165 (33)
Server 4	0.79	476 (100)	—	476 (100)	—	476 (100)	—	476 (100)

5%. In the next set of experiments, the power was varied from 100 to 33%. These values were chosen from the literature review which pointed at the fact that maximum power consumption in a server is between 0 and 50% of CPU utilization as discussed in Section 2. The minimum flow rates used in this study were identified from the literature on similar cold plates from the same manufacturer and comparable heat fluxes (Fernandes, 2015; Hoang et al., 2020a).

The first experiment performed was a base-case experiment, where the maximum power of 480 W was given to all four TTVs and a maximum coolant flow rate was provisioned by the pump. The inlet temperature was kept constant at 35°C using the glycol chiller. This case is represented by column 2 in Table 4. A steady state was obtained when the variation between the TTV temperature was within 1°C of the TTV temperature as shown by the inlet temperature thermocouple. For the next set of experiments, the power to server 1 was reduced to 5% of the maximum power to depict a server at ideal or minimum workload conditions. The coolant flow rate was then reduced to 0.24 lpm using the FCD, which led to a corresponding change in the flow rate in the remaining three cooling loops as seen in column 3 in Table 4. Similarly, for the next two experimental runs, the power to the TTVs was reduced until only one TTV assembly was operating at maximum power utilization.

TABLE 4: Percentage and absolute change in flow rates for different experimental conditions

Server no.	Experiment 1	Experiment 2	Experiment 3	Experiment 4
	% Flow-rate change			
Server 1	0.0	0.0	0.0	0.0
Server 2	0.0	25.6	0.0	0.0
Server 3	0.0	23.0	56.4	0.0
Server 4	0.0	23.0	60.3	140.2
Server no.	% Flow rates (lpm)			
	Experiment 1	Experiment 2	Experiment 3	Experiment 4
Server 1	1.00	0.24	0.25	0.24
Server 2	0.88	1.10	0.26	0.25
Server 3	0.93	1.15	1.46	0.31
Server 4	0.80	0.98	1.28	1.91

Corresponding temperature variation was monitored for coolant inlet and outlet temperatures to each of the cooling loops and the TTV surface. The results for temperature variation in the TTVs corresponding to the flow-rate change, as described in Table 4, for each server are shown in Table 5. Column 2 shows the base-case temperatures for each TTV at maximum power. A major outcome of the dynamic flow control based on power utilization, as seen in Table 5, was that the maximum temperature on the TTV reduces for the case when only one server is running at 100% utilization. A 6.2% reduction in the TTV temperature was seen for server 4 when the remaining servers were operated at ideal power conditions. This means that the FCD and dynamic flow control strategy not only optimizes the flow to each server, it can also improve the cooling of the servers operating at maximum workload in a given rack. This trend was consistently seen for all the experimental runs at different inlet temperatures.

Similarly, for higher inlet temperatures of 40 and 45°C, the power to each server was varied and corresponding temperature variations were recorded for non-uniform power distribution across the rack. The maximum temperature for the servers was reduced for each case as seen in Tables 6 and 7, as the number of servers at lower utilization increased. A maximum reduction of 5.7 and 4.7% was observed for 40 and 45°C coolant inlet temperatures, respectively, for 100% and ideal workload conditions. The experiments were again repeated at 35% workload conditions, but the results are not shown in the manuscript to avoid repetition. For this case, a reduction of 6.2, 5.5, and 4.2% was obtained in the maximum TTV temperature at a coolant temperature of 35, 40, and 45°C, respectively.

The results for TTV temperature variation for the case of 35°C inlet coolant temperature at ideal and 100% power utilization are compiled in Fig. 6. The trend of reducing maximum temperatures for the servers operating at 100% workload is evident from this plot. To further comprehend the improvement in the thermal performance of the entire system using dynamic flow control, the variation in thermal resistance was analyzed. As the TTVs were embedded in a polycarbonate fixture, it was assumed that the heat was only conducted through the liquid in the cold plate. As seen in Fig. 7, a consistent reduction in the thermal resistance was observed for all the cold-plate assemblies with varying non-homogeneity in power distribution per each TTV/server. A maximum of 27% drop in thermal resistance was observed for server 4 cold plate

TABLE 5: Percentage and absolute change in temperatures for different experimental conditions for 35°C inlet temperature

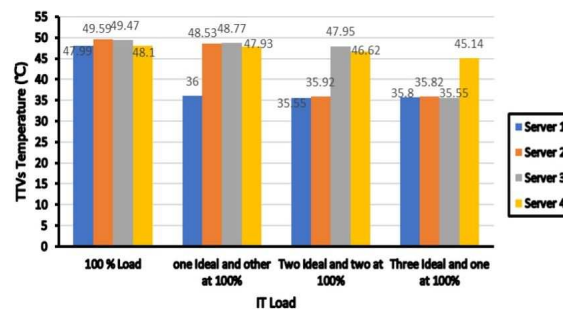
Server no.	Experiment 1	Experiment 2	Experiment 3	Experiment 4
	% TTV temperature change			
Server 1	0.0	0.0	0.0	0.0
Server 2	0.0	2.1	0.0	0.0
Server 3	0.0	1.4	3.1	0.0
Server 4	0.0	1.5	3.1	6.2
	TTVs temperature at 35°C			
Server 1	47.99	37.18	36.48	36.14
Server 2	49.59	48.53	36.46	36.11
Server 3	49.47	48.77	47.95	36.34
Server 4	48.10	47.40	46.62	45.14

TABLE 6: Percentage and absolute change in temperatures for different experimental conditions for 40°C inlet temperature

Server no.	Experiment 1	Experiment 2	Experiment 3	Experiment 4
% TTV temperature change				
Server 1	0.0	0.0	0.0	0.0
Server 2	0.0	2.1	0.0	0.0
Server 3	0.0	1.6	3.1	0.0
Server 4	0.0	1.8	3.2	5.7
TTVs temperature at 40°C				
Server 1	52.69	41.47	40.82	40.66
Server 2	54.23	53.11	40.91	40.65
Server 3	54.34	53.49	52.65	40.94
Server 4	52.81	51.87	51.12	49.78

TABLE 7: Percentage and absolute change in temperatures for different experimental conditions for 45°C inlet temperature

Server no.	Experiment 1	Experiment 2	Experiment 3	Experiment 4
% TTV temperature change				
Server 1	0.0	0.0	0.0	0.0
Server 2	0.0	1.6	0.0	0.0
Server 3	0.0	1.3	2.6	0.0
Server 4	0.0	1.1	2.7	4.7
TTVs temperature at 45°C				
Server 1	57.15	46.54	45.55	45.48
Server 2	58.62	57.69	45.65	45.57
Server 3	58.93	58.19	57.42	45.94
Server 4	57.25	56.62	55.70	54.58

**FIG. 6:** Variation of TTV temperature for each server at different IT loads at 35°C inlet coolant temperature

when it is the only server operating at maximum power utilization. This further bolsters the claim of enhanced performance of DLC using dynamic flow-rate control.

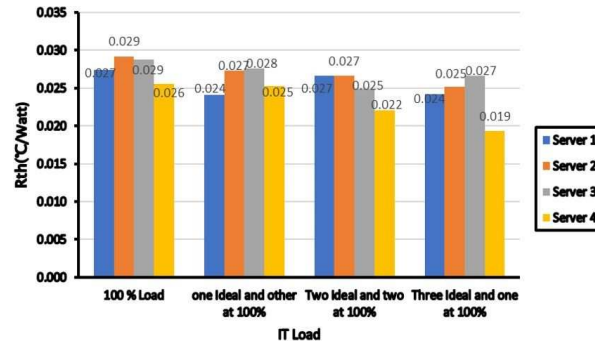


FIG. 7: Variation of thermal resistance for different IT loads and varying flow rates for all four servers at 35°C inlet coolant temperature

It has been well researched that temperature uniformity enhances package reliability. To validate the effect of dynamic flow control in enhancing information technology equipment (ITE) reliability, improvement in temperature uniformity on the TTV surface was assessed as seen in Fig. 8. A reduction of approximately 58% was observed in the temperature difference across the cold plate for the case where only one server operates at the maximum workload. This implies that a real electronic package will experience a significantly lower thermal gradient if the flow rate is provisioned based on the instantaneous package heat flux.

To determine the impact of reliability enhancement in package-level reliability using dynamic cooling, the literature on failure rates in silicon-based microelectronics was reviewed (Lall, 1997). The results obtained for TTV temperatures for different experimental cases were compared with the mean time to failure (MTTF) results from the literature (Lakshminarayanan and Sriraam, 2014). The reviewed study discusses the MTTF for the most common failure mechanisms in the microelectronics industry and mathematical models to determine the MTTF for each failure mechanism. The results of the TTV temperatures from experiment 1 (baseline case) were compared with the MTTF from the reviewed study using Eq. (1). As discussed earlier in the results, a 6.2% change in the maximum temperature was obtained at the same server when operating in dynamic-cooling mode. Based on the maximum temperatures on server 4 in experiment 1 and experiment 4 as shown in Table 5, MTTF accelerates by a factor of 1.15 when a constant flow rate is provided to all the servers. The same variation in MTTF was obtained

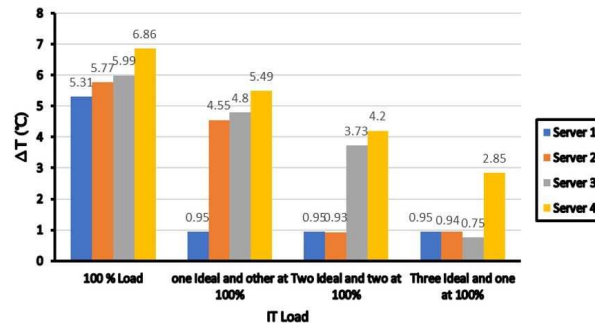


FIG. 8: Variation of the temperature across all the cold plates for the experimental condition with servers at ideal and 100% power utilization at 35°C inlet coolant temperature

for other cases at the two other inlet temperatures tested which quantify the ability of dynamic cooling at the rack level in improving ITE reliability.

4. CONCLUSION AND FUTURE WORK

Constant package miniaturization and the corresponding increase in their heat flux has made DLC the most efficient cooling technique, especially for high-powered GPUs. Current liquid-cooled data centers operate under a constant coolant flow rate through the cold plates irrespective of the instantaneous utilization of the processing units. This can lead to large thermal gradients on the surface of the electronic packages leading to premature thermo-mechanical failures. A further improvement in the efficiency and implementation of this cooling technique is therefore possible.

The current investigation proposed and experimentally validated the concept of dynamic cooling at the rack level for improving ITE reliability. This was done by assessing the temperature uniformity at the surface of the TTVs and temperature difference across the cold plates by dynamically varying the coolant flow rates at the rack level. The variation in coolant flow rate was decided based on the utilization of the custom-made TTVs used to mimic computational load in a high-powered server. Four such TTVs were placed in a 42U rack at different locations on custom-made fixtures to place the heaters and the cold plates. The coolant flow rates and the power to the TTVs were varied to obtain the TTV temperatures, thermal resistance, and temperature difference across the cold plate. The results obtained concluded that both the thermal resistance and the temperature difference across the cold plate were reduced by dynamically varying the flow rate using the FCD for each server. Additionally, varying the flow rates based on the instantaneous power consumed by the server reduced the maximum temperatures on the TTVs.

Past work on FCD design and development and power-savings study using CFD has hinted at potentially large savings in pumping power at the rack level (Shahi et al., 2021; Kasukurthy et al., 2020; Kasukurthy, 2019). Subsequent studies on this work will determine the thermo-hydraulic impact of FCD and present a working control strategy along with the detailed working of the FCD at the rack level. DLC can thus be further improved to accommodate higher heat fluxes and yield significant power savings at the same time as compared to other cooling techniques like single-phase and two-phase immersion cooling (Bansode et al., 2019; Shahi, 2016; Shahi et al., 2020; Hoang et al., 2020b). It can thus be concluded from the results that dynamic liquid cooling using an FCD can improve the package reliability by reducing the maximum package temperature and further improving the temperature uniformity on the package surface.

REFERENCES

- Alkharabsheh, S., Fernandes, J., Gebrehiwot, B., Agonafer, D., Ghose, K., Ortega, A., Joshi, Y., and Sammakia, B., A Brief Overview of Recent Developments in Thermal Management in Data Centers, *J. Electr. Packag.*, vol. **137**, no. 4, Article ID 040801, 2015.
- Bansode, P.V., Shah, J.M., Gupta, G., Agonafer, D., Patel, H., Roe, D., and Tufty, R., Measurement of the Thermal Performance of a Custom-Build Single-Phase Immersion Cooled Server at Various High and Low Temperatures for Prolonged Time, *J. Electr. Packag.*, vol. **142**, no. 1, Article ID 011010, 2019.
- Barroso, L.A. and Hölzle, U., The Case for Energy-Proportional Computing, *Computer*, vol. **40**, no. 12, pp. 33–37, 2007. DOI: 10.1109/MC.2007.443
- Beaty, D.L., Liquid Cooling—Friend or Foe, *ASHRAE Trans.*, vol. **110**, pp. 643–652, 2004.

- Chainer, T.J., Schultz, M.D., Parida, P.R., and Gaynes, M.A., Improving Data Center Energy Efficiency with Advanced Thermal Management, *IEEE Trans. Compon. Packag. Manufact. Technol.*, vol. 7, no. 8, pp. 1228–1239, 2017.
- Chu, R.C., Simons, R.E., Ellsworth, M.J., Schmidt, R.R., and Cozzolino, V., Review of Cooling Technologies for Computer Products, *IEEE Trans. Device Mater. Reliab.*, vol. 4, no. 4, pp. 568–585, 2004.
- Ellsworth, M.J., Jr. and Iyengar, M.K., Energy Efficiency Analyses and Comparison of Air- and Water-Cooled High-Performance Servers, *Proc. of the ASME 2009 InterPACK Conf. Collocated with the ASME 2009 Summer Heat Transfer Conf. and the ASME 2009 3rd Intl. Conf. on Energy Sustainability*, ASME 2009 InterPACK Conf., San Francisco, CA, pp. 907–914, July 19–23, 2009.
- Ellsworth, M.J., Jr. and Iyengar, M.K., Energy Efficiency Analyses and Comparison of Air and Water Cooled High Performance Servers, *ASME 2009 InterPACK Conf.*, San Francisco, CA, pp. 907–914, 2009.
- Fernandes, J.E., Minimizing Power Consumption at Module, Server and Rack-Levels within a Data Center through Design and Energy-Efficient Operation of Dynamic Cooling Solutions, PhD, The University of Texas at Arlington, 2015.
- Helsin, K., 2014 Data Center Industry Survey, Uptime Institute, Seattle, WA, Report, 2014.
- Hoang, C.H., Tradat, M., Manaserh, Y., Ramakrisnan, B., Rangarajan, S., Hadad, Y., Schiffres, S., and Sammakia, B., Liquid Cooling Utilizing a Hybrid Microchannel/Multi-Jet Heat Sink: A Component Level Study of Commercial Product, *ASME 2020 Intl. Technical Conf. and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems*, Virtual, 2020a.
- Hoang, C.H., Khalili, S., Ramakrisnan, B., Rangarajan, S., Hadad, Y., Radmard, V., Sikka, K., Schiffres, S., and Sammakia, B., An Experimental Apparatus for Two-Phase Cooling of High Heat Flux Application Using an Impinging Cold Plate and Dielectric Coolant, *36th Semiconductor Thermal Measurement, Modeling & Management Symp. (SEMI-THERM)*, San Jose, CA, pp. 32–38, March 16–20, 2020b.
- Iyengar, M., Energy Consumption of Information Technology Data Centers, *J. Electr. Cooling*, vol. 16, no. 4, pp. 28–31, 2010.
- Kasukurthy, R., Challa, P.S., Palanikumar, R.R., Manimaran, B.R., and Agonafer, D., Flow Analysis and Linearization of Rectangular Butterfly Valve Flow Control Device for Liquid Cooling, *IEEE Intersociety Conf. on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm)*, San Diego, CA, pp. 683–687, May 29–June 1, 2018.
- Kasukurthy, R., Rachakonda, A., and Agonafer, D., Design and Optimization of Control Strategy to Reduce Pumping Power in Dynamic Liquid Cooling, *ASME J. Electr. Packag.*, vol. 143, no. 3, Article ID 031001, 2021.
- Kumar, A., Shahi, P., and Saha, S.K., Experimental Study of Latent Heat Thermal Energy Storage System for Medium Temperature Solar Applications, in *Proc. of the 4th World Congress on Mechanical, Chemical, and Material Engineering (MCM'18)*, Madrid, Spain, Paper no. HTFF 152, August 16–18, 2018.
- Lakshminarayanan, V. and Sriraam, N., The Effect of Temperature on the Reliability of Electronic Components, *IEEE Intl. Conf. on Electronics, Computing and Communication Technologies (CONECCT)*, Bangalore, pp. 1–6, 2014.
- Lall, P., Pecht, M., and Hakim, E.B., *Influence of Temperature on Microelectronics and System Reliability: A Physics of Failure Approach*, 1st ed., London: CRC Press, 1997.
- Lall, P., Pecht, M., and Hakim, E.B., Characterization of Functional Relationship between Temperature and Microelectronic Reliability, *Microelectr. Reliab.*, vol. 35, no. 3, pp. 377–402, 1995.
- McFarlane, R., Will Water-Cooled Servers Make Another Splash in the Data Center?, *Tech Target Network, Search Data Center*, Newton, MA, accessed from <https://searchdatacenter.techtarget.com/tip/Will-watercooled-servers-make-another-splash-in-the-data-center>, 2012.

- Patterson, M.K., Krishnan, S., and Walters, J.M., On Energy Efficiency of Liquid Cooled HPC Datacenters, *IEEE Intersociety Conf. on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm)*, Las Vegas, NV, pp. 685–693, May 31–June 3, 2016.
- Schmidt, R., Liquid Cooling is Back, *Electr. Cooling*, vol. **11**, no. 3, pp. 34–38, 2005.
- Shahi, P., Agarwal, S., Saini, S., Niazmand, A., Bansode, P., and Agonafer, D., CFD Analysis on Liquid Cooled Cold Plate Using Copper Nanoparticles, *ASME 2020 Intl. Technical Conf. and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems*, Virtual, October 27–29, 2020.
- Shahi, P., Deshmukh, A.P., Hurnekar, H.Y., Saini, S., Bansode, P., Kasukurthy, R., and Agonafer, D., Design, Development, and Characterization of a Flow Control Device for Dynamic Cooling of Liquid-Cooled Servers, *ASME J. Electron. Packag.*, vol. **144**, no. 4, Article ID 041008, 2022.
- Shahi, P., Experimental Analysis of PCM Based Thermal Storage System for Solar Power Plant, M.S., Indian Institute of Technology-Mumbai, Mumbai, India, 2016.
- Shahi, P., Saini, S., Bansode, P., and Agonafer, D., A Comparative Study of Energy Savings in a Liquid-Cooled Server by Dynamic Control of Coolant Flow Rate at Server Level, *IEEE Trans. Compon. Packag. Manufact. Technol.*, vol. **11**, no. 4, pp. 616–624, 2021.
- Shehabi, A., Smith, S.J., Horner, N., Azevedo, I., Brown, R., Koomey, J., Masanet, E., Sartor, D., Herrlin, M., and Lintner, W., United States Data Center Energy Usage Report, Lawrence Berkeley National Laboratory (LBNL), Berkeley, CA, Report no. LBNL-1005775, 2016.
- Stansberry, M. and Kudritzki, J., Uptime Institute 2012 Data Center Industry Survey, Uptime Institute, Seattle, WA, accessed September 2021, from https://www.etherworks.com.au/index.php?option=com_k2&Itemid=230&id=91_2276d1f6c67a40bd451926a7ab5ccc49&lang=en&task=download&view=item, 2013.