Cold Plate Based Dynamic Liquid Cooling at Data Center and Chip Level

Dissertation

by

Pardeep Shahi

Submitted in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy at

The University of Texas at Arlington

August 2022

Supervising Committee:

Dr. Dereje Agonafer
Dr. Abdolhossein Haji-Sheikh
Dr. Miguel A. Amaya
Dr. Sunand Santhangopalan
Dr. Saket Karajgikar

Acknowledgments

To begin with, I owe huge gratitude to my parents who gave me the opportunity to move to the US to pursue my dream of higher education, and to all the sacrifices they have made for me. I cannot thank Dr.Agonafer enough for his guidance, mentorship, and valuable words of wisdom as an advisor throughout my Doctoral degree programs.

I would like to thank Dr. Haji Sheikh, Dr. Miguel Amaya, Dr. Sunand Santhangopalan, and Dr. Saket Karajgikar for being the thesis committee members and providing me with their valuable input in improvising my research work.

My peers and colleagues have also played a vital role throughout my time in the EMNSPC research group. I have been very privileged to have my roommates Patik Bansode and Satyam Saini as my best research partners and for their immense contribution to my doctoral research. In addition to this, I have been able to work with many other like-minded and brilliant individuals within the EMNSPC team. My heartfelt regards also go toward Dr. Ali Heydari and Jeremy Rodriguez who has also been my mentor during my internship at NVIDIA and valuable advisor.

Abstract

Cold Plate Based Dynamic Liquid Cooling at Data Center and Chip Level

The University of Texas at Arlington, 2022

Supervising Professor: Dr.Dereje Agonafer

The rising demand for high-performance central and graphical processing units has resulted in the need for more efficient thermal management techniques like direct-to-chip liquid cooling. Direct Liquid Cooling using cold plates is one of the most efficient and investigated cooling technology since the 1980s. Major data and cloud providers like IBM, Microsoft, and Google are actively deploying liquid-cooled data center infrastructure due to rising computational demands but its performance and efficiency can be further be enhanced using dynamic cooling technologies.

At the chip level since the early ’60s, based on Moore’s law, transistor density has been doubling every generation resulting in increased power density. Eventually, in the early ’90s, we moved from constant voltage to constant electric field and corresponding constant power for a given area during technology changes. Dennard’s model of voltage scaling and corresponding constant power ceased, ending improved performance gains in the early 2000s that again required techniques to mitigate increased power and corresponding temperature. The performance gain is being achieved by using multi-core processors, leading to non-uniform power distribution and localized high temperatures making cooling very challenging.

At rack level the coolant distribution to multiple racks housing such computing systems is an important factor for efficient cooling. Most liquid-cooled data centers provision their IT equipment with a constant coolant flow rate through multi-channel heat sinks mounted on the processing units. A redundant cooling flow rate is provided based on the maximum anticipated

heat loads from the processing units to be cooled. This, in turn, consumes significant pumping power even when the computing systems are working at their lowest possible IT loads.

This study will address the issues as mentioned above for future direct to chip liquid cooled data centers. First the benchtop experiments were performed on open compute (OU) server. In this study 2OU server is used to show the effect of variable flow rate (0.2, 0.4, 0.6 LPM) on the core temperature, DIMMs (Dual In-line Memory Modules) temperature, Platform Controller Hub (PCH) temperature, cooling power at Ideal, 50% and 100% IT load at inlet temperatures of 25, 30, 35, 40 and 45 º C, which falls within the ASHRAE liquid-cooled envelope, W4. For rack level dynamic cooling the novel active Flow control device (FCD) is designed to control the coolant flow rates at the server level. The dynamic cooling will result in pumping power savings by controlling the flow rates based on server utilization. The proposed FCD design contains a V-cut ball valve connected to a micro servo motor. The valve position is varied to change the flow rate through the valve by servo motor actuation based on pre-decided rotational angles. FCD working was validated by varying flow rates and pressure drop across the device by varying the valve position using both CFD and experiments at bench top. Further the experiments were performed at rack level using this FCD and control strategy was tested for pumping power saving and maximum energy saving of 87% was achivied, when entire rack was at idle condition. For chip level dynamic cooling to reduce thermal gradient on the chip, the passive bimetallic based cold plate is designed. The pro

posed dynamic cold pate design is a 3-part assembly that is divided into three main parts. The bottom part of the plate contains 4 different sections of parallel copper microchannels through which the coolant flows, extracting the heat. The middle part of the plate has both inlet and outlet passage with bimetallic strips to control flow rate in each section based on coolant temperature in that section, and the top section of the plate is the plastic sealed cover plate. Dynamic cold plate working was validated by using CFD and it was observed that

a maximum of 62% temperature gradient can be reduced on the chip and 40% of Rth was reduced.

The experiments were performed at UTA in collaboration with NVIDIA. A cold plate-based liquid-cooled data center was developed. Liquid-to-liquid heat exchangers used in liquid-cooled data centers are also referred to as coolant distribution units (CDUs). Most of these CDUs selected by the data center operator are based on the heat load of the data center and the available head with that CDU. A 450-kW liquid-cooled CDU is used, and propylene glycol 25% is used as a coolant. Typical CDUs are designed to operate at 20 to 30% of the rated heat load to achieve a stable secondary coolant supply temperature. The present study will investigate the operations of CDU at very low heat loads, like 1% to 10% of the CDU's rated capacity. At these low loads, large fluctuations in secondary side supply temperature were observed. This large fluctuation can lead to the failure of the 3-way valve used in CDUs at the primary side. In this paper, a control strategy is developed to stabilize the secondary supply temperature within ± 0.5 °C at very low loads using the combination of a flow control valve on the primary side and PID control settings within the CDU.

In another study that was done with NVIDIA an in-depth analysis of hydraulic transients when rack-level flow control valves are used with and without flow control. The operating conditions of the CDU are varied for different parameters such as a constant flow rate, a constant differential pressure, and a constant pump speed. Furthermore, the hydraulic transient is examined when the cooling loop modules are decommissioned from the rack one by one. The effect of this step-by-step decommissioning is assessed on the CDU operation and other racks. The pressure drop-based control strategy has been developed to maintain the same flow rate in the remaining servers in the rack when some cooling loop modules are decommissioned.

Table of Contents

List of Illustrations

List of Tables

Chapter 1 Introduction

1.1 Introduction

Data centers are the buildings that host hundreds of thousands of servers. Each server consists of electronics that constantly produce heat and need to be removed for safe and reliable operations. Multiple cooling technologies developed to remove the heat from the servers from air-cooled servers where fans are used with heat sinks and worked on the principle of forced convection to 2-phase immersion cooling, where the entire server is dipped into the 2-phase coolant and heat is removed by boiling the liquid from liquid to vapors. The need to move from air to liquid or 2 phase cooling arises because of a continuous increase in power densities of chips as the number of transistors are increasing as shown in Figure 1-1 below

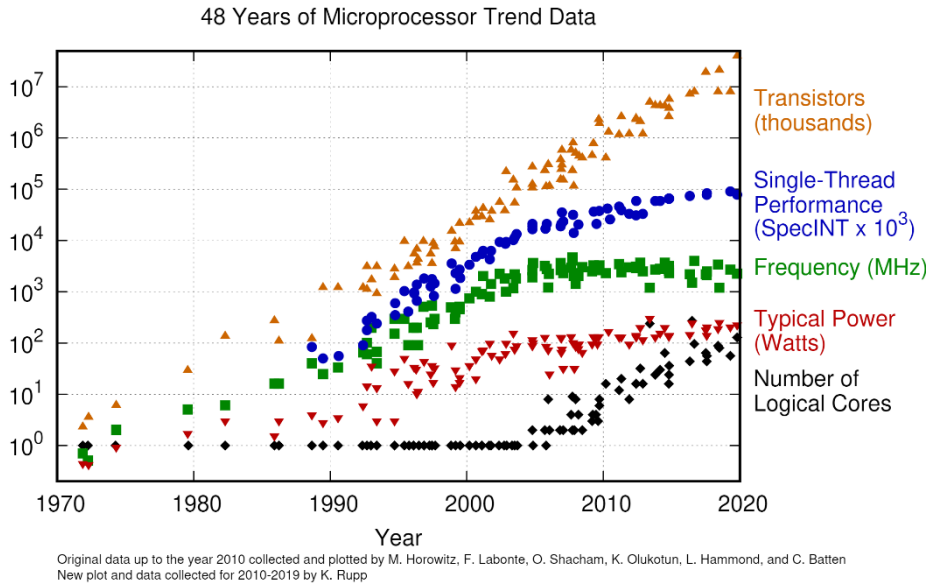


Figure 1-1: Microprocessor trends in recent years [1]

These power densities are increasing because of new technologies like the internet of things (IoT), Bitcoin mining, High-performance computing (HPC), online streaming, gaming, cloud computing, etc. The amount of energy data centers are consuming today is around 2% of

the total electricity consumption of the United States and is projected to consume around 73 billion kWh in 2020 [1]. it is predicted that by 2050, the power consumption of the data centers reaches about 180 billion kWh.

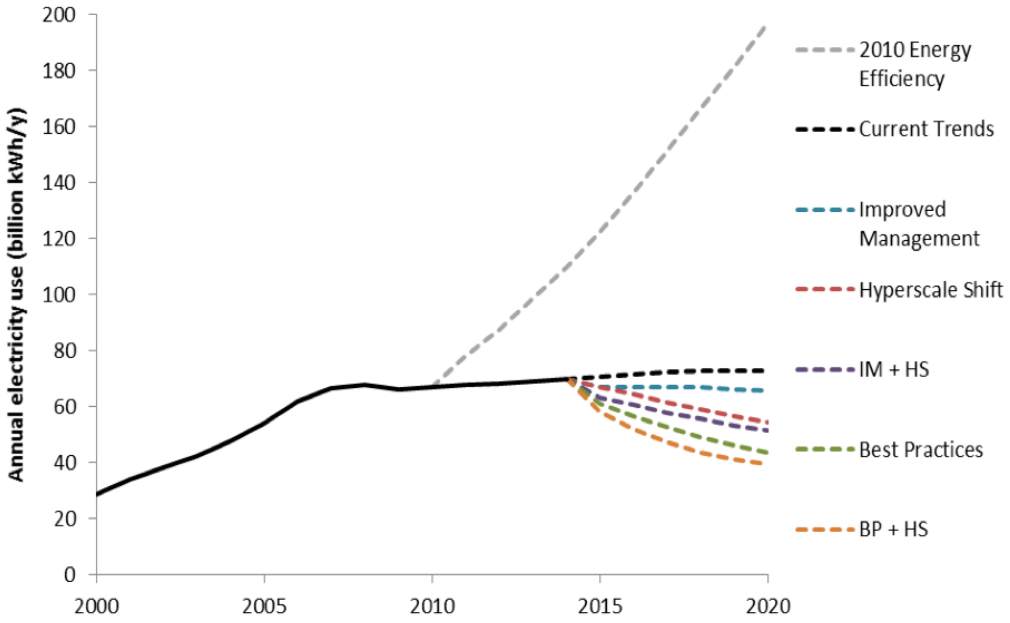


Figure 1-2: Presented trend in data center energy consumption

In usual data centers, 45-55% of energy goes for IT, and 30-40 % goes into cooling. In terms of energy costs, cooling and electricity infrastructure contribute to 70-80% of the capital costs [2]

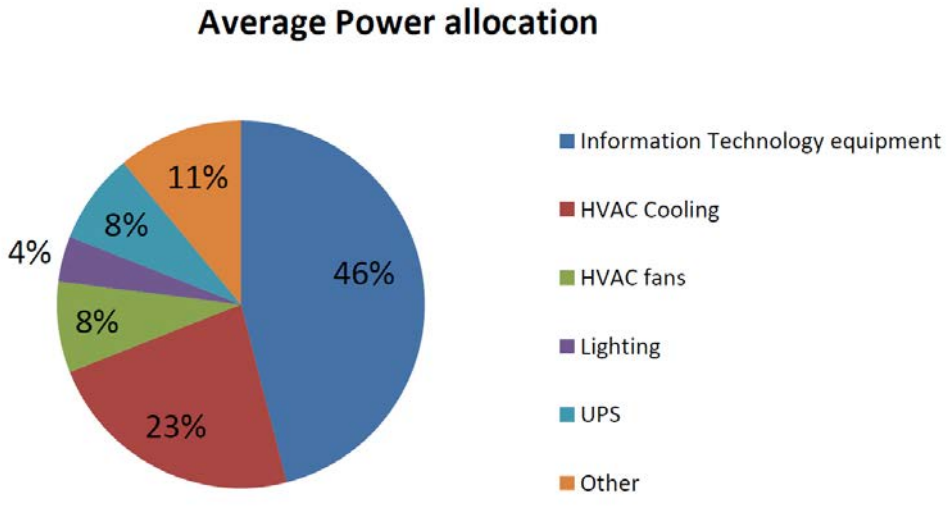


Figure 1-3: Energy Consumption Impact of Mechanical Equipment and Systems [3]

Traditionally, air cooling has been widely utilized to dissipate heat loads from the servers using heat sinks and fans. However, low thermal mass and relatively high PUE limit its implementation for cooling the latest high-performance CPUs and GPUs. Several reports suggest that traditional air-cooled data centers expend approximately 30-40% of energy for cooling purposes [4]. Recent trends have predicted while the global data center energy trends are optimistic, the average PUE value still hovers around 1.58, which is mostly attributed to inefficient cooling methods employed [5].

Several methods are being explored to improve inefficiencies in data center cooling. Among all these methods, liquid cooling provides a highly efficient solution to meet the rising cooling demands. This is mainly due to the higher thermal conductivity and thermal mass of coolants.Air cooling being traditional and most common method in cooling the electronic components many researchers have focused on improvising the current trends in air cooling. Improvised chassis design has been proved to be effective design approach in improving the cooling of major heat dissipitating components [6-8]. Researches on the improving the performance and feasibility of rear door heat exchangers[9], efforts are made to reduce site visit by calibrating Data center CFD models [10] and use of Artificial Intelligence has been implemented to predict a control strategy in a data center for multiple air conditioning units[11], various experimental[12-15] and CFD analysis[16][17] are performed to collect data on the role of contaminants and effects of relative humidity[18][19] in air and the decrease in cooling efficiency caused by air quality. With increase in the density of data centers, air quality also plays a significant role in the efficiency of data center cooling, experiment.

Immersion cooling has been one of the emerging technology where the server/electronic components are immersed in a dielectric fluid which have better thermal properties. In single phase immersion cooling, experiments have been conducted on the thermal performance of the server[20], custom-build[21][20], open compute servers[23] and CFD

analysis on the optimization of server to improvise the cooling in immersion[24], numerical analysis are performed on the used of different fluids such as mineral oil and Al203 nanofluid[25] and CFD analysis over two phase immersion cooling fluids such as FC-72[26]. Also with it's benefits, in immersion cooling reliability is also a concern and research on the PCB properties when immersed in fluid, where the oil plays an important role in affecting the material parameters of the components[27-29]. To further enhance the cooling optimization are performed at heat sink for forced and natural convection in an immersion cooling.[30][31].

At chip level, reliability testing on the impact of long term aging of the TIM shows at high temperature TIM shows degradation in material properties[32-33]. Effective modeling shows the warpage in multi chip modules.[34]. The PCBs with higher stiffness have a lower reliability[35-36]. Study shows Copper content in the PCB is key parameter with the thickness of the PCBs[37]. The change in die size can reduce the delamination issue between TSVs and silicon.[38-40]. The coupled power and thermal cycles showed reduced life of the packages[41-45].

Liquid cooling is further divided into 2 main categories one is Direct chip liquid cooling (DCLC).and immersive. This is further divided in subcategories as shown below in Figure 1-4

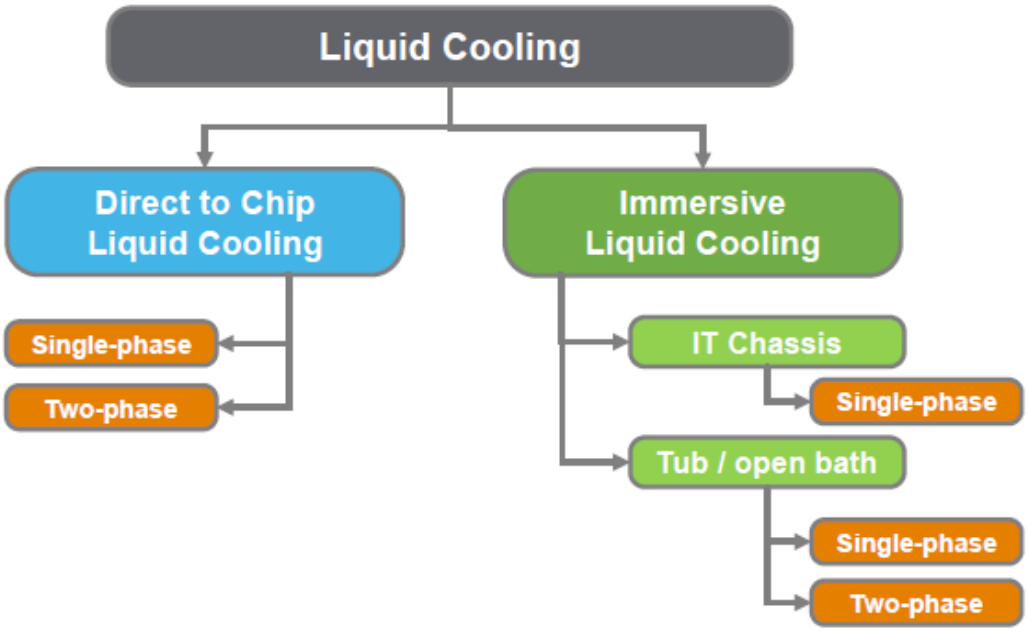


Figure 1-4 Subcategories of Liquid Cooling [5]

A typical DCLC liquid-cooled server contains a manifold and cold plate arrangement instead of heat sinks and fans. The manifolds distribute the coolant to each of the servers which then gets distributed to single or multiple cold plates, mounted on the processors. The earliest works in DLC date back to the 1960s but with the growth in CMOS technology in the 1990s, the heat loads in the electronic packages were significantly reduced and thus making air cooling yet again a viable technology [46-49]. However, with further developments in the CMOS technology, air cooling is once again reaching its limitations and the focus is again being shifted to liquid cooling.

Liquid cooling can be deployed in data centers in two main ways. The first one can be a hybrid air and liquid cooling arrangement where the auxiliary components are cooled by air and the primary high-power heat-dissipating components using liquid cooling. An advantage of this method is that it can be directly used in an air-cooled data center setup with minor changes around the computing systems. If the entire server is liquid-cooled, thermal pathways like cold rails for memory cooling are usually designed within the enclosure to ensure adequate

cooling for all the components. Further improvisations are done in liquid cooling by deployment of control strategies to improvise the cooling of high heat density servers [50-52], with the enhancement of liquid cooling, development of flow control devices[53] shows significant pumping power savings[54],development of liquid to liquid heat exchanger and hydraulic characteristics of the liquid cooled data center proves significant increase in cooling of data center.[55-56]. Reliability studies over liquid cooled cold plates and at rack level proves liquid cooling to be a more reliable and an viable approach in cooling of high heat disspiting components. [57-59] and the calculated heat capture ratio[60] for a liquid and air cooled server is more than for traditional air cooled server.

1.2 Scope of Work

The present study is divided into 9 chapters. The first chapter is the introduction and motivation. The second chapter gives insights into the literature review. Chapter 3-9 will address the issues mentioned above in the abstract for future direct-to-chip liquid-cooled data centers. First, the benchtop experiments were performed on an open compute (OU) server to investigate the effect of varying coolant flow rates on the chip and the auxiliary components. In this study 2OU server is used to show the effect of variable flow rate (0.2, 0.4, 0.6 LPM) on the core temperature, DIMMs (Dual In-line Memory Modules) temperature, Platform Controller Hub (PCH) temperature, cooling power at Ideal, 50% and 100% IT load at inlet temperatures of 25, 30, 35, 40 and 45 ⁰ C, which falls within the ASHRAE liquid-cooled envelope, W4.

For rack-level dynamic cooling, the novel active Flow control device (FCD) is designed to control the coolant flow rates at the server level. The dynamic cooling will result in pumping power savings by controlling the flow rates based on server utilization. The proposed FCD design contains a V-cut ball valve connected to a micro servo motor. The valve

position is varied to change the flow rate through the valve by servo motor actuation based on pre-decided rotational angles. FCD working was validated by varying flow rates and pressure drop across the device by varying the valve position using both CFD and experiments at the bench top. Further, the experiments were performed at rack level using this FCD, and the control strategy was tested for pumping power saving.

For chip-level dynamic cooling to reduce thermal gradient on the chip, the passive bimetallic-based cold plate is designed. The proposed dynamic cold pate design is a 3-part assembly that is divided into three main parts. The bottom part of the plate contains 4 different sections of parallel copper microchannels through which the coolant flows, extracting the heat. The middle part of the plate has both inlet and outlet passage with bimetallic strips to control the flow rate in each section based on coolant temperature in that section, and the top section of the plate is the plastic sealed cover plate. Dynamic cold plate working was validated by using CFD.

Typical CDUs are designed to operate at 20 to 30% of the rated heat load to achieve a stable secondary coolant supply temperature. The present study will investigate the operations of CDU at very low heat loads, like 1% to 10% of the CDU's rated capacity. At these low loads, large fluctuations in secondary side supply temperature were observed. This large fluctuation can lead to the failure of the 3-way valve used in CDUs at the primary side. A control strategy is developed to stabilize the secondary supply temperature within ± 0.5 °C at very low loads using the combination of a flow control valve on the primary side and PID control settings within the CDU.

The operating conditions of the CDU are varied for different parameters such as a constant flow rate, a constant differential pressure, and a constant pump speed. Furthermore, the hydraulic transient is examined when the cooling loop modules are decommissioned from

the rack one by one. The effect of this step-by-step decommissioning is assessed on the CDU operation and other racks. The pressure drop-based control strategy has been developed to maintain the same flow rate in the remaining servers in the rack when some cooling loop modules are decommissioned.

Chapter 2 Literature Review

Till now the most common way of removing heat from IT equipment is by air cooling. The heat generated by the server is typically 70-80% by CPUs and the remaining is generated by auxiliary components. The latest use of GPUs has further increased the heat generated by the servers.

Liquids have a much greater capacity to capture heat by unit volume. This allows the liquids to remove heat more efficiently and allows the chips to work faster (i.e., increase clock speed). As mentioned in the previous section DCLC liquid cooling was first introduced by IBM in the 1960s and it's played a significant role till then with some ups and downs in its journey. Despite the many advantages mentioned above the DCLC need to be improved further.

Most of the microprocessors these days are multichip modules (MCMs) and each module has its own power this gives rise to non-uniform temperature distribution and processors experience locally elevated temperatures known as hot spots [61] which poses a significant challenge to the cooling requirements of high power devices.

Further, to reduce chip temperature concept of distributed pumping emerged but Sahini et al. compared distributed pumping with centralized pumping and observed that distributed pumping consumes much more energy compared to centralized pumping [62]. Cooling chips by forcing coolant to pass through the very close channels on the back side of the silicon wafer was studied by Tuckeman and pease [63]. This study observed that the heat transfer coefficient is directly proportional to the width of the channel in laminar flows. This research was further taken by Zhang et al. [64] who performed experiments on the aluminum-based liquid-cooled heat sink with microchannels. The experiments were performed on 12 x 12 mm and 10 x 10 mm chips and thermal resistance curves were obtained. The spreading resistance was further developed from the Lee et al. model [65]. Qu et al. [66] further investigated both numerical

and experimental liquid-cooled cold plates by varying pressure on the cold plate, porosity, and aspect ratio. The thinner channels were found to be advantageous. Further improvement in DCLC was investigated by Fernandes et al. who investigated multi-design variable optimization of cold plates with the objective of reducing pumping power-saving [67]. Fernandes also worked on the dynamic cold plate where he divided the cold plate into multiple sections and each section has its own control strategy [68] as shown in Figures 2-1 to reduce the hot spots and pumping power saving.

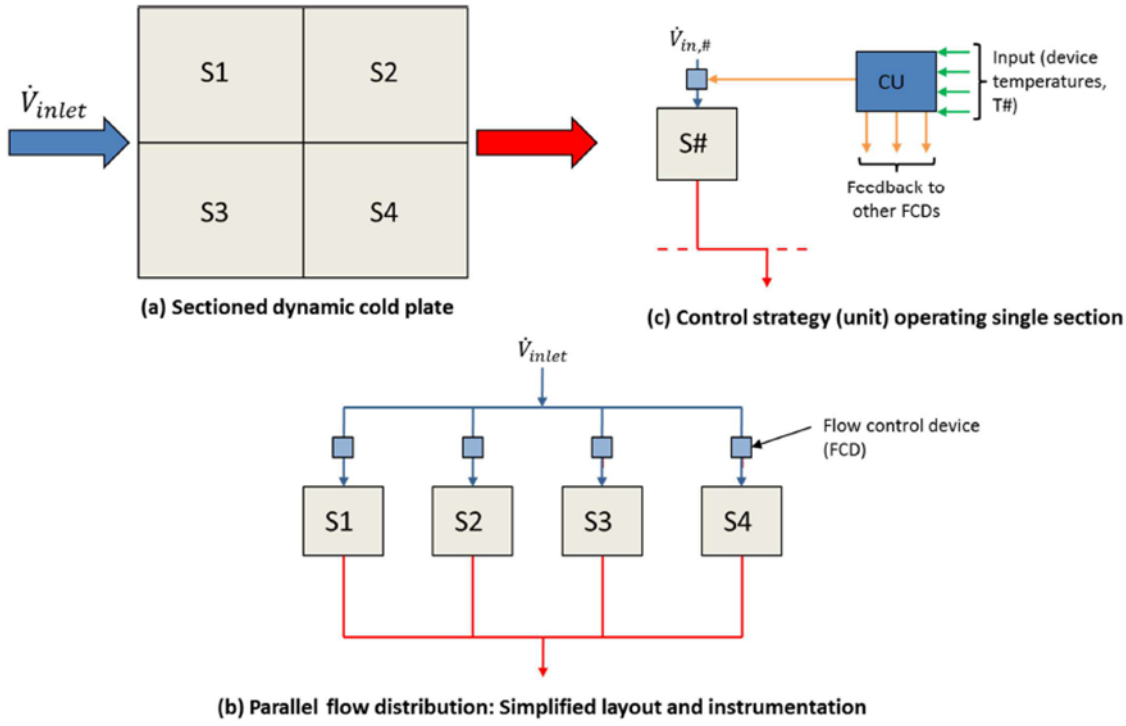


Figure 2-1: Conceptual layout for a cold plate with segregated flow control [8]

Typically, a constant flow rate is provisioned to the ITE in liquid-cooled data centers irrespective of the operating workloads of the servers. Significant savings in the pumping power can be attained if the flow rates are varied in proportion to the instantaneous server workload. Dynamic cooling has been explored in past for air-cooled data centers where control strategies for automatic CRAC (Computer Room Air Conditioning) units are experimentally validated. An improvement of 70% in data center energy performance was achieved by varying

the parameters like CRAC fan speeds and electronically actuated vent tiles [69]. Other approaches include real-time monitoring of aisle temperatures and using machine learning algorithms to predict cooling requirements based on server workloads [70-73]. On the package architecture side, thermally aware workload routes, workload migration, and varying server fan speed reduce power consumption [74].

For liquid-cooled data centers to implement a dynamic cooling flow control device (FCD) need to integrate with ITE for dynamics variation of the coolant flow rate at the server level based on the server workload. Kasukurthy et al. worked on passive FCD design [75] as shown in Figures 2-2.

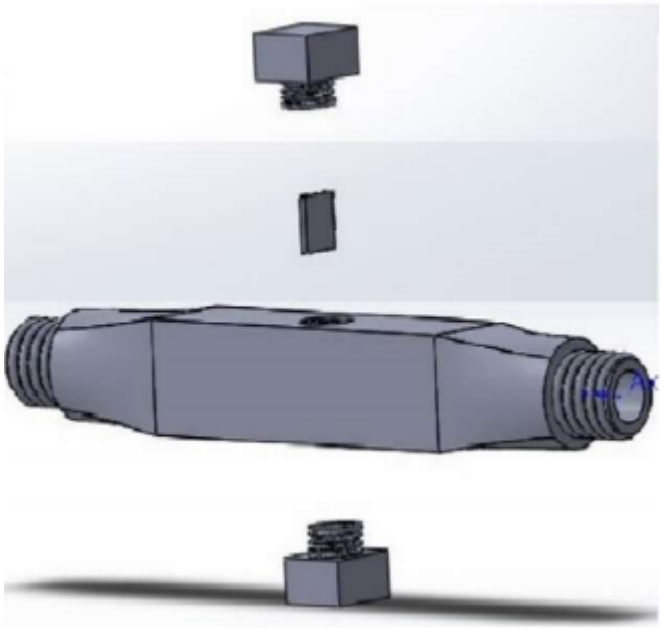


Figure 2-2: Passive Flow Control Device Design [24]

A control strategy for such an FCD was recently published by Kasukurthy et al. [76-79] using CFD where savings in pumping power of 64% was predicted.

Chapter 3 An Experimental Study Demonstrating Pumping Power Savings at Rack Level Using Dynamic Cooling

3.1 Abstract

Data center energy demands continue to rise with the increasing utilization of high-performance computing applications. As a result, a corresponding increase in power densities at the package level has necessitated advancements in IT equipment thermal management technologies. Cold plate-based liquid cooling allows high heat flux dissipation owing to very high convective heat transfer coefficients. A disadvantage of traditional liquid cooling in data centers is that a constant flow rate is provisioned to servers irrespective of their workloads. The present research investigates, empirically, the concept of provisioning targeted coolant flow rate to a server with a radiator fan arrangement. An air-cooled Open Compute Server was retrofitted with cold plates on and sealed completely to have no ambient interaction with the environment. The cold plates were provisioned with a targeted amount of coolant based on instantaneous CPU workload at different coolant inlet temperatures. The server design was able to keep the temperatures of all the components like DIMMs, PCH, and CPUs below their critical temperature values. The results show that a maximum cooling percentage reduction of 24% is achieved at maximum flow rate and minimum CPU utilization. The results also show that cooling power savings are also reduced at higher CPU workloads and coolant inlet temperatures.

3.2 IT Equipment Used

The server used in this study is a Winterfell hybrid cooled server with Intel v2.0 motherboard which has two CPUs and up to 16 Dual In-line Memory Modules (DIMMs) of installable memory as shown in Fig.3-1 (Miller, 2012). Each CPU has a rated Thermal Design Power (TDP) of 115 W. The server is retrofitted with active cold plates on the two CPU units which indirectly cool the CPUs and a fan-assisted radiator that re-circulates air within the server cooling the rest of the heat-dissipating auxiliary components as shown in Fig.3-2. The fans in the server circulate ambient air within the server cooling responsible for cooling the auxiliary components like DIMMs, Platform Controller Hub (PCH), Voltage Regulator Device (VRD), and Hard Disk Drives (HDDs). The fans are 4-wire pulse width modulation (PWM) controlled with a 975-10000 rpm range.

The baffle shown in Fig.3-2 forces the airflow through the DIMMS present on the left-hand side (LHS) as the airflow comes from the fans. In essence, the server becomes mainly liquid-cooled as the coolant entering the chassis passes through the radiator and two cold plates in series before leaving to the enclosed coolant reservoir. Each cold plate has a distributed pump (circled) integrated into it as shown in Fig.3-2. The flow is driven across the system utilizing these pumps. The rated power consumption and speed of the distributed pump are 3.6 W and 6250 rpm. For the current study, the chassis was populated with 8 DIMM cards of 2 GB each dissipating a TDP of approximately 4W. The physical experimental setup with its components is shown in Fig.3-3.

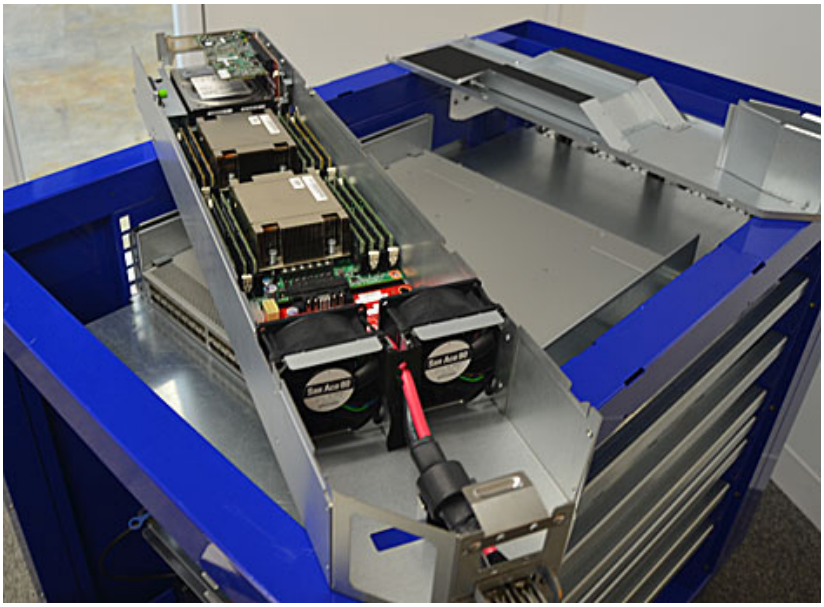


Figure 3-1: Air-cooled version of the original open compute server used in the study

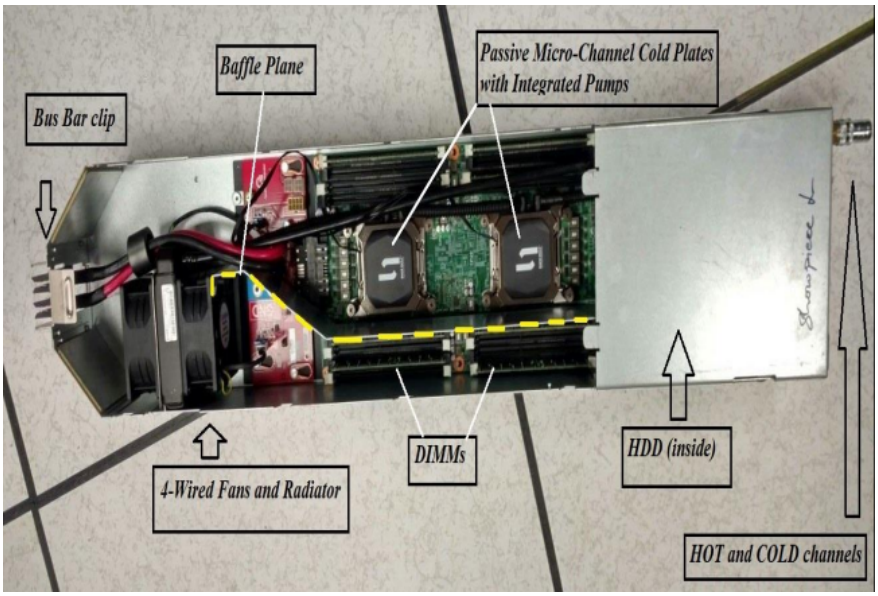


Figure 3-2: Modified air-cooled server retrofitted with cold plates and a baffle (shown with yellow dashed line) to separate cold and hot air paths within the server

Figure 3-3: Experimental setup showing the server cooling loop, pump control circuitry, and instrumentation for data acquisition

The primary cooling loop in the setup consisted of a coolant inlet from a custom-made reservoir towards the server as shown in the setup schematic in Fig.3-4. 20% inhibited ethylene glycol solution with deionized water was used as the coolant in current experiments due to ease of availability. A summary of relevant temperature-dependent thermophysical properties of the coolant used is shown in Table 1.

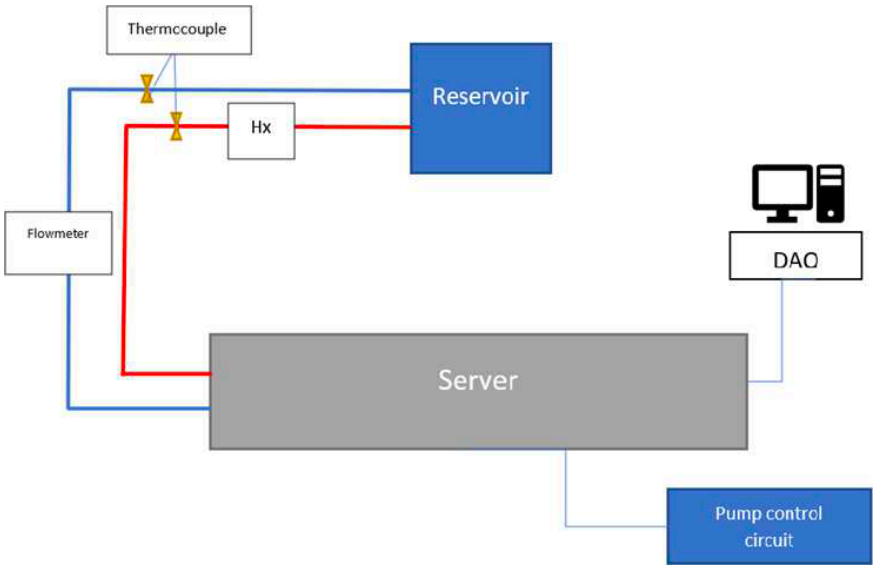


Figure 3-4: Schematic of Experimental Setup

Table 1: 20% Ethylene Glycol thermophysical properties at the experimental temperature setpoints

Property	Value at 25°C	Value at 35°C	Value at 45°C
Specific Heat (kJ/kgK)	3.82	3.84	3.86
Thermal conductivity (W/mK)	0.505	0.514	0.523
Density (kg/m3)	1032.32	1028.32	1024.59
Viscosity (Pa.s)	1.41 x 10-3	1.17 x 10-3	9.73 x 10-4

It should however be noted that while ethylene glycols offer better thermal properties, their higher toxicity as compared to propylene glycol and the inherent biocide behavior of propylene glycol are a few reasons they are not used widely in the industry. A liquid-to-liquid heat exchanger was used to fix and maintain the desired temperature setpoints. An Assured Automation DM-P series digital low flow meter was used to measure the flow rate of the coolant which was calibrated as per the field use with an accuracy of $\pm$1%. A commercially available Thermal Interface Material (TIM) was used to achieve a good thermal contact (k = 11 W/mK) between the CPU heat spreader and cold plate base.

3.3 Methodology

The distributed pumps were controlled using a dc power supply to maintain two extreme cases of minimum and maximum attainable flow rates based on the pump operating voltage. Table 2 provides a summary of various testing parameters that were used in the present investigation and their respective values used. K-type thermocouples were used to continuously monitor the inlet and outlet temperatures of the coolant using an Agilent Bench link datalogger. A WattsUp power meter was to continuously monitor the power consumption change in the server during the experimental run and the pump power was directly obtained from the dc power supply. To stress the server to the desired workload, an open-source stress tool for

windows, Prime95, was used. It tests the computer for stability issues by stressing the CPU to its maximum limit by running Lucas-Lehmer iteration indefinitely and only terminates a stress test when it encounters an error and informs the user that the system may be unstable. Prime95's stress test feature can be configured to test various system components by changing the Fast Fourier Transform (FFT) size. CoreTemp was used for monitoring real-time CPU core temperature when the server utilization varies. CoreTemp is completely motherboard independent and the temperature readings are a good estimate of the temperature values as the software is collecting the data directly from a Digital Thermal Sensor (DTS). DTS is located near the hottest part of each core. This sensor does not rely on an external circuit located to report temperature, all core temperature values are stored in the processor. Typical maximum case temperatures corresponding to 90°C DTS temperature have a value of 75-78°C.

Table 2: Experimental Parameters

S.No.	CPU+MEM Utilization	Inlet temperature	Flow rate
1	10%	25°C	0.2 lpm
2	50%	35°C	0.6 lpm
3	100%	45°C	

PCH and DIMM temperatures were monitored using two open-source sensor software. A total of 18 experimental runs were conducted by varying CPU utilization and flow rates for the three inlet temperatures chosen. After attaining the desired temperature in the cooling loop, the CPUs were stressed and the data logging was started to record core temperatures once the coolant temperature in the loop reached a steady state. The temperatures were recorded for 45 minutes and the maximum value of core, PCH, and DIMM temperatures were reported for each run.

3.4 Error Analysis

Each sensor in the study was calibrated using standard calibrating equipment and procedures. The K-type thermocouples used to measure the fluid temperature were calibrated using a Fluke 7109A portable calibration bath between a temperature of 0-100°C using a two-point calibration method. Table 4 shows the error calculation quantified from the calibration process for pressure sensors and the thermistors. It was observed that the pressure sensors were very precise after factory calibration and did not need additional calibration. A two-point calibration method was used for the thermocouples by calculating the error in the temperature reading. The calibration equation obtained was directly used as input in the DAQ software as gain and offset values. To calibrate the ultrasonic flow sensors, a Coriolis mass flow meter was used by placing the flow sensor in the same closed-loop along with the Coriolis flow meter. A Coriolis mass flow meter was used for calibration as its readings are universal since it measures the mass flow rate rather than just the flow rate with a typical accuracy for liquids between 0.05-0.1%. The results for calibration and error analysis for both the flow meter and the thermocouples used are shown in Table 3

Table 3: Summary of error analysis for the sensors used in the experiments

Reference Temperature (°C)	**Measured Temperature (°C)**	**% Error**
10	9.8	2
90	89.3	0.8
Reference Flow Rate at Coriolis (lpm)	**Measured Flow Rate at Sensor (lpm)**	**Error %**
0.2	0.207	3.5
0.5	0.519	3.8
0.7	0.721	3
1	1.03	3

3.5 Results

The post-processing and data analysis in the current research work included reporting core temperatures, and critical component temperatures, and calculating the cooling power consumption as a percentage of total IT power consumed by the server. Fig.3-5 shows the variation of maximum core temperatures with varying inlet temperatures between 25°C-45°C and at three different CPU utilization of 10%, 50%, and, 100%. The maximum core temperature reported at 45°C inlet temperature, 100% workload, and the minimum flow rate was 79°C which is well below the maximum junction temperature specified by the manufacturer for the processor. It was also noted that the core temperatures for 50% CPU utilization at 0.2 lpm and 100% CPU utilization at 0.6 lpm attained almost similar values for all inlet temperatures as can be seen in Fig.3-5. A similar case was observed for the cooling percentage curve as a function of the total IT power where the curve almost overlaps for the aforementioned parameters. The given data is sufficient to conclude that even the minimum attainable flowrate should be sufficient to maintain safe junction temperatures. Another important observation made from Fig.3-5 is that the core temperature at any utilization varies by 2-4°C only although the lpm is increased by three times. This further bolsters the claim that a target flow delivery based on server workload can yield energy savings with a slight increment in core temperatures. For low CPU workloads, where the core temperatures may not be that high, the reliability deterioration may be negligible due to an increase in the CPU core temperature.

Fig.3-6 shows the variation of total cooling power as a percentage of IT power for varying flow rates, CPU utilization, and inlet temperatures. The cooling percentage (sum of pump power and fan power) is defined as the percentage of power consumed for cooling the server when compared to total server power consumption. It can be seen that only 7% of total power is consumed by the fans in the extreme case of maximum inlet temperature and CPU

utilization. A notable conclusion drawn from Fig.3-6 is that the cooling percentage stays constant between 25°C-35°C for all CPU utilization and both the flow rates. This implies that there is no change in the fan RPM even at 35°C. Due to a lack of data points, it can't be substantially stated at what temperature the exact hike in cooling power occurs. The constant part of the graph in Fig.3-6 also implies that it would be safe to operate the servers at higher coolant inlet temperatures without reaching critical temperatures even at minimum flow rates.

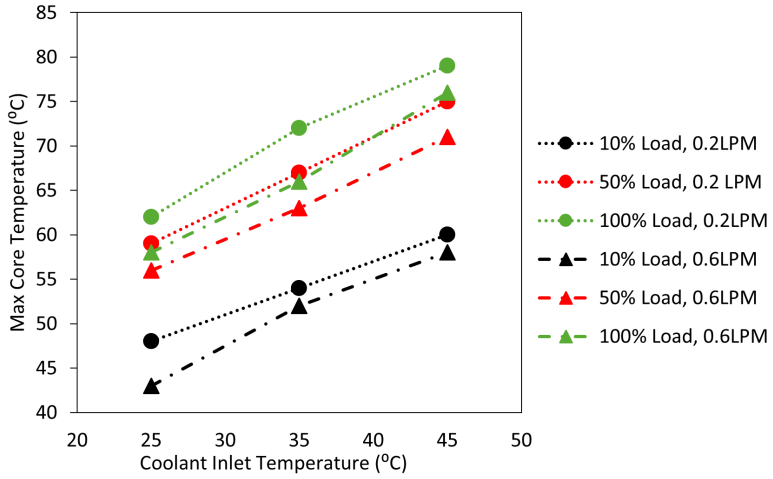


Figure 3-5: Maximum core temperatures at various CPU utilization and coolant inlet temperatures

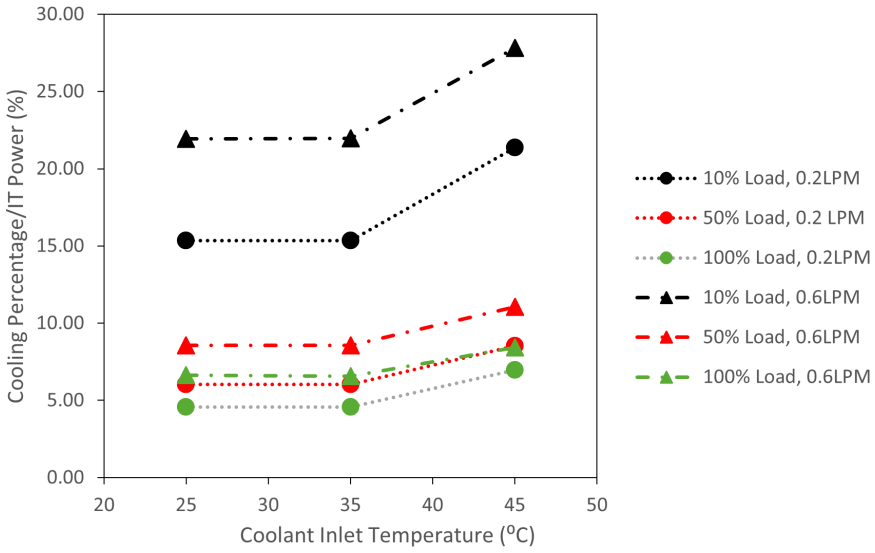


Figure 3-6: Variation of the ratio of cooling power percentage and percentage of IT power for different inlet temperatures

The data for variation in PCH and DIMM temperatures were obtained using two open-source sensor data software available for Windows OS as shown in Fig.3-7 and Fig.3-8. The maximum DIMM temperature obtained was around 62°C and the maximum PCH temperature reported was 69°C at minimum flow rates and maximum CPU utilization. All other values of component temperatures lay between these maximum points, represented by the shaded area on both the graphs, so were not reported for convenience. The maximum difference in temperatures for the extreme cases presented in both the figures presented is 5°C in the case of PCH and almost 8°C for DIMMs. The values of these maximum temperatures were well below the specified limits of safe and reliable operation. These results also bolster the fact that isolating the server from the ambient environment will not have a detrimental effect on the cooling efficiency. It's also evident from Fig.3-7 and Fig.3-8 that the trends followed by DIMMs and PCH temperatures are similar to that of the cooling percentage and maximum core temperatures respectively.

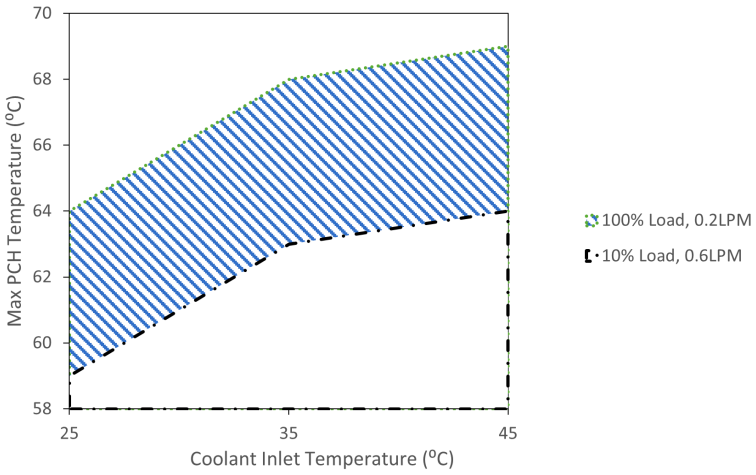


Figure 3-7: Maximum PCH temperatures at various inlet temperatures and CPU utilization

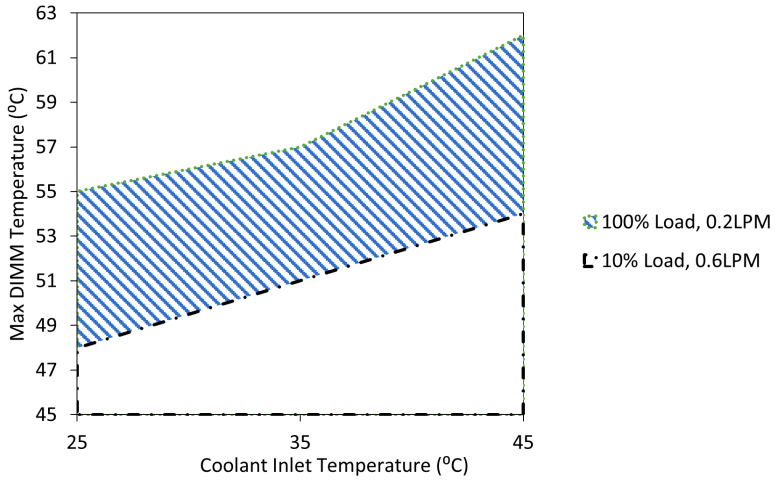


Figure 3-8: Maximum DIMM temperatures at different inlet temperatures

Fig.3-9 describes the variation of cooling percentage with the percentage change in percentage IT power at 45°C coolant inlet temperature. For 0.6 lpm the cooling percentage drops by almost 17% and approximately 12% for 0.2 lpm as the CPU utilization changes from idle to 100% workload. The difference in cooling percentage at minimum CPU utilization is found to be approximately 7% which reduces to 2% at maximum CPU utilization. The trend obtained shows that as the CPU utilization and inlet temperature increase, the cooling power savings reduces to almost 5-7% for both the flow rates. It can, thus, be concluded that there exists a trade-off value between the cooling power savings and equipment reliability due to higher component temperatures.

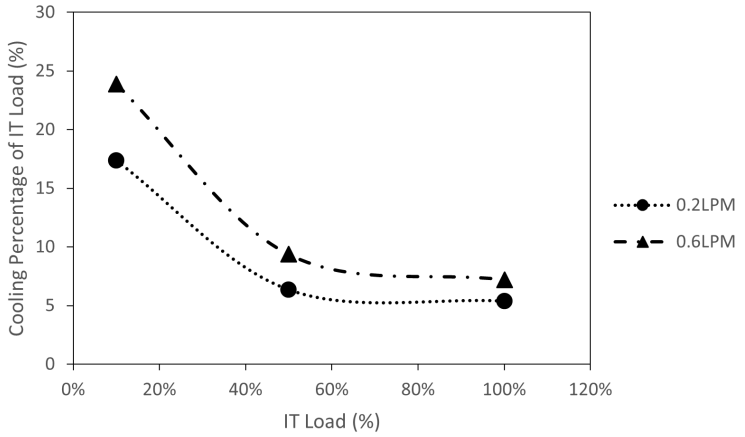


Figure 3-9: Cooling percentage vs total IT power percentage for different flow rates at a coolant inlet temperature of 45°C

The results obtained from the above experiments, as mentioned earlier, are the first step towards a novel concept of dynamic cooling or target flow delivery at the server level. This will be accomplished by using an actively controlled Flow Control Device (FCD) integrated at the outlet of each server in a rack as shown in Fig.3-10 (Shahi et al., 2022). This study details the design and development process of the FCD and characterizes its flow characteristics of the FCD. This device works on a temperature and a pressure-based control strategy. The device itself has an integrated v-cut ball valve attached to a shaft and servo motor assembly. The servo motor rotates the ball valve to desired angles based on the variation in outlet temperature of the warm coolant at the server. A server working at higher utilization will have a higher outlet temperature which will trigger the servo to rotate the ball valve, thus, allowing increased coolant flow rate to the server. Similarly, the coolant flow rate to servers working at a lower utilization will be reduced. The pressure drop changes across the rack due to increased flow resistance because caused by valve rotation gets compensated by varying the pump rpm. This overall optimization of pup power gives yields a lower pump power consumption per rack. For a data center with hundreds of such racks, the cost savings due to this energy optimization strategy will be significant. The experimental setup of a pseudo-rack arrangement integrated with 3-D printed FCDs is being developed by the authors to report pumping power savings for a single rack. The data obtained from these experiments will be used to derive an approximate value of savings that can be achieved at the data center level. A pumping power savings of 64% have also been reported using the above-mentioned control strategy computationally (Kasukurthy, 2019, Kasukurthy et al., 2021).

3.6 Conclusion

The main objectives of the present investigation were to test the concept of dynamic cooling by varying the flow rates of a server that is isolated from the ambient environment. CPU utilization, coolant inlet temperatures, and flow rates were varied and core temperatures

were constantly examined. The effect of an increase in inlet coolant temperature and recirculating inside air of the server was studied by monitoring the PCH and DIMM temperatures. The cooling efficiency by varying the pumping power at the component level was studied to substantiate the fact that controlling the flow at the server level based on the server workload utilization can yield energy savings. Both PCH and DIMM temperatures were successfully maintained well below the critical temperature limits. The maximum temperature obtained at PCH in the extreme case of bare minimum flow rate and 100% CPU utilization was 69°C. The hot circulating air inside the server even at high inlet temperatures kept the DIMMs temperature to a maximum of 61°C. No change in cooling power consumption was observed for till 35°C after which the cooling power increases. Higher-resolution experimental readings between 35°C-45°C will give a better insight into what exact value of temperature this change happens. The cooling power savings reduce with an increase in inlet temperature, as expected, and become constant for higher IT loads.

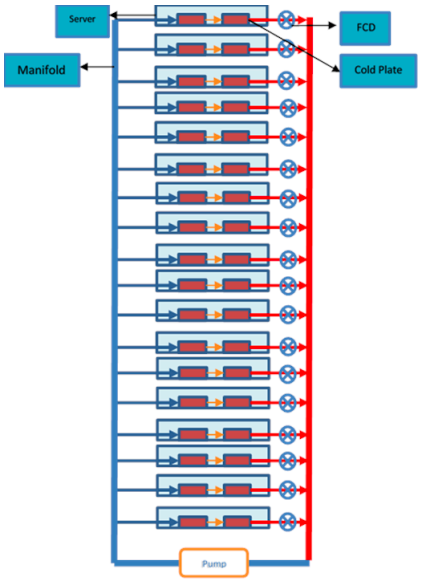


Figure 3-10: The operational strategy for rack level dynamic cooling using Flow Control Device

The proposed concept of dynamic cooling of servers that are isolated from the surrounding environment can offer significant advantages in terms of cooling and improved

reliability. The data center room will still require marginal cooling due to heat radiated from the server chassis which can be cooled using outside air throughout the day with standard filtration techniques. The biggest issue of the impact of gaseous and particulate contamination related to free air-cooled data centers (Saini et al., 2022) on IT equipment can be minimized using the proposed cooling technique. This is for the first time the concept of dynamic cooling at the rack level is being introduced to further improve energy savings in liquid-cooled data centers.

Chapter 4 Design, Development, and Characterization of a Flow Control Device for Dynamic Cooling Liquid-Cooled Servers

4.1 Abstract

Transistor density trends till recently have been following Moore's law, doubling every generation resulting in increased power density. The computational performance gains with the breakdown of Moore's law were achieved by using multi-core processors, leading to non-uniform power distribution and localized high temperatures making thermal management even more challenging. Cold plate-based liquid cooling has proven to be one of the most efficient technologies in overcoming these thermal management issues. Traditional liquid-cooled data center deployments provide a constant flow rate to servers irrespective of the workload, leading to excessive consumption of coolant pumping power. Therefore, a further enhancement in the efficiency of implementation of liquid cooling in data centers is possible. The present investigation proposes the implementation of dynamic cooling using an active flow control device to regulate the coolant flow rates at the server level. This device can aid in pumping power savings by controlling the flow rates based on server utilization. The FCD design contains a V-cut ball valve connected to a micro servo motor used for varying the device valve angle. The valve position was varied to change the flow rate through the valve by servo motor actuation based on pre-decided rotational angles. The device operation was characterized by quantifying the flow rates and pressure drop across the device by changing the valve position using both CFD and experiments. The proposed FCD was able to vary the flow rate between 0.09 lpm to 4 lpm at different valve positions.

4.2 Design and Development

The primary motivation behind this work was to develop a flow control device that can contribute toward pumping power savings with a simple, cheap, and reliable design. Air-cooled data centers use actively controlled dampers to supply the target value of airflow rates to servers. Target delivery of flow rate not only optimizes the coolant transport energy cost but also reduces over-provisioning of coolant, thereby, reducing energy expenditure related to heat removal. The purpose of this paper is to propose the design of a flow control device to control the flow of liquid in a cooling loop that can be directly integrated with liquid-cooled servers to save the energy required to pump the coolant. Another motivation behind the design of this FCD was the ease of manufacturing at a large scale and the ability to be manufactured using 3-D printing.

During the conceptual design phases, many design concepts were taken into consideration for an optimal flow control design and as per a previous study made in this area using a butterfly valve as shown in Figure 4-1 where a prototype is including a servo motor attached at the top to control the flow of the device.

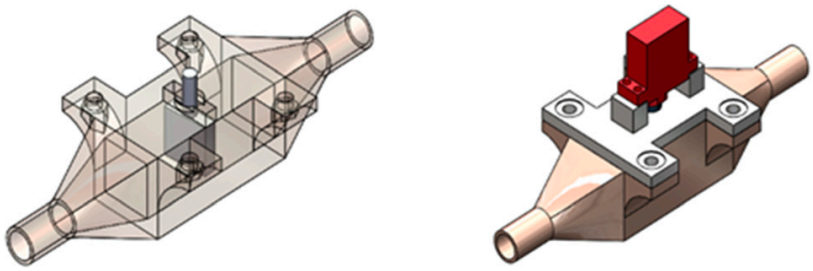


Figure 4-1: CAD model of the first concept design tested for the FCD where a Butterfly Valve design was used for the control valve

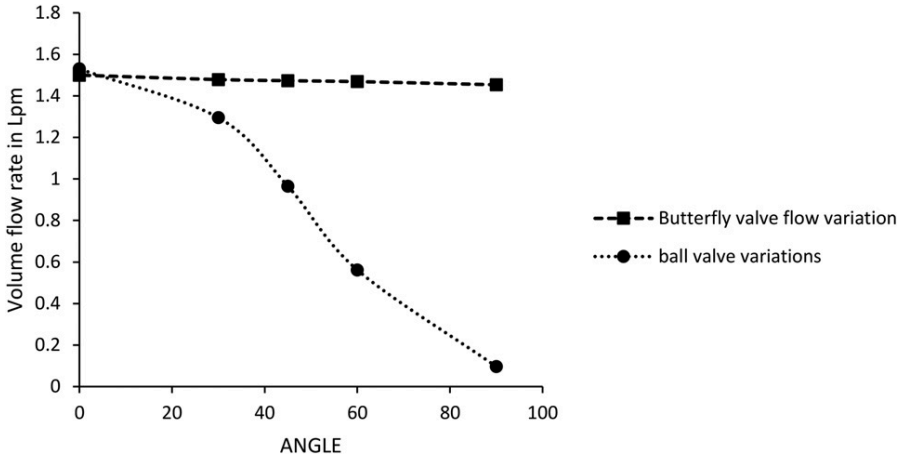


Figure 4-2: Comparison of flow rate variation between the final design of the FCD and the initial design, with the final design showing the desired flow rate variation with changing valve angle.

The model was first tested using computational fluid dynamics testing (CFD) and found that variation in the flow is very small as shown in Figures 4-2. This variation led us to design a newer concept design to suit high variation in flow and high-pressure applications. Efforts were made to study the best possible flow control characterization system that can be best suited for data center purposes. From this study, it was found that ball valves with V cut are the most common use flow control device in the industry. Since the opening in the ball valve cannot produce precise control overflow, it is mainly used for on/off purposes. The recent introduction of the V port ball valve significantly improved the linear response of the flow and introduce wide ranges in terms of flow rate. After analysis and experimentation on the butterfly valve design, it was known that the variation inflow is very small, and hence a new design was chosen which had a ball valve with a V cut. The device is constructed of four parts: two symmetrical casing parts, one ball valve with a V cut, and a shaft sealing for a watertight seal as shown in Figure 4-3.

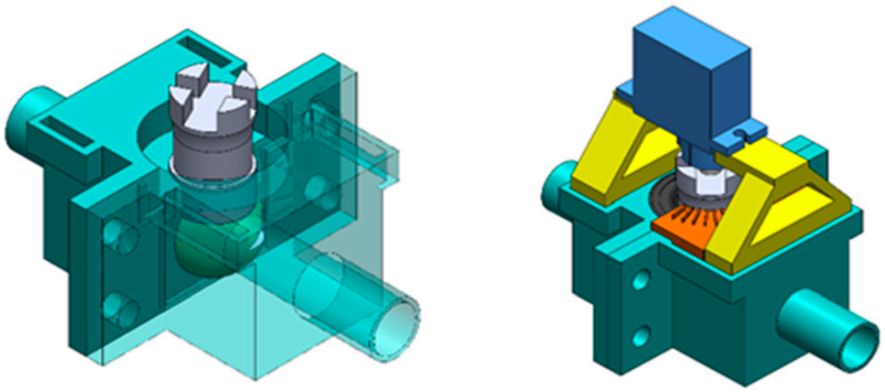


Figure 4-3: CAD model of the final FCD design

After a proper study of the ball valve flow control device, a design was chosen. The inlet and outlet ports of the device were determined by considering standard pipe diameters, 3/8’ and a 10 mm ball valve. The detailed design of FCD was made using SolidWorks and the data file was exported to an ANSYS FLUENT, a commercial CFD package, for further analysis including the determination of the flow characteristics of the FCD designs. After obtaining the required results, the SolidWorks design file was exported to a slicing software and was manufactured using extrusion 3D printing technology. The advantage of 3D printing is that it has a low manufacturing time compared to conventional manufacturing techniques and the production cost is very low. Also, design changes can be made without considering the tooling costs and extra lead time, and related high costs.

4.3 Computational Modeling

Over the years, CFD has become an important tool to ascertain the design flaws and optimize them further rather than going through intensive experimentation with manufactured models. To visualize the 3-D flow inside the FCD and analyze the flow rate variation by changing the angle of the ball valve, momentum, energy, and, mass conservation equations were solved using ANSYS FLUENT [24].

After preparing the CAD model on SolidWorks, the model was imported to the CFD software where it was first meshed using the ANSYS Mesh tool and then solved in the CFD solver. The following sections will detail the model setup and solution process in detail.

4.3.1 Meshing

After importing the geometry to the CFD package, the meshing of the flow domain was carried out. A good quality mesh is essential for the accuracy of results in CFD modeling. An efficient mesh is composed of a sufficient number of elements near the walls, especially for fluid flows, to resolve the shear effects in the near-wall region. Since the flow variation occurs only in the internal cavity of the device, it was not necessary to include all the components in the meshing and CFD solution. For this purpose, the fluid volume was extracted and the solid bodies of the design were suppressed. This was repeated for various cases to be simulated in this study, generating four different tetrahedral meshed volumes for four angles studies for flow rate variation. To verify the quality of the mesh, the element skewness and orthogonality were monitored. For a mesh to be considered to be of good quality, the value of orthogonality should be closer to 1, and mesh skewness should be close to 0. The minimum element size was set to 0.3mm and the maximum element size was 0.6mm. The growth rate was set to the default of 1.2. for inflation, with the number of inflation layers as 10, and the growth rate was set to a default value of 1.2. In all of the simulation cases studied, average orthogonality of 0.8 and average skewness of 0.2 was reported. The optimum mesh element count of 739,416 was used based on the mesh independence study as detailed in Table 4. The final mesh characteristics used in the study are summarized in Table 5. A summary of boundary conditions used for setting up the CFD and FEA simulation domain is shown in Figure 4-4

Table 4: Mesh Independence Results at 45° open-angle at 1 lpm flow rate

Number of Elements	Pressure Drop (KPa)
315,897	10.98
563,217	11.63
739,416	12.97
935,569	13.0

Table 5: CFD mesh characteristics for extracted fluid volume

Volume	5898.8 mm3
bodies	3
Active bodies	1
Nodes	142942
Element	739416

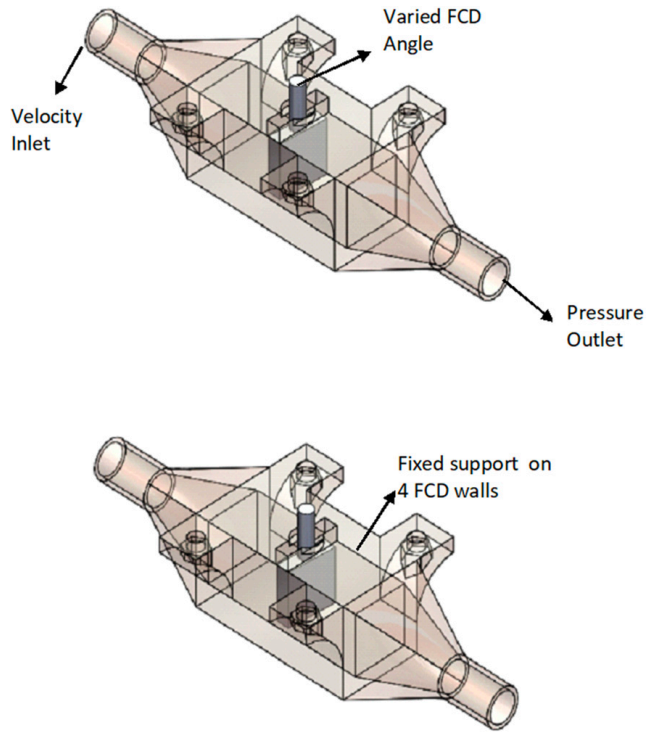


Figure 4-4:Summary of boundary conditions used in CFD (top) and FEA (bottom)

4.3.2 Computational Fluid Dynamics and Finite Element Analysis Setup

As the purpose of the CFD simulation was hydraulic verification of the FCD, the CFD simulation was done by deactivating the energy equation. This further deactivates all the thermal inputs possible in the simulation and saves considerable simulation time as well. The

pressure drop and flow rate characterizations were carried out at a constant temperature of 22°C due to which a constant value of the thermo-physical properties was used in the CFD solver. A pressure-based solver algorithm was used to simulate a velocity field through the FCD by a correction in the pressure equation in the continuity equation. Various turbulence models have been implemented in the literature based on specific applications. Literature shows that, for cases with internal flows, $k-\omega$ turbulence model is preferable. Standard and Shear Stress Transport (SST) are the two available variants of the $k-\omega$ model. It was believed that since flow across the valve in various positions will create large adverse pressure gradients, the SST $k-\omega$ model should be used because of its superiority in predicting such flows. This has also been validated in literature where a comparison of various turbulence models was validated for a butterfly valve. This study showed that the SST $k-\omega$ model predicted the values of the drag coefficient across the valve with the least error when compared to the standard experimental results. Also, the CFD results become less sensitive to grid size when using this turbulence model. The general governing equations of continuity and momentum used by the solver to are given below in equations (1) and (2).

Continuity equation:

$$\frac{\partial \rho}{\partial t} + \nabla.(\rho u) = Sm \tag{1}$$

Momentum:

$$\frac{\rho d\vec{u}}{dt} = -\nabla p + \nabla \cdot (\mu \nabla \vec{u}) + \vec{f} \tag{2}$$

Transport Equation for the SST $k-\omega$:

$$\frac{\partial}{\partial t}(\rho k) + \frac{\partial}{\partial x_i}(\rho k u_i) = \frac{\partial}{\partial x_j}\left(\frac{\partial k}{\partial x_j}(\Gamma_k)\right) + \tilde{G}_k - Y_k + S_k \tag{3}$$

$$\frac{\partial}{\partial t}(\rho\omega) + \frac{\partial}{\partial x_i}(\rho\omega u_i) = \frac{\partial}{\partial x_j}\left(\frac{\partial \omega}{\partial x_j}(\Gamma_\omega)\right) + G_\omega - Y_\omega + D_\omega + S_\omega \qquad (4)$$

In equations (3) and (4), $\tilde{G}_k$ represents the generation of turbulence kinetic energy due to mean velocity gradients and G_ω represents the generation of ω. Γ_k andΓ_ω represent the effective diffusivity of k and ω, respectively. Y_k and Y_ω represent the dissipation of k and ω due to turbulence and D_ω represents the cross-diffusion term. S_k and S_ω are user-defined source terms which were not utilized in the current CFD modeling setup. After choosing the relevant turbulence models, the fluid boundary conditions were chosen based on the anticipated experimental test conditions. The coolant properties of 25% Propylene Glycol (PG25) are used in the simulation to determine the variation in flow rates and pressure drop with changes in the angle of the valve. A summary of the thermophysical properties of the temperature-dependent properties of the coolant is shown in Table 6. In the boundary conditions, the inlet velocity was calculated using the flow rate for the designed area. The outlet was set as a static pressure outlet with 0Pa gauge pressure. The walls of the device were treated as adiabatic walls. A steady-state simulation with convergence criteria of 1e-3 for continuity and 1e-6 for 3-D velocities, k, and ω values was used.

Table 6: Thermo-physical properties of 25% Propylene Glycol in temperature range of interest

Temperature (°F)	**Density (kg/m^3)**	**Viscosity (cP)**	**Specific Heat (J/kg°C)**	**Thermal Conductivity (W/mK)**
70	1023	2.45	3914	0.470
80	1020	2.05	3930	0.475
90	1018	1.74	3943	0.480
100	1015	1.49	3955	0.486
120	1009	1.14	3985	0.489
140	1002	0.90	4010	0.501
160	995	0.73	4039	0.506

Table 7: Material properties of 3-D printing material used for Structural analysis of the

Density	1.04e-006 kg/mm^3
Youngs modulus	239.. MPa
Poisson's Ratio	0.399
Bulk modulus	3943.9 MPa
Shear Modulus	854.18 MPa
Tensile Yield Strength	41.4 MPa
Tensile Ultimate Strength	44.3 MPa

A stress test was performed in ANSYS Static structural to find the working limits of the flow control device for anticipated flow rates. Internal pressure loads on the walls of the FCD were imported from the results of CFD simulations. A tetrahedral mesh with physical preference set to mechanical element order was chosen and was set to program control. The minimum mesh element size was set to 0.5mm with a target mesh quality of 0.0500. The mesh smoothing was set to medium, resulting in 912689 nodes and 534030 mesh elements. The ambient temperature of the environment temperature was set to 22°C. A fixed support boundary was given on the sidewalls and the pressure load was imported from the CFD simulation running for the highest flowrate tested, which was 4 lpm. The material properties of ABS plastic were selected as shown in Table 7. The solver uses the relation given in equation (5) for solving linear static structural analysis problems. The basic assumptions in this theory include a linear elastic behavior and small deflections of the material.

$$[K]\{x\} = \{F\} \qquad (5)$$

Here [K] represents the stiffness matrix of the entire geometry, {x} represents the unknown nodal displacements and {F} represents the external forces or applied loads.

4.4 Experimental Setup

Experimentation was carried out in a controlled environment. The cooling liquid used for the loop is PG25 and care was taken to avoid contamination to get precise flow rates. A simplified schematic of the experimental setup is shown in Figure 4-5.

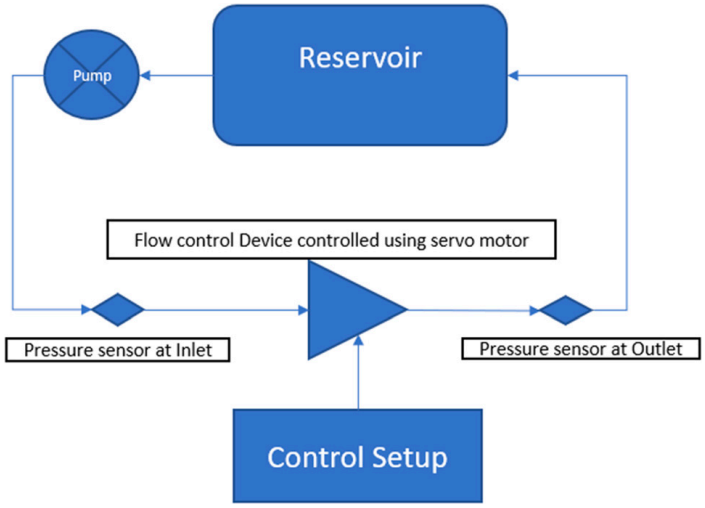


Figure 4-5: Schematic of the experimental setup used for characterizing the flow control device

Pressure sensors were attached near the inlet and outlet of the flow control device to determine the pressure drops for various valve angles. The flow rates were obtained using ultrasonic flow sensors which use ultrasonic frequencies to calculate the flow rates. The loop also consists of a sealed coolant reservoir to hold the cooling liquid. A description of the models and sensor accuracy is shown in Table 8. A 15 W DC pump was used to circulate the fluid through the entire test loop. Variation in flow rates across the FCD was achieved using changing the input voltage to the pump. The pressure and flow sensors were connected to the Data Acquisition system (DAQ). The DAQ is used for sampling electric signals such as voltage, current, and frequency and converting these electrical signals into numerical values such as temperature and pressure using the software. The flow and pressure sensors supplied a voltage of 20V from a different DC power source. The angle of the valve was changed with the help of a servo motor which is controlled by the Arduino microcontroller. The Arduino was

programmed using Arduino IDE in such a way that a change in the resistance value of the potentiometer will change the angles of the valve. Observations were recorded at valve angles of 0, 30, 45 & 90 degrees. Data regarding pressure drop from pressure sensors and variation in flow rate with a change in valve angle were also recorded. The final fabricated FCD and its integration in a cooling loop are shown in Figure 4-6.

Table 8: 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	20V – 30V	0.02L/min - 20L/min	± 0.003ml/min
Honeywell Pressure sensors	4.75V – 5.25V	0 – 50 Psi	± 0.25%
K –Type Thermocouple	_	0 – 400 ˚C	± 0.75%

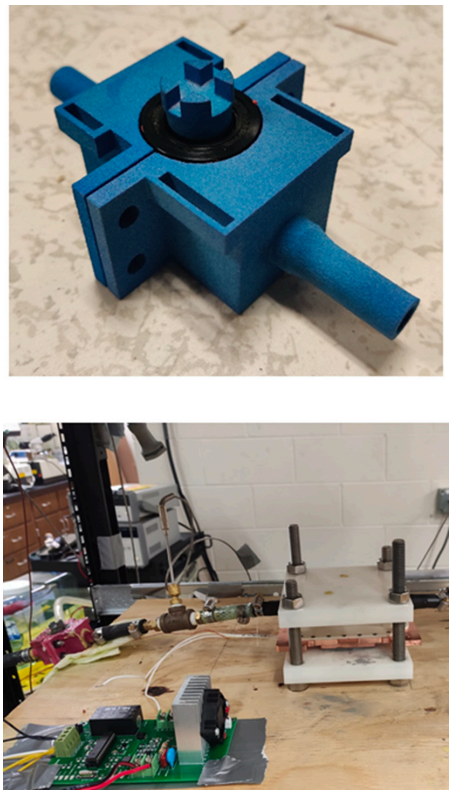


Figure 4-6: Final assembly of the 3-D printed FCD (top) and the integration of the FCD with one of the TTV and cold plate assembly in the rack (bottom)

4.5 Results and Discussion

4.5.1 CFD and FEA Results

After achieving convergence with an appropriate mesh as discussed earlier with the desired accuracy, the CFD results were analyzed in the form of the pressure drops across the device for different valve angles. This variation in the valve angle was repeated for the three chosen values of flow rates of 0.6, 1.0, and 1.5 lpm. The working of the FCD design was verified from the CFD results where a change in flow rate was obtained at all angles varied in the simulations. Figures 4-7,4- 8, and 4-9 depict the variation of velocity at valve angles of 0, 75, and 45 degrees respectively. The results obtained from the simulations were compared with experimental data and showed a close agreement in the flow rate variation trend with changing valve angles. As opposed to the first design of the FCD, it was observed that with the closing of the valve the flow rate decreased across the valve, and the pressure drop increased. The observations obtained for a particular condition had similar values with average error possibilities of 14%. The data in Table 9 summarizes flow characteristics when the flow rate is 1 lpm when the valve is fully open at a 0° angle.

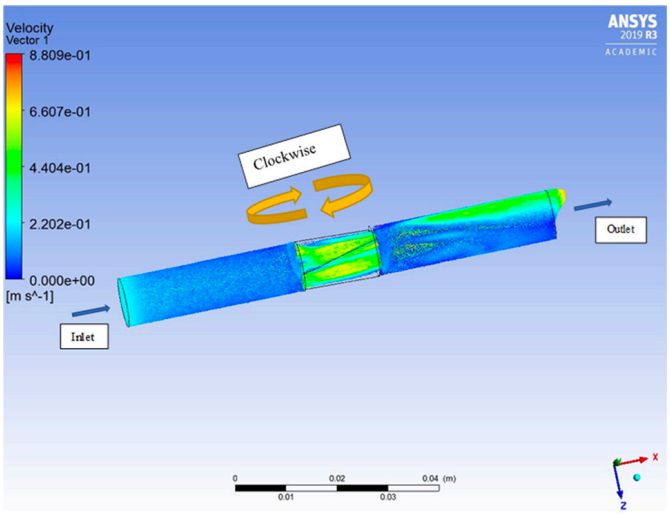


Figure 4-7: Flow characteristics of the FCD showing velocity profile inside the device cavity when the valve is fully opened

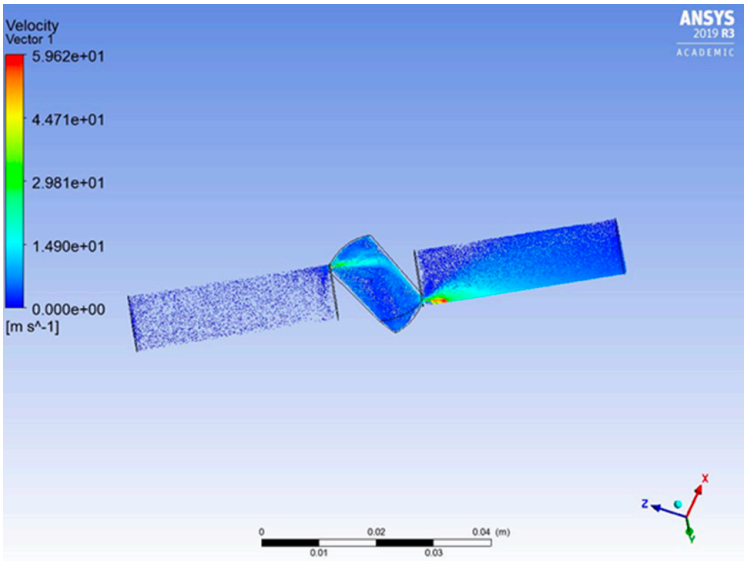


Figure 4-8: Flow characteristics of the FCD showing velocity profile inside the device cavity at a valve opening angle of 75°

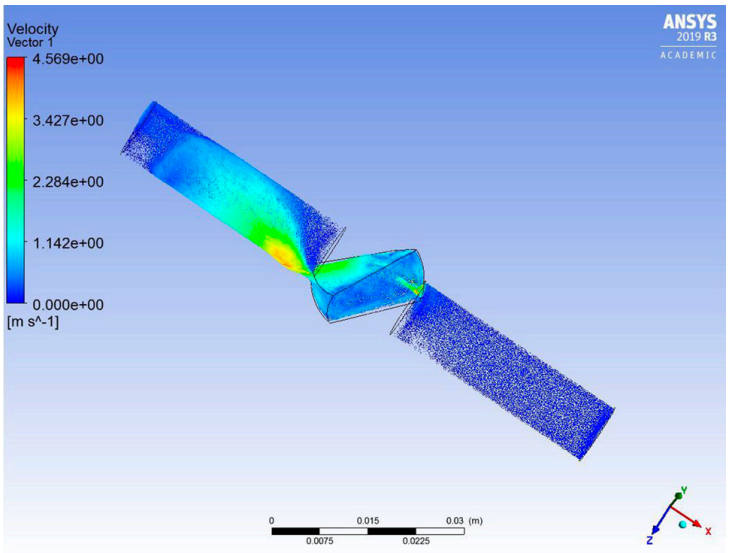


Figure 4-9: Flow characteristics of the FCD showing velocity profile inside the device cavity at a valve opening angle of 45°

The stress analysis was done using the Static Structural module in ANSYS Workbench. The CFD data of the forces exerted by the fluid flow on the valve were exported to the structural module of the software. This was done to verify the structural integrity (based on material fill ratio in 3-D printing) of the device at high lpm can lead to leakages or the device or the device bursting under pressure. The ambient conditions for the simulation setup were kept at a constant

temperature of 22°C. A fixed support boundary condition was given on the device's outer walls and the internal pressure force was imported from the CFD simulation for the highest flow rate, 4 lpm, to identify the maximum stresses and deformations. The results of the FEA simulations are shown in Figures 4-10, 4-11, and 4-12. The value for maximum deformation, observed near the inlet port, was 7.9 x 10-4 mm with maximum equivalent stress of 0.94 MPa, and the average deformation was 1.13 x 10-4 with average stress of 0.076 MPa. The average factor of safety for the device was obtained to be 15. The results from the FEA analysis verify that the FCD body with current dimensions and material will not fail at pressure conditions at the

maximum flow rate of 4 lpm. This, however, needs validation experimentally at different temperatures for prolonged periods and cyclic flow rates (internal pressure) to ascertain the device lifetime and failure rates.

Table 9: Summary of CFD results for pressure drop across the FCD at different valve angles

At 1.0 Lpm @ fully open				
v=0.2122 m/s	**Inlet pressure (kPa)**	**Outlet pressure (kPa)**	**Flow Rate (Lpm)**	**Pressure drop (kPa)**
Angle				
0	12.148	10.125	0.99	1.53
30	13	8	0.825	3.34
45	16	2.37	0.512	12.97
90	20	1.43	0	15.9

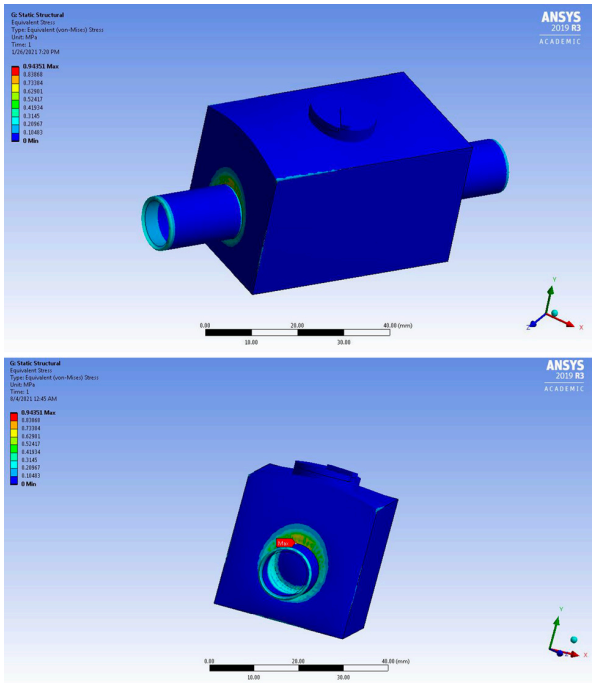


Figure 4-10: Equivalent(von-mises) stress on the FCD walls at a flow rate of 4 lpm in isometric (top) and front view (bottom)

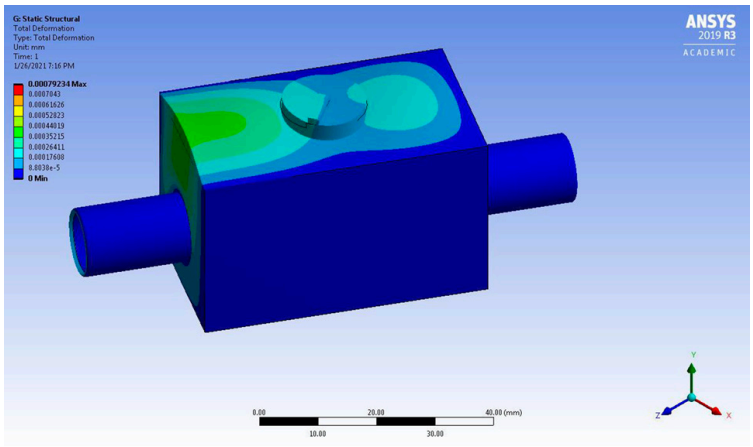


Figure 4-11: Contours for total deformation in FCD material with a maximum deformation towards the inlet and around the valve shaft

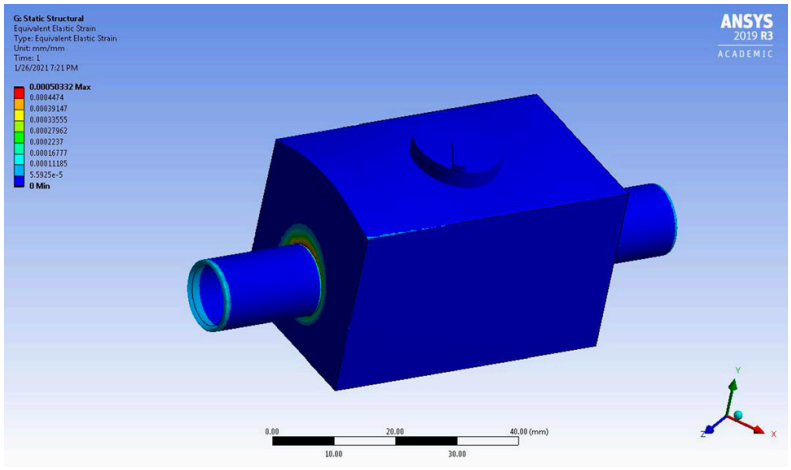


Figure 4-12: Contours for the equivalent elastic strain on the FCD from analysis

4.5.2 Experimental Results

To ascertain the results obtained for the hydraulic performance of the proposed FCD, an experimental study was carried out under the same flow conditions. The pressure drop characteristics and flow rate variation at different valve angles are shown in Figure 4-13, Figure 4-14, Figure 4-15 and summarized in Table 10. The results showed that for a maximum flow rate of 4.0 lpm, a minimum flow rate of 0.2 lpm can be achieved when the FCD is in the fully closed position. It was observed that the maximum change in flow rate occurs between the range of 30°-90°. The largest variation is obtained for the higher value of flow rates but reduces and becomes rather constant for flow rates less than 0.5 lpm.

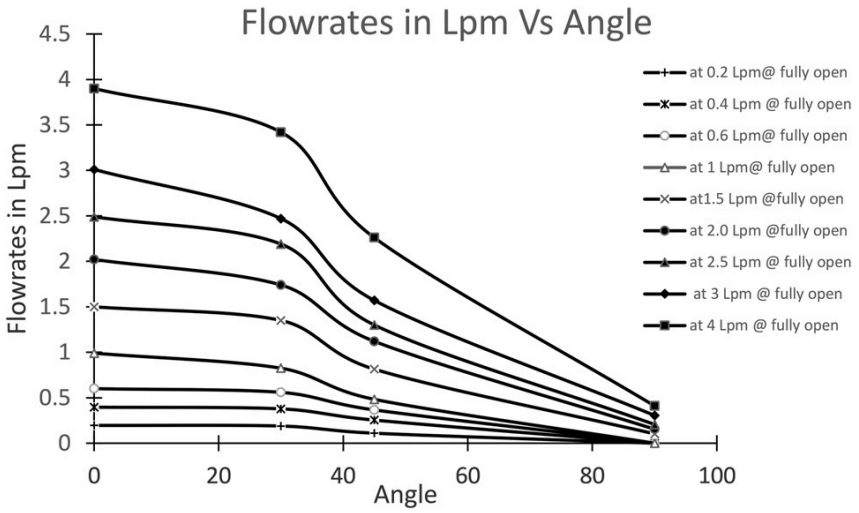


Figure 4-13: Variation of the flow rate through the FCD with the changing angle of the valve

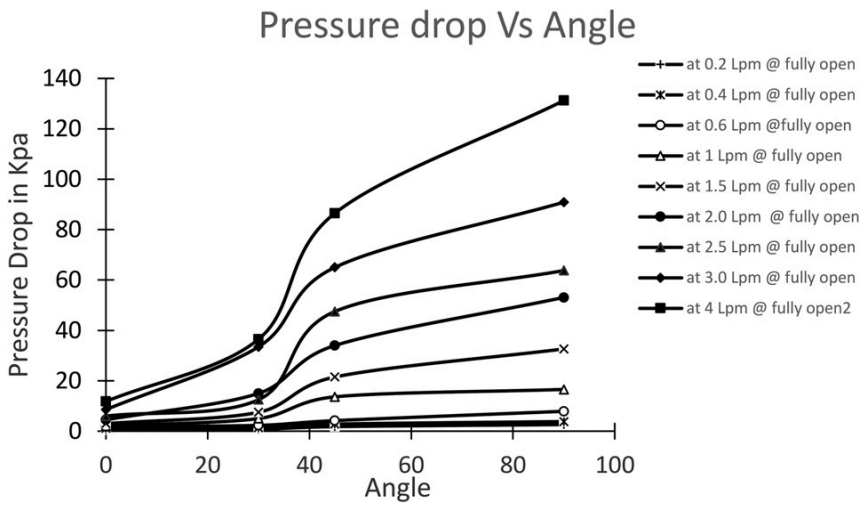


Figure 4-14: Pressure drop variation across the FCD with changing valve angle

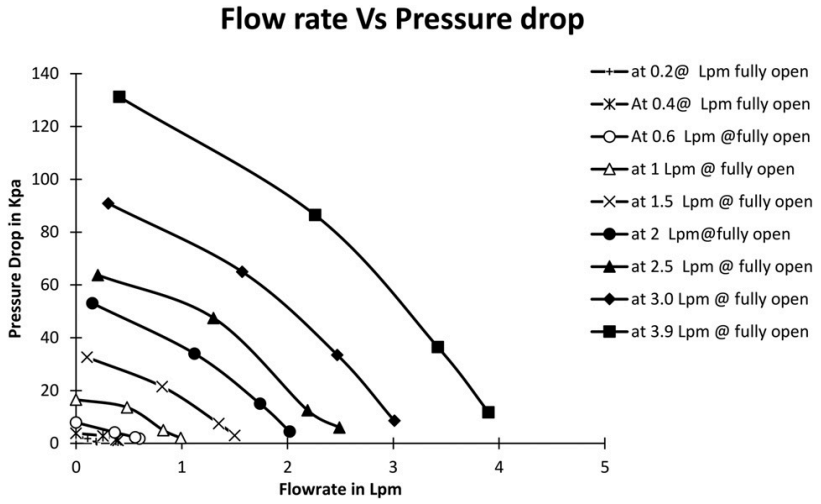


Figure 4-15: Device impedance curve with varying flow rates with the valve is in a fully open condition

A comparison of results from the experimental study and CFD analysis is shown in Figure 4-16, Figure 4-17 and, Figure 4-18. The comparison of results of flow rates with varying angles of the ball valve shows that a maximum discrepancy of 20% was observed in both the results at a flow rate of 1 lpm. This variation reduces with increasing the angle of the valve or when the valve is fully opened. A large variation in the results at fully closed conditions could be due to unresolved flow disturbances due to the current mesh size. A similar trend is observed in the pressure drop characteristics with varying angles of the valve. This is expected as there

is a linearly inverse relationship between pressure drop and flow rate. The maximum error between the pressure drop variation was 16% for at 1 lpm flow rate when the valve is fully closed. The error between the results when the valve was fully opened was within 3% for both pressure drop and flow rate. This is because there could be little to no flow turbulence created by the valve and the CFD code was successfully able to replicate the experimental flow conditions. Based on the plotted graphical data, it was observed that with a change in angle, there is a drop in the flow rate, and the pressure drop across the valve increases. The major variation in pressure drops and flow occurs between the angle of 30o to 90o angle. Similar trends were seen in experimental and CFD data when calculated for a starting flow rate of 0.6 Lpm, 1 Lpm & 1.5 Lpm when the valve is fully open.

Table 10: Summary of experimental results for pressure drop variation with changing valve angle

At 1Lpm @ fully open					
	Inlet Pressure (kPa)	**Outlet pressure (kPa)**	**Flow Rate (Lpm)**	**Experimental Pressure drop (kPa)**	**CFD Pressure drop (kPa)**
Angle					
0	11.98	10.425	1.5296	2.023	1.53
30	12.456	9.125	0.865	5	3.34
45	16.979	4.005	0.512	13.63	12.97
90	18.563	2.663	0	18.57	15.9

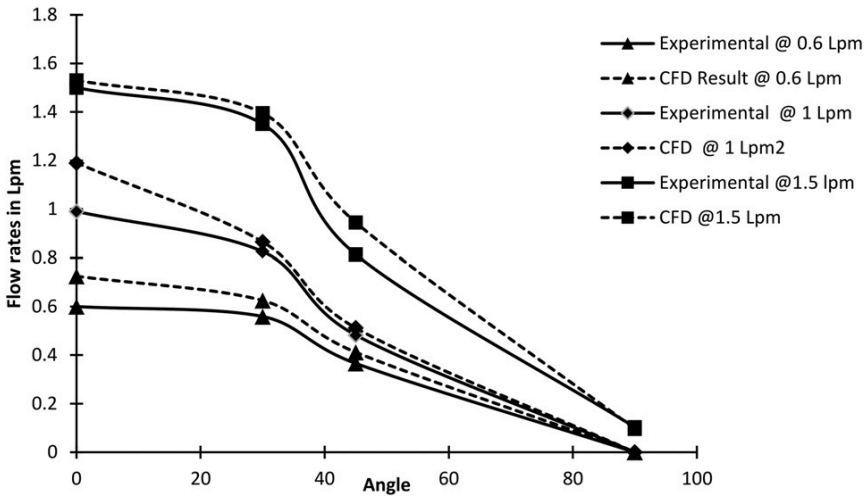


Figure 4-16: Comparison of the trends in CFD and experimental result for flow rate variation with changing valve angle

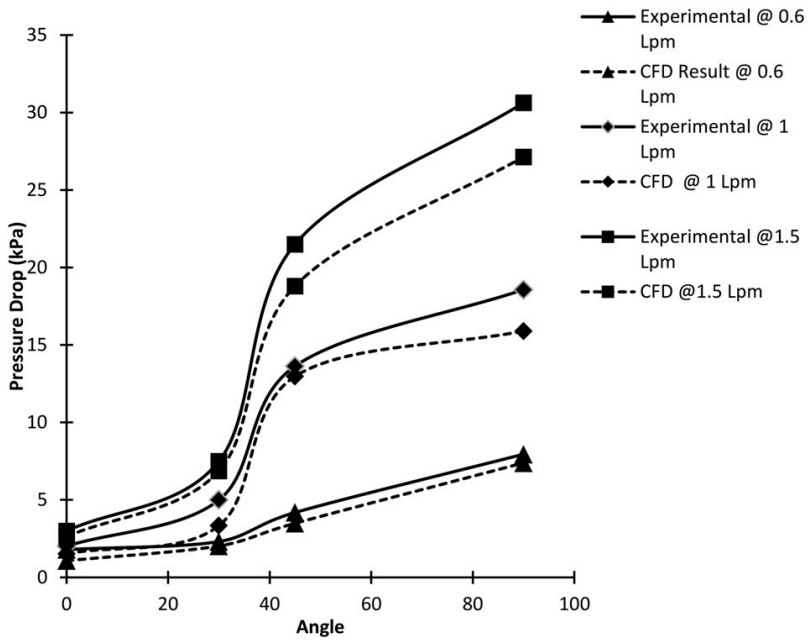


Figure 4-17: Comparison of CFD and experimental results for variation in pressure drop with changing valve angle

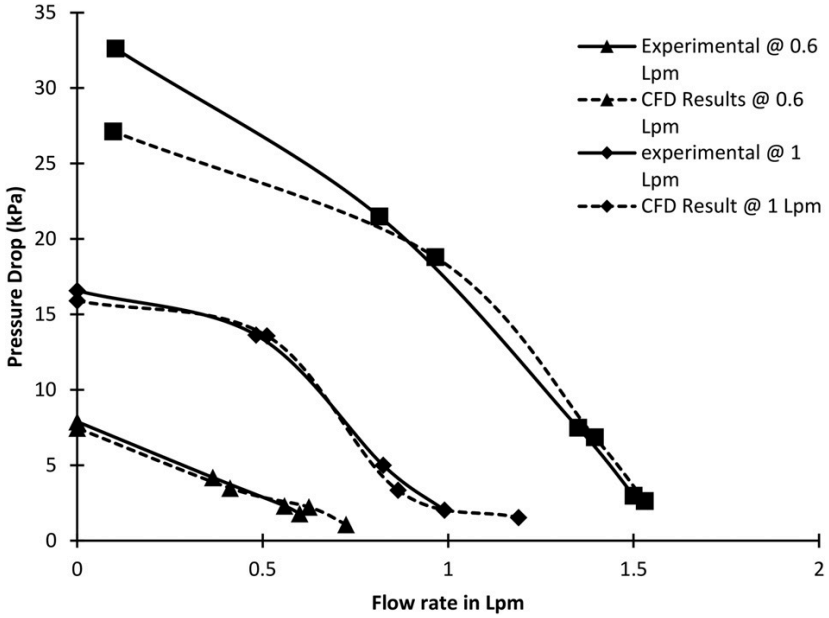


Figure 4-18: Comparison of the impedance curve obtained from CFD analysis and experimental data

4.6 Conclusion and Future Work

DLC using cold plates continues to be the most popular method to overcome the challenges of thermal challenges in high-performance computing systems. While DLC by itself provides an energy-efficient way of dissipating high heat fluxes, further enhancements are possible such as reducing energy consumption related to coolant transport within the data center. Traditional DLC only provides a constant flow rate, typically using a centralized pumping system, that delivers a constant flow rate to each server irrespective of the instantaneous workloads of the servers. DLC also provides notable advantages and better equipment reliability of ITE as issues related to contamination and fan acoustics [29-33]. It also provides higher heat flux values than single-phase immersion cooling and overcomes the issues related to material compatibility and vapor pressure in 2-phase open bath immersion cooling and 2-phase cold plates.

The FCD design proposed in this study will enable workload-based provisioning of the coolant to each of the servers in a rack. A recent study by the authors on an open compute server retrofitted with cold plates with distributed pumps showed considerable power savings when the flow control at the server level is dynamically varied. The proposed FCD will therefore yield considerable pumping power savings by optimizing the pumping power of the rack-level pumps. Out of the two final designs developed, the FCD design with a V-cut ball valve produced the desired flow rate variation that would be sufficient to cool heat loads in the current high-powered electronic packages. The results from the CFD study and experimental testing show a good agreement in depicting the hydraulic performance data of the FCD. A minimum flow rate of 0.09 lpm and a maximum flow rate of 4 lpm was achieved by varying the angle of the valve inside the FCD cavity. Future work on this subject will investigate FCD implementation at the server and rack level for flow control at the rack level to assess the thermal and flow characteristics along with the maximum achievable power savings. The future

work on the FCD will also present an overview of the possible cost savings not only due to reduced pumping power but also provide an estimate of a simple payback period after device deployment. Further studies will also focus on device reliability under operating conditions to assess the device lifetime in the field under transient operating conditions.

.

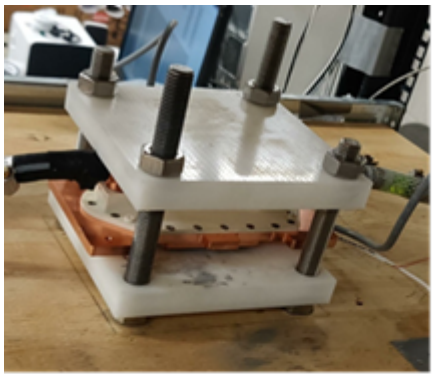
Chapter 5 Assessment of Pump Power Savings at Rack Level for Dynamic Direct-to-Chip Liquid Cooling Using a Novel Flow Control Device

5.1 Abstract

A substantial increment in usage of network-based services such as high-performance computing and cloud storage is being matched by developments in high-power density chips. This also requires more efficient cooling technologies than traditional air cooling due to its heat transfer limitations. Direct-to-chip liquid cooling is one of the most popular techniques for cooling high heat fluxes due to the higher thermal performance of water-based coolants. Typical liquid-cooled data centers provide the servers with a constant and redundant coolant flow rate to the servers irrespective of server workload utilization leading to excessive pump power consumption. This investigation proposes the implementation of an active flow control device to resolve this problem and quantifies the maximum power savings possible for different power levels in servers in a rack. This device is capable of target delivery of coolant to the cold plates based on the instantaneous workloads of processing units. The 3-D printed flow control device contains a V-cut ball valve that varies the open area inside the device to regulate the flow rate. Four custom-made Thermal Test Vehicles, each representing a server, in a rack were used to mimic workloads. The TTV assembly was placed at four different levels in a standard 19-inch Information Technology Equipment rack in test fixtures mounted with cold plates. The flow regulation to each of the TTVs was varied based on the power dissipated by each TTV. A maximum pump power savings of 87% was achieved for a case where all TTVs operate at idle, with 15% utilization.

5.2 Experimental Setup and Methodology

A test fixture using 3-D printed polycarbonate plates was developed based on the design and measurements of cold plates and the TTV assembly as shown in Figure 5-1. The entire test fixture consisted of two polycarbonate plates between which the cold plate and TTV assembly are fixed. The TTV assembly on the bottom of the cold plates consisted of a ceramic and Thermal Interface Material (TIM) with an insulation baseboard at the bottom of the heater. Appropriate pressure was provided on the spring-loaded cold plate screws to achieve maximum heat transfer from the heat source/TTVs to the cold plate. A schematic of the TTV in the fixture assembly is shown in Figure 5-2, depicting the ceramic heater and its bottom plate which acts as insulation. Throughout the manuscript, the entire heater, TIM, and bottom plate assembly are referred to as Thermal Test Vehicle (TTV), and the entire cold plate heater fixture assembly is referred to as a single server.



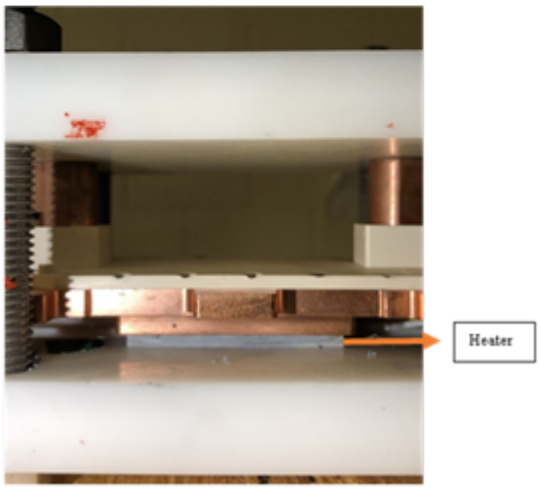


Figure 5-1: Test fixture designed for housing the TTV assembly and cold plate used for the experiment

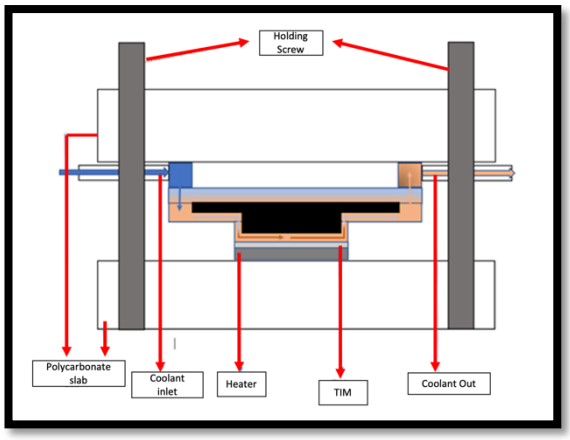


Figure 5-2: Schematic of the TTV and cold plate assembly in the fixture

Figure 5-3 shows the CAD model of the FCD used in the study integrated with a micro-servo motor (labeled as 1) for valve rotation. The FCD controls the coolant flow rate using a V-cut ball valve (labeled as 6) with an aperture to restrict the coolant flow through the device cavity. The entire FCD assembly was 3-D printed using polycarbonate plastic material. As its thermal conductivity value is very low (~0.14W/mK) the device itself doesn't participate in heat transfer. The valve opens and closes based on the shaft rotation (labeled as 4) and changes the flow rate of the coolant. The rotation is accomplished by the torque provided by the servo motor (labeled as 1) to rotate the shaft by a pre-decided angle that reduces or increases the flow rate through each TTV loop. This increase or reduction in the flow rate directly corresponds to the heat dissipation or the percentage utilization of the TTVs. Therefore, at any instant of time within the IT rack, the TTVs dissipating more heat receive a larger flow rate of coolant and those at lower heat flux receive a lesser volume of coolant. It is also ensured that the valve design provides a constant minimum coolant flow rate through the aperture at its lowest percentage open area. This minimum open area corresponds to the TTVs or a real server in an idle state or at the lowest possible work utilization. This allows a sufficient volume flow rate to keep the processing units at a safe and reliable operating temperature all the time.

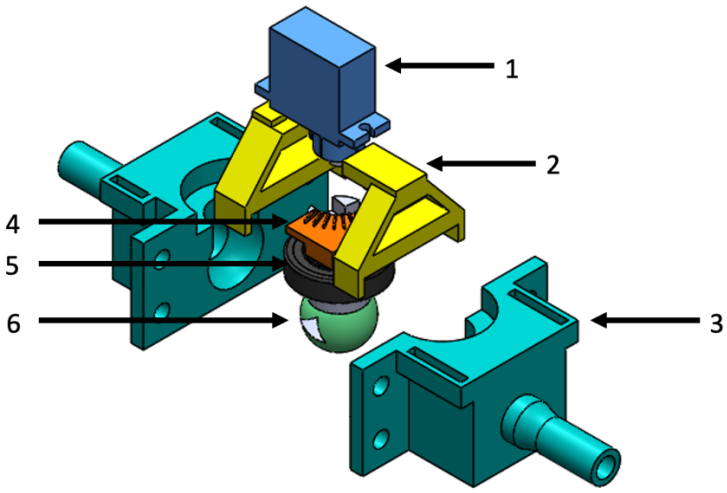


Figure 5-3: CAD model of the Flow Control Device used

As shown in Figure 5-4, the valve control strategy for power savings is divided into two parts, a temperature-based control at the TTV or server level and pressure-based control at the rack level. At the TTV or server level, temperature sensors located at the outlet of each of the cold plate manifolds are used to vary the valve angle based on the cold plate outlet temperature. The outlet temperature for the TTV with higher wattage will be higher than the TTVs with lower power. Thus, the FCD attached to a TTV loop with a higher instantaneous power is controlled to increase the open area to allow a greater flow rate through the cold plates for more cooling and vice versa. At the rack level, the pressure control-based strategy maintains a constant pressure drop across the rack. When the FCD valves to one of the TTVs operating at lower power are partially closed, it increases the overall pressure drop in the rack manifold. After this, the pressure drop in the rack is brought back to the initial value when the FCDs were fully opened. This is accomplished by directly varying the pulse width modulation (PWM) to the dc pump to maintain a constant pressure drop value across the entire rack, similar to the case when all the FCDs are in fully open condition. When the pressure drop is equalized, the new pump power consumption is read from the dc power supply and compared to the original, maximum pump power to calculate the reduction in the power value.

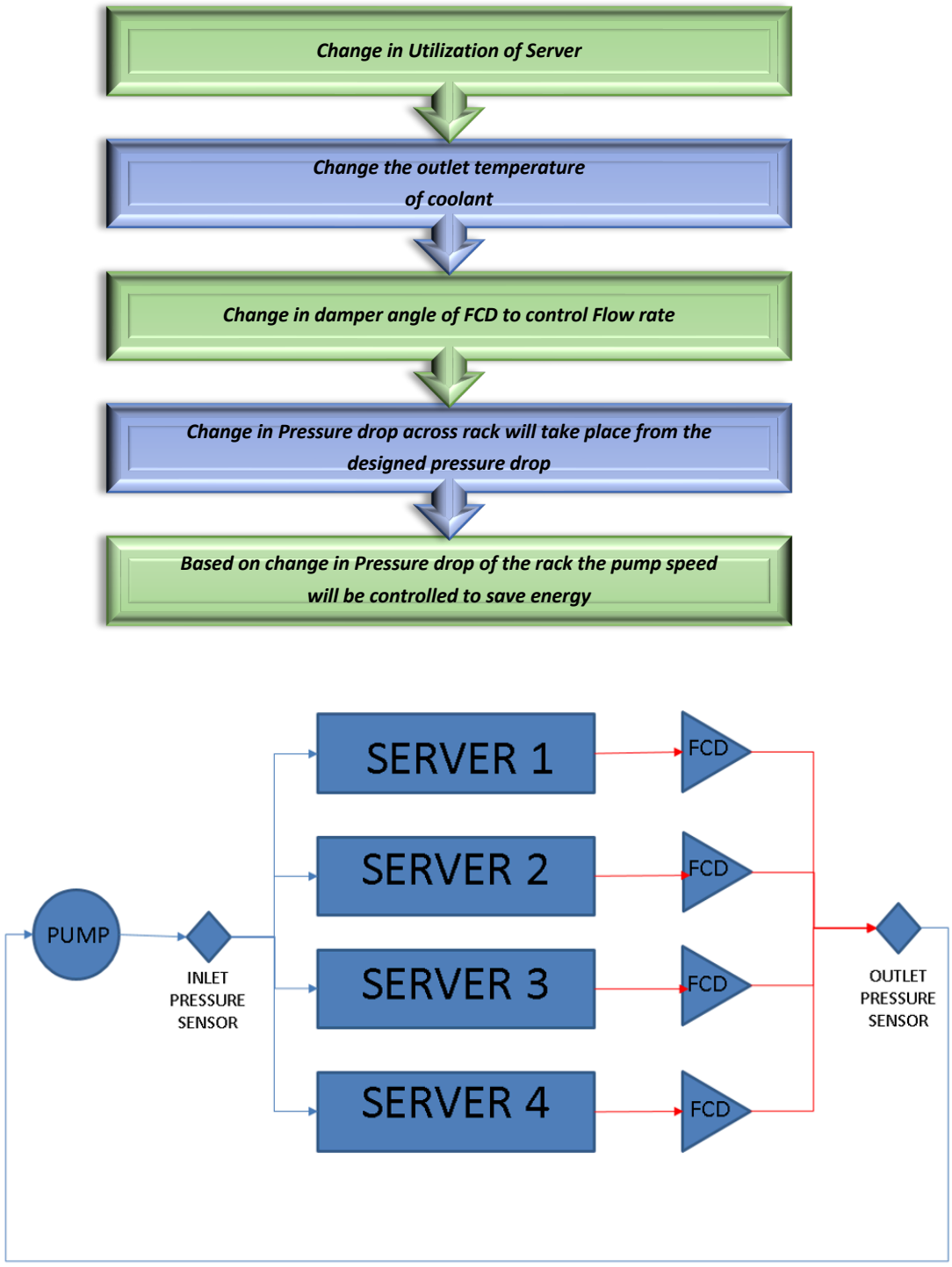


Figure 5-4: Description of temperature and control strategy used for controlling the pump

For the TTV build-up, a Watlow ceramic heater was used as the heat source in the stack-up in these experiments. The dimensions of the heater were 50.4 X 50.4 mm with a maximum power dissipation value of 2000 watts and an embedded K-type thermocouple. An off-the-shelf

cold plate with a base dimension of 55mm x 55mm was used in this study. The maximum power utilization by the TTV developed in the manuscript was kept to 430 watts per TTV due to limits on the available power from the power source. The idle or low power utilization conditions were reproduced by providing 50 watts to the TTVs. A thermal paste was applied between the heater surface and the cold plate as a thermal interface material (Arctic MX-4 TIM) with thermal conductivity of 8.5 W/mK to ensure good thermal contact.

Each of the four cooling loops in the rack was connected to the inlet and the outlet of the rack-level manifold. The manifold was connected to the pump and reservoir arrangement placed on the top of the rack. The sensors used in the setup consisted of ultrasonic flow meters for each of the cooling loops, pressure sensors measuring rack-level and TTV cooling loop pressure drop, and K-type thermocouples measuring coolant outlet temperatures for FCD control. The 3-D printed FCDs were placed at the outlet of each of the cooling loops. The specifications of the above-mentioned components and sensors are listed in Table 11. The entire manifold and cooling loop assembly were then mounted on a 42U rack. A dc powered pump was connected to the reservoir providing a maximum flow rate of 3.8 lpm between an operating voltage of 8-12 volts. A PolyScience liquid to liquid cooler was used as a heat exchanger during the experiments, rejecting the heat to the facility's chilled water supply. The sensor data was directly read in the Agilent data acquisition software from the DAQs (Data Acquisition Units). 25% propylene glycol (PG25) was used as the coolant in this study owing to its excellent anti-microbial and heat transfer capabilities. Figure 5-5 shows the sensor placements across a single TTV assembly and the final rack-level setup is shown in Figure 5-6.

Table 11: Details of sensor accuracy, measurement range, and operating voltages of the sensors used in the experiment

Sensor	Operating voltage/current (mA)	Range of measurement	Accuracy
Keyence FDX-A1	20V – 30V	0.02L/min -20L/min	+ 0.003ml/min
Keyence Pressure sensors	4-20	0-1000 kPa	+ 0.25%
K – Type Thermocouple	_	0 – 400 ˚C	+ 0.75%

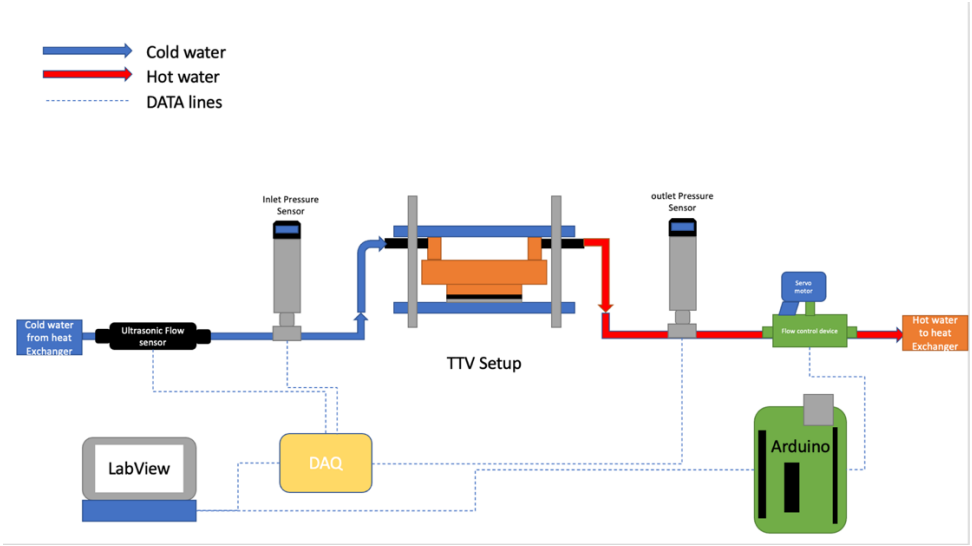


Figure 5-5: Schematic of the experimental arrangement for a single TTV showing the cold plates, test fixture, pressure, and flow sensors, and, the Flow Control Device

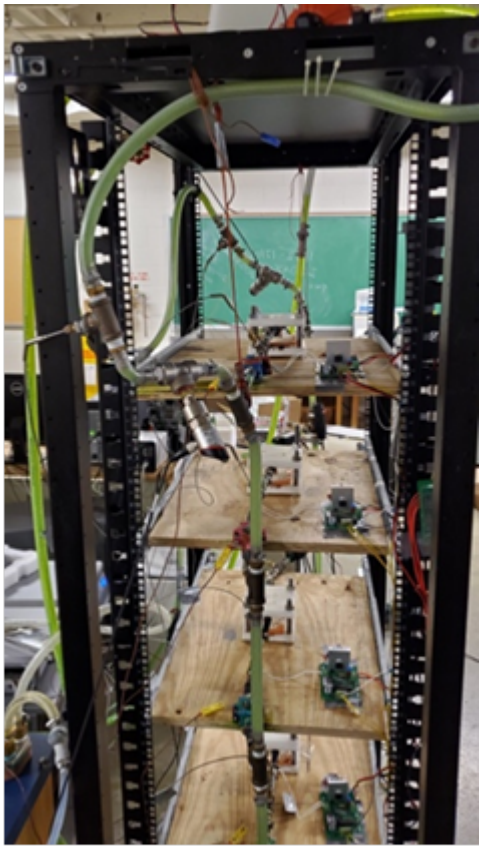


Figure 5-6: Experimental setup showing the arrangement of sensors, manifolds, cold plates, and the Flow Control Device

5.2.1 Experimental Procedure

Figure 5-7 shows the schematic of the voltage dimmer circuit that was used to control the ac voltage used for powering the TTVs. The control circuit is a Triode for Alternating Current (TRIAC) arrangement which varies the voltage according to the PWM input signal provided by the attached Arduino. The TRIAC arrangement works as a bidirectional three-terminal switch that can be used for phase control or voltage control applications. The TRIAC board used in the current study also contained a zero-cross controller embedded in the board circuitry. The dimmer controller is powered by an external 5V dc input signal and the input for the percentage of voltage dimming is provided using a PWM signal supplied by an Arduino circuit. The PWM control input was directly given from LabVIEW VI which provides the duty cycle as an input signal to Arduino. The input ac load used for powering the TTVs was 120-

volt ac at 50Hz. The maximum power value of 430 watts was obtained at the maximum duty cycle of 100% and the minimum of 50 watts was obtained at a 15% duty cycle.

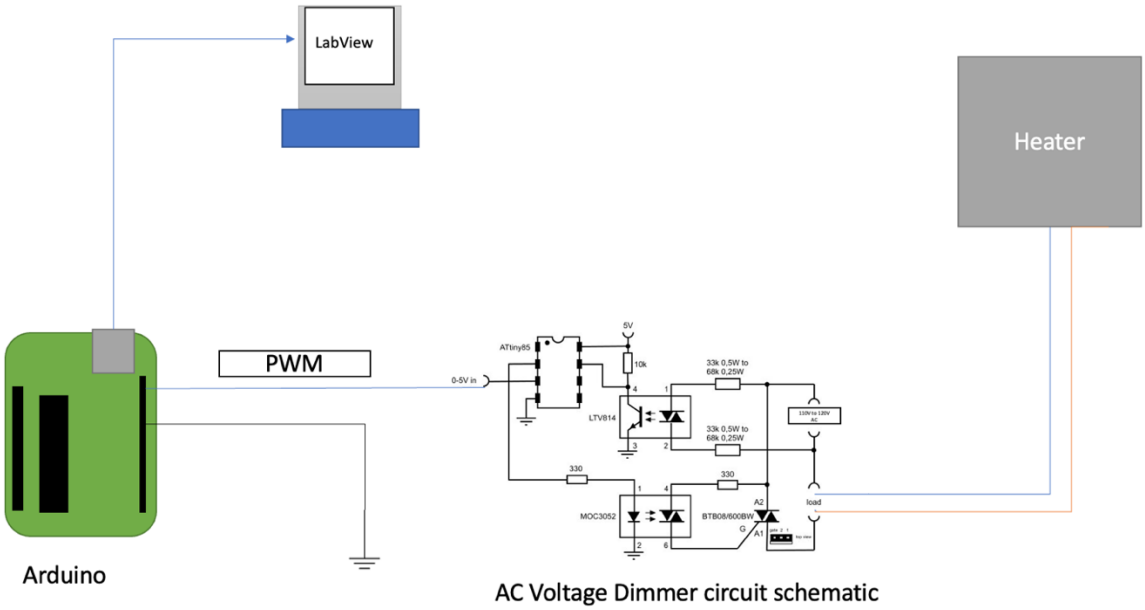


Figure 5-7: Schematic of the data acquisition and ac dimming circuit used for controlling the heater power

Figure 5-8 shows the schematic of the final experimental setup. As discussed earlier, two power levels of approximately 50 watts and 430 watts were used to mimic 100% utilization and idle workload cases. The power values chosen in the study have been decided based on the power levels of current high-power CPUs and GPUs. The experiments were conducted for two values of coolant inlet temperatures, 35ºC and 45ºC, representing ASHRAE W3 and W4 liquid cooling envelopes. A baseline case was first run at both inlet temperatures where the maximum power was given to each of the TTV assemblies. In this baseline case, the FCD ball valve was kept in a fully open position to allow the maximum flow rate through the device cavity for all for TTVs. The pump power value was obtained directly from the current and voltage readings from the power supply and the rack level pressure drop, flow rates, and TTV temperature values from Agilent software.

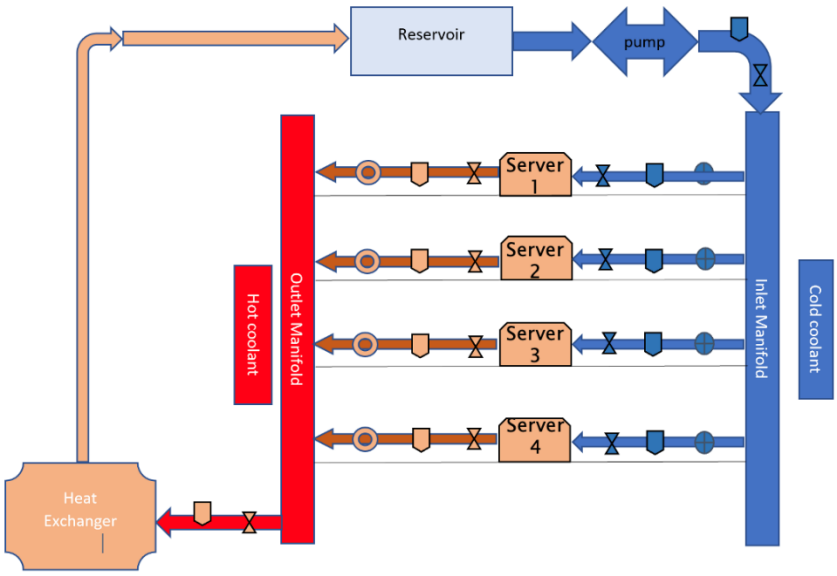


Figure 5-8: Schematic showing the final TTV (annotated as Server) arrangement in the experimental setup along with sensor instrumentation

After this, the power to the TTVs was reduced to 50 watts one by one until all the TTVs operated at minimum power value. For each reduced power level case, the servo motor was powered to reduce the percentage of opening inside the FCD, by rotating the ball valve to a desired angular position. This reduced the flow rate to the TTVs at lower power utilization and increased the pressure drop within the rack manifold. After closing the valve for the low-power TTV, the pressure drop in the entire manifold was re-established back to the same value as in the case where all valves are in fully open condition. This was done by controlling the pump by PWM control. The new, reduced value of the pump operating power was noted, and the percentage reduction was reported for each case. Before the power savings experiments were conducted, the authors also characterized the FCD for maximum flow rates possible at various open positions of the valve. A summary of the flow rates and pressure drop values corresponding to different valve angles is shown in Figure 5-9, where 90° represents the fully closed position. A maximum flow rate of 4 lpm is obtained in the current FCD design when the valve is fully opened and a flow rate of 0.2 lpm passes through the FCD in the fully closed

position (corresponding to idle workloads in servers). For the current study, only extreme cases of fully opened and fully closed FCD valve (maximum TTV power dissipation) positions are tested as these bounds represent the maximum and minimum pump power savings

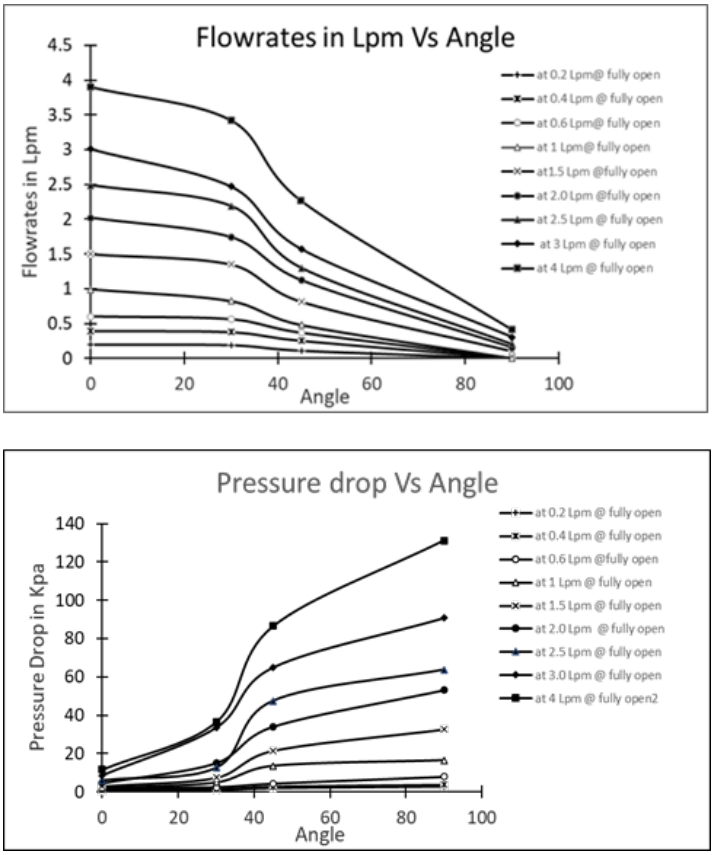


Figure 5-9: Flow rate and pressure drop variation of the coolant through the flow control device for various valve angular positions (Shahi et al., 2021)

5.2.2 Sensor Calibration

Each sensor in the study was calibrated using standard calibrating equipment and procedures. In this experiment, Keyence GP-M010 pressure sensors were used and calibrated using Fluke P5510-2M Pneumatic Comparison Test Pump. On the left side of the test pump was the GP-M010 pressure sensor and on the right side, the reference pressure gauge was mounted. The hand pump was used to increase the pressure in the test rig and to keep the test pump pressurized, the rotating knob was closed. The error in the reading of the sensor and the reference gauge was recorded for error analysis. Similarly, the Keyence clamp-on microflow

sensors (Keyence FDX-A1) were calibrated with a Coriolis flowmeter. The k-type thermocouples used to measure the fluid temperature were calibrated using a Fluke 7109A portable calibration bath between a temperature of 0-100°C using a two-point calibration method as shown in Figure 5-10. Table 12 shows the error calculation quantified from the calibration process for pressure sensors and the thermistors. It was observed that the pressure sensors were very precise after factory calibration and did not need additional calibration. A two-point calibration method was used for the thermocouples by calculating the error in the temperature reading. The calibration equation obtained was directly used as input in the DAQ software as gain and offset values. To calibrate the ultrasonic flow sensors, a Coriolis mass flow meter was used by placing the flow sensor in the same closed-loop along with the Coriolis flow meter. A Coriolis mass flow meter was used for calibration as its readings are universal since it measures the mass flow rate rather than just the flow rate with a typical accuracy for liquids between 0.05-0.1%.

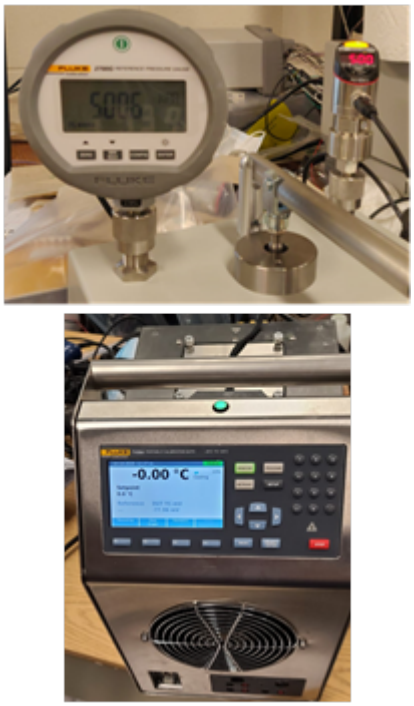


Figure 5-10: Pneumatic pressure calibration gauge (top) and temperature sensor calibration bath (bottom) used for pressure sensors and thermocouples

Table 12: Percentage error calculation for GP-M010 pressure sensors and thermocouple

Reference Guage (KPa)	**GP-M010 pressure sensors (KPa)**	**% Error**
150.3	150	0.2
200.2	200	0.1
300.6	300	0.2
400.5	400	0.124
500.6	500	0.12
Reference Temperature (°C)	**Measured Temperature (°C)**	**% Error**
10	9.8	2
90	89.3	0.8
Reference Flow Rate at Coriolis (lpm)	**Measured Flow Rate at Sensor (lpm)**	**% Error**
0.5	0.513	2.6
1	0.936	4
1.5	1.528	1.8
2	1.912	0.6

5.3 Results and Discussion

Based on the details discussed in the experimental procedure, five experimental cases were performed for the inlet temperature values of 35ºC and 45ºC. The power levels were varied from a maximum of 430W to a minimum of 50W in 4 different experimental runs after the baseline case. Table 13 shows the summary of the baseline results for the coolant inlet temperature of 45ºC. This case depicts the testing condition when all the TTVs are at maximum or 100% workload utilization. The pressure drop in the system for this case was at 5kPa and the pump power consumed was approximately 18.5W as read from the dc power source. The maximum flow rate through the FCD, in this case, was 1.28 lpm. The baseline results represent the case of typical data centers where no flow control or dynamic cooling is implemented and are presented in the topmost part of the tables in this section, as seen in Table 14. The pump power in these results is the maximum power value consumed due to redundant flow rate provisioning. The test data collection was started by assuming that a steady state is reached when the variation in the outlet temperatures from the cold plate is within 0.5°C. Figure 5-11

shows the variation of pump power for all the experimental cases analyzed at the inlet temperatures of 35ºC and 45ºC.

Table 13: Pressure drop and flow rate summary for the baseline case when all fcd vales are fully opened

TTV No.	Power (W)	Pressure Drop (kPa)	TTV Temperature (°C)	Flow Rate (lpm)	FCD Position	Pump Power (W)
1	430	5	57.8	1.28	open	18.5
2	430	5	58.5	1.14	open	
3	430	5	60.9	1.11	open	
4	430	5	58.3	1.09	open	

Table 14: Summary of pump power when 1 TTV is at idle load at 45 °C inlet temperature

TTV No.	Power (W)	TTV Temperature (°C)	Flow Rate (lpm)	FCD Position	Pressure Drop (kPa)	Pump Power (W)
1	50	47.6	1.28	open	5	18.
2	430	59	1.15	open		
3	430	61.4	1.12	open		
4	430	58.9	1.1	open		
After Closing FCD Valve					**Pressure Drop (kPa)**	**Pump Power (W)**
1	50	48.5	0.55	closed	7	18.5
2	430	57.5	1.41	open		
3	430	59.8	1.36	open		
4	430	57.4	1.32	open		
After Pressure Drop Balancing					**Pressure Drop (kPa)**	**Reduced Pump Power (W)**
1	50	48.8	0.41	closed	5	12.1
2	430	59.1	1.15	open		
3	430	61.4	1.14	open		
4	430	58.7	1.08	open		

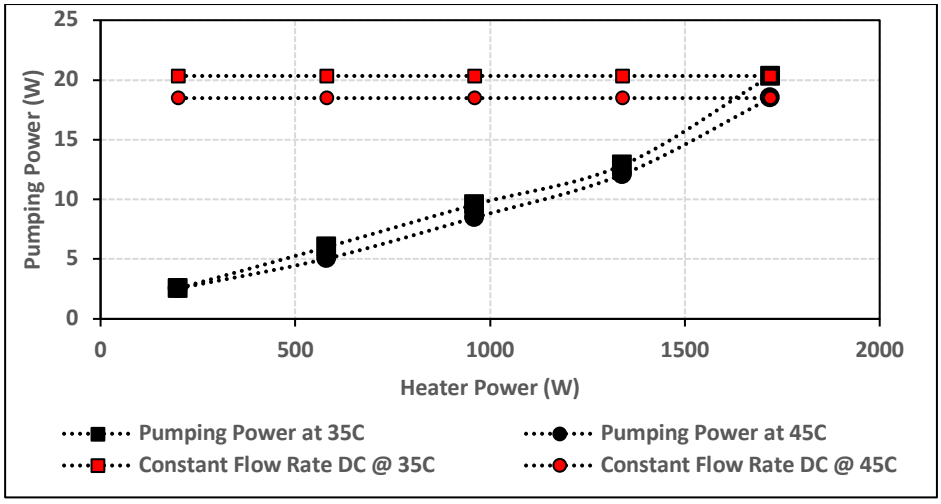


Figure 5-11: Variation of dynamically controlled pump power for different power levels vs pump power consumption at constant pump power for different coolant inlet temperatures

After acquiring the data for the baseline case, the power to the first TTV was reduced to the assumed idle workload of 50 watts. The pump power consumed, manifold pressure drops, and TTV temperature data were obtained and the FCD corresponding to this TTV was closed. The middle part of Table 14 shows the increased pressure drop case after the FCD valves are controlled due to lower coolant outlet temperatures. At this position, the ball valve allows a minimum flow rate to pass through the device cavity. As seen in Table 14, this increased the pressure drop of the system to 7kPa and increased the flow rate to the TTVs corresponding to which the FCD was open. Figure 5-12 shows the variation of rack pressure drops as the number of closed FCDs increases for both 35ºC and 45ºC inlet temperatures. The next step was to bring the system to its initial state by controlling the pump using a PWM control. This represents the case of a data center that uses FCDs integrated with the rack manifold assembly. The results of pump control showed that the pump now required approximately 12.1 watts to maintain similar temperatures and flow rates in the remaining TTVs. This yielded a pump power saving of approximately 34% from the baseline case.

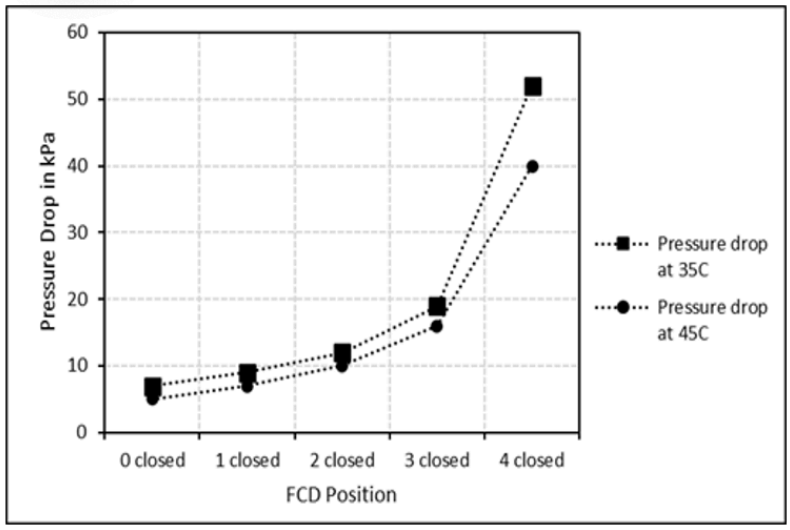


Figure 5-12: Variation of the pump power with changing total rack power at a different inlet temperature

In the last case, all the TTVs were brought down to the minimum power of 50 watts each. The values of the rack pressure drop and TTV temperatures are summarized in Table 15. The maximum power consumed by the pump for this case was approximately 20 watts. The flow rates to each of the TTVs were then controlled by closing all the valves of the FCDs increasing the rack pressure to 40 kPa. After the flow balancing was performed by reducing the pump power, the rack pressure dropped to the initial value of 5kPa. The pump power after flow balancing dropped down to 2.56W. This means that to maintain the almost same surface temperature of the TTVs using an optimized flow rate, an almost 87% reduction in pump power can be obtained.

Table 15: Summary of pump power when all TTV is at idle load at 45 °C inlet temperature

TTV No.	**Power (W)**	**TTV Temperature (°C)**	**Flow Rate (lpm)**	**FCD Position**	**Pressure Drop (kPa)**	**Pump Power (W)**
1	50	46.4	1.3	open	5	20.2
2	50	46.7	1.16	open		
3	50	47.3	1.08	open		
4	50	46.5	1.07	open		
After Closing FCD Valve					**Pressure Drop (kPa)**	**Pump Power (W)**

1	50	45.5	1.33	Close	40	21.1
2	50	45.9	0.83	Close		
3	50	46.5	0.92	Close		
4	50	45.6	0.87	Close		
After Pressure Drop Balancing					**Pressure Drop (kPa)**	**Reduced Pump Power (W)**
1	50	45.4	0.45	Close	5	2.6
2	50	47	0.27	Close		
3	50	47.8	0.25	Close		
4	50	46.2	0.28	Close		

The above-mentioned procedure was also repeated for a lower inlet temperature value of 35ºC. As seen in Table 16, for the baseline case a maximum power of 20.3 watts with all the FCD valves in a fully open position. A baseline pressure drop of 6kPa was obtained. Table 17 shows the case when one TTV in the rack was reduced to minimum power. The total pump power consumption for this case was obtained to be 20.3 watts. As done in the previous cases, the FCD corresponding to the low-power TTVs was closed to provide a minimum flow rate through the attached cold plate. After pump flow balancing was performed, it was observed that a reduced pump power of 12.9 watts was obtained. This is a reduction of 37% as compared to the case where a constant flow rate is used for cooling.

Table 16: Summary of pump for baseline when valve FCD is fully open for 35 °C inlet temperature

TTV No.	Power (W)	TTV Temperature (°C)	Flow Rate (lpm)	FCD Position	Pressure Drop (kPa)	Pump Power (W)
1	430	47.5	1.28	open	6	20.3
2	430	47.8	1.16	open		
3	430	50	1.1	open		
4	430	47.9	1.08	open		

Table 17: Summary of pump power when 1 TTV is at idle load at 35 °C inlet temperature

TTV No.	Power (W)	TTV Temperature (°C)	Flow Rate (lpm)	FCD Position	Pressure Drop (kPa)	Pump Power (W)
1	50	37.3	1.26	open	6	20.3
2	430	48.7	1.16	open		
3	430	50.9	1.12	open		
4	430	48.7	1.08	open		
After Closing FCD Valve					**Pressure Drop (kPa)**	**Pump Power (W)**
1	50	41.7	0.16	close	9	20.3
2	430	47.1	1.48	open		
3	430	49.2	1.38	open		
4	430	47	1.39	open		
After Pressure Drop Balancing					**Pressure Drop (kPa)**	**Reduced Pump Power (W)**
1	50	43	0.12	close	6	12.9
2	430	48.2	1.2	open		
3	430	50.5	1.13	open		
4	430	48.2	1.16	open		

Table 18 shows the summary for pump power reduction for the case where all the TTVs operate at minimum power or workload utilization. After operating the FCD valves for all the devices, the rack pressure drop increased to 52 kPa. The flow balancing by pump control after this step reduced the system pressure back to its initial value and the pump power was reduced to 2.5 watts or 87.5% power savings. The summary of pump power savings for both the inlet temperatures considered in this study is shown in Figure 5-13. It was concluded that the percentage power savings using dynamic cooling at rack level variation was within 2% for both the temperature values.

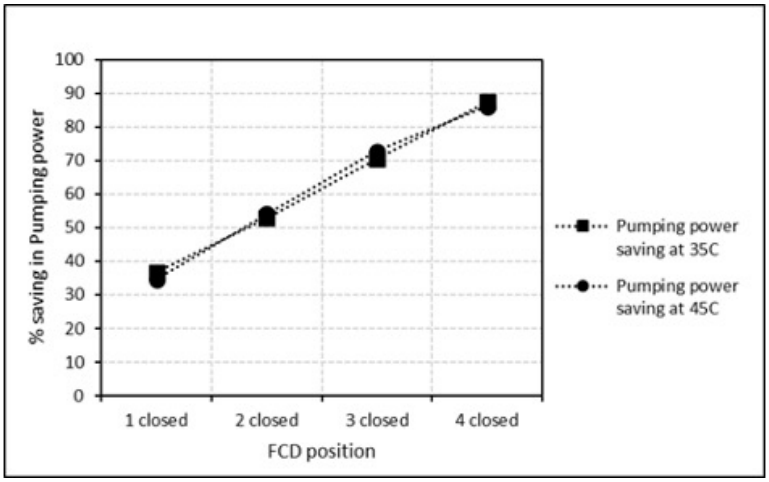


Figure 5-13: Change of pump power savings at different inlet temperatures for various FCD position cases

Table 18: Summary of pump power when all TTV is at idle load at 35 °C inlet temperature

TTV No.	Power (W)	TTV Temperature (°C)	Flow Rate (lpm)	FCD Position	Pressure Drop (kPa)	Pump Power (W)
1	50	37	1.29	open	6	20.3
2	50	37	1.18	open		
3	50	37.4	1.08	open		
4	50	37	1.08	open		
After Closing FCD Valve					**Pressure Drop (kPa)**	**Pump Power (W)**
1	50	37.2	1.19	Close	52	20.3
2	50	37.2	1.1	Close		
3	50	37.7	0.97	Close		
4	50	37.3	0.67	Close		
After Pressure Drop Balancing					**Pressure Drop (kPa)**	**Reduced Pump Power (W)**
1	50	38.5	0.3	close	6	2.6
2	50	39.7	0.2	Close		
3	50	40.7	0.2	Close		
4	50	37.1	0.3	Close		

5.4 Conclusion

Direct-to-chip liquid cooling has gained significant popularity, especially for the thermal management of high-performance computing platforms. It also offers a significantly simpler option to transition from air cooling with minimum changes in the existing air-cooled infrastructure. Energy-efficient air-cooling strategies like airside economization offer lower PUE values but have inherent contamination-related reliability issues, high failure rates, especially in areas with high PM2.5 and sulfur content [29]. Dynamic cooling strategies have been implemented and discussed widely in the literature where predictive learning and PID-controlled coolant delivery of the coolant have been used to optimize the coolant delivery to the hot devices. For single-phase direct-to-chip cold plates, actively controlled FCD can be a reliable and cost-effective approach to accomplish this.

A pseudo-rack test setup with four custom-made TTVs and cold plates was built to mimic four servers in a rack. The proposed FCD was tested at two different power and temperature values with a baseline case depicting maximum power dissipation from all the TTVs. The TTV temperatures, rack pressure drops, and pump power values were acquired for the experimental run. After this, both the power and flow rates (closing the FCD valve to allow minimum flow rate) to one of the TTVs were reduced. The results show that a maximum of 88% savings in pump power was achieved for the case with an inlet temperature of 35ºC. the power savings were reduced by 2% when the inlet temperature was increased to 45ºC. Thus, dynamic control of server flow rate based on server workload utilization can yield substantial power savings. It was also seen that for the TTV operating at minimum power, the average difference in the TTV temperature at maximum and reduced optimized flow rate (after closing the FCD valve) was only 1ºC.

Chapter 6 Numerical Investigation on Effect of Target Coolant Delivery in Liquid-Cooled Microchannel Heat Sinks

6.1 Abstract

Rising demand for high-performance processors and increased difficulty in their thermal management have resulted in the need for advanced and efficient cooling technologies like direct-to-chip liquid cooling. Typical CPUs/GPUs today have multiple cores and at any given time, not all the cores in the package are utilized which creates a non-uniform heat distribution or a thermal gradient across the processor. This can lead to reliability issues due to localized thermal transients and cyclic thermo-mechanical stresses. A practical solution to solve this issue is investigated in this paper by proposing a dynamic cold plate and analyzing its thermal performance using CFD modeling. The proposed cold plate design consists of four different fin sections. The flow rate to each of these fin sections is regulated passively using bimetallic strips that respond to the outlet temperature of the coolant. A cold plate that can dissipate a maximum of 360W was first designed and optimized using multi-parametric optimization in OptiSLANG. Based on the anticipated heat loads, a bimetallic strip was selected and analyzed for maximum deflections using FEA using a range of cold plate operating temperatures. The results show that the proposed cold plate design can reduce the thermal resistance by a maximum of 42%. Furthermore, a maximum reduction of 62% was observed in the temperature difference value of the power sources as compared to the baseline design.

6.2 Cold Plate Design and Optimization

This section will discuss the design approach followed for the dynamic cold plate concept and final design. The approaches discussed include a preliminary approach based on a miniaturized flow control device and the final design approach using bimetal strips. The cold plate-fin design optimization process and the selection process of the bimetal strip for the final cold plate design are also described.

6.2.1 Nitinol Spring Approch

Previous work done on the dynamic cold plate by our research group focused on the development of a cold plate design using a self-regulated temperature-sensitive material. In this design, a miniaturized flow control device was made from Nitinol, which is a Nickel-Titanium shape memory alloy that responds to temperature change. This idea of implementing a self-regulated miniatured Flow Control Device (FCD) is discussed in recent literature (Kasukurthy et al., 2018). In this study, the miniaturized FCD is integrated with a nitinol spring that is connected to a damper inside the FCD cavity. Based on the outlet coolant temperature (coming out from the hotter section of the cold plate), the spring deflects the damper allowing more coolant to enter that section of the cold plate. Thus, the device proposed the idea of improved temperature distribution on the processor case. The major bottleneck in the design and development of this device was the lack of material reliability data as well as difficulties in miniaturizing the device that can be integrated inside a cold plate successfully. Thus, the authors decided to integrate bimetallic strips with the cold plate design that addresses the material reliability, miniaturization, and any related operational issues. Further details about the new cold plate design are discussed in the upcoming sections.

6.2.2 Cold Plate Fin Optimization

Liquid-cooling using cold plates for high-performance computing servers has gained a lot of momentum and significant attention has been paid to optimizing the cold plate performance. Traditional cold plate designs for data center applications comprise parallel mini or micro-channel fins embedded inside a cavity. These fins are typically made up of copper. With water-based coolants, cold plates typically offer a very high value of convective heat transfer coefficients. Significant efforts are made to optimize the manufacturing processes and design geometries of the cold plate fins to improve the heat transfer and reduce cold plate pressure drops. For the current cold plate design, ANSYS optiSLang (ANSYS Inc., 2019), a commercially available design optimization tool, was used to optimize the baseline cold plate geometry.

The baseline design of the cold plate is divided into three main sub-assemblies. The middle plate assembly of the cold plate is modified in such a way that the V-shape bimetal can be fitted into it and flow leakage is avoided. For that purpose, the supporting V structures are patterned as shown in Figure 6-1. The gap between two supporting structures acts as a fluid path and can be regulated with help of the bimetal deflection. The bottom of the middle plate has a slit to enter the water into the fins as shown in the figure. The bottom cold plate is divided into four identical fin sections that are separated to avoid the mixing of the hot and cold fluid streams. The flow path of the coolant in the typical single inlet and single outlet cold plates with parallel channels results in non-uniform flow. This is primarily due to the jetting action of fluid at the entrance of the micro-channel passage. Providing the inlet in the middle of the cold plate and the sectioning of the fin cavity results in a more uniform flow distribution through the fin cavity as will be shown in the result section. The thickness of the entire cold plate assembly is maintained close to 22mm to meet the space constraints for final assembly within a 1U server. The fin passages and inlet and outlet are kept perpendicular to the flow

direction as shown in the complete 3D CAD model in Figure 6-1. The baseline design of the cold plate operates like the traditional cold plates but the fin cavity is optimized for the best thermal and hydraulic performance using multi-design variable optimization. For optimization of the cold plate-fin cavity, the objective functions used that control geometry optimization of the fin parameters was minimizing the pressure drop and reducing the temperature difference between the cold plate inlet and outlet. The variables parametrized in the optimization and their optimization range is as listed below:

Fin thickness: 0.1 -1 mm

Fin Spacing: 0.2 - 1 mm

Inlet Velocity: 0.029 - 1.91 m/s

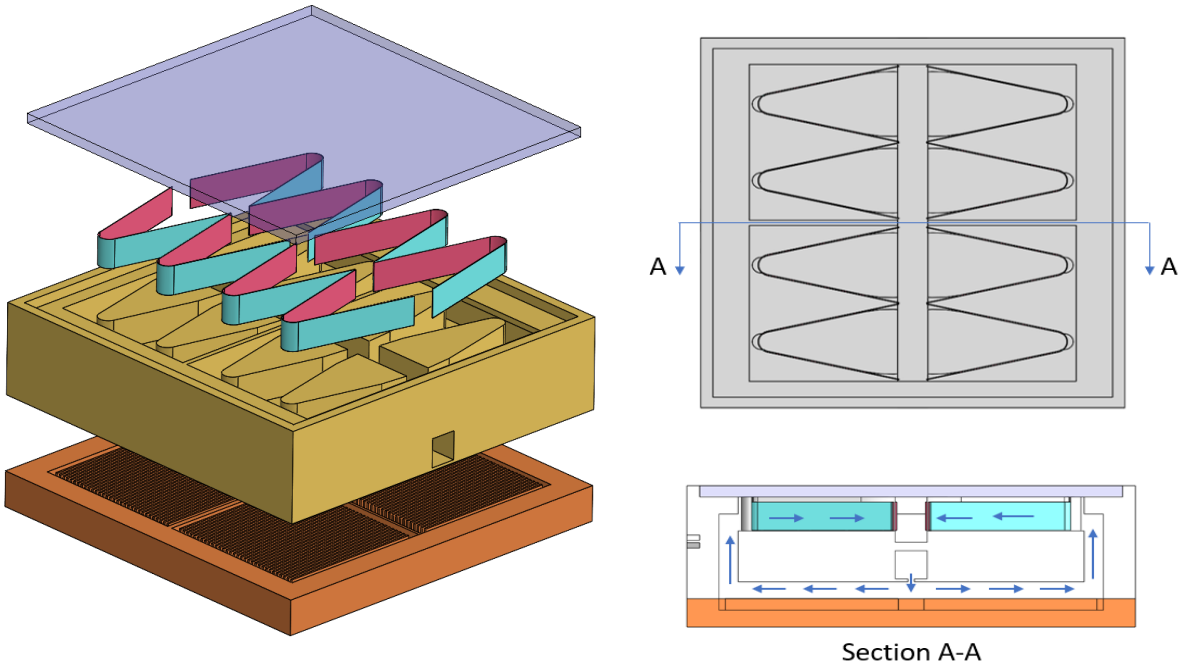


Figure 6-1: (Left) Exploded view of the cold plate CAD geometry showing the top, middle and bottom plates (Right) Sectional view of the cold plate showing the movement of the coolant through various parts inside the cold plate

Baseline steady-state thermal simulations of the cold plate are run using ANSYS Icepak where the input parameters for optimization are provided as input and the objective functions

are defined. These design variables and objective functions are integrated with optiSLANG, where the design points are generated based on the range of the defined variables/design parameters. The objective functions set for the current optimization study were minimum pressure drop and minimum source temperature below the cold plate. Each of these generated design points is then iteratively run in Icepak and the results of the objective function values are generated. A summary of the results of optimized design points for the objective functions of minimum pressure drop and source temperature values is shown in Figure 6-2. Figure 6-2 is a Pareto plot where each black dot represents the solution of a design point of the optimization study. The design points shown on this graph depict the best values of the fin parameters for which both the pressure drop and the temperature difference are minimum. The final optimized CAD of the old plate bottom plate is shown in Figure 6-3. Based on the obtained data, the best values of the fin parameters is as given below:

Fin Thickness: 0.48 mm

Fin Spacing: 0.48 mm

Inlet Velocity: 0.4 m/s

Inlet Pressure: 135.46 Pascal

Maximum Temperature: 33.83 °C

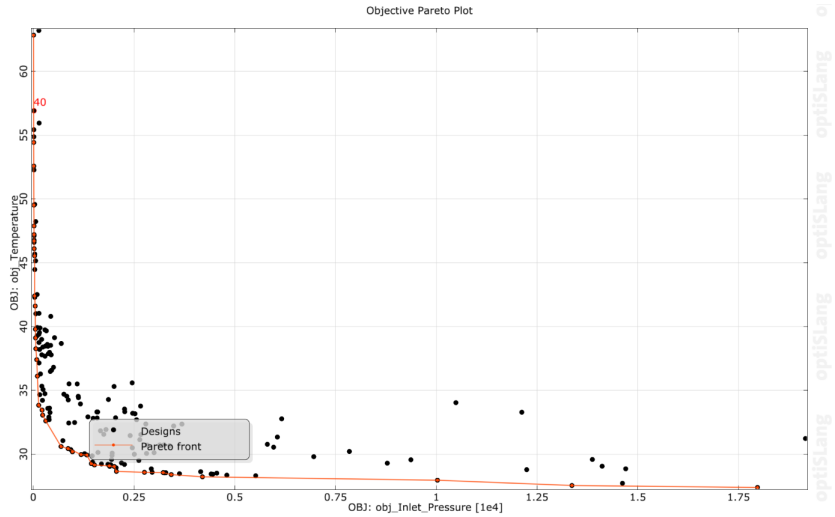


Figure 6-2: Pareto plot showing the values of the objective functions (temperature and pressure drops) for various design points simulated during multi-objective optimization

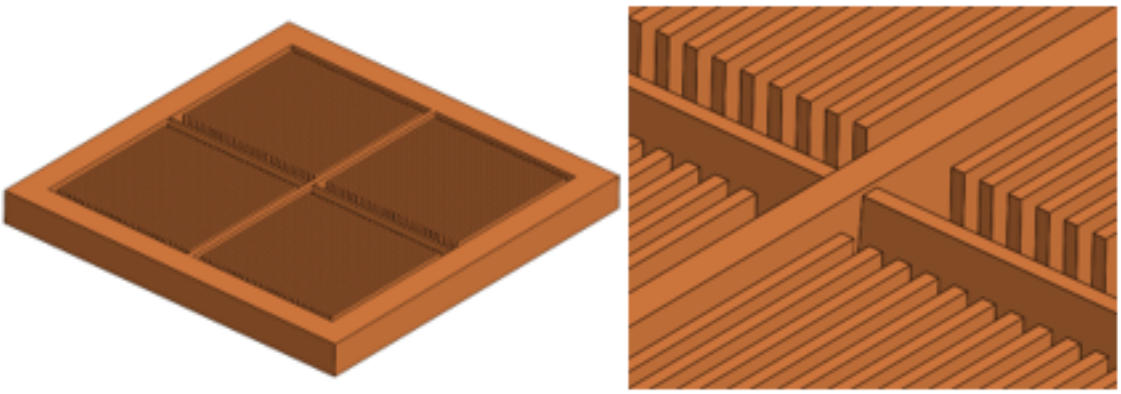


Figure 6-3: CAD model of the optimized fin geometry showing the four fin sections in the bottom plate of the cold plate

6.2.3 Bimetallic Strip Approach

The next step was to integrate the bimetallic strips within the cold plate design. Before this, a detailed analysis of the bimetallic strip was done to quantify the amount of deflection bimetal will undergo based on the anticipated outlet temperature from each section of the cold plates. Bimetal is a composite material and is usually found in the form of a sheet or a strip. It is manufactured with a special engineering technique that involves the bonding of two or more

metallic layers having different Coefficient of Thermal Expansion (CTE) values. As the temperature at the strips varies, both the materials expand or contract differently causing a deflection as shown in Figure 6-4. This property of the bimetal, also known as thermostatic material is widely used. These materials are employed in many industries including automotive, HVAC, space science, and even the military. The deflection of the bimetal is directly proportional to the CTE and working temperature and inversely proportional to the total thickness of the strip. It is also a function of the modulus of elasticity of both the materials and their thicknesses.

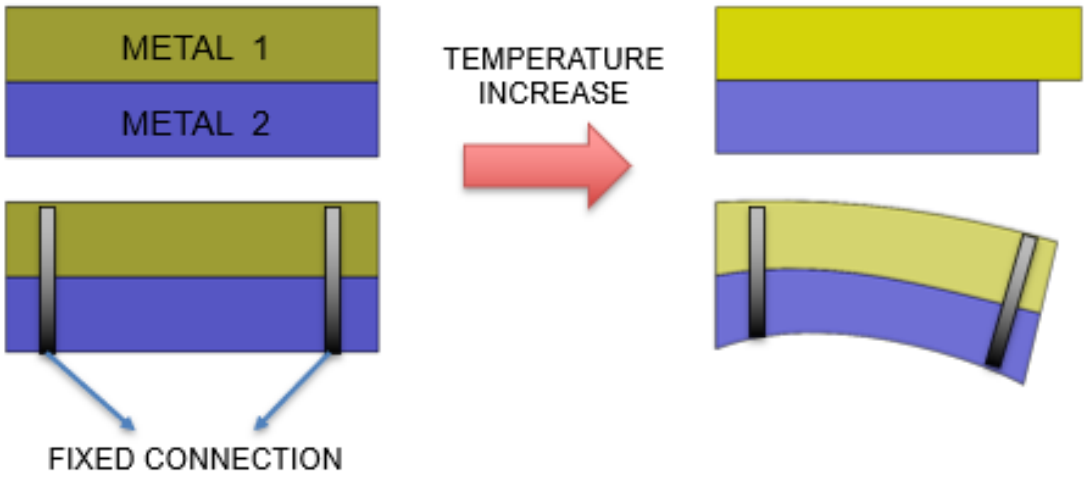


Figure 6-4: Overview of the working principle of bimetal strip showing the deflection due to temperature increase

The V-shaped bimetallic strips are located in the middle plate of the cold plate assembly, where warm coolant returns from the bottom plate. These bimetallic strips deflect based on the outlet temperatures from different fin sections at the bottom, thus, varying the volume flow rate passing through them. When the temperature of the coolant coming out from any of the four sections rises due to a local increase in heat flux from the processor region below the cold plate, the bimetallic strip for that region will respond by increasing the percentage of open area for more coolant to flow through it. At the same time, if other regions on the processor are experiencing low localized heat flux, the bimetallic strip for that section

will deflect and partially close, thereby, reducing the coolant flow rate to that section. This process will, thus, reduce the overall temperature gradient over the surface of the processer, unlike traditional cold plates.

6.2.4 Bimetallic Strip Deflection Analysis

Many bimetal materials are available in the market that can be selected as per the deflection requirements specific to the application. The deflection of these materials is based on two material properties: thermal expansion coefficient and flexivity. Flexivity is the factor that defines both CTE and deflection. Table 19 shows the list of the bimetals that have been referred to from the Emsclad bimetal manufacturer for selection for the cold plate design. Out of several available bimetals, GB14 is a 2-layer bimetal in which alloy GB has 10.2E-6 F-1 CTE and alloy 14 has 3.5E-6 F-1. To accommodate the bimetal strips into the cold plate, a V-shape bimetal geometry was chosen, as it allows a larger bimetal strip length. This is because the bimetal deflection is directly proportional to the thickness of the layer. Therefore, a larger length of the bimetal will result in a higher deflection. To obtain the optimum length, thickness, and deflection combination for the bimetal strips, steady-state thermal simulations were done. As shown in Figure 6-5, a steady-state thermal loading from 25°C to 70°C in ten-step intervals is provided to V shape bimetallic strips. The material properties were directly referred to from the company datasheet and imported to the ANSYS engineering materials list. A rectangular mesh of around 79,002 elements and 559,920 nodes with an element size of 0.08 mm was created on bimetal. Table 20 shows the mesh sensitivity analysis of the bimetal strip for thermal simulations. By using such geometry relative longer bimetal strips can be adopted into a cold plate. Figure 6-6 represents a comparative study of length and relative deflection in mm.

Table 19: List of the bimetal from EMSclad data sheet (EMSclad)

Material Type	ASTM Flexivity, F x 10E-7	Maximum sensitivity temp, F	Deflection temp. range, F	Modulus of elasticity, E	Resistivity	Density	ASTM Type
F7OR	147	0 to 300	-100 to 500	24.5	70	0.299	TM27
F9OR	148	0 to 300	-100 to 500	25.0	90	0.298	TM28
F100R	149	0 to 300	-100 to 500	25.0	100	0.297	--
F125R	148	0 to 300	-100 to 500	25.0	125	0.297	--
F55R20	130	100 to 500	-100 to 700	22.0	54	0.300	--
G7	61	0 to 800	-100 to 1000	27.5	440	0.280	--
GB2	128	100 to 550	-100 to 1000	26.0	445	0.295	--
GB5	75	300 to 800	-100 to 1000	26.0	342	0.296	--
GB14	100	0 to 300	-100 to 1000	26.0	511	0.294	--
J1	134	0 to 300	-100 to 500	19.0	110	0.310	--
J7	56	0 to 500	-100 to 500	22.0	106	0.300	--
LA1	158	0 to 300	-100 to 700	25.0	475	0.292	TM29

Table 20: Mesh sensitivity study for (top) bimetal steady-state thermal simulation and (bottom) cold plate thermal simulation

Number of Elements	Max. Deflection at 70°C
45236	2.054 mm
79002	2.2296 mm
120120	2.2297 mm
145638	2.2298 mm
Number of Elements	**Thermal Resistance @ 0.6 lpm (°C/W)**
126098	0.041
341732	0.036
695713	0.034
1186324	0.034

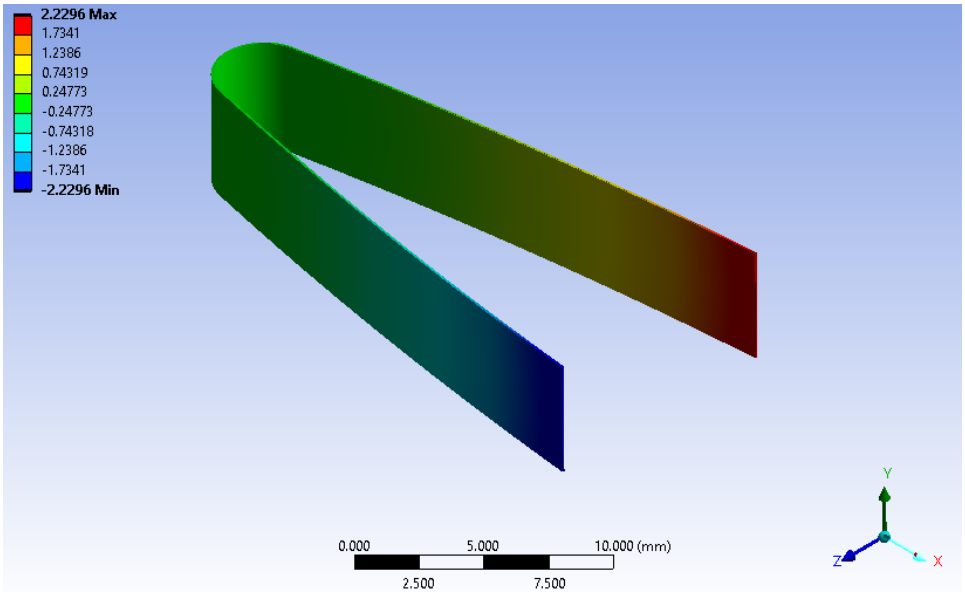


Figure 6-5: GB14 deflection(z-axis) in V shape. 50mm length and 0.1 thickness

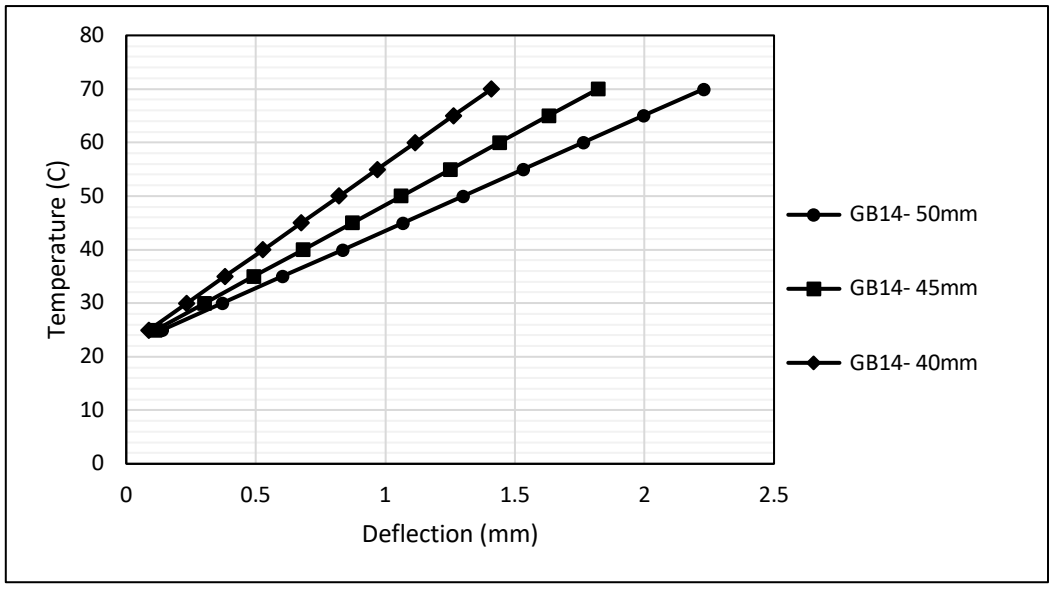


Figure 6-6. Comparison between the different lengths of the bimetal to respective deflection

6.3 Computational Modeling

Since it is extremely difficult to simulate a three-way fluid-structure-thermal interaction using CFD, steady-state simulations of the cold plate assembly with bimetals were done in ANSYS Icepak. The total deflections of the bimetal strip were first calculated using steady-state thermal simulations as discussed above. The CFD was then run as separate cases by using

the deflection values from the steady-state simulations and the improvement in temperature drop across the cold plate was calculated for each of these cases in Icepak. This was done by manually modeling blocks (where the bimetal is located) to create the same effect as in the case of bimetal deflection corresponding to different coolant temperatures. Each of these blocks represents one end of the bimetal and as the directional deformation of the bimetal is known, a precise opening length can be given for different cases of coolant temperature. As an example, if the temperature in a specific section of the cold plate is 45°C and the corresponding bimetal directional deflection is 1.2mm, a value of 1.2 mm is used as an open length and the rest is blocked. The exact values of the temperatures at the bimetal have been obtained from the simulation case where no bimetal is present.

The quality of the mesh in the CFD study defines the quality of the results, and good quality and optimum element count mesh are characterized by grid-independent results. For conjugate heat transfer problems, as in this case, care should be taken to have a sufficient number of nodes on the fluid region between the fins and on the solid as well. To achieve this, per-object meshing properties are applied to all the fins where a user-defined number of elements is selected in the x, y, and z-direction of the fins and also the fluid regions. Figure 6-7 represents meshing on the fins. For the global geometry, a minimum mesh element size was set to 0.04 mm and around 1.58 million elements were produced.

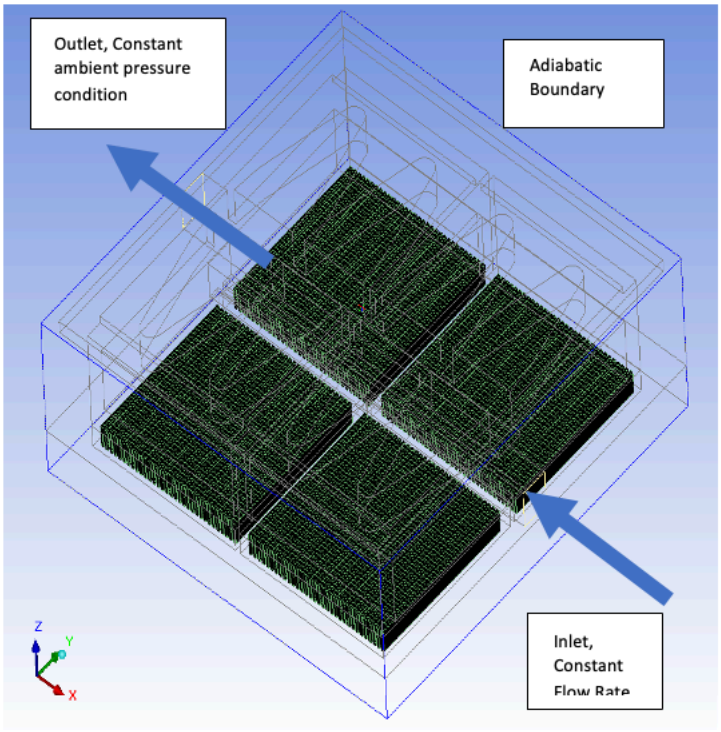


Figure 6-7: Summary of the boundary conditions used for the CFD simulations

For this study, Laminar flow equations were considered as the flow regime inside the inlet, outlet, and cold plate fins remain laminar. Below are the governing equations used by the software to generate a converged solution.

Mass Conservation,

$$d\rho/dt+\nabla.(\rho v)=0 \quad (1)$$

Momentum Conservation,

$$d/dt\ (\rho v)+\nabla.(\rho vv)=-\nabla P+\nabla(\tau)+\rho g \quad (2)$$

In the above equation, P is referred to as static pressure, ρg is known as body force due to gravitation, and τ is stress tensor.

Energy Conservation,

$$d/dt\ (\rho h)+\nabla.(\rho hv)=\nabla[(k+k\_t\ )\nabla T]+S\_h \quad (3)$$

Where k and kt are the molecular and turbulent transport conductivity and Sh is a term that includes heat sources due to volumetric definitions. Ansys Icepak involves simple conductive equations to solve solid conductions models which can be written as,

$$d/dt\ (\rho h)=\nabla.(k\nabla T)+S\_h \quad (4)$$

In this equation k represents thermal conductivity, T is the temperature, h is enthalpy (sum of flow energy and internal energy) and ρ is the density.

6.4 Results

The results reviewed in this section are divided into two parts: results for cold pate without bimetal strips (baseline simulations) and CFD results with bimetal strips added to the cold plate. A summary of different cases in terms of source power simulated is shown in Table 21. The heat source in the CFD simulations is modeled as a 2-D heat source with a user-defined power value as shown in Figure 6-8. The total heat dissipation value has been chosen to represent the future experimental work that will be carried out to validate this numerical study. The value of 10W represents the idle processor workload condition which is typically around 10% of the total TDP of processors. Each heat source is symmetrically located under each fin cavity section, below the cold plate base. A total of five different cases were simulated for the cold plate with bimetal (WB) and without bimetal (NB). A minimum source power of 10 watts and a maximum source power of 90 watts were used in different combinations with five different inlet flow rates between the range of 0.6lpm to 1.5 lpm with an increment of 0.3 lpm.

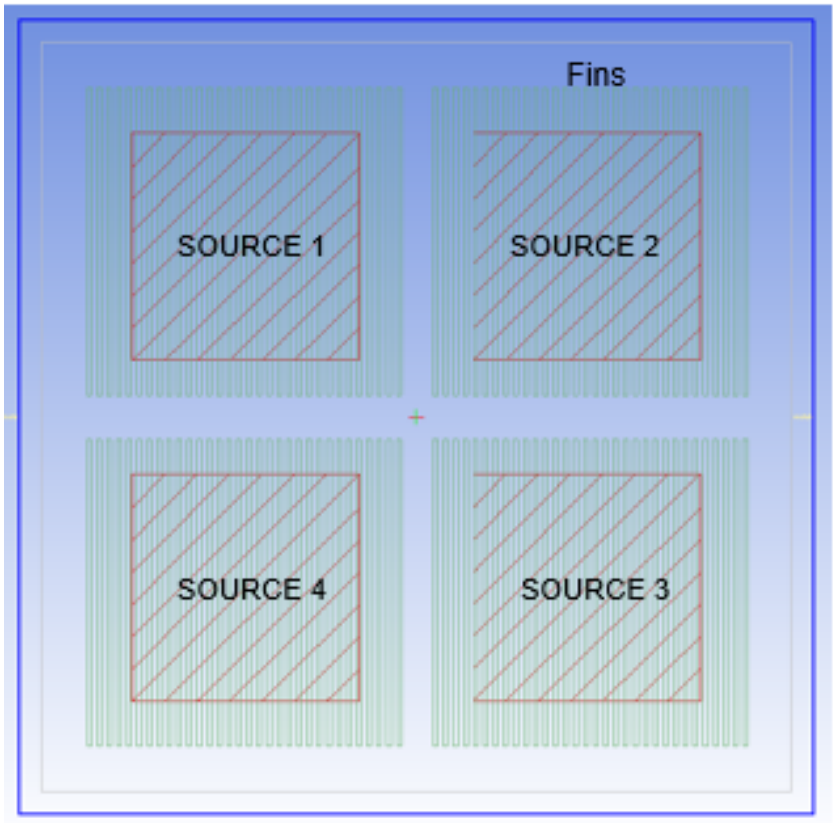


Figure 6-8: 2-D heat sources modeled in ANSYS Icepack to represent non-homogeneous thermal profile on a chip

6.4.1 CFD Result for Non-Bimetallic Condition

The first set of simulations of the cold plate design was done with a no-bimetal (NB) condition as a baseline design case for the cold plate. The results from each of the five cases as described in Table 21 of NB simulations at a 1 lpm coolant flow rate. The temperature at the top surface of the middle plate was also captured for temperature input to bimetal deflection value. Figure 6-9 represents different zones defined to capture top surface coolant temperature for bimetal deflection. As explained in the previous example of the simulation procedure, these zone temperatures will predict the deflection of the bimetal accurately at an instance. Table 22 represents the different zone powers for the bimetal at the 1 lpm coolant flow rate case. A similar trend in the temperature was observed for all other lpm. The temperature contours, as seen in Figure 6-10 show that temperature values for source 2 and source 3 are large as

compared to the other two. This was also noted for the other flow rate cases that were simulated. The possible explanation for this behavior is that the coolant enters the cold plate from right to left, as seen in Figure 6-10. This causes better flow movement and a more uniform distribution in the fin sections for sources 1 and 4. Hence, as seen in the thermal profiles there is a lower temperature distribution in those fin sections. CFD data shows that the difference in source temperature for the baseline case was approximately 12°C as seen from the data in Table 23 for the NB condition in case 1 when all the heat sources are at peak power values. When the power value to two of the sources is reduced to 10W, the temperature difference among the sources still stays at approximately 12°C for the NB condition. The maximum temperature difference of approximately 16.7°C occurs between the heat sources for case 3 where the heat sources of the same power dissipation are located on either side of the inlet section.

Table 21: No Bimetal- Top Surface Temperature at 1 lpm for all cases

Zone No.	Temperatures at 1 lpm-No Bimetal (°C)				
	CASE1	CASE2	CASE3	CASE4	CASE5
Zone 1	29.2	26.0	28.4	26.0	28.3
Zone 2	30.7	26.5	29.9	27.7	28.8
Zone 3	30.7	26.2	29.8	28.7	28.0
Zone 4	34.9	28.1	32.9	32.2	29.3
Zone 5	39.8	30.0	37.9	35.7	30.7
Zone 6	43.1	33.0	38.7	38.3	32.7
Zone 7	38.7	36.1	29.6	29.9	35.7
Zone 8	34.7	32.6	27.9	28.6	31.7
Zone 9	30.1	28.7	26.5	27.0	27.7
Zone 10	29.6	28.3	25.6	27.5	27.5
Zone 11	29.8	29.0	26.3	28.5	26.8
Zone 12	28.6	27.9	25.9	27.7	26.1

Table 22: Different power conditions for each heat source for the CFD analysis

	Source 1	Source 2	Source 3	Source 4
Case 1	90 W	90 W	90 W	90 W
Case 2	10 W	10 W	90 W	90 W

Case 3	90 W	90 W	10 W	10 W
Case 4	10 W	90 W	10 W	90 W
Case 5	90 W	10 W	90 W	10 W

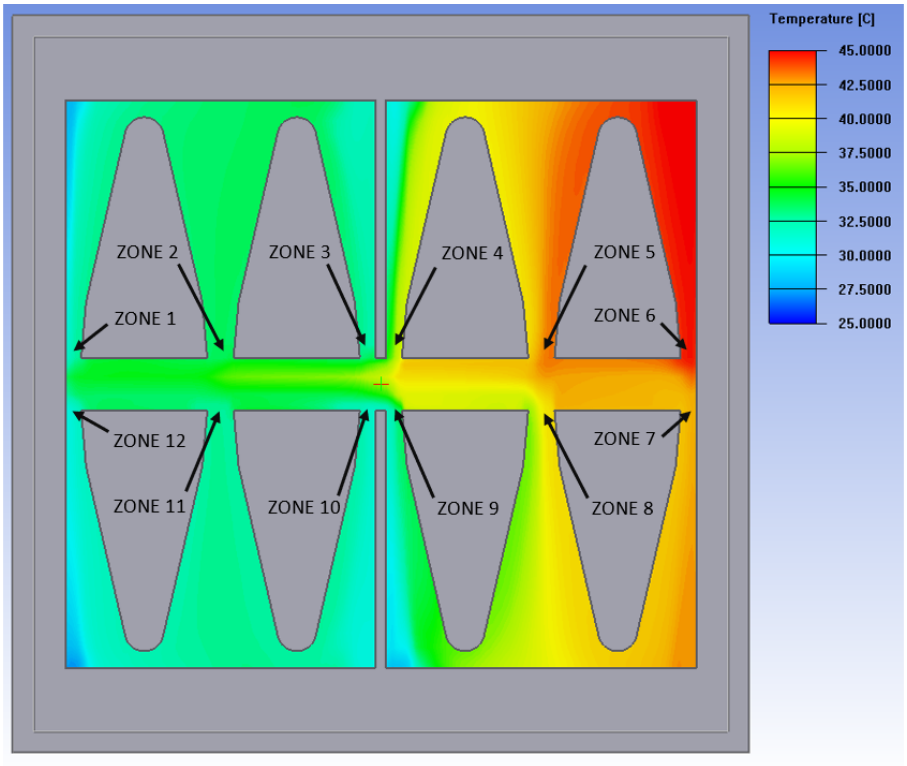
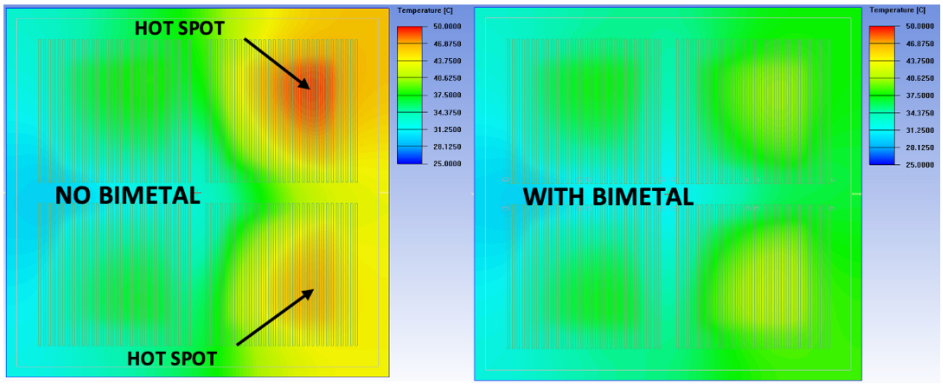


Figure 6-9: Locations of the zone defined for temperature probes in the CFD model



(a) (b)

Figure 6-10: Comparison of the temperature contours for CASE 1(a) showing the Base temperature for cold plate without bimetal and (b) with Bimetal (dynamic cold plate case)

Table 23: Source Temp Comparison, No Bimetal VS With Bimetal at 1 lpm

	Flow Rate (1 lpm)								Delta T (°C)	
	SOURCE 1 (°C)		SOURCE 2 (°C)		SOURCE 3 (°C)		SOURCE 4 (°C)			
	NB	WB	NB	WB	NB	WB	NB	WB	NB	WB
Case 1	37.9	37.3	48.9	40.6	46.6	41.8	36.7	37.1	12.2	4.7
Case 2	28.1	27.7	33.3	29.5	42.6	37.2	35.4	35.5	14.5	9.5
Case 3	36.6	35.6	44.9	36.4	33.5	29.1	28.2	27.7	16.7	8.7
Case 4	30.8	36.5	44.0	36.7	33.5	29.3	34.6	27.7	13.1	8.9
Case 5	35.4	35.4	33.6	29.9	42.0	37.3	29.9	29.1	12.2	8.2

6.4.2 CFD Results comparison between no-bimetal and with-bimetal cases

After performing all the simulations for no bimetal condition on the cold plate for different power distribution conditions at different lpm, simulations for the case after integrating the bimetal were performed. As the simulation in the software tool are only conjugate heat transfer simulations and structural movement cannot be simulated, a way to replicate the motion of the bimetal strip was used. This was done by creating blocks that represent the same percentage of open area at any given temperature as the bimetal strips. As an example, if the bimetal strip allows a 3 x 3 mm area for the coolant to pass through at 35°C, the same can be created by using 3-D blocks that allow the same area for the coolant to pass through.

Figure 6-10 shows a comparison of the temperatures on the base of the cold plate for case 1 at 1 lpm. Similar temperature behavior was seen at all other coolant inlet flow rate conditions. It was observed that the implementation of the bimetal strip or the dynamic cold plate design significantly reduces the difference in the temperature distribution on the base of the baseline cold plate design. Table 23 shows a comparison between source temperatures between the baseline cold plate design against the dynamic cold plate design with bimetal. Figure 6-11 shows the comparison of the temperature difference across the base of the cold

plate for all the cases with and without bimetal. A maximum reduction of up to 60% in the temperature difference of the base plate was seen and this is further evaluated by calculating the thermal resistance value for each case. This shows that the proposed dynamic cold plate design using bimetal strips can significantly improve the temperature homogeneity on the case surface of high-power CPUs or GPUs. An improvement in this temperature distribution can also be vital in enhancing the reliability of these devices and preventing failures caused due to temperature non-uniformity on and around chip surfaces.

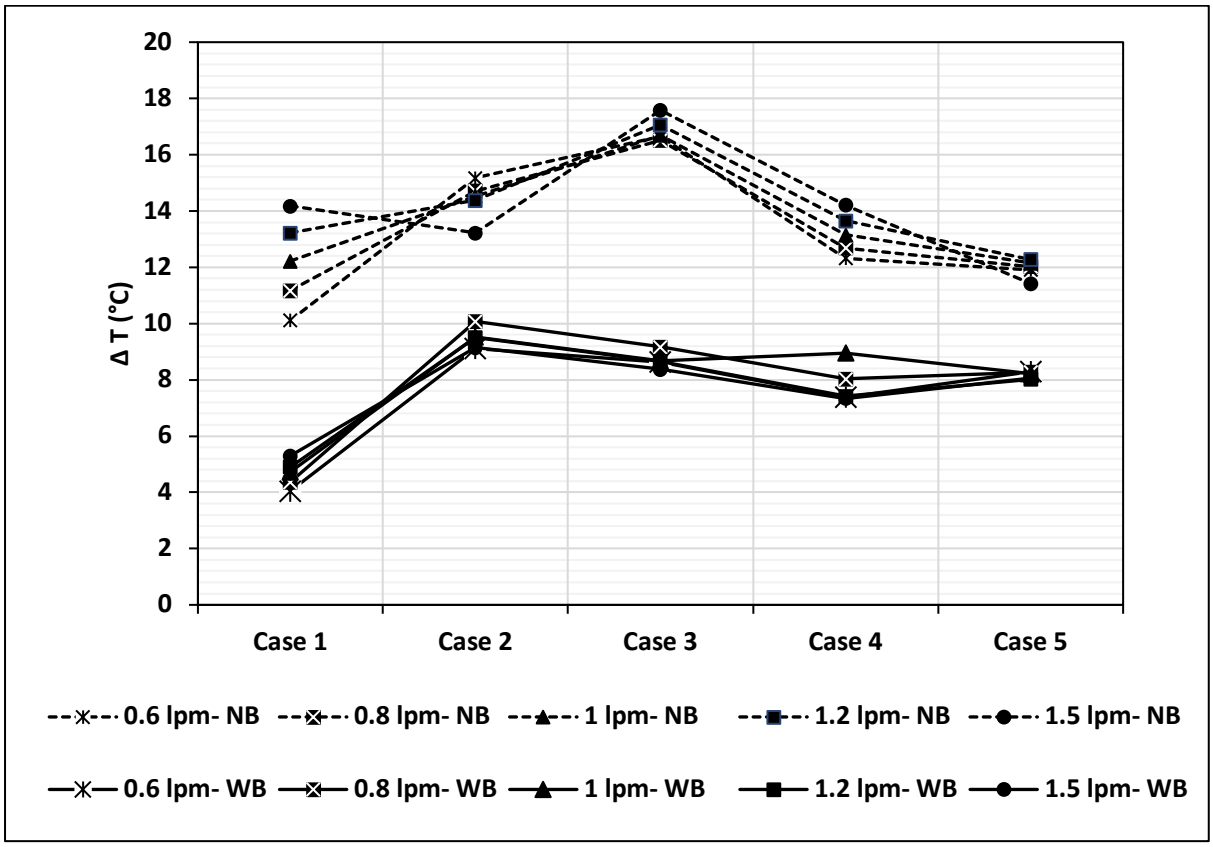


Figure 6-11: Variation of the temperature difference across the cold plate for all cases at different coolant lpm for the baseline cold plate and dynamic cold plate

To calculate the Rth value for each case following equation was used,

$$R\_th = ((T\_base - T\_in))/Q \quad (5)$$

Where Tbase is taken from the average of the base temperature at each case, Tin is the water inlet temperature which is 25°C for all the cases. Q represents the total load applied on

the cold plate and taken from Table 22 for the respective case. Figure 6-12 represents thermal resistance variation for different power distribution cases between the baseline cold plate design and the dynamic cold plate design. It was observed that up to 42% of the reduction in the thermal resistance value can be achieved by implementing dynamic control of the flow at the cold plate level. A comparison of the variation of the thermal resistance trend for the baseline cold plate design and dynamic cold plate also shows that the difference in the thermal resistance values converges at higher flow rates for the dynamic cold plate designs. The same trend is the opposite for the baseline cold plate design without bimetal strips.

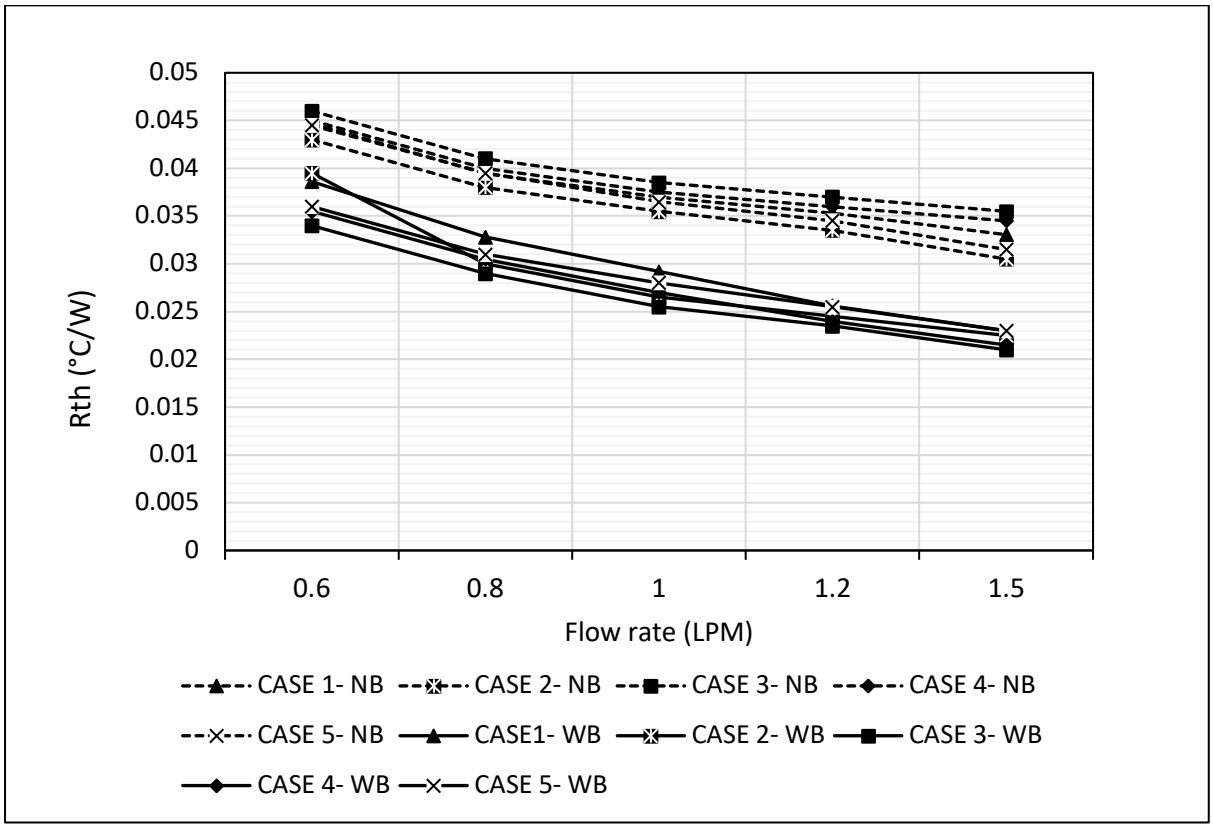


Figure 6-12: Variation of the thermal resistance for the baseline and dynamic cold plate design at different lpm

6.5 Conclusion and Future Work

As the power densities at the chip level continue to increase due to higher processing demands, data center administrators are moving towards liquid-based cooling technologies for

better thermal management. Direct-to-chip liquid cooling using cold plates allows easy retrofitting in air-cooled servers and can be used to dissipate very high heat fluxes. However, the challenge of non-uniform heat distribution over the chip surface persists which has been addressed in this study by proposing a dynamic cold plate design. An in-depth numerical study was conducted for design optimization of the fin geometry and selection of the best-suited bimetals to be integrated inside the cold plate. The results of thermal enhancement using the dynamic cold plate were discussed by comparing the pressure drop, thermal resistance, and temperature difference across the inlet and the outlet with the baseline design of the cold plate.

The results obtained from the comparison of the proposed dynamic cold plate to the baseline design show that a reduction of a minimum of 30% to a maximum of 62% is achieved in the value of temperature gradient across the cold plate. This reduction in the thermal gradient on the cold plate base also leads to a reduction of up to a maximum of 40% in the thermal resistance value of the cold plate. Studies have shown that temperature causes as much as 55% of failures in electronic devices. It can thus be hypothesized that by implementing the proposed cold plate design This significant improvement in the base temperature also directly affects the thermal resistance value of the cold plate which was found to be reduced by 40%. As discussed previously in this paper that temperature and thermal mismatches are one of the primary causes of failures among electronic devices. It can be postulated from the results obtained using CFD that some package-level reliability enhancement can be expected by implementing the proposed cold plate solution. However, this would require a significant experimental testing effort to quantify the improvement. The authors are actively working on more bimetal strip designs and materials to obtain comparisons for the best possible dynamic cold plate design. Work is also being done to automate the opening area of the cold plate using a MATLAB code integrated with a commercial data center CFD tool, 6SigmaET. This will allow visualization of the dynamic response of the cold plate to a local hotspot and generate more accurate results

in terms of performance enhancement of the cold plate. Different inlet and outlet configurations are also being studied using numerical analysis. The final cold plate design will be manufactured for validation of the CFD data

Chapter 7 A Control Strategy for Minimizing Temperature Fluctuation in High Power Liquid to Liquid CDU's Operated at Very Low Heat Loads

7.1 Abstract

The rising demand for high-performance central and graphical processing units has resulted in the need for more efficient thermal management techniques like direct-to-chip liquid cooling. Direct Liquid Cooling using cold plates is one of the most efficient and investigated cooling technologies since the 1980s. Major data and cloud providers are actively deploying liquid-cooled data center infrastructure due to rising computational demands. Liquid to liquid heat exchangers used in liquid-cooled data centers is also referred to as coolant distribution units (CDUs). Most of these CDUs selected by the data center operator is based on the heat load of the data center and the available head with that CDU. In this study, three 52U racks with six high-power TTV-based servers (Thermal Test Vehicles) in each rack were designed and deployed. Each server consists of eight GPU TTVs and six NV switch heaters. A 450-kW liquid-cooled CDU is used, and propylene glycol 25% is used as a coolant. Typical CDUs are designed to operate at 20 to 30% of the rated heat load to achieve a stable secondary coolant supply temperature. The present study will investigate the operations of CDU at very low heat loads, like 1% to 10% of the CDU's rated capacity. At these low loads, large fluctuations in secondary side supply temperature were observed. This large fluctuation can lead to the failure of the 3-way valve used in CDUs at the primary side. In this paper, a control strategy is developed to stabilize the secondary supply temperature within ± 0.5 °C at very low loads using the combination of a flow control valve on the primary side and PID control settings within the CDU.

7.2 Experimental Setup

For the experimental setup, three 52U liquid-cooled racks make up the data center facility. Each rack contains six direct-liquid-cooled (DLC) servers. The server is made up of a thermal test vehicle, which consists of eight heaters behaving as actual GPUs and six heaters behaving as NV switches. The liquid cooling loop is mounted on top of it. The centralized (CDU) is used to circulate the flow in the data center room. Coolant circulates through cooling loops, absorbing heat from servers. This heated fluid is then returned to the Coolant Distribution Unit (CDU) via a flow network, where it is cooled by primary side chilled water provided by the facility.

The fluid is distributed by the CDU using Y-connections to the row manifolds and the racks via individual rack manifolds. For the present experimental setup, each of the three racks is placed on the adjacent sides and is connected to only one row manifold, as shown in Figure 7-1. A temperature sensor and a flow sensor are fitted on both the return as well as the supply side of the CDU to measure temperature and flow rate across the system. Pressure sensors are also placed on both sides to measure the pressure drop across the system.

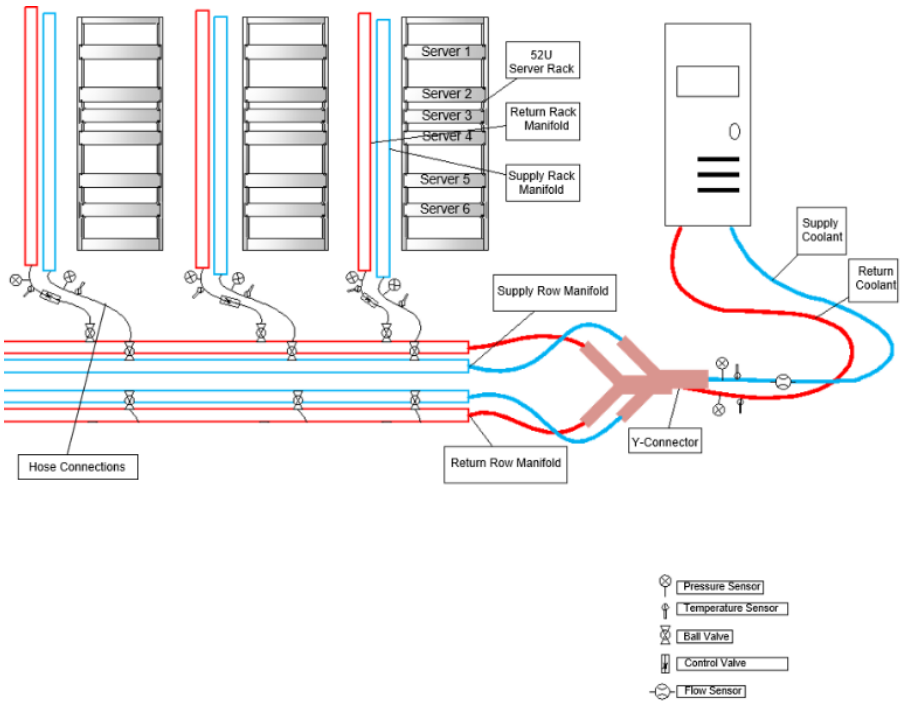


Figure 7-1: Schematic Layout of the Experimental Setup

For each of the connections from the row manifold to the rack manifold, ball valves are installed. There are control valves installed at the return side of each rack manifold that regulate the flow between the row and rack manifolds. These control valves regulate flow and provide information such as delta T and flow rate across the rack. There are pressure sensors, temperature sensors, and flow sensors placed on both the supply and return sides of each rack to take readings. These valves are connected to the workstation using an ethernet port for data logging purposes. A data acquisition (DAQ) device is linked to all the sensors installed in the system, and a serial USB connection is used to connect the DAQ devices to the workstation. The CDU is also connected to the workstation through an ethernet cable. The data from all the equipment and sensors is logged and analyzed using LabVIEW. In LabVIEW, control valves and CDU use a Modbus TCP connection, and IP addresses are assigned to CDU and control valves. For taking the readings from the sensors, a time interval of 5 seconds is provided to log every bit of data from the DAQ, control valves, and CDU.

In this paper, in the first experiment, the secondary side heat load (1%–10% CDU rated capacity) was varied for an interval of 1800 seconds, keeping the primary side flow rate and temperature constant. Secondary side supply temperature and valve demand were observed in this experiment. In the second experiment, a hardware-based control strategy was developed. An experiment was performed to prove that this control strategy is working. It shows that even at very low heat loads on the secondary side, the secondary-side supply temperature can be stabilized by varying the primary side flow rates. In the third set of experiments, the PID-based control strategy was developed. The first experiments were conducted to get the PID parameters, and later experiments were conducted to see how these parameters were stabilizing the secondary-side supply temperature with PI and PID experiments.

7.3 Results and Discussion

These experimental results are divided into 3 parts. In the first part, the secondary side heat load was kept very low, around 8.6% of the rated heat capacity of the CDU, and the heat load was reduced to as low as 0.7% of the rated capacity of the CDU. The default setting of CDU was used for these experiments. In the second part, it will be shown how by controlling the primary side flow rate, secondary-side supply temperature fluctuation can be controlled even for very low heat load, and in the third part, proportional integral and derivative (PID) control of CDU was used to control the fluctuation of secondary-side supply temperature.

In the first experiment, the secondary side supply temperature was set at 32 °C and the heat load was kept at 39 kW for 1800 seconds, and then it was reduced to 3.2 kW in steps as shown in Table 24 below. The primary side flow rate was kept at around 30 GPM, and the primary side supply temperature was kept at 7 °C.

Table 24: Variation of Heat Load w.r.t Time

Time (Sec)	Load (kW)
0-1800	39
1800-3600	26
3600-5400	13
5400-7200	6.5
7200-9000	3.2

Figure 7-2 shown below shows the secondary-side supply temperature fluctuations at different heat loads and fluctuations in the secondary supply temperature from 34 °C to 22 °C. This fluctuation amplitude reduces a little bit and the frequency reduces as the heat load reduces and all the parameters were kept the same. This happens because when the heat load is reduced, it takes more time to increase the supply temperature.

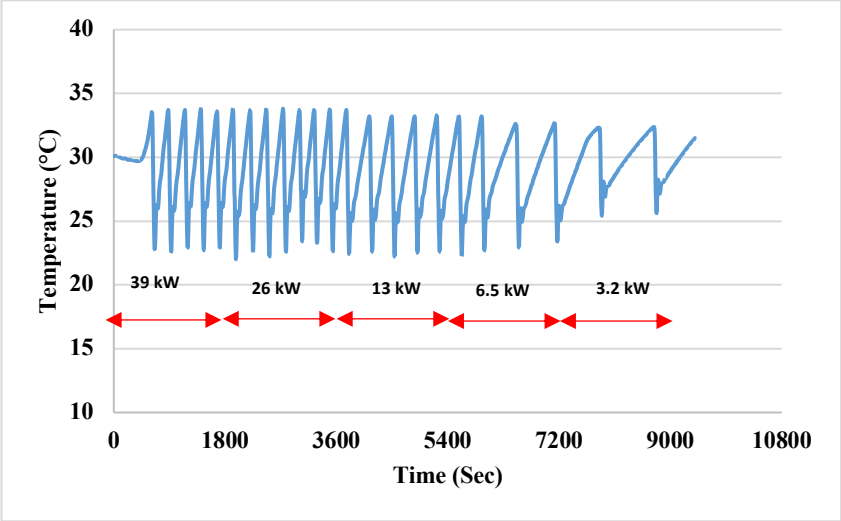


Figure 7-2: CDU Secondary Supply Temperature(°C) Vs Time (Sec) at different heat loads

Figure 7-3 shown below shows the primary side 3-way valve cooling demand with the change in secondary side heat load as shown above. The primary side 3-way valve shows a high level of fluctuations from 0 to 100% opening for 39 to 26 kW load with a high frequency, and this fluctuation frequency reduces as the secondary side heat load reduces further from 26

kW to 3.2 kW. For a 3.2 kW load, the maximum valve opening demand goes up to 80%. This high frequency of opening and closing of the primary side 3-way valve can lead to its failure as these valves are designed for certain thousands of openings and closings.

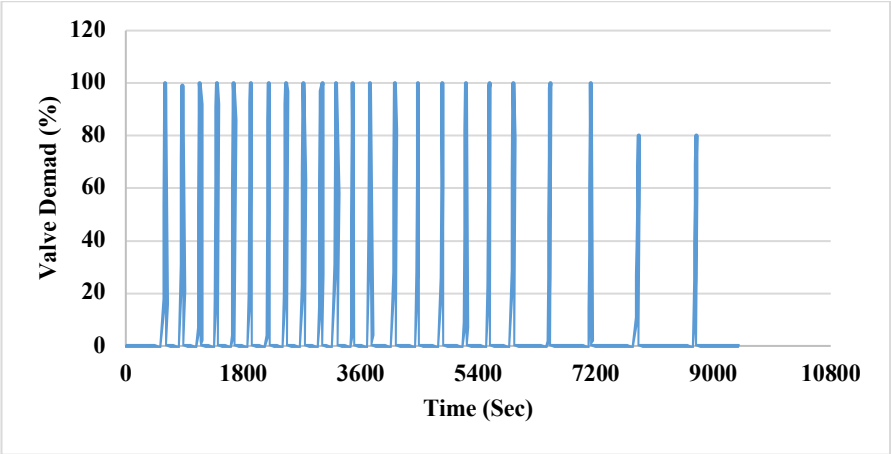


Figure 7-3: CDU Control Valve Demand(%) VS Time(Sec) at varying heat loads

To overcome this problem the control strategy of controlling the primary side flow rate which will help to stabilize the secondary-side supply temperature was developed. The control strategy is shown below in Figure 7-4

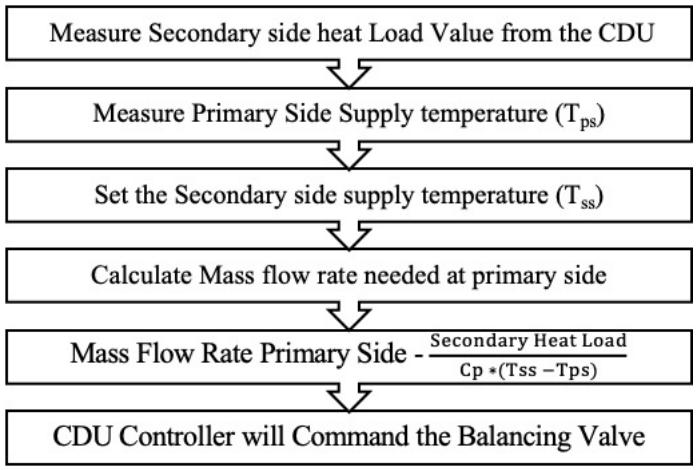


Figure 7-4: Flow Chart of the Control Strategy

Based on this control strategy the CDU will continuously measure the secondary load, primary supply temperature, and secondary set point temperature and calculate the mass flow rate

needed at the primary side to maintain the set secondary supply temperature using the equation (1) as shown below.

Mass Flow Rate Primary Side- Secondary Heat Load/(Cp *(Tss -Tps)) (1)

To verify this control strategy the experiment was performed at the default CDU setting and the Load varied from 39 kW to 6.5 kW. The primary side flow rate varied according to the mass flow rate calculated using the equation shown above. The secondary supply set temperature for the experiment was kept at 25 °C. Figure 7-5 is shown below secondary side supply temperature at different heat loads and primary side mass flow rates

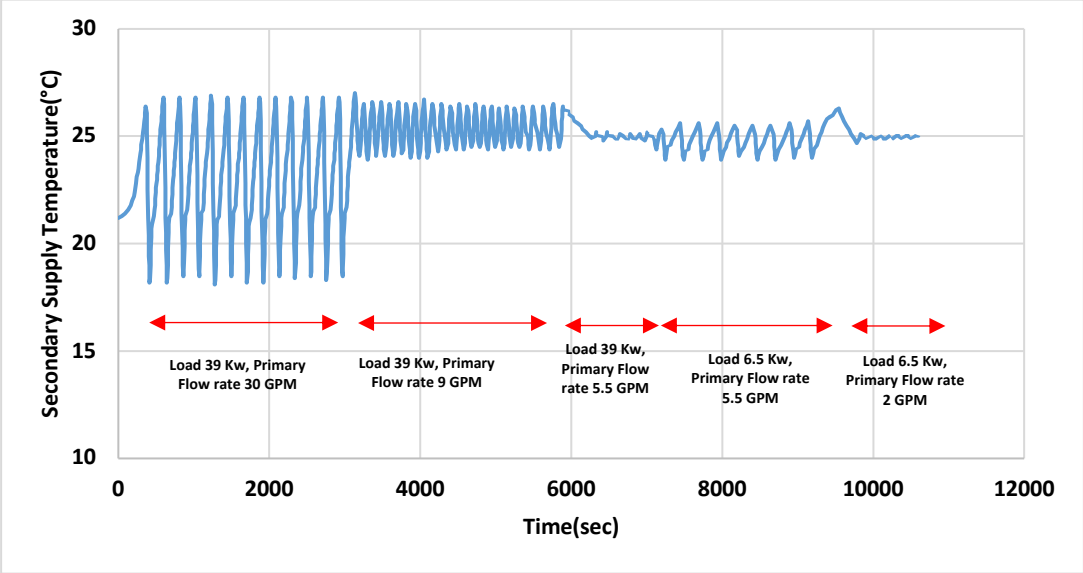


Figure 7-5:CDU Secondary Supply Temperature VS Time

In this experiment, as shown above in Figure 7-5, for the first 3000 seconds, the heat load was kept at 39 kW and the primary flow rate at 30 GPM. With default settings, the secondary side fluctuations can be seen from 27 °C to 18 °C. Then the flow rate on the primary side was reduced to 9 GPM from 3000 seconds to 6000 seconds, keeping the secondary heat load the same at 39 kW. The secondary side fluctuation saw a significant reduction in fluctuations between 26 and 23 °C. From 6000 seconds to 7500 seconds, the primary side flow rate was further reduced to 5 GPM and the secondary side supply temperature fluctuations were

completely dampened down, so the supply temperature just varied between □ 0.5 °C. From 7500 seconds to 9500 seconds, the heat load was reduced from 39 kW to 6.5 kW, keeping the primary flow rate at 5.5 GPM. As soon as the load was dropped to 6.5 kW, some fluctuations were observed, but these fluctuations were taken care of by reducing the primary flow rate to 2 GPM from 9500 seconds to 10500 seconds. This shows that, based on the developed control strategy, the secondary-side supply temperature can be stabilized.

Furthermore, the software-based PID experiments were performed to stabilize the secondary-side supply temperature. In these experiments, the first step is to follow the procedure to calculate the required parameters. The following steps were followed according to the Zeigler-Nichols method. In the first step in CDU, settings set the integral reset time and derivative reset time to 0 seconds. Set the proportional band to a higher value, such as 20 °C, and so on. The secondary side supply temperature will stabilize at a higher temperature than the set temperature. Then start reducing the proportional band setting until the supply temperature starts to oscillate at a constant rate. In the first set of experiments, the secondary supply temperature was set at 32 °C, the primary flow rate at 30 GPM, and the heat load at 39 kW.

In Figure 7-6 shown below, the same steps were followed as explained above. The first integral time and derivative time were set to 0, and then, as shown in the figure below, the proportional band (PB) value changed from 20 to 6 °C. At the PB value of 10, constant oscillation starts taking place. Oscillation time peak to peak was calculated for the PB 10, and it comes out to be 60 seconds. Then this oscillation value and proportional band value were used to calculate the new PB value, integral reset time, and derivative reset time.

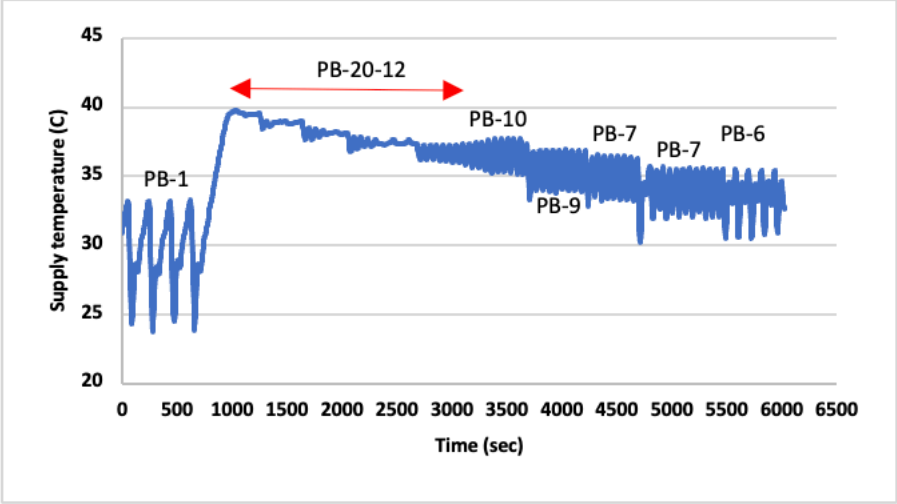


Figure 7-6: CDU-Secondary Supply Temperature(ºC) VS Time(Sec) at different PB values

For systems that have steady or slow-changing heat loads, PI control alone is sufficient. For the PI setting new proportional band, the setting was calculated as 2.2 times the proportional band setting at which the system becomes unstable and the integral reset time was 0.83 times the oscillation time. Based on these equations, the new proportional band value was calculated as 22 °C and the integral rest time as 49.8 sec.

New experiments were performed based on these changed settings of PI. The heat load remained at 39 kW, the secondary supply temperature was set to 32° C, and the primary flow rate was set to 30 GPM. In the Figure 7-7 shown above, up to 300 seconds of CDU operation was operated at default settings, and after that, the PI setting as calculated above was set to CDU, and it was observed that the 32°C set temperature was maintained at the secondary side with ± 0.8 °C.

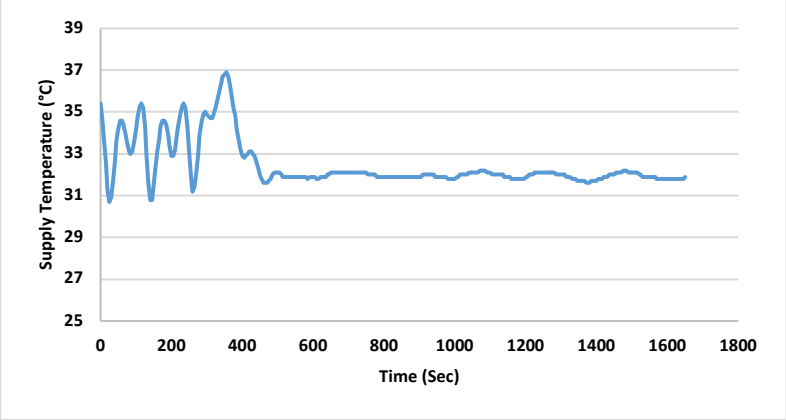


Figure 7-7: CDU Secondary Supply Temperature(ºC) Vs Time with new PI settings

For systems that see high and sudden changing heat loads, PID control is generally the preferred option. For PID settings, the new proportional band is calculated as 1.67 times the proportional band setting at which the system becomes unstable. The integral reset time is 0.5 times the oscillation time, and the derivative reset time is 0.125 times the oscillation time. Based on these calculations, new values were calculated as PB 16.7 °C, integral rest time as 30 sec, and derivative reset time as 7.5 sec.

A new experiment has been performed by keeping the heat load at 39 kW, the primary flow rate at 30 GPM, and the second set temperature at 32 °C. In the Figure 7-8 shown below, after 800 sec, a new PID setting was given to the CDU and the secondary side supply temperature was maintained at 32 °C with a variation of ± 0.5 °C after 1000 sec.

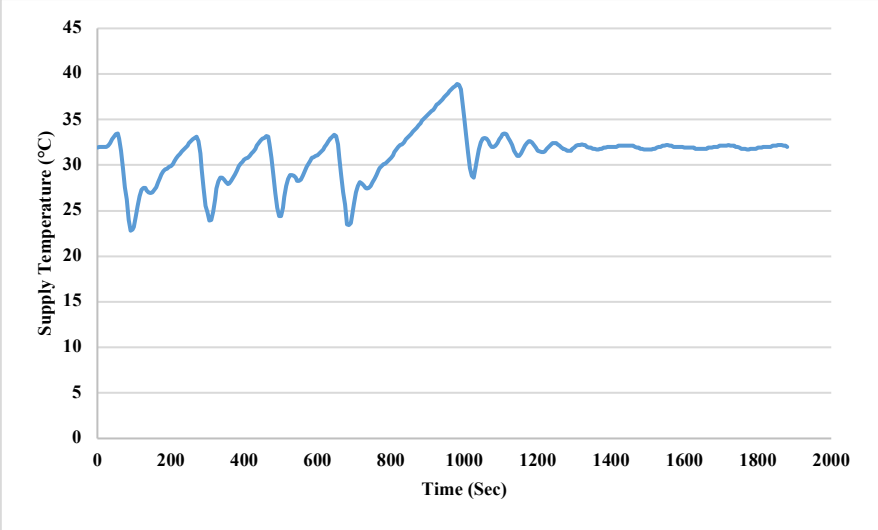


Figure 7-8: CDU-Secondary Supply Temperature(ºC) VS Time(Sec) at constant heat load(39kW) with new PID settings.

Following the determination of the PI or PID settings at 39 kW or 8.6% of the CDU rated capacity, the settings were examined when used for a very low heat load of 6.5 kW or 1.4% of the rated CDU capacity. In these experiments, the heat load was kept at 6.5 kW, the primary flow rate was set at 30 GPM, and the secondary side supply temperature was set at 32°C. Figure 7-9 shown below shows the secondary-side supply temperature getting stable to the secondary side set temperature of 32 °C after 1800 sec with a variation of ± 0.5 °C even with the PI and PID settings obtained at different heat load settings.

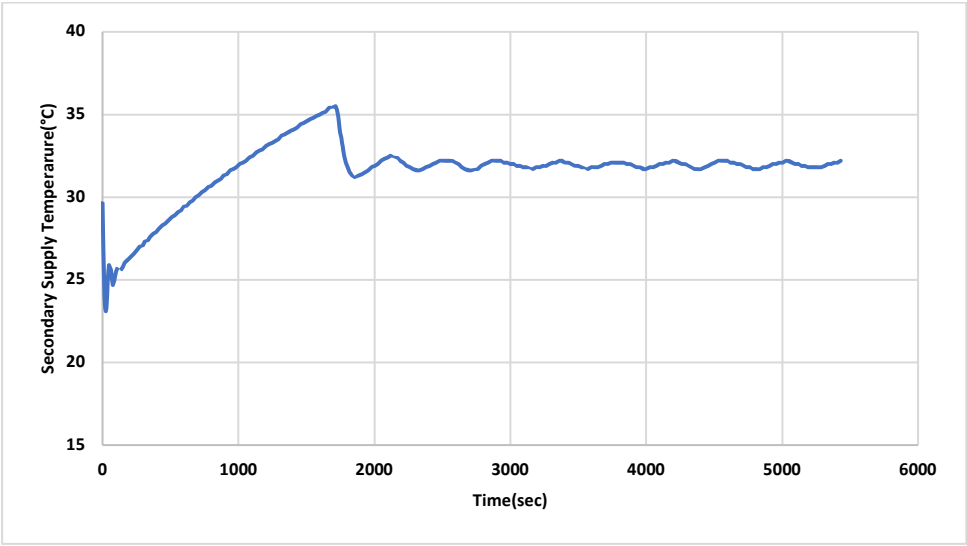


Figure 7-9: CDU-Secondary Supply Temperature(⁰C) VS Time(Sec) at constant heat load(6.5kW) with new PID settings

7.4 Conclusion

Typical CDUs are designed to operate for 20 to 30% of the rated heat load to get the stable secondary side coolant supply temperature. The present study will investigate operations of CDU at very low heat loads like 1% to 10% of CDU rated capacity. Large fluctuations in secondary side supply temperature around 10 □C were observed at these low loads. This large fluctuation can lead to the failure of the 3-way valve used in CDUs on the primary side as well as thermal shocks to the chips in the servers. In this paper, a control strategy is developed to stabilize the secondary supply temperature within ± 0.5 °C at very low heat loads using the combination of a flow control valve at the primary side and PID control settings within the CDU.

Chapter 8 Experimental Study of Transient Hydraulic Characteristics for Liquid Cooled Data Center Deployment

8.1 Abstract

Increasing demands for cloud-based computing and storage, Internet-of-Things, and machine learning-based applications have necessitated the utilization of more efficient cooling technologies. Direct-to-chip liquid cooling using cold plates has proven to be one of the most efficient methods to dissipate the high heat fluxes of modern high-power CPUs and GPUs. While the published literature has well-documented research on the thermal aspects of direct liquid cooling, a detailed account of transient hydraulic investigation is still missing. In this experiment, a total of four 52U racks with four high-power TTV- servers (Thermal Test Vehicles) in each rack were designed and deployed. Each server consists of eight GPU TTVs and six NV switch heaters. Each of the two racks has a different vendor rack manifold and cooling loop modules (CLM). A 450 kW coolant distribution unit (CDU) is used to supply 25% propylene glycol coolant to these racks. Each rack has its own rack-level flow control valve to maintain the same flow rate. The present study provides an in-depth analysis of hydraulic transients when rack-level flow control valves are used with and without flow control. The operating conditions of the CDU are varied for different parameters, such as a constant flow rate, constant differential pressure, and constant pump speed. Furthermore, hydraulic transient is examined when the cooling loop modules are decommissioned from the rack one by one. The effect of this step-by-step decommissioning is assessed on the CDU operation and other racks. The pressure drop-based control strategy has been developed to maintain the same flow rate in the remaining servers in the rack when some cooling loop modules are decommissioned.

8.2 Experimental Setup

For the experimental setup, four 52U racks were used, with each rack consisting of four high power thermal test vehicles (TTV). Each TTV consists of eight heaters mimicking GPUs and six heaters mimicking NV switches. Each server is attached with a cooling loop module to collect the heat dissipated from the heaters. A 450 kW Coolant Distribution Unit (CDU) was used to circulate the coolant across the racks. The coolant used for the experiment consists of 25% propylene glycol.

Two of the four racks are on each side. The cooling loop module attached to the TTV is connected via quick disconnects to the rack manifold for commissioning and decommissioning of the TTV. The coolant was circulated to the system from a 450KW CDU using Y-connectors that supplied the coolant on both sides using row manifolds. For every supply and return of a row manifold, port ball valves were connected. Thermistors and pressure sensors are connected at the inlet and outlet sides of each rack and on the supply and return sides of the CDU to take temperature readings and calculate pressure drop across the rack as well as across the system. From the row manifold, the coolant goes to the rack manifold and then to each TTV in the rack. Flow sensors were connected on the supply side of each TTV and each rack, as well as on the supply side of the CDU. On the return side of each rack, flow control valves are connected to adjust the flow rate at each rack, as shown in Figure 8-1

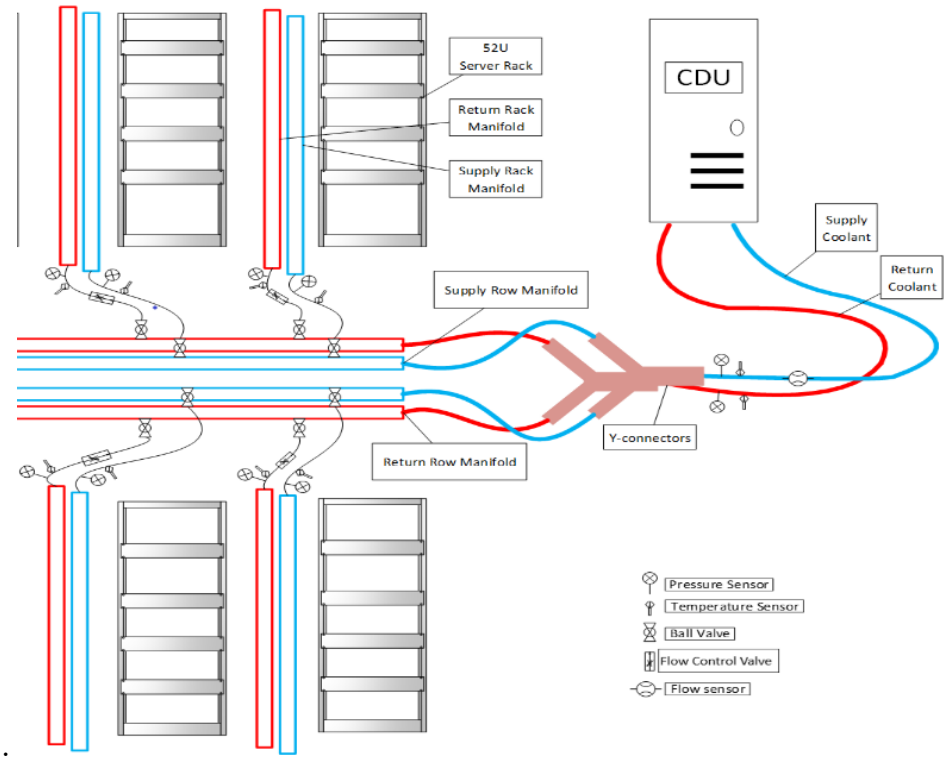


Figure 8-1: Schematic Layout for the Experimental Setup

The experiments are solely concerned with analyzing the transient hydraulic characteristics of liquid-cooled data centers with heterogeneous rack manifolds and cooling loops. No thermal power was given to any of the TTVs. The temperature of the coolant, propylene glycol 25% for these experiments, was maintained at 25°C. The experiments were performed at three CDU operating conditions: constant pump speed, constant differential pressure, and constant flow rate conditions. The present investigation provides an in-depth analysis of hydraulic transients when a rack-level flow control valve is used. In the first set of experiments, in all the CDU conditions, initially, rack-level flow control valves were kept fully open, and then the set flow rate of 40 lpm was instructed to the flow control device to maintain the same flow rate in all the racks. The purpose of these experiments was to understand the compatibility of rack-level flow control devices with the CDU and among themselves at different CDU operating conditions.

Other sets of experiments were performed where servers/TTVs were commissioned and decommissioned from the rack. These experiments were performed to understand how rack-level flow control devices and CDUs react when the servers/TTVs are commissioned and decommissioned from the rack. In the third set of experiments, a control strategy was introduced to maintain the same flow rate going to each server/TTV in the rack during the commissioning and decommissioning of servers/TTVs from the rack.

8.3 Results and Discussion

Experimental results are divided into 4 parts. In the first part, the operating condition at the CDU side was chosen as constant pump speed. In the second part, the CDU operating condition was chosen as constant differential pressure. In the third part, as a constant flow rate, and in the fourth part, a control strategy was introduced to keep the same flow rate in the remaining servers in the rack when the few TTVs with liquid cooling modules are decommissioned from the rack.

8.3.1 CDU operating Condition- Constant Pump Speed

In the first part of the experiments, as described above, the CDU was operated at constant pump speed. The speed of the pump was increased until all the racks were operating at a designed flow rate of a minimum of 40 lpm with the rack level flow control valve completely opened as shown in Figure 8-2 below. Figure 8-2 shows that at time 0, all the racks were getting different flow rates as all the racks with the servers were heterogeneous in nature. Rack 1,2, and 4 show very high flow rates because they have the least pressure drop, whereas rack 3 shows the lowest flow rate because it has the highest pressure drop. After some time, rack-level flow control valves were operated and set to a design flow rate of 40 lpm.

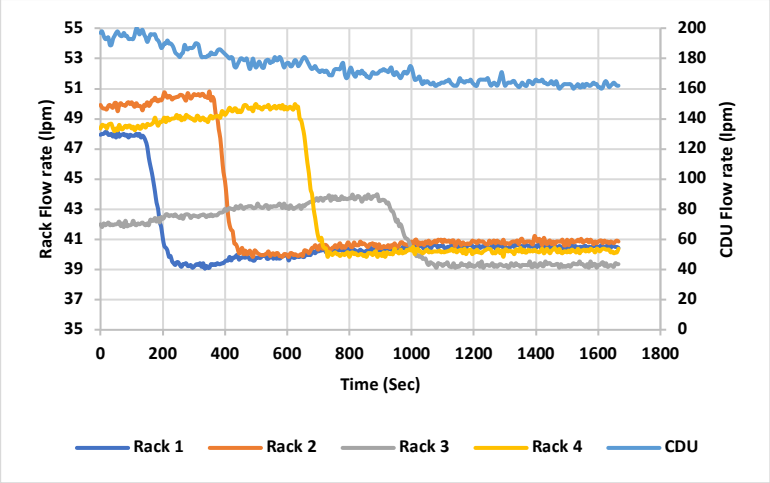


Figure 8-2: Flow rate vs time for rack and CDU at constant pump speed

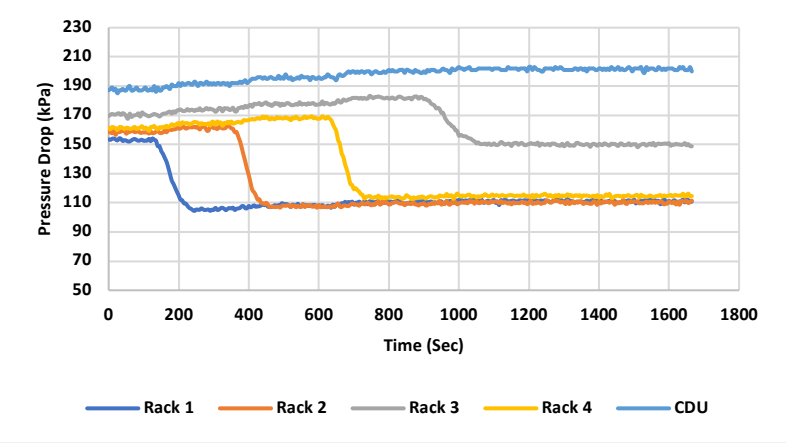


Figure 8-3: Pressure drop vs time graph for rack and CDU

Figure 8-3 shows the pressure drop of the racks before and after the rack level flow valve was operated. It also shows the change in the pressure drop increasing across the system from 190 kpa to 207 kpa as the control valves were operated. A corresponding change in system flow rate from 200 lpm to 160 lpm can be observed in Figure1. The rack-level flow control valve position change can be seen in Figure 8-4 below.

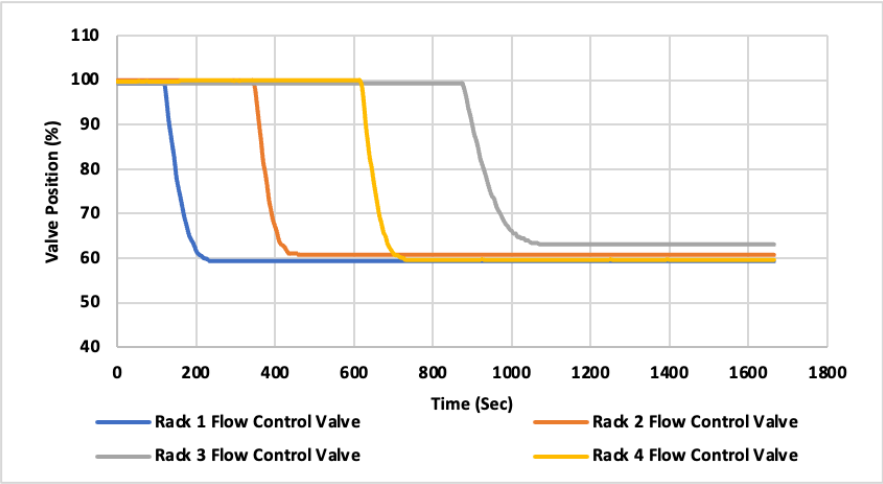


Figure 8-4: Valve position vs time graph of rack level flow control valve at constant pump speed

8.3.1.1 Hydraulic Transient for TTVs Decommissioning or Commissioning at Constant CDU Pump Speed

Once the 40 lpm flow rate was maintained in all the racks at constant pump speed as shown in Figure 8-2, the next decommissioning experiments were performed. In these experiments, 2 TTVs were decommissioned and then put back to understand how the rack level flow control valve and CDU respond to the decommissioning of TTVs.

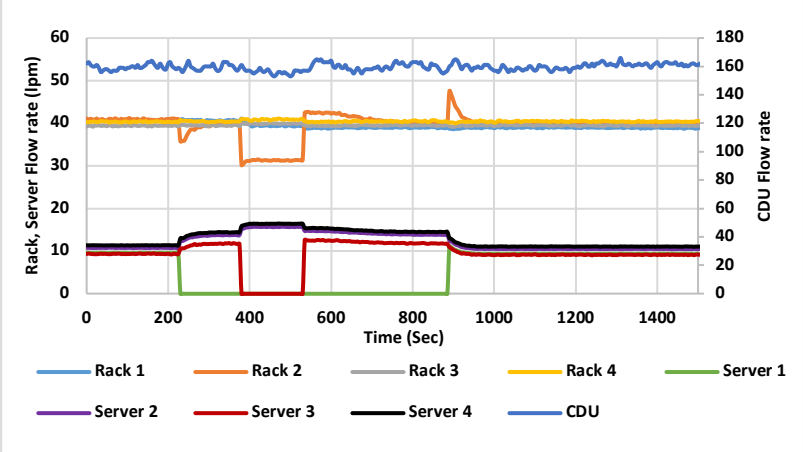


Figure 8-5: Flow rate Vs time graph of servers, racks and CDU

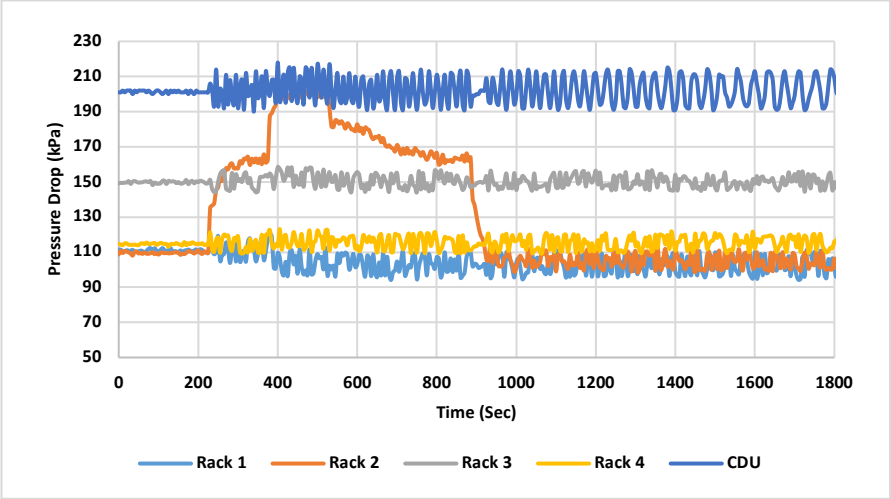


Figure 8-6: Pressure drop Vs time graph of racks and CDU

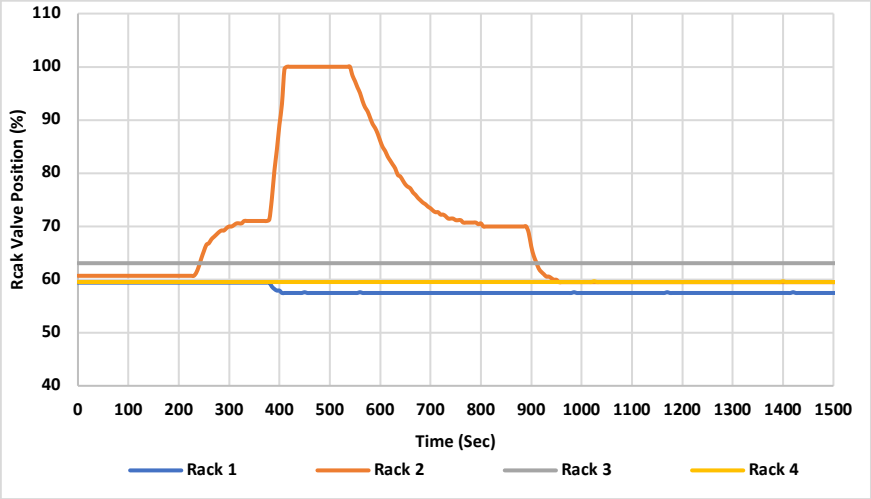


Figure 8-7: Rack Valve position Vs time graph

Figure 8-5 shows the flow rate going to four TTVs (servers) in rack 2, the flow rate of all the racks, and the system, i.e. the CDU flow rate. Till 200 seconds, the flow rate going to each of the TTVs in rack 2 was approximately 10 lpm. The flow rate going to racks was 40 lpm and the flow rate of CDU was 160 lpm. At 230 seconds, the TTV one was decommissioned from the rack. The flow rate going to other TTVs was increased to 12 to 14 lpm till 380 seconds,

as shown in the graph above. The corresponding rack 2 flow rate first dropped at 230 seconds because the pressure drop across that rack was increased suddenly, as shown in Figure 8-6, as the flow rate going to the remaining TTVs in the rack was increased. At 240 seconds, the rack control valve starts opening the damper as the flow rate set for it was 40 lpm. The damper valve position change can be observed in Figure 8-7 below, where the valve damper position change till 240 seconds for rack 2 was 59.6%, and when TTV 1 was removed, it started opening the valve position to meet the rack flow rate of 40 lpm and reached upto 70% valve position.

At 300 seconds, the flow rate of rack 2 again reaches 40 lpm, with the rack flow control valve position at 70%. At 375 seconds, TTV 2 was decommissioned and the flow rate going to the other 2 TTVs in rack 2 increased to 15 lpm each, approximately, and at the same time, the flow rate going to the rack dropped to 30 lpm at 380 seconds because the pressure drop of rack 2 was increased as shown in Figures 8-5 and 8-6. Meanwhile, the flow control valve opens itself to 100% to get the set flow rate of 40 lpm, which can be seen in Figure 8-7 from 380 seconds to 410 seconds. It can be observed that as the rack flow control valve opened itself to 100% at 410 seconds, the maximum flow rate it was able to deliver was 31 lpm.

At time 530 seconds the 2nd TTV that was decommissioned is put back into the rack 2 and a sudden jump in flow rate to 42.5 lpm was observed as flow rate going to TTVs now divided into 3 TTVs instead of 2 and pressure drop of the rack drop from 200 kPa to 182 kPa. Meanwhile, rack flow control valve starts changing the position of the damper from 100% to 70% to bring the flow rate of the damper down to 40 lpm.

At time 885 seconds 1st TTV put back into the rack similar trend as explained above when 2nd TTV put back was observed bringing rack level flow control device back to 59.5% of damper opening and maintaining the flow rate of 40 lpm for rack 2.

8.3.2 CDU Operating Condition of Constant Differential Pressure

In the second set of experiments, the CDU was operated at constant differential pressure. The value of differential pressure was increased until the desired flow rate of a minimum of 40 lpm in each rack was obtained while the rack level flow control valves were completely open. The set constant differential pressure for these experiments was 178 kPa.

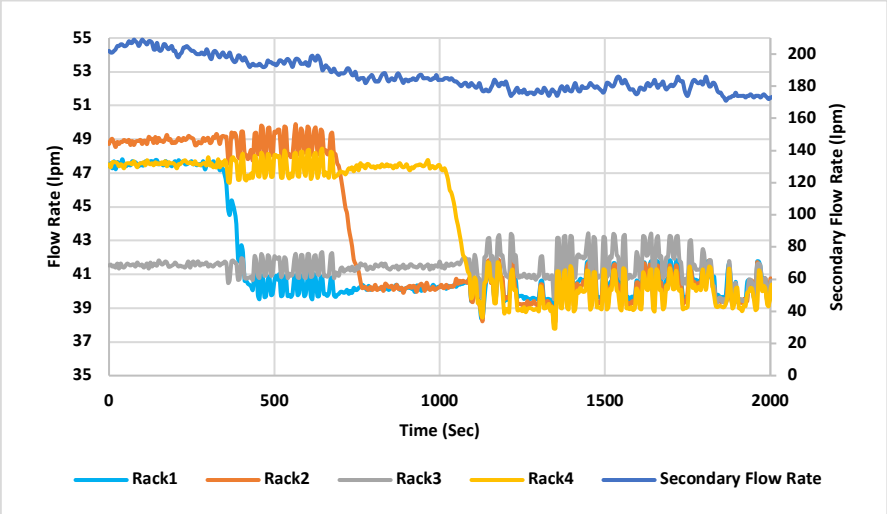


Figure 8-8: Flow rate Vs time graph of racks and CDU for constant differential pressure

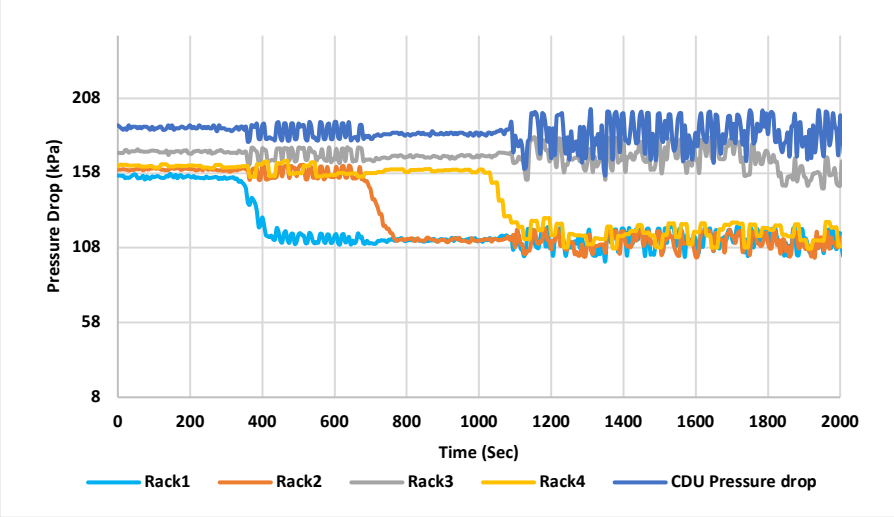


Figure 8-9: Pressure drop Vs time graph of racks and CDU

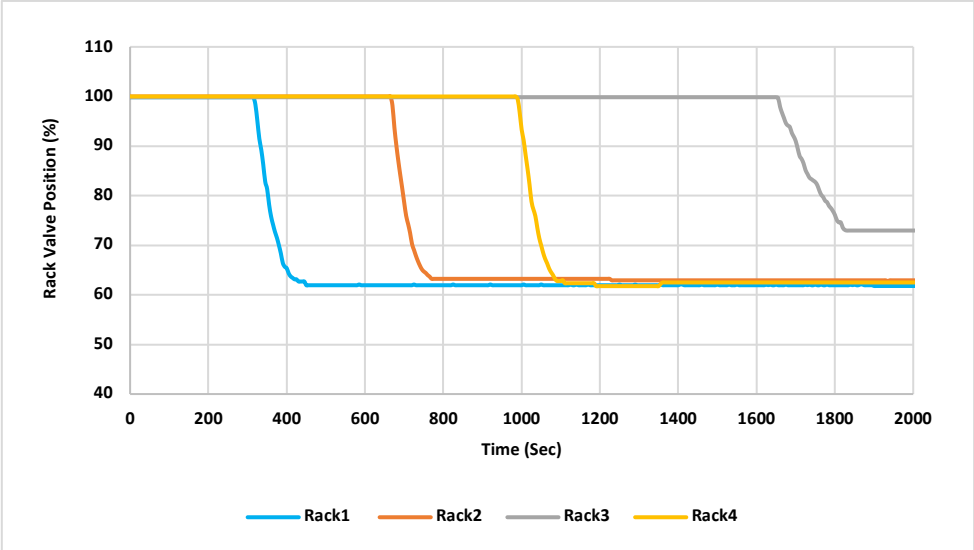


Figure 8-10: Rack Valve position Vs time graph for constant Differential pressure

Thr graph shown in Figure 8-8 depicts the flow rate of racks and CDU. It follows the same trend as explained in section 3.1 above. Flow rate of CDU change from 200 lpm to 160 lpm approximately when the rack flow control valves are operated to bring the flow rate to each rack to 40 lpm each.

Figure 8-9 shows the corresponding change in pressure drop, where it was observed that rack 3 pressure drop is more or less was equal to system or CDU pressure drop. This leads to the observation that CDU pressure drop is usually equal to the highest pressure drop rack in constant pressure differential case, which was not the case when CDU was operated at constant pump speed. Figure 8-10 shows the rack level valve position during this experiment.

8.3.2.1 Hydraulic Transient for TTVs Decommissioning or Commissioning at Constant Differential pressure Condition at CDU

After maintaining a 40 lpm flow rate in all the racks for constant differential pressure conditions at CDU, TTV decommissioning experiments were performed. In these experiments, three TTVs were decommissioned one at a time and then reinstalled in the rack, with the hydraulic behavior of the racks, CDU, and remaining TTVs being observed.

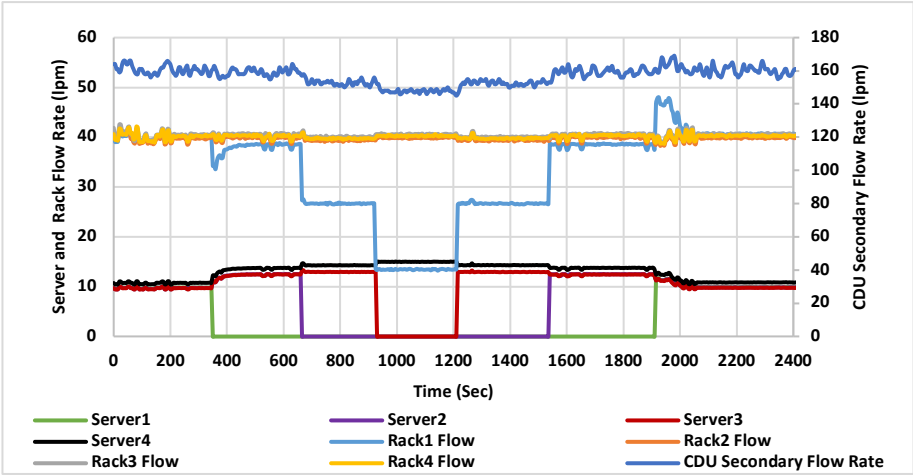


Figure 8-11: Flow rate Vs time graph of servers, racks and CDU

Figure 8-11 depicts the flow rate of the servers, racks, and CDU. The graph follows mostly the same trend as the constant pump speed case. Although it was observed that there was an increase in flow rate in the remaining servers in the rack, it was 1 to 2 lpm less than that of a constant pump speed case when 2 TTVs were decommissioned. Second, it was observed that the CDU flow rate dropped from 160 lpm to 150 lpm while decommissioning of three TTVs took place. That means the CDU tried to control the flow rate going to the remaining TTVs in the rack and act as self-regulating, but still, the flow rate going to TTV was much higher than needed. We have tried to resolve this issue in the fourth part of the results, where a control strategy is developed to maintain the same flow rate in the remaining TTVs in the rack even if the TTVS are decommissioned or put back into the rack

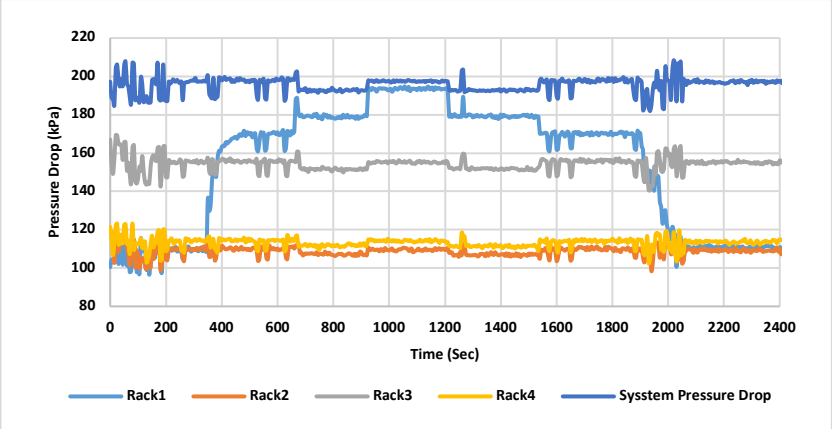


Figure 8-12: Pressure drop Vs time graph of racks and CDU

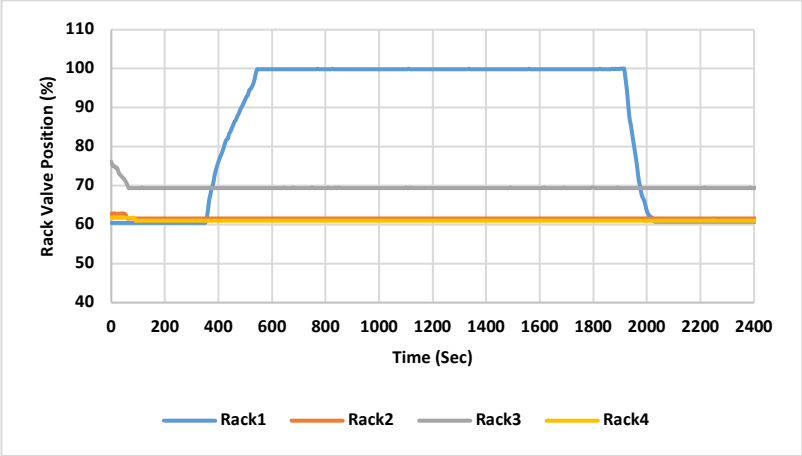


Figure 8-13: Rack Valve position Vs time graph

Figures 8-12 and 8-13 show the pressure drop of racks, CDU, and rack flow control valve positions respectively. The pressure drop of rack 1 increased as the TTVs were removed from rack 1, but the pressure drop of CDU and other racks was maintained. For rack-level flow control valves. Rack one flow control valve position was initially maintained at 60% to supply the flow rate of 40 lpm to rack 1, but when TTV 1 was decommissioned, it was opened to 100% but could only supply 38 lpm to rack 1, and then remained 100% open for the rest of the time until TTV 1 was reinstalled.

8.3.3 CDU Operating Condition of Constant Flow Rate

In the third set of experiments, CDU was operated at constant flow rate condition. First of all, the flow rate was increased to the point where each rack was at minimum 40 lpm when the rack level flow control valve is opened 100% and then the flow control valves is instructed to maintain 40 lpm in each rack. Figure 8-14 shows the flow rate of racks and CDU. Rack 3 was getting minimum flow rate because it was having maximum pressure drop at time = 0 sec. First rack 1 then rack 2 and rack 4 bring to 40 lpm by operating the rack flow control valve. Meanwhile, flow rate going to rack 3 reached 53 lpm, and then it came down to 40 lpm.

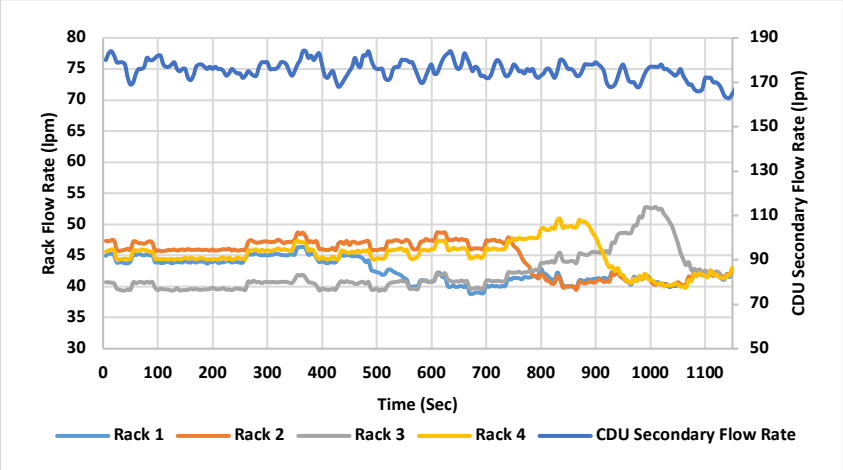


Figure 8-14: Flow rate Vs time graph of racks and CDU for constant Flow rate condition

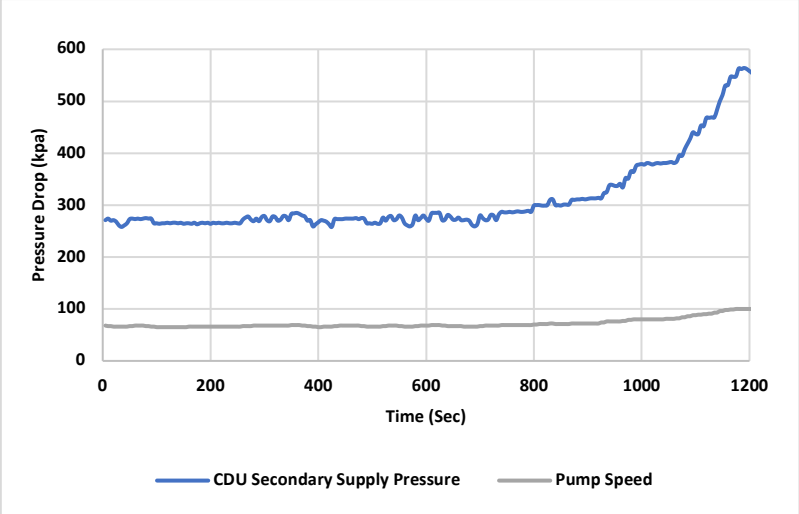


Figure 8-15: CDU Pressure drop, Pump Speed Vs time graph of racks and CDU for constant Flow rate condition

Figure 8-15 depicts the CDU pressure drop and pump speed during a constant flow rate condition at CDU. From the graph, it can be observed that CDU pressure reaches a very high level of 500 kPa and pump speed is 100%. This can be attributed to the feedback control mechanism of the CDU to maintain the flow rate of 160 lpm. In our experimental design, to modulate the flow rate and simultaneously control the rack control valve, the CDU inherently increases the pump speed to meet the desired flow rate condition. This causes very high supply pressure, with pump speed reaching 100%. However, the experiment could be continued further due to the extreme conditions to maintain the constant flow rate.

8.3.4 Novel Control Strategy

As seen in sections above, when the cooling loops modules or TTVs are decommissioned, there is an increase in the flow rate going to other remaining TTVs in the rack. This high flow rate can lead to leakage as well as erosion problems in cold plates. Out of the three cases discussed above, constant differential pressure at CDU gave the best results,

where the least flow rate increase was observed in the remaining TTVs in the rack when a few TTVs were decommissioned.

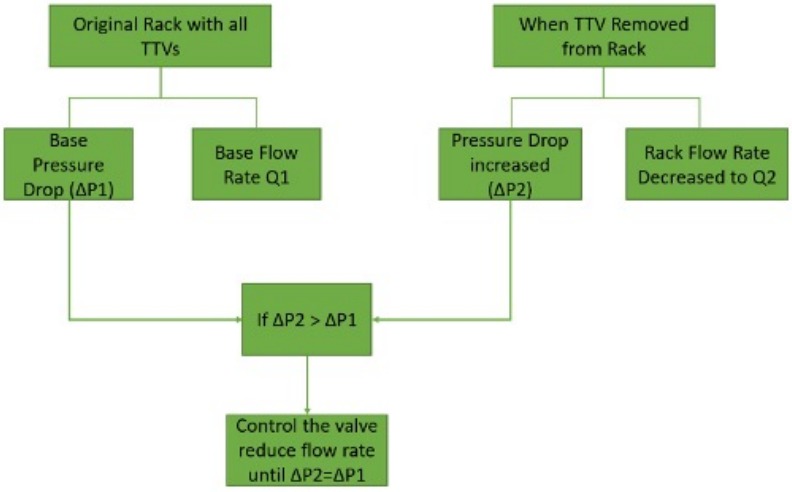


Figure 8-16: Control Strategy when TTVs is decommissioned from the rack

The pressure drop based control strategy was developed to over come this problem. The base pressure drop ΔP1, which is the pressure drop of the rack when all the TTVs are inside the rack and all the TTVs are getting the designed flow rate based on the thermal design power of the server. Once the ΔP1 is known for the original rack and TTV is removed, the pressure drop ΔP2 of that rack will increase as shown in the sections above. To overcome this, we need to instruct the rack level flow control valve to reduce the flow rate until the pressure drop ΔP2 comes equal to the base pressure drop of ΔP1 as shown in the Figure 8-16 above.

Similarly, Figure 8-17 shows the condition when the TTV is placed back into rack, the pressure drop of rack decreased to ΔP3, so the flow rate going to rack needs to be increased to match the base pressure drop of ΔP1.

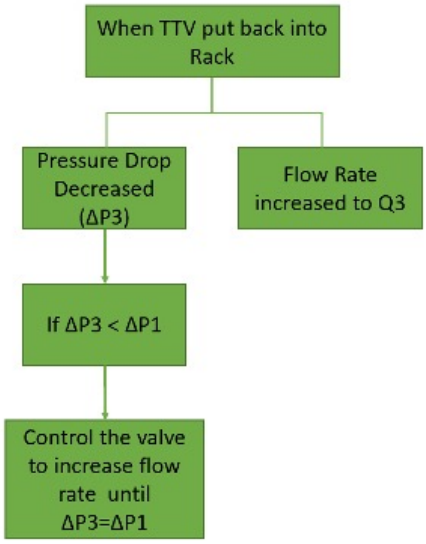


Figure 8-17: Control Strategy when TTVs is commissioned back into rack

The experiments discussed in the following sections were conducted to prove the control strategy. Figures 8-18, 8-19, and 8-20 show the flow rate, pressure drop, and rack control valve position vs time graphs. At time = 0 sec, the flow rate going to each TTV, which was represented as servers, was 10 lpm, the flow rate going to each rack was 40 lpm, and the CDU flow rate was at 160 lpm. At t = 310 sec, a sudden jump was observed in the pressure drop of rack 1, from which TTV 1 was removed, and the rack pressure drop changed from 110 kPa to 170 kPa and the corresponding increase in flow rate in the remaining TTVs in the rack reached from 10 lpm to 12.5 lpm each. At time t = 620 sec, the rack control valve is instructed to send 30 lpm to the rack instead of 40 lpm as the remaining TTVs in the rack were 3 instead of 4. As soon as the rack flow control valve maintains 30 lpm to the rack and the flow rate going to each server again comes back to 10 lpm. At the same time, t=700 sec, the pressure drop of the rack comes back from 170 kpa to 110 kPa, which proves the control strategy shown in Figure 8-15.

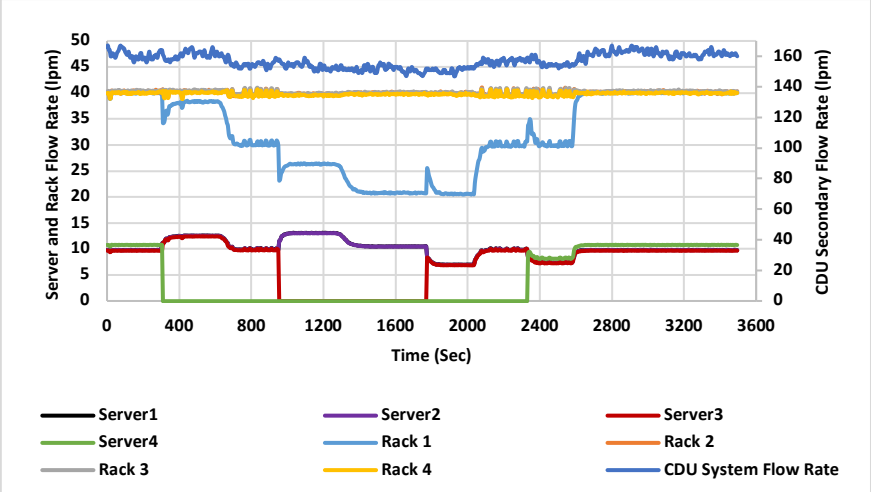


Figure 8-18: Flow rate Vs time graph for TTVs, Racks and CDU

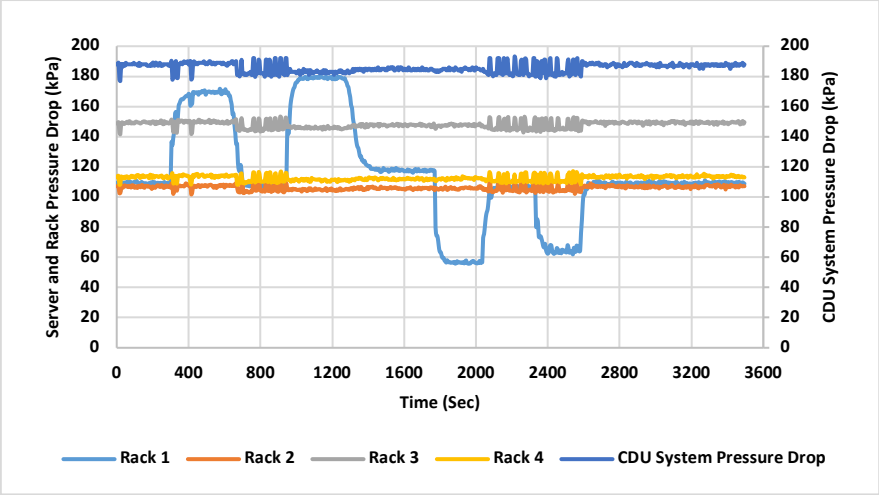


Figure 8-19: Pressure Drop Vs time graph for Racks and CDU

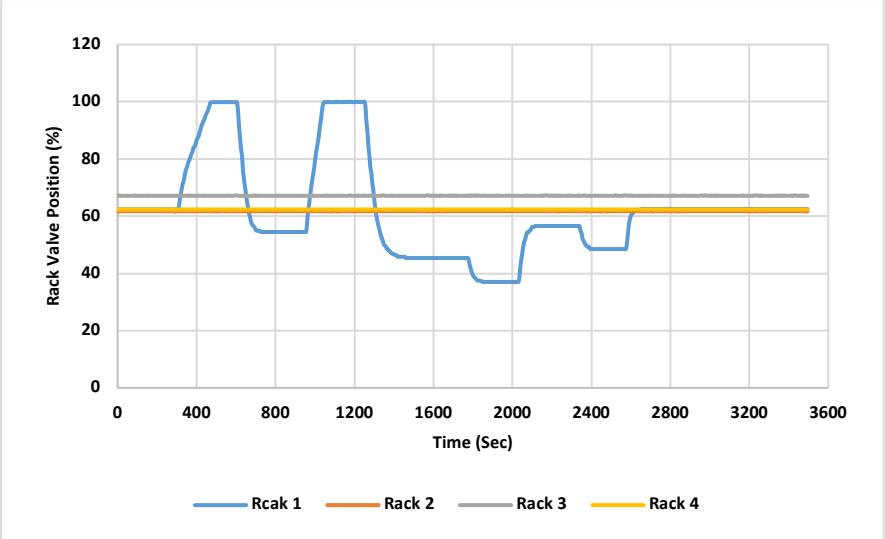


Figure 8-20: Rack Flow Control Valve position Vs time graph

A similar trend can be seen when the TTV 2 was removed from the rack at time t = 995 sec. At time t= 1775 sec, TTV 2 was put back into the rack. The flow rate goin to each TTV decreased to 6.8 lpm When the rack control valve was instructed at time t= 2040 sec to send 30 lpm to the rack, the pressure drop of the rack increased from 58 kPa to 110 kPa which proves the control strategy explained in Figure 8-17 and maintains the flow rate of 10 lpm to each TTV. A similar trend was observed when the TTV 1 was put back into the rack

8.4 Conclusion and Future Work

The present investigation provides an in-depth analysis of hydraulic transients when rack-level flow control valves with and without flow control are opened. The operating conditions of the CDU varied from a constant flow rate, a constant differential pressure, and a constant pump speed. Further hydraulic transients are investigated when the cooling loop modules are decommissioned from the rack one by one. The influence of this step-by-step decommissioning is assessed on the CDU operation and other racks.

The experimental strategy followed in this paper helps us to conclude that, out of the three operating conditions at the CDU level, constant differential pressure works best and

shows better compatibility with rack-level flow controlled devices. The decommissioning of TTVs from the rack shows an increase in the flow rate going to the remaining TTVs in the rack, but the constant differential pressure condition at the CDU shows the least increase in flow rate as compared to other conditions. The constant flow rate condition is not at all compatible with rack level flow control devices. The constant pump speed condition gave better results than the constant flow rate condition at CDU, but it needs to be investigated more. The pressure drop-based control strategy was developed to maintain the same flow rate in the remaining servers in the rack when some cooling loop modules were decommissioned.

Chapter 9 REFERENCES

[1] A. Shehabi *et al.*, "United States Data Center Energy Usage Report," 2016.

[2] H. Geng, *Data Center Handbook*. Wiley Blackwell, 2014.

[3] Andy Lawrence, April 2020, "Data center PUEs flat since 2013", Accessed January 7, 2020, Global Uptime Institute Survey, https://journal.uptimeinstitute.com/data-center-pues-flat-since-2013/

[4] Tony Day, Paul Lin, Rober Bunger, "Liquid Cooling Technologies for Data Centers and Edge Applications", white paper 265, Schneider Electric.

[6] Modi, Himanshu. "Computational Study of Air-Cooled Servers with Improved Ducting and Chassis Re-design." PhD diss., The University of Texas at Arlington, 2020.

[7]. Modi, Himanshu, Pardeep Shahi , Lochan Sai Reddy Chintaparthy, Gautam Gupta, Pratik Bansode, Vibin Simon and Dereje Agonafer. "Experimental Investigation of the Impact of Improved Ducting and Chassis Re-design of a Hybrid-Cooled Server", ASME 2022 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems, InterPACK, 2022.

[8] Modi, Himanshu, Uschas Chowdhury and Dereje Agonafer, "Impact of Improved Ducting and Chassis Re-design for Air-Cooled Servers in a Data Center", ITHERM 2022 The 21st Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems, 2022.

[9] Simon, Vibin Shalom, Himanshu Modi, Krishna Sivaraju, Pratik Bansode, Satyam Saini, Pardeep Shahi, Saket Karajgikar, Veerendra Mulay and Dereje Agonafer, "Feasibility study of Rear Door Heat Exchanger for a High Capacity Data Center", ASME 2022 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems, InterPACK, 2022.

[10] Singh, Saurabh, Kourosh Nemati, Vibin Simon, Ashwin Siddarth, Mark Seymour, and Dereje Agonafer. "Sensitivity Analysis of a Calibrated Data Center Model to Minimize the Site Survey Effort." In 2021 37th Semiconductor Thermal Measurement, Modeling & Management Symposium (SEMI-THERM), pp. 50-57. IEEE, 2021.

[11] Simon, Vibin Shalom, Ashwin Siddarth, and Dereje Agonafer. "Artificial neural network based prediction of control strategies for multiple air-cooling units in a raised-floor data center." In 2020 19th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), pp. 334-340. IEEE, 2020.

[12] Saini, Satyam, Jimil M. Shah, Pardeep Shahi, Pratik Bansode, Dereje Agonafer, Prabjit Singh, Roger Schmidt, and Mike Kaler. "Effects of Gaseous and Particulate Contaminants on Information Technology Equipment Reliability—A Review." Journal of Electronic Packaging 144, no. 3 (2022).

[13] Saini, Satyam, Pardeep Shahi, Pratik Bansode, Jimil M. Shah, and Dereje Agonafer. "Simplified and Detailed Analysis of Data Center Particulate Contamination at Server and Room Level Using Computational Fluid Dynamics." Journal of Electronic Packaging 144, no. 2 (2022): 024501.

[14] Saini, Satyam. "Airflow path and flow pattern analysis of sub-micron particulate contaminants in a data center with hot aisle containment system utilizing direct air cooling." PhD diss., The University of Texas at Arlington, 2018

[15] Thirunavakkarasu, Gautham, Satyam Saini, Jimil Shah, and Dereje Agonafer. "Air flow pattern and path flow simulation of airborne particulate contaminants in a high-density data center utilizing airside economization." In International Electronic Packaging Technical Conference and Exhibition, vol. 51920, p. V001T02A011. American Society of Mechanical Engineers, 2018.

[16] Saini, Satyam, Kaustubh K. Adsul, Pardeep Shahi, Amirreza Niazmand, Pratik Bansode, and Dereje Agonafer. "CFD Modeling of the Distribution of Airborne Particulate Contaminants Inside Data Center Hardware." In International Electronic Packaging Technical Conference and Exhibition, vol. 84041, p. V001T08A005. American Society of Mechanical Engineers, 2020.

[17] Saini, Satyam, Pardeep Shahi, Pratik Bansode, Ashwin Siddarth, and Dereje Agonafer. "CFD investigation of dispersion of airborne particulate contaminants in a raised floor data center." In 2020 36th Semiconductor Thermal Measurement, Modeling & Management Symposium (SEMI-THERM), pp. 39-47. IEEE, 2020.

[18] Shah, Jimil M., Roshan Anand, Satyam Saini, Rawhan Cyriac, Dereje Agonafer, Prabjit Singh, and Mike Kaler. "Development of a technique to measure deliquescent relative humidity of particulate contaminants and determination of the operating relative humidity of a data center." In International Electronic Packaging Technical Conference and Exhibition, vol. 59322, p. V001T02A016. American Society of Mechanical Engineers, 2019.

[19] Shah, Jimil M., Roshan Anand, Prabjit Singh, Satyam Saini, Rawhan Cyriac, Dereje Agonafer, and Mike Kaler. "Development of a Precise and Cost-Effective Technique to Measure Deliquescent Relative Humidity of Particulate Contaminants and Determination of the Operating Relative Humidity of a Data Center Utilizing Airside Economization." Journal of Electronic Packaging 142, no. 4 (2020).

[20] Shinde, Pravin A., Pratik V. Bansode, Satyam Saini, Rajesh Kasukurthy, Tushar Chauhan, Jimil M. Shah, and Dereje Agonafer. "Experimental analysis for optimization of thermal performance of a server in single phase immersion cooling." In International Electronic Packaging Technical Conference and Exhibition, vol. 59322, p. V001T02A014. American Society of Mechanical Engineers, 2019.

[21] Bansode, Pratik V., Jimil M. Shah, Gautam Gupta, Dereje Agonafer, Harsh Patel, David Roe, and Rick Tufty. "Measurement of the thermal performance of a custom-build single-phase immersion cooled server at various high and low temperatures for prolonged time." Journal of Electronic Packaging 142, no. 1 (2020): 011010.

[22] Bansode, Pratik V., Jimil M. Shah, Gautam Gupta, Dereje Agonafer, Harsh Patel, David Roe, and Rick Tufty. "Measurement of the thermal performance of a single-phase immersion cooled server at elevated temperatures for prolonged time." In International Electronic Packaging Technical Conference and Exhibition, vol. 51920, p. V001T02A010. American Society of Mechanical Engineers, 2018.

[23] Shah, Jimil M., Chinmay Bhatt, Pranavi Rachamreddy, Ravya Dandamudi, Satyam Saini, and Dereje Agonafer. "Computational form factor study of a 3rd generation open compute server for single-phase immersion cooling." In International Electronic Packaging Technical Conference and Exhibition, vol. 59322, p. V001T02A017. American Society of Mechanical Engineers, 2019.

[24] Gandhi, Dhruvkumar, Uschas Chowdhury, Tushar Chauhan, Pratik Bansode, Satyam Saini, Jimil M. Shah, and Dereje Agonafer. "Computational analysis for thermal optimization of server for single phase immersion cooling." In International Electronic Packaging Technical Conference and Exhibition, vol. 59322, p. V001T02A013. American Society of Mechanical Engineers, 2019.

[25] Niazmand, Amirreza, Prajwal Murthy, Satyam Saini, Pardeep Shahi, Pratik Bansode, and Dereje Agonafer. "Numerical analysis of oil immersion cooling of a server using mineral oil and Al2O3 nanofluid." In International Electronic Packaging Technical Conference and Exhibition, vol. 84041, p. V001T08A009. American Society of Mechanical Engineers, 2020

[26] Niazmand, Amirreza, Tushar Chauhan, Satyam Saini, Pardeep Shahi, Pratik Vithoba Bansode, and Dereje Agonafer. "CFD simulation of two-phase immersion cooling using FC-

72 dielectric fluid." In International Electronic Packaging Technical Conference and Exhibition, vol. 84041, p. V001T07A009. American Society of Mechanical Engineers, 2020

[27] Shah, Jimil M., Richard Eiland, Pavan Rajmane, Ashwin Siddarth, Dereje Agonafer, and Veerendra Mulay. "Reliability Considerations for Oil Immersion-Cooled Data Centers." Journal of Electronic Packaging 141, no. 2 (2019): 021007.

[28] Ramdas, Shrinath, Pavan Rajmane, Tushar Chauhan, Abel Misrak, and Dereje Agonafer. "Impact of immersion cooling on thermo-mechanical properties of PCB's and reliability of electronic packages." In International Electronic Packaging Technical Conference and Exhibition, vol. 59322, p. V001T02A011. American Society of Mechanical Engineers, 2019.

[29] Shah, Jimil M., Keerthivasan Padmanaban, Hrishabh Singh, Surya Duraisamy Asokan, Satyam Saini, and Dereje Agonafer. "Evaluating the Reliability of Passive Server Components for Single-Phase Immersion Cooling." Journal of Electronic Packaging 144, no. 2 (2022).

[30] Saini, Satyam, Tushar Wagh, Pratik Bansode, Pardeep Shahi, Joseph Herring, Jacob Lamotte-Dawaghreh, Jimil M. Shah, and Dereje Agonafer. "A Numerical Study on Multi-objective Design Optimization of Heatsinks for Forced and Natural Convection Cooling of Immersion Cooled Servers." Journal of Enhanced Heat Transfer.

[31] Shah, Jimil M., Ravya Dandamudi, Chinmay Bhatt, Pranavi Rachamreddy, Pratik Bansode, and Dereje Agonafer. "CFD Analysis of Thermal Shadowing and Optimization of Heatsinks in 3rd Generation Open Compute Server for Single-Phase Immersion Cooling." In International Electronic Packaging Technical Conference and Exhibition, vol. 59322, p. V001T02A015. American Society of Mechanical Engineers, 2019.

[32] Misrak, Abel, Tushar Chauhan, Pavan Rajmane, Rabin Bhandari, and Dereje Agonafer. "Impact of aging on mechanical properties of thermally conductive gap fillers." Journal of Electronic Packaging 142, no. 1 (2020).

[33] Chauhan, T., A. Misrak, R. Bhandari, P. Rajmane, A. S. M. R. Chowdhury, K. B. Sivaraju, M. Abdulhasansari, and D. Agonafer. "Impact of thermal aging and cycling on reliability of thermal interface materials." In SMTA International, pp. 118-123. 2019.

[34] Rajmane, Pavan, Karthikeyan Dhandapani, Mark Schwarz, and Ahmer Syed. "Investigation of the Factors Affecting the Warpage Prediction of Multi-Chip Package." In 2021 IEEE 71st Electronic Components and Technology Conference (ECTC), pp. 1514-1520. IEEE, 2021.

[35] Rahangdale, Unique, Rahul Srinivas, S. Krishnamurthy, Pavan Rajmane, Abel Misrak, A. R. Sakib, Dereje Agonafer, Alok Lohia, Steven Kummerl, and Luu T. Nguyen. "Effect of PCB thickness on solder joint reliability of Quad Flat no-lead assembly under Power Cycling and Thermal Cycling." In 2017 33rd Thermal Measurement, Modeling & Management Symposium (SEMI-THERM), pp. 70-76. IEEE, 2017.

[36] Rahangdale, Unique, Pavan Rajmane, Abel Misrak, and Dereje Agonafer. "A Computational Approach to Study the Impact of PCB Thickness on QFN Assembly Under Drop Testing With Package Power Supply." In International Electronic Packaging Technical Conference and Exhibition, vol. 58097, p. V001T01A019. American Society of Mechanical Engineers, 2017.

[37] Rajmane, Pavan, Hassaan Ahmad Khan, Aniruddha Doiphode, Unique Rahangdale, Dereje Agonafer, Alok Lohia, Steven Kummerl, and Luu Nguyen. "Failure mechanisms of boards in a thin wafer level chip scale package." In 2017 16th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), pp. 1099-1105. IEEE, 2017.

[38] Rajmane, Pavan, Fahad Mirza, Hassaan Khan, and Dereje Agonafer. "Chip package interaction study to analyze the mechanical integrity of a 3-D TSV package." In International

Electronic Packaging Technical Conference and Exhibition, vol. 56895, p. V002T02A009. American Society of Mechanical Engineers, 2015.

[39] Rahangdale, Unique, Pavan Rajmane, Aniruddha Doiphode, and Abel Misrak. "Structural integrity optimization of 3D TSV package by analyzing crack behavior at TSV and BEOL." In 2017 28th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC), pp. 201-208. IEEE, 2017.

[40] Rahangdale, Unique, Pavan Rajmane, Aniruddha Doiphode, A. R. Sakib, Abel Misrak, Alok Lohia, and Dereje Agonafer. "Damage progression study of 3D TSV package during reflow, thermal shocks and thermal cycling." In 2017 16th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), pp. 1119-1125. IEEE, 2017.

[41] Rahangdale, Unique, Bhavna Conjeevaram, Aniruddha Doiphode, Pavan Rajmane, Abel Misrak, A. R. Sakib, Dereje Agonafer, Luu T. Nguyen, Alok Lohia, and Steven Kummerl. "Solder ball reliability assessment of WLCSP—Power cycling versus thermal cycling." In 2017 16th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), pp. 1361-1368. IEEE, 2017.

[42] Denria, Jyotirmoy, Pavan Rajmane, and Dereje Agonafer. "Board level solder joint reliability assessment study of Megtron 6 vs FR-4 under power cycling and thermal cycling." In 2018 17th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), pp. 1289-1295. IEEE, 2018.

[43] Chaudhari, Mugdha, A. S. M. R. R. Chowdhury, U. Rahangdale, A. Misrak, P. Rajmane, A. Doiphode, and D. Agonafer. "Reliability assessment of bga solder joints-Megtron 6 vs FR4 printed circuit boards." Proceedings of SMTA International. Rosemont, IL, USA: SMTA (2019): 136-141.

[44] Rahangdale, Unique, Pavan Rajmane, Abel Misrak, and Dereje Agonafer. "Reliability Analysis of Ultra-Low-K Large-Die Package and Wire Bond Chip Package on Varying Structural Parameter Under Thermal Loading." In International Electronic Packaging Technical Conference and Exhibition, vol. 58097, p. V001T01A020. American Society of Mechanical Engineers, 2017.

[45] Rajmane, Pavan. "Multi-Physics Design Optimization of 2D and Advanced Heterogenous 3D Integrated Circuits." PhD diss., The University of Texas at Arlington, 2018.

[46] Chu, R. C., Simons, R. E., Ellsworth, M. J., Schmidt, R. R., and Cozzolino, V., 2004, "Review of Cooling Technologies for Computer Products," IEEE Trans. Device Mater. Reliab., 4(4), pp. 568–585.

[47] Ellsworth, M. J., Campbell, L. A., Simons, R. E., Iyengar, M. K., Schmidt, R. R., and Chu, R. C., 2008, "The Evolution of Water Cooling for Large IBM Large Server Systems: Back to the Future," 11th Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), Orlando, FL, May 28–31, pp. 266–274.

[48] McFarlane, R., 2012, "Will Water-Cooled Servers Make Another Splash in the Data Center?," Tech Target Network, Search Data Center, Newton, MA, accessed Feb. 26, 2012, https://searchdatacenter.techtarget.com/tip/Will-watercooled-servers-make-another-splash-in-the-data-center

[49] Schmidt, R. R., 2005, "Liquid Cooling Is Back," Electronics Cooling, 11(3), (epub).

[50] Shahi, Pardeep, Satyam Saini, Pratik Bansode, and Dereje Agonafer. "A comparative study of energy savings in a liquid-cooled server by dynamic control of coolant flow rate at server level." IEEE Transactions on Components, Packaging and Manufacturing Technology 11, no. 4 (2021): 616-624.

[51] Modi, Himanshu, Pardeep Shahi, Gautam Gupta, Pratik Bansode, Satya, Saini, Vibin Simon, Akiilessh Sivakumar, Amrutha Rachakonda and Dereje Agonafer, "Transient CFD

Analysis of Dynamic Liquid-Cooling Implementation at Rack Level", ASME 2022 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems, InterPACK, 2022

[52] Heydari, Ali, Pardeep Shahi, Vahideh Radmard, Bahareh Eslami, Uschas Chowdhury, Chandraprakash Hinge, Lochan Sai Reddy Chinthaparthy, Harold Miyamura, Himanshu Modi, Dereje Agonafer and Jeremy Rodriguez "A Control Strategy for minimizing temperature fluctuations in high power liquid to liquid CDUs operated at very low heat loads", ASME 2022 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems, InterPACK, 2022.

[53] Shahi, Pardeep, Apruv Pravin Deshmukh, Hardik Yashwant Hurnekar, Satyam Saini, Pratik Bansode, Rajesh Kasukurthy, and Dereje Agonafer. "Design, Development, and Characterization of a Flow Control Device for Dynamic Cooling of Liquid-Cooled Servers." Journal of Electronic Packaging 144, no. 4 (2022).

[54] Shahi, Pardeep, Satyam Saini, Pratik Bansode, Rajesh Kasukurthy, and Dereje Agonafer. "Experimental Study Demonstrating Pumping Power Savings at Rack Level Using Dynamic Cooling." Journal of Enhanced Heat Transfer 29, no. 6 (2022).

[55] Heydari, Ali, Pardeep Shahi, Vahideh Radmard, Bahareh Eslami, Uschas Chowdhury, Akiilessh Sivakmar, Akshay Lakshminarayana, Harold Miyamura, Gautam Gupta, Dereje Agonafer and Jeremy Rodriguez " Experimental Study of Hydraulic Characteristics for Liquid Cooled Data Center Deployment", ASME 2022 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems, InterPACK, 2022

[56] Heydari, Ali, Pardeep Shahi, Vahideh Radmard, Bahareh Eslami, Uschas Chowdhury, Satyam Saini, Pratik Bansode, Harold Miyamura, Dereje Agonafer and Jeremy Rodriguez "Liquid to Liquid Cooling for High Heat Density Liquid Cooled Data Centers", ASME 2022

International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems, InterPACK, 2022

[57] Shahi, Pardeep, Amith Mathew, Satyam Saini, Pratik Bansode, Rajesh Kasukurthy, and Dereje Agonafer. "Assessment of Reliability Enhancement in High-Power CPUs and GPUs Using Dynamic Direct-to-Chip Liquid Cooling." Journal of Enhanced Heat Transfer 29, no. 8 (2022).

[58] Shahi, Pardeep, Sarthak Agarwal, Satyam Saini, Amirreza Niazmand, Pratik Bansode, and Dereje Agonafer. "CFD analysis on liquid cooled cold plate using copper nanoparticles." In International Electronic Packaging Technical Conference and Exhibition, vol. 84041, p. V001T08A007. American Society of Mechanical Engineers, 2020.

[59] Kumar, Ashish, Pardeep Shahi, and Sandip K. Saha. "Experimental study of latent heat thermal energy storage system for medium temperature solar applications." MCM'18 (2018): 16-18.

[60] Chinthaparthy, Lochan Sai Reddy, Vibin Simon, Pardeep Shahi, Akiilessh Sivakumar, Satyam Saini, Himanshu Modi, Pratik Bansode and Dereje Agonafer, “CFD analysis of Heat capture ratio in a hybrid cooled server”, ASME 2022 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems, InterPACK, 2022.

[61] Mahajan, Ravi, Chia-pin Chiu, and Greg Chrysler. "Cooling a microprocessor chip." Proceedings of the IEEE 94, no. 8 (2006): 1476-1486.

[62] Sahini, Manasa, Chinmay Kshirsagar, Mathan Kumar, Dereje Agonafer, John Fernandes, Jacob Na, Veerendra Mulay, Patrick McGinn, and Michael Soares. "Rack-level study of hybrid cooled servers using warm water cooling for distributed vs. centralized pumping systems." In 2017 33rd Thermal Measurement, Modeling & Management Symposium (SEMI-THERM), pp. 155-162. IEEE, 2017.

[63] Tuckerman, David B., and Roger Fabian W. Pease. "High-performance heat sinking for VLSI." IEEE Electron device letters 2, no. 5 (1981): 126-129.

[64] Zhang, H. Y., D. Pinjala, T. N. Wong, K. C. Toh, and Y. K. Joshi. "Single-phase liquid cooled microchannel heat sink for electronic packages." Applied Thermal Engineering 25, no. 10 (2005): 1472-1487.

[65] Song, Seaho, Van Au, and Kevin P. Moran. "Constriction/spreading resistance model for electronics packaging." In Proceedings of the 4th ASME/JSME thermal engineering joint conference, vol. 4, pp. 199-206. 1995.

[66] Qu, Weilin, and Issam Mudawar. "Experimental and numerical study of pressure drop and heat transfer in a single-phase micro-channel heat sink." International journal of heat and mass transfer 45, no. 12 (2002): 2549-2565.

[67] Fernandes, John, Saeed Ghalambor, Dereje Agonafer, Vinod Kamath, and Roger Schmidt. "Multi-design variable optimization for a fixed pumping power of a water-cooled cold plate for high power electronics applications." In 13th InterSociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems, pp. 684-692. IEEE, 2012.

[68] Fernandes, John Edward. "Minimizing Power Consumption At Module, Server And Rack-levels Within A Data Center Through Design And Energy-efficient Operation Of Dynamic Cooling Solutions." (2015).

[69] Boucher, Timothy D., David M. Auslander, Cullen E. Bash, Clifford C. Federspiel, and Chandrakant D. Patel. "Viability of dynamic cooling control in a data center environment." (2006): 137-144.

[71] Hazelwood, Kim, Sarah Bird, David Brooks, Soumith Chintala, Utku Diril, Dmytro Dzhulgakov, Mohamed Fawzy et al. "Applied machine learning at facebook: A datacenter infrastructure perspective." In 2018 IEEE International Symposium on High Performance Computer Architecture (HPCA), pp. 620-629. IEEE, 2018.

[72] Arghode, Vaibhav K., Vikneshan Sundaralingam, and Yogendra Joshi. "Airflow management in a contained cold aisle using active fan tiles for energy efficient data-center operation." Heat Transfer Engineering 37, no. 3-4 (2016): 246-256. [73] S. Khalili, G. Mohsenian, A. Desu, K. Ghose and B. Sammakia, "Airflow Management Using Active Air Dampers in Presence of a Dynamic Workload in Data Centers," 35th Semiconductor Thermal Measurement, Modeling and Management Symposium (SEMI-THERM), San Jose, CA, USA, 2019, pp. 101-110..

[74] Xu, Hong, Chen Feng, and Baochun Li. "Temperature aware workload management in geo-distributed datacenters." In 10th International Conference on Autonomic Computing (ICAC 13), pp. 303-314. 2013.

[75] Kasukurthy, Rajesh, Pencala S. Challa, Rishi Ruben Palanikumar, Barath Ragul Manimaran, and Dereje Agonafer. "Flow Analysis and Linearization of Rectangular Butterfly Valve Flow Control Device for Liquid Cooling." In 2018 17th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), pp. 683-687. IEEE, 2018.

[76] Kasukurthy, Rajesh. "Design and Optimization of Energy Conserving Solutions in Data Center Application." PhD, The University of Texas at Arlington (2019).

[77] Kasukurthy, Rajesh, Amruthavalli Rachakonda, and Dereje Agonafer. "Design and optimization of control strategy to reduce pumping power in dynamic liquid cooling." Journal of Electronic Packaging 143, no. 3 (2021).

[78] Shahi, Pardeep,Hardik Hurnekar,Apurv Deshmukh, Satyam Saini, Pratik Bansode, Rajesh Kasukurthy and Dereje Agonafer, “Assessment of Pump Power Savings at Rack Level for Dynamic Direct-to-Chip Liquid Cooling Using a Novel Flow Control Device”, Journal of Enhanced Heat Transfer, 2022.

[79] Shahi, Pardeep, Apurv Deshmukh, Hardik Hurnekar, Satyam Saini, Pratik Bansode and Dereje Agonafer, “Numerical Investigation on Effect of Target Coolant Delivery in Liquid-Cooled Microchannel Heat Sinks”, Journal of Enhanced Heat Transfer, 2022.