

University of Texas at Arlington

MavMatrix

Mechanical and Aerospace Engineering
Dissertations

Mechanical and Aerospace Engineering
Department

2023

EVALUATION AND OPTIMIZATION OF THERMAL SOLUTIONS IN AIR AND LIQUID COOLING SYSTEMS FOR DATACENTERS

Vibin Shalom Simon

Follow this and additional works at: https://mavmatrix.uta.edu/mechaerospace_dissertations



Part of the [Aerospace Engineering Commons](#), and the [Mechanical Engineering Commons](#)

Recommended Citation

Simon, Vibin Shalom, "EVALUATION AND OPTIMIZATION OF THERMAL SOLUTIONS IN AIR AND LIQUID COOLING SYSTEMS FOR DATACENTERS" (2023). *Mechanical and Aerospace Engineering Dissertations*. 391.

https://mavmatrix.uta.edu/mechaerospace_dissertations/391

This Dissertation is brought to you for free and open access by the Mechanical and Aerospace Engineering Department at MavMatrix. It has been accepted for inclusion in Mechanical and Aerospace Engineering Dissertations by an authorized administrator of MavMatrix. For more information, please contact leah.mccurdy@uta.edu, erica.rousseau@uta.edu, vanessa.garrett@uta.edu.

**EVALUATION AND OPTIMIZATION OF THERMAL
SOLUTIONS IN AIR AND LIQUID COOLING SYSTEMS FOR
DATACENTERS**

by

Vibin Shalom Simon

DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

At The University of Texas at Arlington

August 2023

Supervising Committee:

Dr. Dereje Agonafer

Dr. Abdolhossein Haji-Sheikh

Dr. Miguel A. Amaya

Dr. Amir Ameri

Dr. Saket Karajgikar



Copyright © by Vibin Shalom Simon

2023 All Rights Reserved

Acknowledgments

I wish to convey my heartfelt gratitude to my beloved family, with a special mention to my father, Mr. Francis Simon, and my mother, Mrs. Srividhya Hariharan for their unwavering support, love, and affection through every stage of my doctoral journey.

I extend my sincere and heartfelt gratitude to Professor Dr. Dereje Agonafer for his support and encouragement which has been a cornerstone throughout the course of my PhD journey. I am forever indebted to his generosity and compassion towards me. I would also like to thank him for providing the privilege of working in industry-funded projects, as well as enabling my participation in diverse conferences that have significantly expanded my professional network and horizons.

I am deeply appreciative of my committee members Dr. Abdolhossein Haji Sheikh, Dr. Miguel Amaya, Dr. Amir Ameri, and Dr. Saket Karajgikar for their meticulous scrutiny, constructive feedback, and constant encouragement. I've been lucky to connect with amazing individuals at UTA, including EMNSPC lab and alumni. Also, I would like to thank all the industry mentors in the Industry Advisory Board (IAB) consortium, Mestex, Meta, Future Facilities (Cadence), Intel, and NVIDIA for giving me the opportunity to work with them over the years of my PhD program.

I'm truly grateful to my research partners and friends, Dr. Ashwin Siddarth, Dr. Pardeep Shahi, Dr. Satyam Saini, and Pratik Bansode, for their invaluable contributions that have been pivotal to the progression and accomplishment of my research endeavors.

Most of all, I am grateful to have an esteemed circle of friends, Ivneet, Anto, Aditya, Akiilessh, Aswin, Abdul, Sasi, Lokesh, Keerthi, Nallamani, Aravind, Theliban, Walid, and Shanmugam. Their positive impact on my life during this period is immeasurable, and I am truly fortunate to have such wonderful friends by my side.

August 13, 2023

Abstract

EVALUATION AND OPTIMIZATION OF THERMAL SOLUTIONS IN AIR AND LIQUID COOLING SYSTEMS FOR DATACENTERS

Vibin Shalom Simon, Ph.D.

The University of Texas at Arlington, 2023

Supervising Professor: Dr. Dereje Agonafer

A data center is a specifically constructed establishment designed to accommodate computing, storage, and transmission equipment, thereby facilitating essential business operations spanning diverse sectors of the economy. This facility serves as the designated physical infrastructure for housing servers, networking components, and data storage systems.

In the context of traditional data centers, a substantial portion of energy, specifically around 45% to 55%, is directed towards supporting information technology (IT) operations. Additionally, a considerable allocation, varying from 30% to 40%, is attributed to the facilitation of cooling processes.

The scope of this dissertation study comprises nine distinct chapters, each contributing to addressing the challenges to establish energy-efficient and optimized data centers. The research commences by focusing on energy-efficient control strategies for hyperscale data centers through scaled-down Computational Fluid Dynamics (CFD) modeling. This approach involves validating the model against real data center datasets and conducting simulations tailored to specific needs. Leveraging machine learning techniques, this investigation aims to anticipate control setpoints for optimizing airflow and chilled water control at the data center room level. As the research progresses, the study addresses the limitations of CRAH units in sustaining high-density rack setups. Consequently, the feasibility and energy implications of implementing Rear Door Heat Exchangers (RDHx) are explored. This includes modeling a

single rack with RDHx, assessing both passive and active modes. Moving forward, the study expands to assess the energy efficiency of RDHx across data center deployment.

Furthermore, in response to the challenges posed by hybrid cooled data centers, an investigation delves into the heat capture ratio for servers equipped with cold plates. The study quantifies heat capture by different mediums, considering varying temperatures and flow rates. It also delves into the impact of immersing hybrid cooled servers under diverse ambient conditions.

The following research also evaluates strategies to optimize heat sinks for immersion cooling environments, considering a wide range of fluids and design parameters. This includes simulation-driven optimization of cold plates' configurations, enabling improved performance for processors characterized by high Thermal Design Power (TDP). Moreover, the study delves into the enhancement of air-cooled heat sinks for single-phase immersion cooling setups, considering fluid thermo-physical properties for optimal design.

Through these interconnected chapters, the dissertation endeavors to provide a comprehensive understanding of energy-efficient and optimized data centers, encompassing a spectrum of strategies and technologies aimed at addressing the challenges associated with increasing power densities in modern chip architectures.

Table of Contents

CHAPTER 1 INTRODUCTION	1
1.1 INTRODUCTION.....	1
1.2 SCOPE OF WORK	5
REFERENCES	7
CHAPTER 2 LITERATURE REVIEW	13
REFERENCES	15
CHAPTER 3 ENERGY AWARE PRO-ACTIVE CONTROL FOR AIR COOLED DATACENTERS USING MACHINE LEARNING	20
3.1 INTRODUCTION.....	20
3.2 COMPUTATIONAL MODEL	25
3.3 ZONE OF INFLUENCE OF CRAH UNITS	28
3.4 BOUNDARY CONDITIONS AND CASES	28
3.5 ARTIFICIAL NEURAL NETWORK	29
3.6 CONCLUSION	30
3.7 REFERENCES	31
CHAPTER 4 FEASIBILITY AND ENERGY ANALYSIS OF REAR DOOR HEAT EXCHANGER FOR DATA CENTERS	33
4.1 ABSTRACT	33
4.2 INTRODUCTION.....	33
4.3 COMPUTATIONAL MODEL	36
4.4 RESULTS AND OBSERVATION	41
4.4.1 <i>Single Rack Model</i>	41
4.4.2 <i>Data Hall Analysis</i>	44
4.5 CONCLUSION AND FUTURE WORK.....	48
4.6 REFERENCES	49

CHAPTER 5 ENERGY ANALYSIS OF REAR DOOR HEAT EXCHANGERS IN DATA CENTERS WITH SPATIAL	
WORKLOAD DISTRIBUTION.....	51
5.1 ABSTRACT	51
5.2 INTRODUCTION.....	52
5.3 COMPUTATIONAL MODEL	55
5.4 REAR DOOR HEAT EXCHANGER.....	57
5.5 UNIFORM WORKLOAD DISTRIBUTION	60
5.6 SPATIAL WORKLOAD DISTRIBUTION	61
5.7 CONCLUSION	65
5.8 REFERENCES	65
CHAPTER 6 EXPERIMENTAL AND CFD ANALYSIS OF HEAT CAPTURE RATIO IN A HYBRID COOLED SERVER	68
6.1 ABSTRACT	68
6.2 INTRODUCTION.....	69
6.3 EXPERIMENTAL SETUP AND METHODOLOGY.....	70
6.4 RESULTS AND DISCUSSION	77
6.5 CONCLUSION	79
6.6 REFERENCES	80
CHAPTER 7 COMPACT MODELING AND THERMAL ANALYSIS OF IMMERSION BASED HYBRID COOLED SERVER	
.....	82
7.1 ABSTRACT	82
7.2 INTRODUCTION.....	83
7.3 SERVER DESCRIPTION	85
7.4 CFD MODEL.....	87
7.5 TEST CASES.....	90
7.5.1 <i>Immersion Hybrid Cooling</i>	91
7.5.2 <i>Boundary Conditions</i>	91
7.5.3 <i>Flow Pattern</i>	91

7.6 RESULTS AND OBSERVATIONS.....	92
7.7 COOLING CAPABILITY PREDICTIONS	94
7.8 CONCLUSION	96
7.9 REFERENCES	97
CHAPTER 8 PARAMETRIC MULTI-OBJECTIVE OPTIMIZATION OF COLD PLATE FOR SINGLE-PHASE IMMERSION	
COOLING.....	101
8.1 ABSTRACT	101
8.2 INTRODUCTION.....	102
8.3 COMPUTATIONAL MODEL AND EXPERIMENTAL VALIDATION	104
8.4 RESULTS AND OBSERVATION	112
<i>8.4.1 Thermal performance</i>	<i>112</i>
<i>8.4.2 Flow performance</i>	<i>115</i>
8.5 CONCLUSION	117
8.6 REFERENCES	117
CHAPTER 9 FLUID BASED OPTIMIZATION SCHEME FOR HEAT SINKS IN SINGLE-PHASE IMMERSION COOLING	
UNDER NATURAL CONVECTION	121
9.1 INTRODUCTION.....	121
9.2 COMPUTATIONAL MODEL	127
9.3 METHODOLOGY.....	130
9.4 RESULTS AND DISCUSSION.....	131
9.5 CONCLUSION	133
REFERENCES.....	133
BIOGRAPHICAL INFORMATION	138

List of Illustrations

Figure 1-1 HYPERSCALE DATA CENTER (META)	1
Figure 1-2 RISE IN CHIP POWER WITH YEARS	2
Figure 1-3 ENERGY CONSUMPTION IMPACT OF MECHANICAL EQUIPMENT AND SYSTEMS [3]	3
Figure 3-1 DISTRIBUTION OF ENERGY CONSUMPTION IN A TYPICAL DATA CENTER	21
Figure 3-2 SEQUENCE OF OPERATION	24
Figure 3-3 DATA CENTER CFD MODEL (TOP VIEW) IN 6SIGMA ROOM	26
Figure 3-4 1D FLOW-NETWORK FOR CHILLED WATER LOOP	26
Figure 3-5 INFLUENCE OF CRAH UNIT (a) 1 & 8 (b) 3 & 6 (c) 2 & 7	28
Figure 3-6 Feature Selection	29
Figure 3-7 ANN MODEL WITH INPUT AND OUTPUT PARAMETERS	30
Figure 3-8 ANN ERROR	30
Figure 4-1 EVOLUTION OF THE ASHRAE POWER TRENDS [4]	34
Figure 4-2 LIQUID COOLING LOOP [9]	36
Figure 4-3 FRONT AND BACK OF A TYPICAL OCP RACK	37
Figure 4-4 P-Q CURVE OF A CR 9256 FAN	38
Figure 4-5 ORV2 CABINET WITH RDHX - CFD MODEL	39
Figure 4-6 RESULT PLANE POSITION	42
Figure 4-7 PRESSURE CONTOUR PLOT FOR CASE 1 (LEFT) AND CASE 2 (RIGHT)	42
Figure 4-8 PRESSURE CONTOUR PLOT - CASE 3 (LEFT) AND CASE 4 (RIGHT)	42
Figure 4-9 TYPICAL DATACENTER WITH RDHX INSTALLATION	46

Figure 4-10 CFD RESULTS - TEMPERATURES PLANES.....	46
Figure 4-11 A SECTIONAL PLANE SHOWING THE VELOCITY PROFILE AND AIR MOVEMENT AROUND THE CABINET WITH RDHx	48
Figure 5-1 Air cooling limits: rack power density [3]	53
Figure 5-2 Cost of Equipment vs rack power density [13].....	55
Figure 5-3 Raised floor data center with CRAH units.....	56
Figure 5-4 Data center with RDHx.....	57
Figure 5-5 Traditional rack vs RDHx Rack Schematic (Side View).....	58
Figure 5-6 Temperature Contours for uniform workload distribution.....	61
Figure 5-7 Randomly selected racks to mimic spatial workload distribution in data center.....	62
Figure 5-8 Temperature contours for spatial workload distribution.....	63
Figure 5-9 Velocity contours for spatial workload distribution.....	64
Figure 6-1 Cisco M220 series M3 server.....	70
Figure 6-2 CFD model of Cisco M220 series M3 server.....	72
Figure 6-3 Thermal Resistance (Rth) Curve	74
Figure 6-4 Flow resistance measurement setup	74
Figure 6-5 Flow resistance (PQ) curve	75
Figure 6-6 1-D Flow network model	75
Figure 6-7 Heat capture ratio trend lines for varying flowrates.....	78
Figure 7-1 Server retrofitted with cold plates	86
Figure 7-2 CFD model of a 1U server	87
Figure 7-3 Flow network model for cold plates.....	88
Figure 7-4 Mesh Sensitivity Analysis.....	88

Figure 7-5 Thermal resistance for the cold plate at various coolant inlet temperature and flowrate	89
Figure 7-6 Pressure drop for the cold plate.....	90
Figure 7-7 (a) Air hybrid server (b) Immersion hybrid server.....	92
Figure 7-8 PCH temperature vs Air and Immersion server inlet temperatures	93
Figure 7-9 DIMM temperature vs Air and Immersion server inlet temperatures.....	93
Figure 7-10 CPU temperature vs Air and Immersion server inlet temperatures	94
Figure 7-11 DIMM cooling capability.....	95
Figure 7-12 CPU cooling capability	95
Figure 8-1 MODEL OF COLD PLATE & STACK-UP OF THE COLD PLATE ON THE HEATER.....	104
Figure 8-2 Experimental Setup	109
Figure 8-3 EXPERIMENTAL SETUP SCHEMATIC.....	109
Figure 8-4 Heater temperature - EC100 vs PG25	110
Figure 8-5 MODEL VALIDATION PLOT - HEATER TEMPERATURE VS. FLOW RATE.....	111
Figure 8-6: Temperature contour of cold plate fins	112
Figure 8-7 Temperature contour at X-Plane	112
Figure 8-8 THERMAL RESISTANCE VS. FIN THICKNESS FOR DIFFERENT FIN SPACING (FS) AT A FIXED FIN HEIGHT OF 3 MM.....	113
Figure 8-9 THERMAL RESISTANCE VS. FIN THICKNESS FOR DIFFERENT FIN SPACING (FS) AT A FIXED FIN HEIGHT OF 4 MM.....	113
Figure 8-10 THERMAL RESISTANCE VS. FIN THICKNESS FOR DIFFERENT FIN SPACING (FS) AT A FIXED FIN HEIGHT OF 5 MM.....	114

Figure 8-11 THERMAL RESISTANCE VS. FIN HEIGHT FOR DIFFERENT FIN THICKNESSES AT A FIXED FIN SPACING OF 1 MM	114
Figure 8-12 PRESSURE DROP ACROSS THE COLD PLATE AS A FUNCTION OF FIN THICKNESS AND FIN SPACING	115
Figure 8-13 COP MATRIX OF RESULTS	116
Figure 9-1 Viscosity Vs Fluid Temperature	125
Figure 9-2 Heat sink design variables [25]	126
Figure 9-3 CFD model in Ansys Icepak	127
Figure 9-4 Model Validation	128
Figure 9-5 Thermal resistance for selected immersion compatible fluids	129
Figure 9-6 Thermal resistance for selected immersion compatible fluids	132

List of Tables

Table 3-1 DATA CENTER CONFIGURATION.....	25
Table 3-2 CRAH UNIT SPECIFICATION.....	25
Table 3-3 Boundary Conditions.....	29
Table 4-1 DETAILS OF IT EQUIPMENTS IN THE RACK.....	37
Table 4-2 HX SPECIFICATION.....	40
Table 4-3 RDHx CASES.....	41
Table 4-4 PRESSURE AND TEMPERATURE RESULTS.....	43
Table 4-5 PRESSURE AND TEMPERATURE RESULTS.....	44
Table 5-1 CRAH unit specification and BC.....	59
Table 5-2 CRAH unit specification and BC.....	59
Table 5-3 ITE Specifications.....	60
Table 5-4 CRAH and RDHx-based DC chilled water results.....	64
Table 6-1 Server specification.....	71
Table 6-2 CFD model validation using experimental result.....	72
Table 6-3 Boundary Conditions.....	76
Table 6-4 Heat Capture Ratio (HCR).....	77
Table 6-5 Experimental vs CFD results.....	79
Table 7-1 Critical heat generating components.....	87
Table 7-2 Boundary Conditions for all Test Cases.....	91
Table 8-1 ELECTROCOOL (EC) 100 THERMO-MECHANICAL PROPERTIES	105
Table 8-2 BOUNDARY CONDITIONS.....	106
Table 8-3 List of Equipments.....	107
Table 8-4 VARIABLE INPUT PARAMETERS.....	111

Table 8-5 FIN PARAMETERS OF BASELINE COLD PLATE VS OPTIMIZED COLD PLATE	116
Table 8-6 RTH FOR SAME BOUNDARY CONDITIONS	117
Table 9-1 Thermal resistance and viscosity variation for selected immersion compatible fluids	129
Table 9-2 Fin spacing for different designs of Heat Sink	130

Chapter 1 Introduction

1.1 Introduction

Data center is a purpose-built facility that houses computing, storage and transmission equipments to serve critical business operation across all sectors of economy. Physical facility for servers, networking, and data storage. Key features: Redundant power, backup systems, cooling, and fire suppression. Scalability: Ability to grow computing resources as needed. Importance: Critical for digital infrastructure, supporting data-driven operations and the modern economy.



Figure 1-1 HYPERSCALE DATA CENTER (META)

The substantial surge in data center demand has garnered the interest of a diverse range of investors, encompassing growth capital, buyout, real estate, and notably, infrastructure investors. Within the United States market exclusively, the level of demand, assessed by power consumption to gauge the hosting capacity for servers within a data center, is projected to attain

35 gigawatts (GW) by the year 2030. This projection represents a substantial increase from the 17 GW recorded in 2022. [1] Increase in chip power with years is shown in Figure 1-2 below.



Figure 1-2 RISE IN CHIP POWER WITH YEARS

The equipment within data centers, often comprising numerous servers, necessitates efficient cooling for optimal functionality. In fact, the operational capacity of a data center hinges upon its ability to effectively cool the servers. The closer these servers can be arranged, the higher the productivity per unit of space. Efficient cooling holds paramount importance as a determinant of a data center's profitability, given that cooling constitutes approximately 40 percent of the overall energy consumption of such facilities. The consequences of system downtime due to overheating can be financially significant.

Notably, cooling technology has undergone substantial enhancements over the preceding decade. Many large-scale data centers have transitioned from traditional air-conditioning systems that regulate the temperature of entire rooms to in-row or rotodynamic heater-based cooling configurations. In these designs, fans facilitate the removal of heat emitted by servers, subsequently cooled using water or refrigerants. Despite these

advancements, further progress is imperative, as contemporary sophisticated cooling systems often encounter challenges in managing temperatures exacerbated by global warming.

Within conventional data centers, a significant proportion of energy, approximately 45% to 55%, is allocated to information technology (IT) operations, while a notable share, ranging from 30% to 40%, is dedicated to cooling processes. When considering energy-related expenditures, the combined expenses associated with cooling mechanisms and electricity infrastructure constitute a substantial 70% to 80% of the overall capital costs. [2]

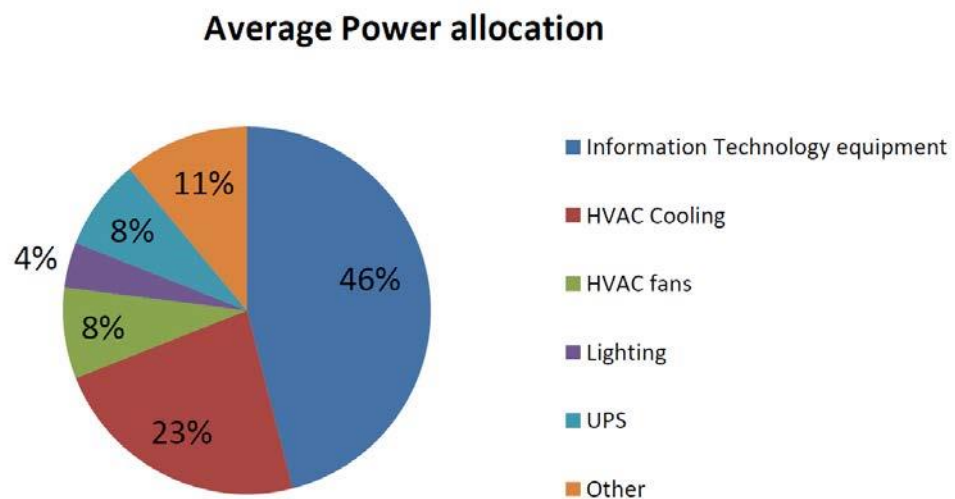


Figure 1-3 ENERGY CONSUMPTION IMPACT OF MECHANICAL EQUIPMENT AND SYSTEMS
[3]

Various approaches are currently under investigation to enhance the effectiveness of cooling systems within data centers. Among these methods, liquid cooling stands out as an exceptionally efficient solution to address the escalating requirements for cooling. This is primarily attributed to the enhanced thermal conductivity and thermal mass exhibited by coolants. While air cooling, a conventional and widely used method for cooling electronic components, persists as a prevailing approach, numerous researchers have directed their efforts toward refining the existing trends in air cooling.

Studies aimed at enhancing the efficiency and viability of rear door heat exchangers have been conducted [4]. Endeavors have been undertaken to minimize on-site visits by fine-tuning Computational Fluid Dynamics (CFD) models for data centers [5]. Moreover, the integration of Artificial Intelligence has been employed to anticipate and establish control strategies for multiple air conditioning units within data centers [6].

Refinements have been implemented in the domain of liquid cooling through the deployment of control strategies aimed at enhancing the cooling efficiency of high heat density servers [7-9]. Concurrently, alongside the advancements in liquid cooling, the development of flow control devices has demonstrated notable reductions in pumping power requirements [10], while the introduction of liquid-to-liquid heat exchangers and the analysis of hydraulic characteristics within liquid-cooled data centers have exhibited substantial improvements in overall cooling efficacy [11-12]. The assessment of reliability pertaining to liquid-cooled cold plates and their application at the rack level has substantiated liquid cooling as a more dependable and viable approach for cooling high heat dissipating components [13-15]. Additionally, the calculated heat capture ratio [16] for both liquid and air-cooled servers surpass that of conventional air-cooled servers, underlining the advantages of liquid cooling.

Immersion cooling has been one of the emerging technologies 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[17], custom-build[18], open compute servers[19] and CFD analysis on the optimization of server to improvise the cooling in immersion[20], numerical analysis are performed on the use of different fluids such as mineral oil and Al₂O₃ nanofluid[21] and CFD analysis over two phase immersion cooling fluids such as FC-72 [22]. Also, with its 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[23-25]. To further enhance the cooling optimization are performed at heat sink for forced and natural convection in an immersion cooling.[26-28].

1.2 Scope of Work

The scope of the dissertation study is divided into 9 chapters. The initial chapter encompasses the introduction and the underlying motivation. Subsequently, the second chapter delves into an in-depth exploration of the existing body of literature. Chapter 3-9 will address the challenges mentioned in the abstract for energy efficient and optimized data centers.

Initially, the research delved into energy-efficient control strategies for a hyperscale data center, employing scaled-down Computational Fluid Dynamics (CFD) modeling. The model's validity was established against a realistic data center dataset, with subsequent simulations tailored to specific requirements. At the data center room level, investigations centered on airflow and chilled water control strategies, complemented by the application of machine learning techniques to anticipate control setpoints. The objective was to achieve optimal functionality for Computer Room Air Handler (CRAH) and chiller units.

As rack density surpasses a certain threshold, CRAH units prove inadequate for supporting workloads, prompting an exploration into the feasibility and energy implications of Rear Door Heat Exchangers (RDHx). This phase involved modeling a single rack equipped with an RDHx, testing both passive and active RDHx modes. Extending this analysis to a broader scope, an energy assessment encompassing data center-level deployment was conducted. A comparison of energy consumption and chiller efficiencies was undertaken, contrasting RDHx-based and CRAH-based data center setups.

In addressing the context of hybrid cooled data centers, an investigation into the heat capture ratio for servers featuring cold plates was performed. This inquiry was pivotal in tackling challenges stemming from varying inlet temperatures of air and coolant, and their

consequential impact on energy consumption. The overarching aim of this study was to quantify the heat captured by liquid as opposed to air, subsequently analyzing the Heat Capture Ratio's (HCR) potential influence resulting from heat exchange between different mediums across varying temperatures and flow rates.

Given the constraints of hybrid cooling, which primarily focuses on high-power consuming components, the study further examined the implications of immersing a hybrid cooled server under diverse ambient conditions. Furthermore, the significance of targeted cooling strategies was underscored. Subsequent analysis entailed the optimization of heat sinks for fluids used in immersion environments, with a forward-looking perspective towards processors characterized by high Thermal Design Power (TDP). Utilizing Ansys Icepack and OptiSLang, CFD simulations of cold plates were conducted to ascertain the optimal configuration for a given fluid and temperature.

Current interests in single-phase immersion cooling have spurred efforts towards enhancing the performance of air-cooled heat sinks within immersion environments. Given the diversity of fluids available and the need to select based on multiple factors, the study entails designing optimized heat sinks for maximum heat transfer efficacy. This study is also anticipated to yield a fluid-based heat sink design. Such an outcome involves the analysis of various thermo-physical properties of fluids, benchmarked against a reference fluid across a spectrum of heat sink designs. Ultimately, this endeavor will empower users to select an appropriate heat sink design based on chosen fluid, mode of heat transfer (natural/forced convection), and the fluid bath's temperature within the Single-Phase Immersion Cooling (SPIC) tank.

References

- [1] <https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/investing-in-the-rising-data-center-economy>
- [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] 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.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] 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
- [9] Heydari, Ali, Pardeep Shahi, Vahideh Radmard, Bahareh Eslami, Uschas Chowdhury, Chandraprakash Hinge, Lochan Sai Reddy Chinthaparthi, 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.
- [10] 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).
- [11] 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).
- [12] Heydari, Ali, Pardeep Shahi, Vahideh Radmard, Bahareh Eslami, Uschas Chowdhury, Akiilessh Sivakumar, 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

- [13] 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
- [14] 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).
- [15] 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.
- [16] Shalom Simon, V, Reddy, LS, Shahi, P, Valli, A, Saini, S, Modi, H, Bansode, P, & Agonafer, D. "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*. Garden Grove, California, USA. October 25–27, 2022. V001T01A013. ASME.
- [17] 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.
- [18] 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.
- [19] 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.
- [20] 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.
- [21] 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.
- [22] 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
- [23] 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
- [24] 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.
- [25] 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.
- [26] 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).
- [27] 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*.
- [28] 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.

Chapter 2 Literature Review

The dynamic nature of workload distribution and arrival within data centers has a direct impact on the airflow dynamics in the hot and cold aisles. Traditional control strategies employed within data center cooling systems make use of an ON/OFF mechanism, employing PID control based on sensor feedback [1]. This mechanism is activated when the temperature setpoint deviates within a predetermined deadband. The existence of this deadband range results in continuous temperature oscillations, which, if beyond the recommended temperature threshold, can contribute to an elevated rate of server failures [2]. As a consequence, relying on reactive control strategies to optimize airflow to Information Technology Equipment (ITE) amidst such fluctuating ambient conditions might not be the most optimal approach.

A significant impediment in the advancement of such methodologies is the scarcity of historical data. This challenge is exacerbated when considering hyperscale cloud data centers, where conducting experimental investigations can entail substantial costs, time investment, and infeasibility. In scenarios like these, Computational Fluid Dynamics (CFD) can serve as a valuable tool to comprehend diverse thermodynamic parameters through data-driven models that have been trained using validated CFD models of actual data centers.

Rear Door Heat Exchangers (RDHx) incorporated into various data center configurations confer several advantages, including: 1) reducing the requirement for a high number of Computer Room Air Handler (CRAH) units, 2) mitigating hotspots by expelling heat in closer proximity to its origin, and 3) facilitating the usage of higher chilled water temperatures while simultaneously adhering to recommended IT inlet temperature ranges. RDHx units were effectively integrated with room-level air conditioning, adopting a hybrid approach [3].

Liquid-based cooling technologies exhibit superior heat transfer coefficients and are progressively employed for dissipating the elevated power densities characteristic of next-generation Central Processing Units (CPUs) and Graphics Processing Units (GPUs) [4]. Direct on-chip liquid cooling is typically employed on CPUs, GPUs, and occasionally on high-power-consuming DIMMs. Analytical models were formulated to explore split-flow microchannel liquid-cooled cold plates with impingement flow to enhance heat dissipation [5][6]. Another study focused on determining the thermal performance thresholds of single-phase liquid cooling, employing an improved effectiveness cold plate model [7]. Optimization of heat sink design was pursued through the incorporation of guided vanes targeting hotspots in liquid cooling systems [8][9]. A comprehensive Computational Fluid Dynamics (CFD) analysis scrutinized the influence of various parameters on the heat capture ratio of liquid cooling in hybrid cooled servers [10]. Experimental investigations assessed the impact of design modifications on the chassis and ducting of a server [11][12]. Transient studies were conducted to refine cooling performance in rack-level direct-to-chip liquid cooling through the application of control strategies [13-15]. These research endeavors collectively aimed to enhance cooling efficiency in hybrid cooled servers, encompassing both air and liquid heat transfer mediums.

For instance, Chen and Chen employed an innovative, multi-objective direction-based algorithm to optimize plate-fin heat sinks alongside impingement fans, utilizing a commercially available Multiphysics tool [16]. The optimization outcome highlighted heightened heat transfer performance with concurrent weight reduction in the optimized parallel plate heat sink configuration. Additionally, methodologies grounded in fuzzy logic were employed to quantitatively evaluate the influence of heat sink design parameters on thermal efficiency [17]. Experimental inquiries targeted pin-fin heat sink design parameters, incorporating fin spacing, pin-fin diameter, and height. Subsequently, Analysis of Variance (ANOVA) was employed to explore the impact of these parameters on pivotal heat sink

attributes encompassing thermal resistance, pressure drop, and average heat transfer coefficient.

In a distinct vein, Chaing and Chang harnessed the Response Surface Method (RSM) to attain optimal design parameters for pin-fin heat sinks, thereby achieving heightened thermal performance [18]. The endeavor to minimize entropy generation rate as a designated objective function was undertaken by Chen and colleagues [19]. In this pursuit, they optimized a plate-fin heat sink for CPU application by deploying a coded genetic algorithm, thereby identifying optimal design parameters for the heat sink. Addressing diverse objective functions, Devi et al. employed a Taguchi-based non-gradient method to minimize three distinct objective functions—namely, radiation emission, thermal resistance, and heat sink mass [20].

References

- [1] Theory and applications of HVAC control systems – a review of model predictive control (MPC)
- [2] Data center power equipment thermal guidelines and best practices ASHRAE TC9.9 (2016)
- [3] R. Schmidt and M. Iyengar, “Server rack rear door heat exchanger and the new ASHRAE recommended environmental guidelines,” in Proc. ASME InterPACK Conf. Collocated ASME Summer Heat Transf. Conf. ASME 3rd Int. Conf. Energy Sustainability, 2009, pp. 851–862.
- [4] P. Shahi, S. Saini, P. Bansode and D. Agonafer, "A Comparative Study of Energy Savings in a Liquid-Cooled Server by Dynamic Control of Coolant Flow Rate at Server Level," in IEEE Transactions on Components, Packaging and Manufacturing

- Technology, vol. 11, no. 4, pp. 616-624, April 2021, doi: 10.1109/TCPMT.2021.3067045.
- [5] Kisitu, Deogratius, and Alfonso Ortega. "Thermal-Hydraulic Analytical Models of Split-Flow Microchannel Liquid-Cooled Cold Plates With Flow Impingement." In International Electronic Packaging Technical Conference and Exhibition, vol. 85505, p. V001T02A015. American Society of Mechanical Engineers, 2021.
- [6] Kisitu, Deogratius. "Approximate Compact Thermal-Hydraulic Analytical Models for Laminar Microchannel Impingement Liquid-Cooled Cold Plates for Data Center Thermal Management." PhD diss., Villanova University, 2021.
- [7] Ortega, Alfonso, Carol Caceres, Umut Uras, Deogratius Kisitu, Uschas Chowdhury, Vahideh Radmard, and Ali Heydari. "Determination of the Thermal Performance Limits for Single Phase Liquid Cooling Using an Improved Effectiveness-NTU Cold Plate Model." In International Electronic Packaging Technical Conference and Exhibition, vol. 86557, p. V001T01A008. American Society of Mechanical Engineers, 2022.
- [8] Gharaibeh, Ahmad R., Mohammad I. Tradat, Srikanth Rangarajan, Bahgat G. Sammakia, and Husam A. Alissa. "Multi-objective optimization of 3D printed liquid cooled heat sink with guide vanes for targeting hotspots in high heat flux electronics." *International Journal of Heat and Mass Transfer* 184 (2022): 122287.
- [9] Gharaibeh, Ahmad R., Yaman M. Manaserh, Mohammad I. Tradat, Firas W. AlShatnawi, Scott N. Schiffres, and Bahgat G. Sammakia. "Using a Multi-Inlet/Outlet Manifold to Improve Heat Transfer and Flow Distribution of a Pin Fin Heat Sink." *Journal of Electronic Packaging* 144, no. 3 (2022): 031017.
- [10] Shalom Simon, Vibin, Lochan Sai Reddy, Pardeep Shahi, Amrutha Valli, Satyam Saini, Himanshu Modi, Pratik Bansode, and Dereje Agonafer. "CFD Analysis of Heat Capture

- Ratio in a Hybrid Cooled Server." In International Electronic Packaging Technical Conference and Exhibition, vol. 86557, p. V001T01A013. American Society of Mechanical Engineers, 2022.
- [11] Modi, Himanshu, Uschas Chowdhury, and Dereje Agonafer. "Impact of Improved Ducting and Chassis Re-design for Air-Cooled Servers in a Data Center." In 2022 21st IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (iTherm), pp. 1-8. IEEE, 2022.
- [12] Modi, Himanshu, Pardeep Shahi, Lochan Sai Reddy Chinthaparthi, Gautam Gupta, Pratik Bansode, Vibin Shalom Simon, and Dereje Agonafer. "Experimental Investigation of the Impact of Improved Ducting and Chassis Re-Design of a Hybrid-Cooled Server." In International Electronic Packaging Technical Conference and Exhibition, vol. 86557, p. V001T01A019. American Society of Mechanical Engineers, 2022.
- [13] Modi, Himanshu, Pardeep Shahi, Akiilessh Sivakumar, Satyam Saini, Pratik Bansode, Vibin Shalom, Amrutha Valli Rachakonda, Gautam Gupta, and Dereje Agonafer. "Transient CFD Analysis of Dynamic Liquid-Cooling Implementation at Rack Level." In International Electronic Packaging Technical Conference and Exhibition, vol. 86557, p. V001T01A012. American Society of Mechanical Engineers, 2022.
- [14] Heydari, Ali, Pardeep Shahi, Vahideh Radmard, Bahareh Eslami, Uschas Chowdhury, Chandraprakash Hinge, Lochan Sai Reddy Cinthaparthi et al. "A Control Strategy for Minimizing Temperature Fluctuations in High Power Liquid to Liquid CDUs Operated at Very Low Heat Loads." In International Electronic Packaging Technical Conference and Exhibition, vol. 86557, p. V001T01A011. American Society of Mechanical Engineers, 2022.

- [15] Shalom Simon, V, Modi, H, Sivaraju, KB, Bansode, P, Saini, S, Shahi, P, Karajgikar, S, Mulay, V, & Agonafer, D. "Feasibility Study of Rear Door Heat Exchanger for a High Capacity Data Center." Proceedings of the ASME 2022 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems. ASME 2022 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems. Garden Grove, California, USA. October 25–27, 2022. V001T01A018. ASME. <https://doi.org/10.1115/IPACK2022-97494>
- [16] Chen, C.T. and Chen, H.I., 2013. Multi-objective optimization design of plate-fin heat sinks using a direction-based genetic algorithm. Journal of the Taiwan Institute of chemical Engineers, 44(2), pp.257-265, <https://doi.org/10.1016/j.jtice.2012.11.012>, 2013.
- [17] Chiang, K.T., Chang, F.P. and Tsai, T.C., 2006. Optimum design parameters of Pin-Fin heat sink using the grey-fuzzy logic based on the orthogonal arrays. International communications in heat and mass transfer, 33(6), pp.744-752, <https://doi.org/10.1016/j.icheatmasstransfer.2006.02.011>, 2006
- [18] Chiang, K.T. and Chang, F.P., 2006. Application of response surface methodology in the parametric optimization of a pin-fin type heat sink. International communications in heat and mass transfer, 33(7), pp.836-845, <https://doi.org/10.1016/j.icheatmasstransfer.2006.04.011>, 2006
- [19] Chen, C.T., Wu, C.K. and Hwang, C., 2008. Optimal design and control of CPU heat sink processes. IEEE Transactions on components and Packaging Technologies, 31(1), pp.184-195, <https://doi.org/10.1109/TCAPT.2008.916855>, 2008

[20] Devi, S.P., Manivannan, S. and Rao, K.S., 2012. Comparison of nongradient methods with hybrid Taguchi-based epsilon constraint method for multiobjective optimization of cylindrical fin heat sink. *The International Journal of Advanced Manufacturing Technology*, 63(9-12), pp.1081-1094, DOI 10.1007/s00170-012-3985-7, 2012

Chapter 3 Energy aware pro-active control for air cooled datacenters using Machine Learning

3.1 Introduction

The rapid expansion of networking-based applications such as the Internet of Things (IoT), growth in video streaming and social media, and increased demand for high-performance servers for computation-intensive applications have posed a major data center thermal management challenge. These advancements leading to data center proliferation have also created a problem of energy efficiency due to increasing server and consequently, data center power consumption. Recent data shows that the data centers account for nearly 1% of global electricity consumption which may increase threefold over this decade [1]. At the same time, data centers also contribute around 0.3% of global carbon emissions, triggering concerns by many environmental protection bodies across the globe [2]. Typical air-cooled data centers rely on refrigeration-based Computer Room Air Conditioning (CRAC), or Computer Room Air Handling (CRAH) units connected to a chilled water-cooling tower. A typical energy consumption distribution for various purposes in a refrigeration-based air-cooled data center is shown in Figure 3-1 [3]. It can be assumed that one of the approaches to enhance energy efficiency could be in the reduction or optimization of cooling power by strategically optimizing the coolant (air) delivery to the servers.

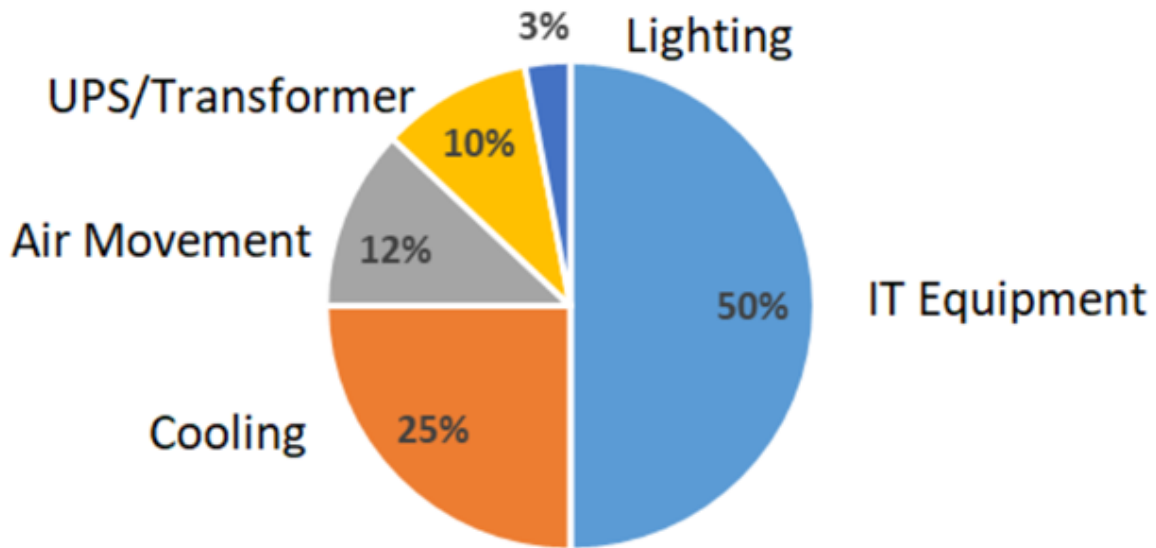


Figure 3-1 DISTRIBUTION OF ENERGY CONSUMPTION IN A TYPICAL DATA CENTER

The workload distribution and workload arrival in data centers are highly dynamic which directly impacts the airflow in hot and cold aisles. Conventional control strategies inside the cooling systems used in data centers utilize an ON/OFF mechanism using PID control [4] from sensor feedback. The mechanism is triggered when the temperature setpoint deviates within a certain pre-decided deadband. This deadband range gives rise to continuous temperature fluctuation, which if outside the recommended temperature range, can lead to higher server failure rates [5]. Thus, utilizing reactive control strategies to optimize the airflow to Information Technology Equipment (ITE) under such highly varying ambient conditions isn't the best approach. Model-based predictive control strategies can alleviate these concerns by predicting the cooling requirements based on historical data of the workloads and consequent cooling requirements [1]. Using this approach, the fluctuations that are typically observed in PID-based cooling system control can be minimized, resulting in more stable cooling control [6, 7]. One of the major bottlenecks in developing such approaches is the lack of historical data. This task can become even more cost and time-intensive, and completely non-feasible in hyperscale cloud data centers if an experimental investigation is carried out. Computation Fluid Dynamics (CFD) can be used in such situations to understand various

thermodynamic parameters using data-driven models trained with validated CFD models of real data centers.

Artificial Neural Network (ANN) has been used in various domains such as trade and finance, audio-visual recognition and even building energy optimization and thermal control [8, 9]. More recent studies also show the application of ANN in predicting supply air temperatures and power management in data centers [10, 11]. Studies have been done to compare ANN and typically used mathematical modeling techniques to prove that ANN can be used for both transient and steady-state thermal management of data centers [12]. Song et al. [13] used ANN for predicting the average airflow rate and airflow temperature coming out of the data center floor tiles. Thus, it can be referred from these studies that data-driven modeling combined with CFD can be used to simulate various field scenarios with limited resources. Machine learning approaches can then be used to develop forecasting models that can proactively respond to various transient scenarios inside the data centers. Sensitivity studies for developing CFD models as established by previous literature has also been considered [14].

To analyze the distribution of air and its distribution efficiency, the most realistic method is to calculate the performance metrics. It is important to assess the required supply of air temperature and flowrate through the ITEs to maintain its operational efficiency based on the datacenter server/room design standards. The datacenter room standards are followed mainly focusing on the ITE reliability/downtime which impacts the energy efficiency at the room level due to excessive cooling provisioning. It is observed that in air cooled data centers, cooling energy consumption contributes to almost 40% of the total energy consumption of the facility. Since datacenter energy consumption has become three folds in the past decade, we must consider optimizing the existing and upcoming datacenters through efficient and smart methods of cooling which directly impacts the PUE.

Chilled water systems are usually designed to operate for a differential temperature (ΔT) between the supply and return water between the chiller and the facility loads. This ΔT will influence air handling units, coil sizes, distribution system, pumping costs, and chiller sizing. A higher ΔT usually affect the chiller evaporator log mean temperature difference (LMTD) and require long tubes or more coil passes, which in turn increases chiller pressure drops (ΔP) that need to be overcome by the pumps. At higher temperature differentials of 10°C to $15^{\circ}\text{C} - \Delta T$, low supply water temperatures (6°C to 8°C), and variable flow with modulating valves, a design strategy could reduce pump energy (lower flowrate) and piping installation cost. The total annual system energy use must be considered to calculate overall energy saving. The control schemes for a chilled water system usually vary with the size and complexity of the system, and with the type and number of pumps used. The chilled water system's flowrate can be controlled from static pressure, but has limited flexibility for operational changes, and might waste energy in over pumping. Condenser flowrate can be controlled from differential pressure (ΔP) at the chiller plant, which has similar disadvantages as the chiller. Temperature setpoint usually doesn't change dynamically at the chiller thereby wasting energy.

The study's main goal is to enable dynamic setpoint modifications based on constantly changing spatial and temporal workloads in white space. Since studying the behavior and adjusting to see the effect of various integrated cooling systems in a working data center is exceedingly challenging, CFD modeling and simulation of such situations help us understand and make changes to operational setpoints and controls. The goal of the research is to optimize the datacenter's air and liquid side energy consumption simultaneously, focusing on the control of airside and liquid-side fluid and thermal network as shown in Figure 3-2.

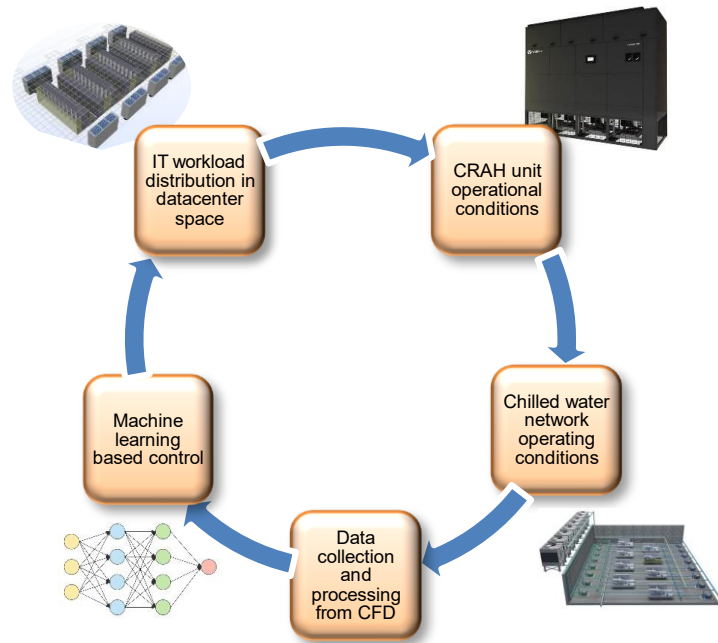


Figure 3-2 SEQUENCE OF OPERATION

A set of scenarios that occur in a realistic datacenter are modeled and simulated using a commercial CFD tool. In this study we investigate the application of machine learning model in predicting the control setpoints for air and liquid side cooling systems. A parametric study is conducted using a validated datacenter CFD model where air cooling unit setpoints are managed based on the responses of internal sensors, and we override them if necessary for optimum energy efficiency. Data for 'n' situations is collected and compared to the standard control technique. Traditional vs proactive controllers' power usage and heat dissipation are calculated and compared. The study also looks at scenarios including recirculation, over/under pressurization, over/under-provisioning, thermal runaway, and cooling unit failure. Machine learning models are chosen depending on the type and size of data acquired, as well as the forecast. A good approach for predicting the most contributing parameter for determining the setpoints is using the random forest method and for optimizing the control, artificial neural network machine learning model is used.

3.2 Computational Model

Two CFD analysis were carried out, one to visualize the zone of influence of the CRAC units in provisioning the ITEs and the second to calculate the total power consumed by the CRAC units in various scenarios also to generate datasets for ANN training. Table 3-1 and 3-2 has the data center configuration and CRAH unit specifications.

Table 3-1 DATA CENTER CONFIGURATION

Configuration of the datacenter	
Floor area	1200 ft ²
No. of rows	8 rows with 10 racks each
Rack arrangement	Hot aisle contained
No. of servers/rack	10 servers/rack
Server dimensions	2U
No. of CRAC units	8 (n+2) configuration
No. of chillers	3 (n+1) configuration

Table 3-2 CRAH UNIT SPECIFICATION

CRAH Unit specifications	
Model	Liebert CW 180
Max. Capacity	200 kW
Max. Flowrate	17,300 CFM
Control setpoints	
Supply Air Temperature	20°C - 27°C
Supply Air Flowrate	60% (10,380 CFM)

The data center is a hot aisle contained with a ceiling height of 2ft. The return of hot air from the IT passes through hot-aisle containment to the Computer Room Air-Conditioning (CRAC) units, the cold air is supplied to the IT through perforated tiles having 50% open area from a 2ft underfloor. A typical datacenter is modeled in a commercial CFD tool 6SigmaRoom as shown in fig 3-3.

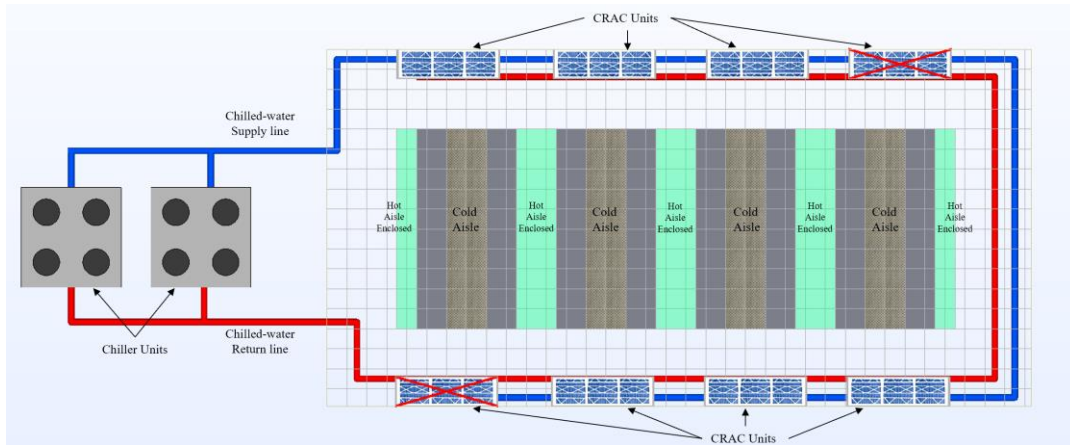


Figure 3-3 DATA CENTER CFD MODEL (TOP VIEW) IN 6SIGMA ROOM

A baseline simulation with idle load is compared with the realistic data from the data center facility to validate the model. Further analysis using the model is expected to represent realistic scenarios in a datacenter. A 1D flow network model (Fig 3-4) is coupled with the room level CFD model to simultaneously control the chilled water-side and air-side parameters of the data center.

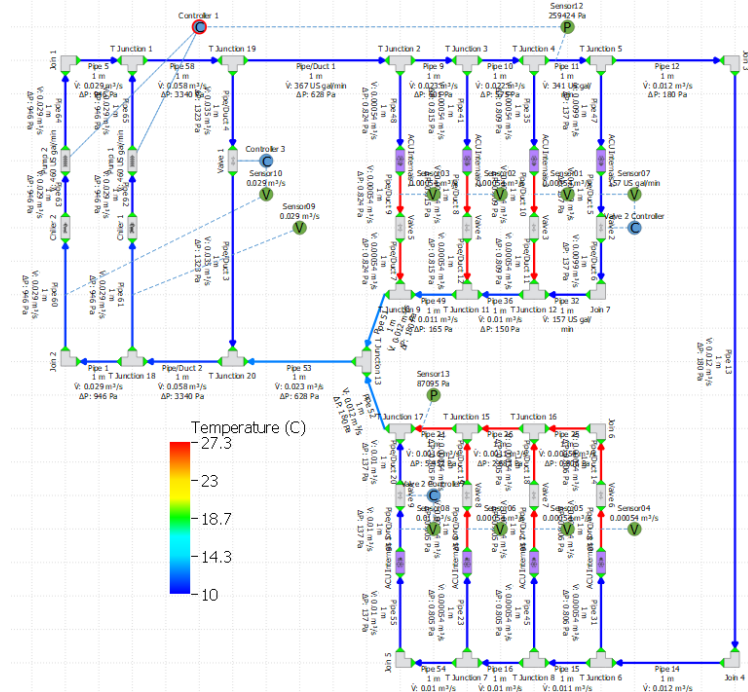


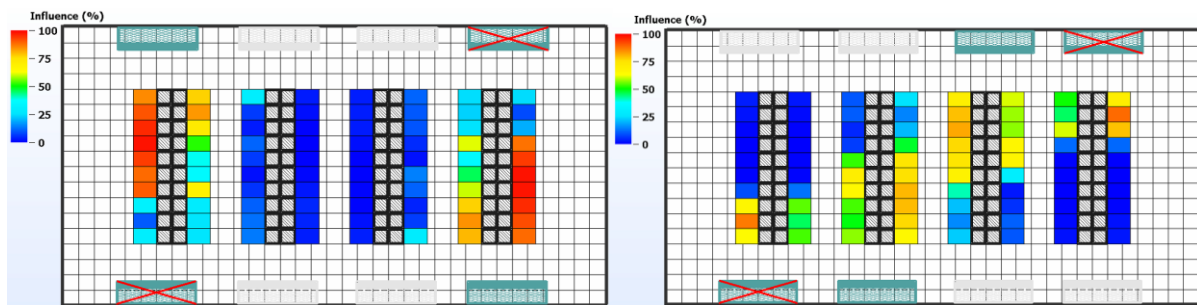
Figure 3-4 1D FLOW-NETWORK FOR CHILLED WATER LOOP

The CFD model to understand the ACU's influence on provisioning the ITE's, we have set up the control network in such a manner that can be Average ITE outlet temperature sensor values based on % of influence are taken as T Return. Similarly, supply temperature T Supply setpoint is set to 22°C such that the ACUs respond when any of the inlet temperature sensors read a value of more than 22°C. The blower speed for the ACU fans were controlled using VFDs and the values are updated for every scenario to capture different scenarios using the same boundary conditions. CRAC unit dimensions and specifications are modeled according to Liebert CW 180, ACU built by Vertiv cooling technologies.

The algorithm will make a small adjustment to the supply water temperature, by increasing/decreasing the temperature and measuring the resulting changes in energy consumed by the chiller, pumps, and ACU. If the supply water temperature is to the left of the optimum point, then total energy used by these systems will decrease. The algorithm raises the temperature a little bit more. On the contrary, if supply water temperature is to the right of the optimum point, the total energy used by the system will increase, then the algorithm reduces the temperature a little bit. Eventually, it will find the optimum temperature, where either increasing or decreasing the supply water temperature increases the total energy used. It will happen over the course of time based on the fluctuating IT load. If water cooled chillers are used, there is a similar trade-off on the condenser water side. If you lower the temperature of the condenser water as it's supplied to the chiller, the amount of energy used by the chiller will be reduced. The cooling tower can't provide water cooler than the outside air wet bulb temperature. To achieve closer to the ambient air wet bulb temperature fans and circulating pumps must spend more energy. This may be achieved by speeding up the fans and pumps, bringing on more tower stages, sometimes both. Therefore, the algorithm decides to adjust the pump flowrate and fan speed based on the ambient conditions.

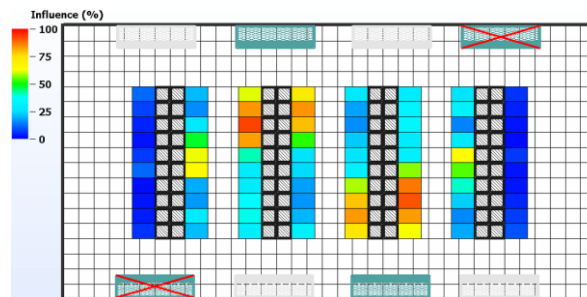
3.3 Zone of Influence of CRAH units

Zone of influence for each CRAH unit can be found using a baseline simulation and a built in feature in 6Sigma Room. The charts depict the influence of provisioning the ITEs located in the data center white space at the rack level. Figure 3-5 (a), (b), and (c) shows the influence of CRAH units 1 & 8, 3 & 6, and 2 & 7.



(a)

(b)



(c)

Figure 3-5 INFLUENCE OF CRAH UNIT (a) 1 & 8 (b) 3 & 6 (c) 2 & 7

3.4 Boundary conditions and cases

A range of input parameters were considered for boundary conditions for the CFD model. Table 3-3 shows the input parameter space for which the simulations were run.

Table 3-3 Boundary Conditions

DC Room model Boundary Conditions	
Rack Power utilization (%)	30, 60, 100
Supply Air Temperature (°C)	22, 26, 30
Coil inlet Temperature (°C)	10, 14

Numerous parameters were recorded for all the scenarios from the CFD simulation. These datasets were fed into a Random Forest regression model to identify the contribution of each parameter to the total CRAH unit supply Air flow rate. The results are shown in fig 3-6. It was found that rack power, rack temperature difference and Plenum pressure were the most contributing factors for the CRAH unit supply air flow rate setpoint. These parameters were selected for training the Machine learning algorithm which is Artificial Neural Network in this case.

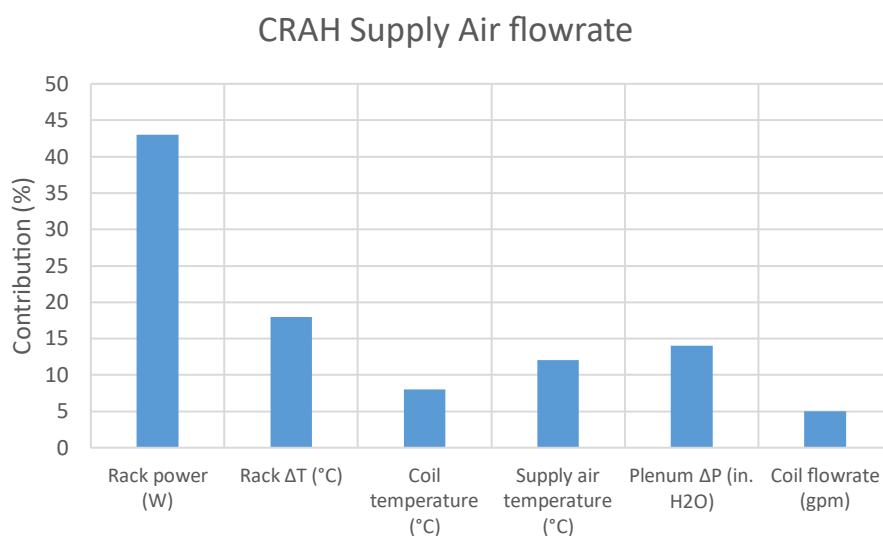


Figure 3-6 Feature Selection

3.5 Artificial Neural Network

The input and output parameters selected from the random forest model were divided into training and validation datasets. 80% of the total data was selected to train the Neural Network and the remaining 20% were selected to validate or test the performance of the

algorithm. The input, output parameters and the neural network model is shown in fig 3-7. Data separation, training and validation was done using python libraries.

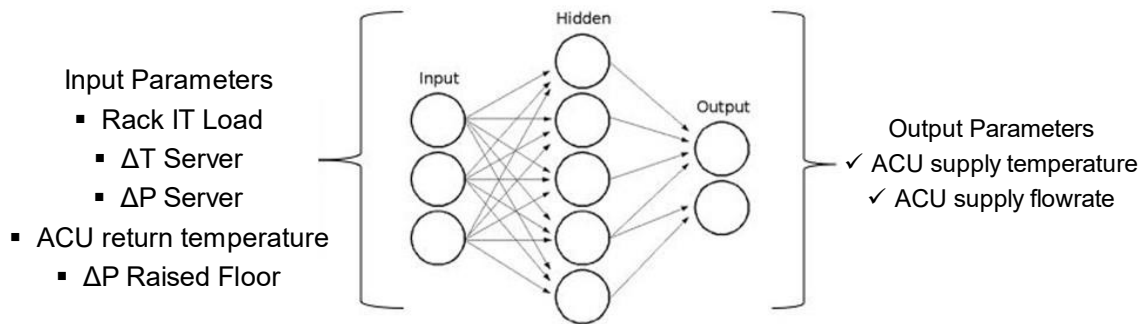


Figure 3-7 ANN MODEL WITH INPUT AND OUTPUT PARAMETERS

The performance of the algorithm was close to 100% accuracy as shown in fig 3-8. More data and non-linearities need to be addressed in future studies to adapt various layouts and configurations of a realistic data center.

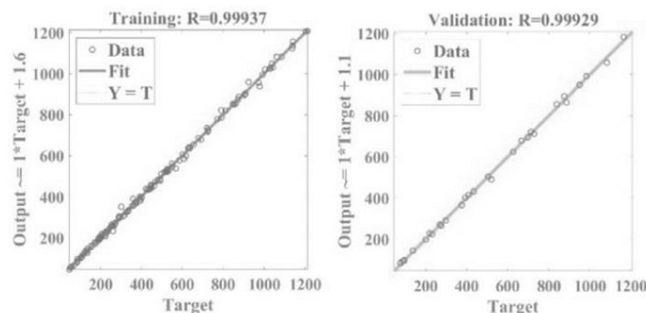


Figure 3-8 ANN ERROR

3.6 Conclusion

CRAH setpoint modification reflects change in ΔT at the chiller. CRAH Supply air temperature and flowrate setpoint predicted by the neural network will be compared to the baseline model working on traditional controls. If the Outside air temperature is favourable, chiller efficiency goes high. – Potential to include OA variable to the neural network training algorithm. The future work includes energy Analysis comparing sensor based and ML based controls.

3.7 References

- [1] Recalibrating global data center energy-use estimates
- [2] “How to stop data centers from gobbling up the world's electricity.
- [3] Huigui Rong, Haomin Zhang, Sheng Xiao, Canbing Li, Chunhua Hu, Optimizing energy consumption for data centers, *Renewable and Sustainable Energy Reviews*, Volume 58, 2016, Pages 674-691, ISSN 1364-0321, <https://doi.org/10.1016/j.rser.2015.12.283>.
- [4] Theory and applications of HVAC control systems – a review of model predictive control (MPC)
- [5] Data center power equipment thermal guidelines and best practices ASHRAE TC9.9 (2016)
- [6] Nonlinear Model Predictive Control for Energy Efficient Cooling in Shopping Center HVAC, 2019 IEEE Conference on Control Technology and Applications
- [7] Y. Ma, F. Borrelli, B. Hancey, B. Coffey, S. Bengoa, P. Haves Model predictive control for the operation of building cooling systems
- [8] C. Deb, L.S. Eang, J. Yang, M. Santamouris Forecasting diurnal cooling energy load for institutional buildings using Artificial Neural Networks *Energy Build.*, 121 (2016), pp. 284-297, [10.1016/j.enbuild.2015.12.050](https://doi.org/10.1016/j.enbuild.2015.12.050)
- [9] Y.T. Chae, R. Horesh, Y. Hwang, Y.M. Lee Artificial neural network model for forecasting sub-hourly electricity usage in commercial buildings *Energy Build.*, 111 (2016), pp. 184-194, [10.1016/j.enbuild.2015.11.045](https://doi.org/10.1016/j.enbuild.2015.11.045)
- [10] N. Kansara, R. Katti, K. Nemati, A.P. Bowling, B. Sammakia Neural Network Modeling in Model-Based Control of a Data Center (2015), [10.1115/IPACK2015-48684](https://doi.org/10.1115/IPACK2015-48684)

- [11] N. Liu, X. Lin, Y. Wang Data center power management for regulation service using neural network-based power prediction 2017 18th International Symposium on Quality Electronic Design, ISQED (2017), pp. 367-372, 10.1109/ISQED.2017.7918343
- [12] Athavale, M. Yoda, Y. Joshi Comparison of data driven modeling approaches for temperature prediction in data centers Int. J. Heat Mass Tran., 135 (Jun. 2019), pp. 1039-1052, 10.1016/j.ijheatmasstransfer.2019.02.041
- [13] Z. Song, B.T. Murray, B. Sammakia Airflow and temperature distribution optimization in data centers using artificial neural networks Int. J. Heat Mass Tran., 64 (2013), pp. 80-90
- [14] S. Singh, K. Nemati, V. Simon, A. Siddarth, M. Seymour, and D. Agonafer, "Sensitivity Analysis of a Calibrated Data Center Model to Minimize the Site Survey Effort," 2021 37th Semiconductor Thermal Measurement, Modeling & Management Symposium (SEMI-THERM), 2021, pp. 50-57.

Chapter 4 Feasibility and Energy Analysis of Rear Door Heat Exchanger for Data Centers

Reprinted with permission © 2022 ASME

4.1 Abstract

Due to increased use of high-performance computing in datacenters to cater to huge workloads, old low-performance compute servers must be replaced endlessly with high-performance compute servers. Traditional air-cooling systems are insufficient to provision and run the servers in optimal conditions as the datacenter thermal footprint or rack density grows, resulting in thermal throttling. To sustain the growing needs, Rear Door Heat Exchangers (RDHx) are deployed in existing datacenters along with peripheral Computer Room Air Handling/Conditioning (CRAH/CRAC) units. RDHx transfers heat from the rear end of the racks and rejects it into the facility's chilled water. This study will demonstrate the suitability of RDHx for low density as well as high density rack applications. A baseline CFD model had a generic datacenter layout with peripheral CRAH/CRAC units and RDHx. Several case studies were conducted by varying the air and liquid inlet temperatures for rack and RDHx, respectively. We also compared active and passive modes of operating RDHx while server fans provide flowrate based on the IT inlet temperature. The paper will also discuss the feasibility of designing a datacenter with only RDHx and no peripheral CRAC/CRAH units while maintaining the thermal envelop. The research will also provide a guideline in implementing RDHx based on the heat load and server design.

4.2 Introduction

Air cooled datacenters typically have two configurations for cooling the IT equipment. Raised floor configuration where CRAH/CRAC units are placed at the periphery in the white

space supplying air through the underfloor plenum which reaches the ITE equipment through the perforated tiles. Non-raised floor configuration where the cold air is supplied from the ceiling or via peripheral units. This configuration typically requires hot aisle containment and a return plenum. In the traditional design of data centers, issues such as hot aisle recirculation or cold air bypass are very commonly observed. Recirculation of the warm air from the ITEs mixing with the ambient air is also a major concern as it leads to higher air inlet temperature for the ITEs. Conventional methods of air cooling have significant disadvantages in heat dissipation due to high energy consumption at the CRAC/CRAH units, humidity excursions, recirculation into the cold aisle, extensive site survey effort etc. [1]. Even though there are optimization strategies for air cooling provisioning, there are limitations due to increased demand in high power consuming devices [2]. The growing rate of worldwide servers is 2.5×10^6 annually based on the industry data, and this growth is expected to continue [3]. Also, as per ASHRAE (Fig. 4-1), heat load per 42U rack shows that rack density is continuously increasing. This makes it difficult for conventional air-cooling methodologies to provision additional capacity [4]. Therefore, there are potential advancements in methods of air cooling; one such method is retrofitting existing datacenters with Rear door heat exchangers (RDHx).

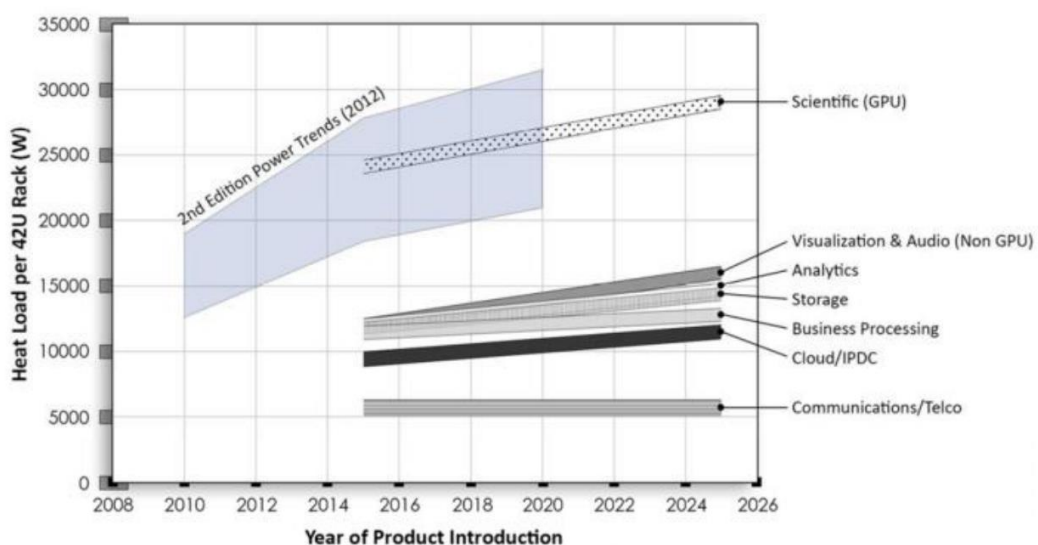


Figure 4-1 EVOLUTION OF THE ASHRAE POWER TRENDS [4]

Hybrid cooling systems are typically self-contained components like the side car, which is an air-liquid heat exchanger installed on the side of a server cabinet. IBM introduced this enclosed heat exchanger, which is designed to remove up to 35 kW [5]. A fully enclosed server cabinet, which commonly uses a V-shaped heat exchanger on the bottom of the cabinet, is another example of a hybrid cooling system [6]. The RDHx (Rear Door Heat eXchanger) is a heat exchanger located at the back of the cabinet, where hot exhaust air leaves the servers. It was demonstrated in [7] that using an RDHx in various data center configurations can have several advantages, including 1) reducing the number of CRAH units required, 2) eliminate hot spots by removing heat closer to the source, and 3) allowing higher chilled water temperatures while still maintaining IT inlet temperatures within recommended ranges. RDHXs were used in tandem with room level air conditioning in a hybrid manner [8]. The CRAC fan speed was optimized using a control mechanism based on the heat load of the servers. They were able to save 47 kWh of energy, or around 6% of the total energy consumed by the CRAC. RDHx needs chilled water from the facility to cool the warm air exiting the ITEs. To provide chilled water for the RDHx, Cooling Distribution Units (CDUs) are deployed in datacenters as shown in Fig. 4-2. CDUs conditions the water to the desired temperature and the flowrate needed at the RDHx to cool the warm air exiting the ITE. In this paper we are focusing on feasibility and considerations in deploying only RDHx without peripheral CRAH/CRAC units in a datacenter. This will enable us to have a room neutral design and eliminate the need for hot aisle containment.

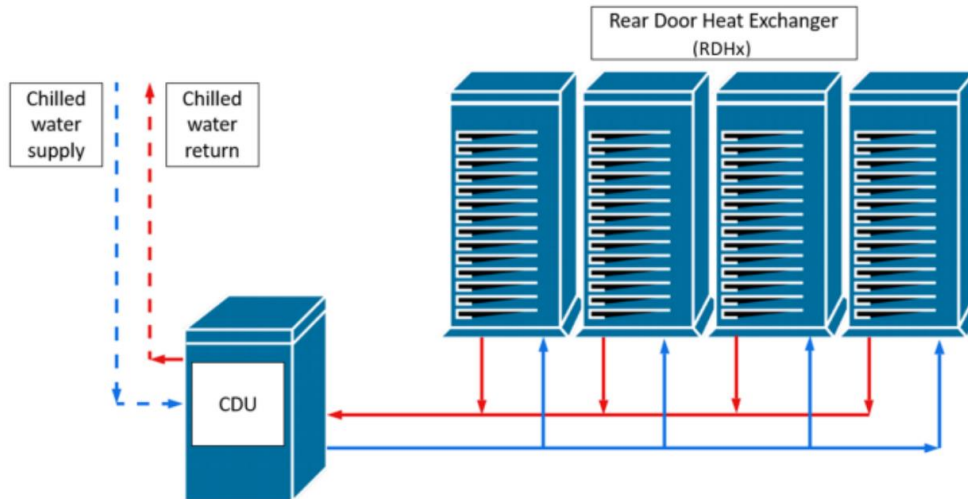


Figure 4-2 LIQUID COOLING LOOP [9]

A commercially available CFD tool from Future facilities, 6SigmaRoom was used to model the ITE rack with the RDHx. The baseline version of the model of the racks were fitted with a commercially available RDHx and simulated for various air and coolant inlet conditions.

4.3 Computational Model

Rear door heat exchangers are typically built to fit at the rear end of the rack which is 2ft wide. A commercially available rear door heat exchanger (68 x 21.3 x 7 in) was chosen to fit at the back of the ORV2 rack (84.8 x 23.6 x 42 in) from Open Compute Project (OCP) [10]. The rack considered for analysis is a representation consisting of a 4-OU IT equipment, power shelves and a TOR switch. All the ITEs are provided with power ratio curves based on a 12°C ΔT . The rack's CFD model is shown in Fig. 4-3 and the details are tabulated in table 4-1.

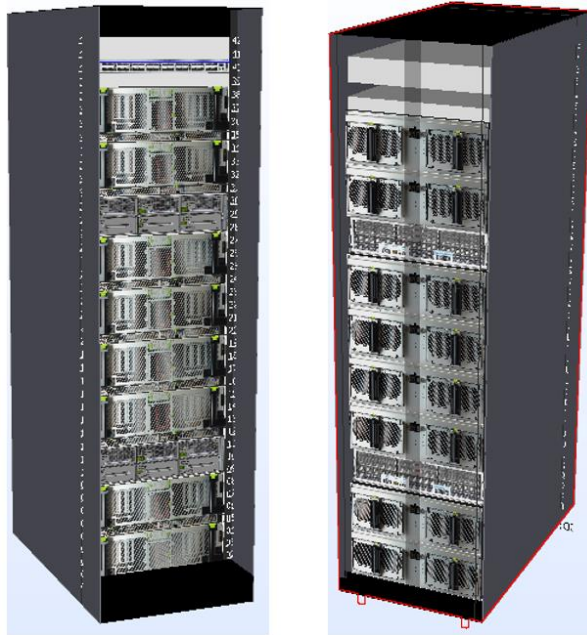


Figure 4-3 FRONT AND BACK OF A TYPICAL OCP RACK

Table 4-1 DETAILS OF IT EQUIPMENTS IN THE RACK

ITEs	Form factor & units per rack	Max. power limit (W)	Airflow Required (CFM)
TOR Switch	1U, 1 per rack	170	23.4
Power Shelf	3U, 2 per rack	315	44.1
ITE	4U, 8 per rack	2500	350

The 4-OU IT server is assumed to be representative of a high-density server cooled by four 9256 CR fans. There are 4 fans per server in parallel in pull configuration. These are powerful fans designed to overcome higher pressure drops as indicated by the PQ curve (pressure drop vs air flowrate curve) of a typical 9256 CR fan in Fig. 4-4. However, it is to be noted that data centers typically have a heterogeneous deployment of servers. It may house servers with fans that can overcome modest pressure drops of one to two inches of water. Thus, it is important to consider the aforementioned aspects while deciding on an RDHx solution.

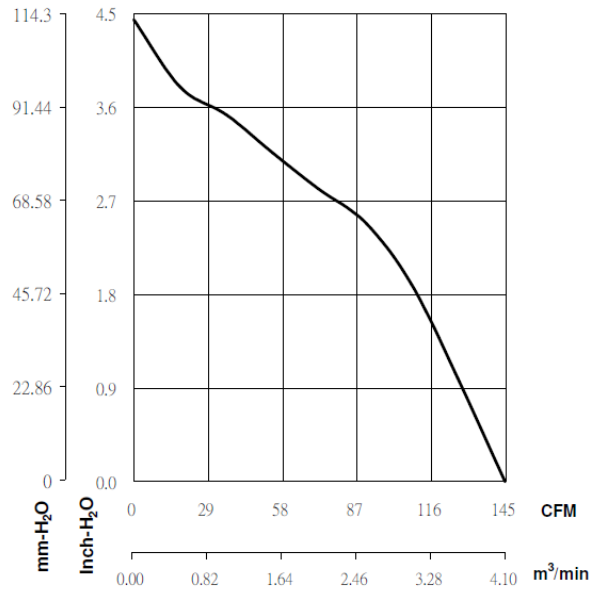


Figure 4-4 P-Q CURVE OF A CR 9256 FAN

The RDHx typically has a single heat exchanger unit with valve connections for both supply and return of facility water. The position of the valves depends on the data center layout and the placement of chilled water pipes. They are either located on the top (for overhead piping) or at the bottom (raised floor configuration). RDHx is connected to the main loop via an isolation valve. The isolation valve helps to disconnect the rack from the chilled water loop without disrupting facility operations at large. The coolant flow through each of the Hx could be modulated by a control valve attached to the RDHx unit. This is necessary to avoid condensation and moisture accumulation when servers are idle. Depending on the design of the coil, and how the air flow is distributed (variation in face velocity at the inlet of RDHx), air may be cooled non-uniformly. This is a common occurrence in a passive RDHx configuration. To counter this, axial fans are placed at the rear of the Hx unit such that it assists to pull the air from the cabinet through the server and cabinet and release it into the data center. Although the addition of fans would consume more power, it helps to maintain a pressure-neutral data center environment for the server fans. Apart from distributing the flow uniformly over the coil, it also helps in overcoming the air-side pressure drop of the heat exchanger. Fig 4-5 shows the CFD model of a typical active RDHx with the cabinet.

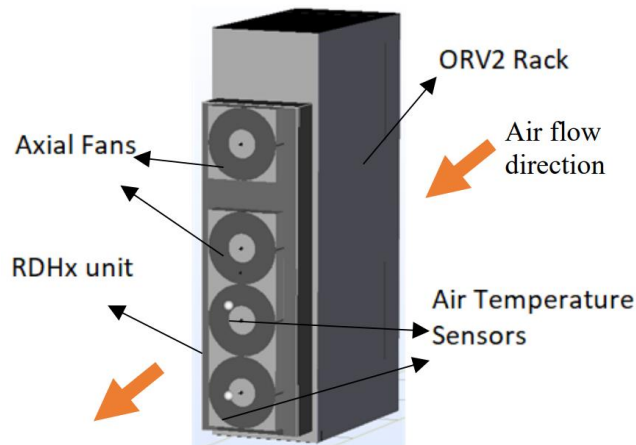


Figure 4-5 ORV2 CABINET WITH RDHX - CFD MODEL

To assess the feasibility of deploying the RDHx at scale for a large data center; a CFD model was first built with single rack placed in a room where the width of the room is same as the width of the rack and has a high ceiling of about 15' from the floor. The model is created in such a way that the air exiting the RDHx can reach the IT inlet only by passing on top of the cabinet and not from the sides. The set-up was to mimic an actual data center deployment, where the racks are placed next to each other in a row. It is assumed that the empty U-slots will have blanking panels and no noticeable gaps exist. A typical 5% leakage is assumed to account for gaps in between the servers, mounting rails etc. In idle scenario, the server inlet air temperature would be the same as exit air temperature of the RDHx.

In a passive RDHx, server fans need to have capacity to deliver pressurized air. If the required pressure is not attained by the server fans to overcome pressure drop through heat exchanger, air will cease to flow and start accumulating in the area between the server fans and the heat exchanger. At certain point, it will flow back into the room through gaps (at elevated temperature due to the heat removed from the server) – defeating the purpose of RDHx. In addition, for a passive RDHx, coolant flow and airflow cannot be controlled to maintain a design ΔT . This could potentially lead to inefficient operation of the overall data center. Passive

RDHx are more suitable for servers that can provide high static pressure at the downstream of the server to have effective flow across the heat exchanger.

In an active RDHx, fans assist the extraction of the air from the cabinet if the internal fans cannot provide enough static pressure at the desired fan speed. The fan speeds are modulated to respond to required air-side ΔT . The control valve on the coolant side modulates the coolant flow to ensure air is cooled to the required temperature. Therefore, the desired temperature of 30°C is given to the controller such that the flowrate across the RDHx adjusts to provide the corresponding amount of cooling.

For the current study, the coolant flowrate for the RDHx is controlled using an internal controller that works on the sensors built in the CFD model. The coolant flowrate controller is modulated to have direct response with respect to the air temperature exiting the RDHx. If the temperature starts to increase from the set-point (30°C in our case study), coolant flow rate will start to increase until the requirement is met, or maximum flowrate is achieved. In the later scenario, that would be the case of a thermal run-away where the coolant flow rate is maximum yet the exit air temperature from the RDHx is above the set-point. Table 4-2 summarizes the heat exchanger specification considered for analysis.

Table 4-2 HX SPECIFICATION

Hx Specification	Values
Coolant flow direction	Top to Bottom
Nominal Cooling Capacity	35 kW
Reference coolant flowrate (maximum)	23 GPM
Hx Effectiveness	0.8
Reference Air flowrate (maximum)	3,000 CFM
Coolant Temperature	26°C
Coolant flowrate	Controlled based on air temperature sensor (30°C)

4.4 Results and Observation

Over the years, CFD has become an important tool to ascertain design flaws and optimize them further rather than going through intensive experimentation with manufactured models. To visualize the 3-D flow inside the cabinet and analyze the flow rate variation, momentum, energy, and mass conservation equations were solved using 6SigmaRoom. The following sections will detail the results and key takeaways after post processing the results.

4.4.1 Single Rack Model

A single rack model was defined in the section before. Three cases of racks installed with passive rear door heat exchanger with varying rack densities are considered for analysis. For various rack power densities, the total required air flowrate requirement is calculated based on approximately 12°C ΔT . For simulations, depending on the case study, rack power is scaled by adjusting the utilization factor. For a 7-kW rack the corresponding air flowrate requirement is 1050 CFM, similarly, for 9 kW and 11 kW the air flow requirements are listed in Table 3. Server fans for the use case are selected to provide adequate flowrate across the servers. When RDHx is in place, the server fans must overcome the pressure drop across the Hx to push the air through it with minimal to no recirculation. Table 4-3 shows the set of simulation runs in CFD to determine the heat exchanger operating conditions and potential impact on server fans. Static or back pressure due to the installation of RDHx is noted.

Table 4-3 RDHx CASES

Cases	RDHx mode	IT Load (kW)	Required flowrate (IT-CFM)
Case 1	Passive	7	1050
Case 2	Passive	9	1350
Case 3	Passive	11	1670
Case 4	Active	11	1670

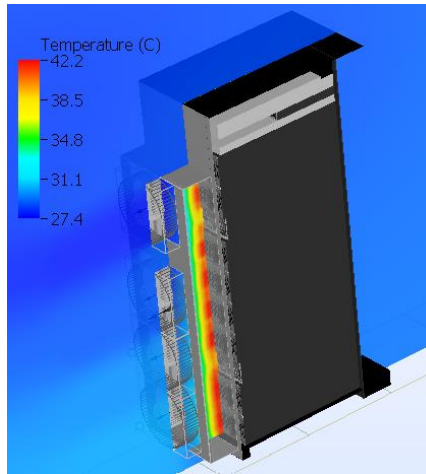


Figure 4-6 RESULT PLANE POSITION

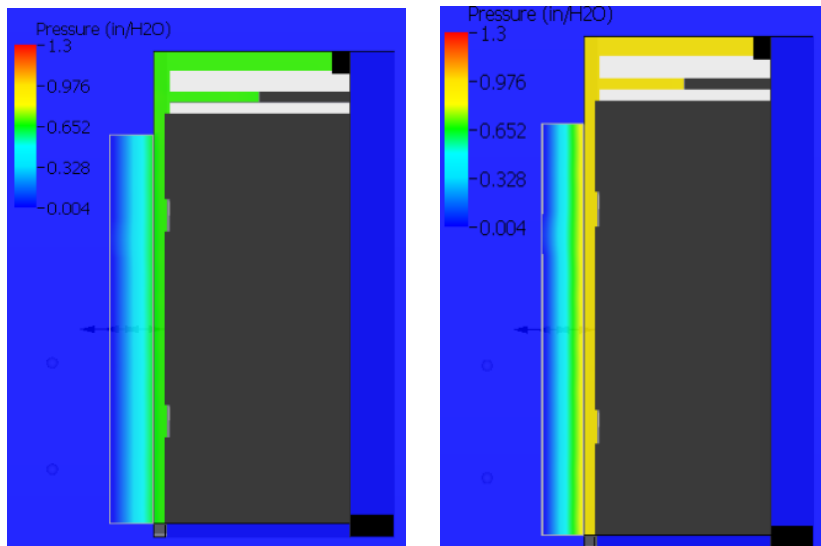


Figure 4-7 PRESSURE CONTOUR PLOT FOR CASE 1 (LEFT) AND CASE 2 (RIGHT)

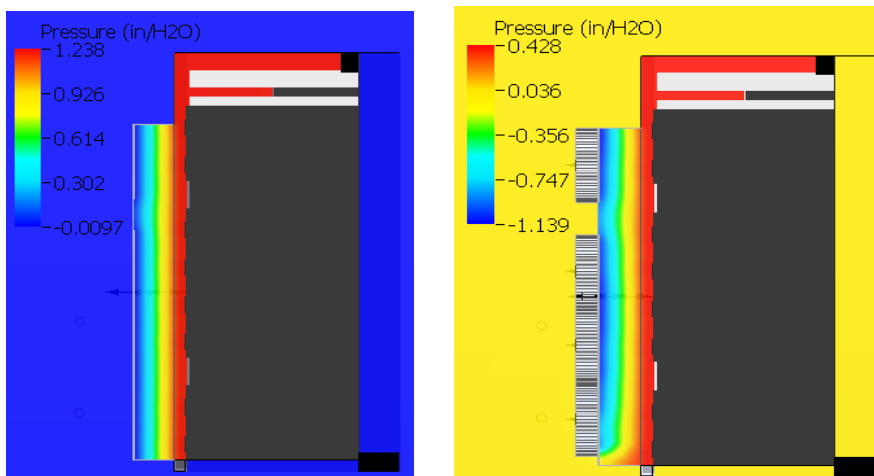


Figure 4-8 PRESSURE CONTOUR PLOT - CASE 3 (LEFT) AND CASE 4 (RIGHT)

Table 4-4 PRESSURE AND TEMPERATURE RESULTS

Cases	Static Pressure (in/H ₂ O)	RDHx -Air Flowrate (CFM)	Mean-IT Inlet Temp (°C)	Coolant Exit Temp (°C)	Coolant flowrate (GPM)
1	0.66	961	29.5	28.9	9.13
2	0.90	1234	29.8	29.3	10.3
3	1.15	1507	30.2	29.7	11.2
4	1.26	1728	29.7	29.7	11.6

The result plane is positioned midway of the rack width as shown in fig 4-6. All the temperature and pressure contours are shown based on the same plane. For case 1 and 2 (fig. 4-7), the contour plots for temperature and pressure show that the rack inlet temperature conditions are met and the static pressure inside the cabinet increases as expected. In case 3 (fig.4-8), mean inlet temperature for ITE exceeded 30°C. Upon further inspection, it can be seen from the pressure plot, that the static pressure between the heat exchanger and the server fans is quite high compared to the previous two cases. It is to be noted that there are other structural elements such as a bus bar at the rear of the cabinet which will also add to overall resistance to airflow. This increased pressure will impact the server fan performance. If capable, server fans would spin at higher RPM to overcome the back pressure (assuming 12°C ΔT across servers). This in turn can increase the server fan power. In the current CFD set-up, the exact impact on fan performance was not captured as black box for servers and fans were used. Authors plan to further evaluate and calculate the increase in fan power for the cases studied. Another way of looking at the results was to understand the amount of hot air recirculation. Airflow through servers was modeled as fixed airflow. However, it was noted that the airflow exiting out of heat exchanger was lower than the server airflow. This can be accounted for some of the airflow escaping through leakage. With increase in static pressure, the volume of hot air that escapes through gaps increases. The results are tabulated in table 4-5. This explains the increase in mean server inlet temperature.

Table 4-5 PRESSURE AND TEMPERATURE RESULTS

Cases	Rack CFM (From table 3)	RDHx -Air Flowrate (CFM)	Leakage/recirculation (CFM)
1	1050	961	89
2	1350	1234	116
3	1670	1507	163
4	1670	1728	-58 (excess airflow through RDHx)

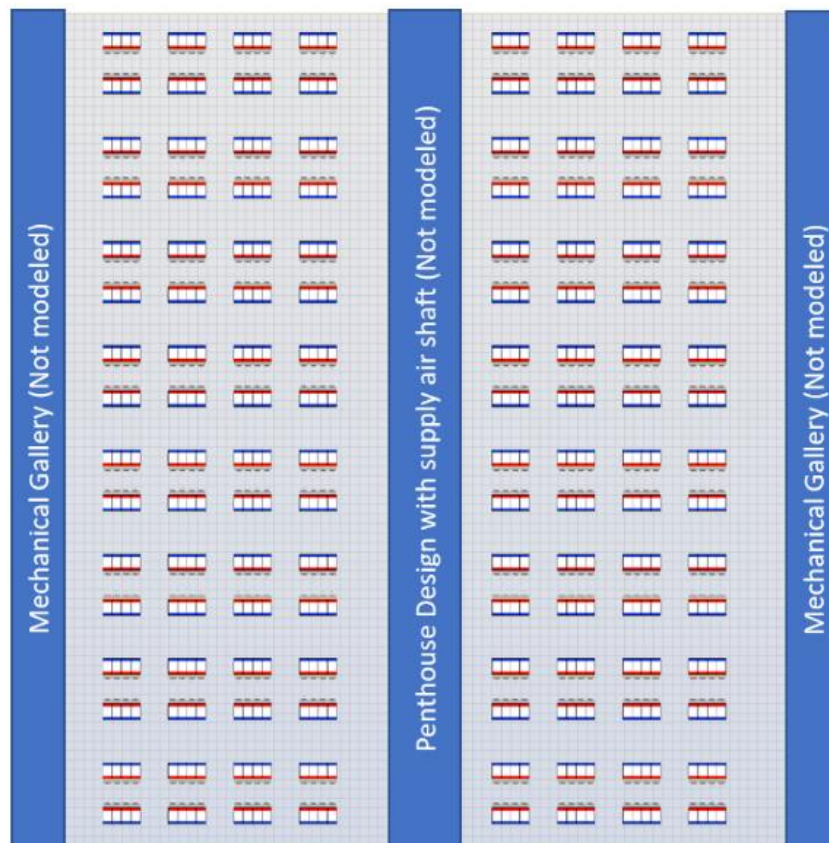
Since the exit air temperature from RDHx exceeded 30°C, rack power greater than 11 kW for passive mode was not analyzed. Instead, a fourth case wherein the RDHx was operated in active mode was analyzed. Fans were controlled to maintain air-side ΔT of 12°C which resulted in about 432 cfm of airflow per fan. It is typical for an active RDHx to have specification of 600W of total fan power for 3,000 cfm of air. Using it as a reference and based on fan laws, it is estimated that under active mode, the rack will consume an additional 500W of power.

At first glance, this may sound like a lot. However, from a data center point of view it may prove beneficial. Previous studies have shown that how addition of RDHx can help lower the cooling load and in turn save power. If a data center can be operated without the need of any peripheral CRAC/CRAH units, there is a potential for savings.

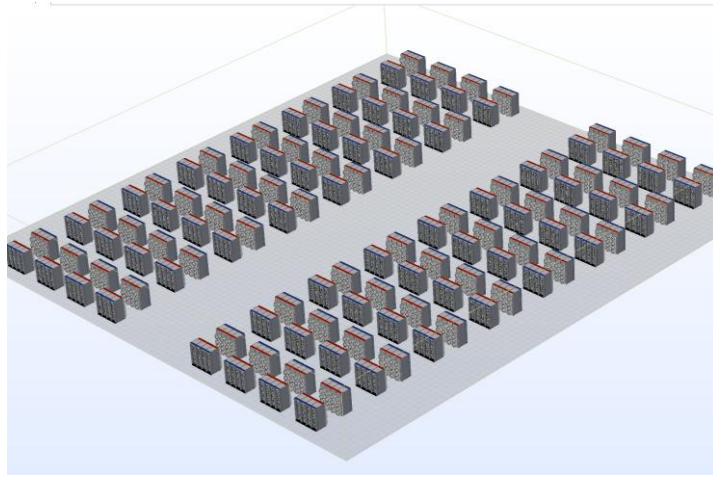
4.4.2 Data Hall Analysis

To understand the impact of RDHx at a large scale, an active RDHx is deployed in a typical datacenter layout without traditional cooling units. These include peripheral CRAC/CRAH units or a penthouse style design with supply shafts based on evaporative cooling. The data hall considered is approximately 160' x 175' x 22' (L x W x H) with a typical hot aisle cold aisle layout. Each row has sixteen rack positions installed in a set of four racks with RDHx. This was intentional so that cold air could circulate in between the racks. Hot aisle

containment was not considered, and the entire data hall is expected to operate as one big cold space (room neutral). To meet the maximum airflow provided by the RDHx fans, rack power was increased to 20 kW, requiring 3050 CFM per rack to dissipate heat. A typical hyper-scale data center layout consisting of over 500 racks was considered for analysis. Fig. 4-9 shows the data center layout with RDHx attached. Rack orientation is indicated by red and blue color indicating rear and front side of the racks respectively. Operating conditions (target RDHX air temperature of 30°C and coolant supply temperature of 26°C) similar to earlier analysis was considered.

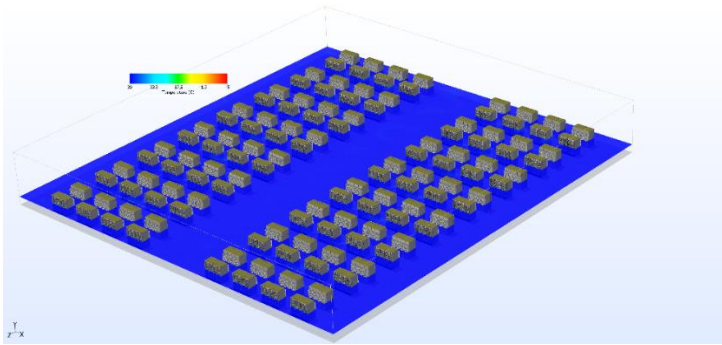


(a) Plan view

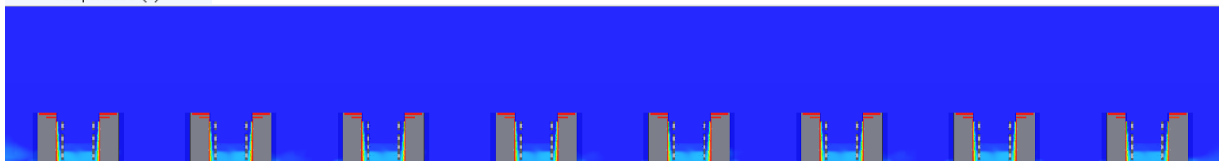
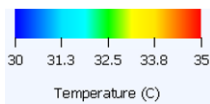


(b) 3-D View

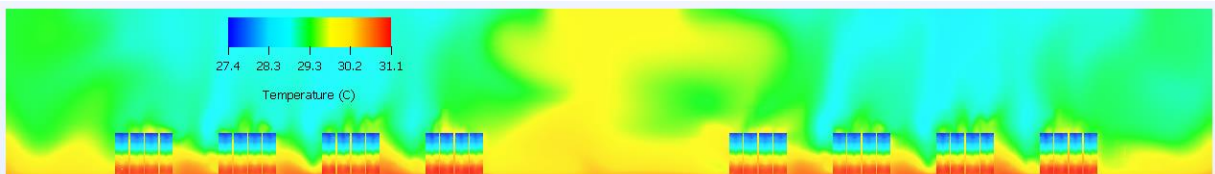
Figure 4-9 TYPICAL DATACENTER WITH RDHx INSTALLATION



(a) Temperature plane -4 ft above the floor (Legend from 26-42°C).



(b) A sectional temperature plane showing elevated temperature of around 1°C around the base of the rack. (Legend from 30-35°C).



(c) A sectional plane aligned with the exit face of RDHx showing spatial variation in temperature

Figure 4-10 CFD RESULTS - TEMPERATURE PLANES

From the results (Fig. 4-10), it can be seen that the RDHx can provide a room neutral solution thereby eliminating the need for any mechanical cooling units such as CRAH/C or penthouse unit with direct evaporative system. It will still require a chiller but given the operating condition of 26°C facility water, there is potential to lower chiller usage depending on the location of the data center where ambient conditions could be favorable. It is to be noted that the legend scale is intentionally changed to get a visual contrast in the results. For plots in figs. 4-10a and 4-10b, the lower limit is selected to be 30°C as that is the control temperature set to which all RDHx fans respond. For plot in fig. 4-10c, scale is lowered to 27.4°C as that is the minimum temperature of the air delivered out of RDHx. Fig 4-10c shows the spatial variation in temperature along the rear face of the RDHx where the processed air enters back into the room via fans. About 3-4°C variation in processed air is observed from top to bottom. This is due to the fact that the facility water connection is defined at the top. This could be further investigated by studying different coil designs; however, authors do believe that spatial variation in temperature is not a risk as the air gets mixed and average server inlet temperature of air is about 29.7°C. The datacenter is designed to meet ASHRAE Class A1 temperature limits. If the mean server inlet temperatures exceed 32°C the solution is considered failed.

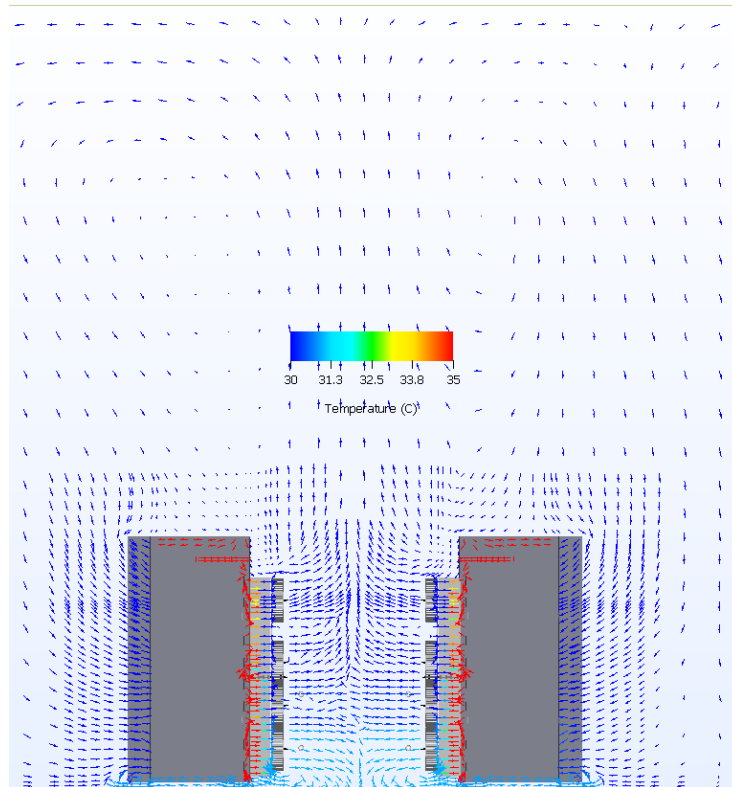


Figure 4-11 A SECTIONAL PLANE SHOWING THE VELOCITY PROFILE AND AIR MOVEMENT AROUND THE CABINET WITH RDHx

Fig. 4-11 shows the vector plot to study how the air moves within the data hall space. After the air exits RDHx, it is forced upward and eventually circulates back into the cold aisle. Considering the density of velocity vectors around the rack, most of the air turns within four feet from the top of the rack. This suggests that there is potential to further lower the height of the data center. In these scenarios, rows were assumed to be placed at a distance greater than what normally is practiced. In the layout analyzed, the hot aisle is six feet wide (normally varies between 4-5) while cold aisle is about twelve feet wide. There is further room for optimization to reduce the overall size of the data center without having any negative impact on RDHx performance.

4.5 Conclusion and Future Work

The impact of using RDHx without auxiliary mode of cooling such as CRAH/CRAC units in datacenters with low- or high-density racks is discussed in this paper. For passive

RDHx implementation although feasible, it should be carefully investigated for its compatibility with ITE. The server fans considered in this analysis assumed that it could overcome higher pressure drop. However, for the same rack density, it is not uncommon to use smaller or single rotor fans which are suitable for pressure drops less than one inch of water. It is expected that the RDHx once installed to be compatible with all future generations of hardware as well. For this reason, implementation of active RDHx may be beneficial. Although additional power needs to be provided for the fans, facility operators no longer require mechanical cooling units (other than a chiller) such as CRAC/H. It also showed the feasibility of RDHx for a high MW data center with rack power of 20kW. With further optimization, additional space savings could be realized which would further reduce the construction cost of the data center. Future studies include optimization and a complete cost benefit analysis of implementing RDHx for high-capacity data centers.

4.6 References

- [1] S. Singh, K. Nemati, V. Simon, A. Siddarth, M. Seymour, and D. Agonafer, "Sensitivity Analysis of a Calibrated Data Center Model to Minimize the Site Survey Effort," 2021 37th Semiconductor Thermal Measurement, Modeling & Management Symposium (SEMI-THERM), 2021, pp. 50-57.
- [2] V. S. Simon, A. Siddarth, and D. Agonafer, "Artificial Neural Network Based Prediction of Control Strategies for Multiple Air-Cooling Units in a Raised-floor Data Center," 2020 19th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), 2020, pp. 334-340, Doi: 10.1109/ITherm45881.2020.9190431.
- [3] Scaramella, J., 2008, "Next-Generation Power and Cooling for Blade Environments," IDC, Technical Report No. 215675.

- [4] IT Equipment Power Trends is authored by ASHRAE Technical Committee (TC) 9.9, Mission Critical Facilities, Data Centers, Technology Spaces and Electronic Equipment.
- [5] R. Schmidt, M. Iyengar, D. Porter, G. Weber, D. Graybill, and J. Steffes, "Open side car heat exchanger that removes entire server heat load without any added fan power," in Proc. 12th IEEE Intersoc. Conf. Thermal Thermomech. Phenomena Electron. Syst. (ITherm), Jun. 2010, pp. 1–6.
- [6] K. Nemati, H. A. Alissa, B. T. Murray, B. Sammakia, and M. Seymour, "Experimentally validated numerical model of a fully-enclosed hybrid cooled server cabinet," in Proc. ASME Int. Tech. Conf. Exhibit. Packag. Integr. Electron. Photon. Microsystem. Collocated, ASME 13th Int. Conf. Nanochannels, Microchannels, Minichannels, 2015, p. V001T09A041.
- [7] F. Douchet, D. Nortershauser, S. Le Masson, and P. Glouannec, "Experimental and numerical study of water-cooled datacom equipment," *Appl. Thermal Eng.*, vol. 84, pp. 350–359, Jun. 2015.
- [8] R. Schmidt and M. Iyengar, "Server rack rear door heat exchanger and the new ASHRAE recommended environmental guidelines," in Proc. ASME InterPACK Conf. Collocated ASME Summer Heat Transf. Conf. ASME 3rd Int. Conf. Energy Sustainability, 2009, pp. 851–862.
- [9] Patterson, M., Fenwick, D., "The State of Data Center Cooling-A review of current air and liquid cooling solutions." Intel white paper, 2008.
- [10] Specification for Open Rack V2, Rev. 12, July 2015.
<https://www.opencompute.org/documents/open-rack-v2-specification-rev12-pdf>

Chapter 5 Energy Analysis of Rear Door Heat Exchangers in Data Centers with Spatial Workload Distribution

Reprinted with permission © 2023 ASME

5.1 Abstract

Data centers have complex environments that undergo constant changes due to fluctuations in IT load, commissioning and decommissioning of IT equipment, heterogeneous rack architectures and varying environmental conditions. These dynamic factors often pose challenges in effectively provisioning cooling systems, resulting in higher energy consumption. To address this issue, it is crucial to consider data center thermal heterogeneity when allocating workloads and controlling cooling, as it can impact operational efficiency. Computational Fluid Dynamics (CFD) models are used to simulate data center heterogeneity and analyze the impact of two different cooling mechanisms on operational efficiency. This research focuses on comparing the cooling based on facility water for Rear Door Heat Exchanger (RDHx) and conventional Computer Room Air Conditioning (CRAH) systems in two different data center configurations. Efficiency is measured in terms of ΔT across facility water. Higher ΔT will result in efficient operation of chillers. The actual chiller efficiency is not calculated as it would depend on local ambient conditions in which the chiller is operated.

The first data center model represents a typical enterprise-level configuration where all servers and racks have homogeneous IT power. The second model represents a colocation facility where server/rack power configurations are randomly distributed. These models predict temperature variations at different locations based on IT workload and cooling parameters. Traditionally, CRAH configurations are selected based on total IT power consumption, rack power density, and required cooling capacity for the entire data center space. On the other hand, RDHx can be scaled based on individual rack power density, offering localized cooling

advantages. Multiple workload distribution scenarios were simulated for both CRAH and RDHx-based data center models. The results showed that RDHx provides a uniform thermal profile across the data center, irrespective of server/rack power density or workload distribution. This characteristic reduces the risk of over- or under-provisioning racks when using RDHx. Operational efficiency is compared in terms of difference in supply and return temperature of facility water for CRAH and RDHx units based on spatial heat dissipation and workload distribution. RDHx demonstrated excellent cooling capabilities while maintaining a higher ΔT , resulting in reduced cooling energy consumption, operational carbon footprint (?), and water usage.

5.2 Introduction

Data centers are centralized facilities that house networking and computing equipment for remote data storage, distribution, and processing. Data Centers are particularly important for cloud service providers for backing up huge amounts of user data and keeping critical applications up and running during nominal and peak demand periods.

Data centers are significant energy consumers, and their energy usage has been increasing rapidly in recent years. According to a report by the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE), data centers consumed approximately 2% of the electricity in the United States in 2020[1]. IT equipment, cooling systems, and power distribution infrastructure are the main contributors to energy usage in data centers. IT equipment, including servers, storage devices, and network equipment, is responsible for the largest portion of energy consumption in data centers. The US Department of Energy estimated that IT equipment consumes between 45-55% of the total energy used in data centers [2]. Cooling systems, such as air conditioners, chillers, and computer room air handlers (CRAHs), are vital for maintaining optimal temperatures and ensuring reliable IT

equipment operation. The same DOE report suggests that cooling systems account for approximately 25-40% of total energy consumption in data centers. Power distribution infrastructure, such as uninterruptible power supplies (UPS) and power distribution units (PDUs), completes the energy usage picture in data centers [2]

“Air-based cooling systems lose their effectiveness when rack densities exceed 20 kW, at which point, liquid cooling becomes the viable approach.”, Vertiv [3]

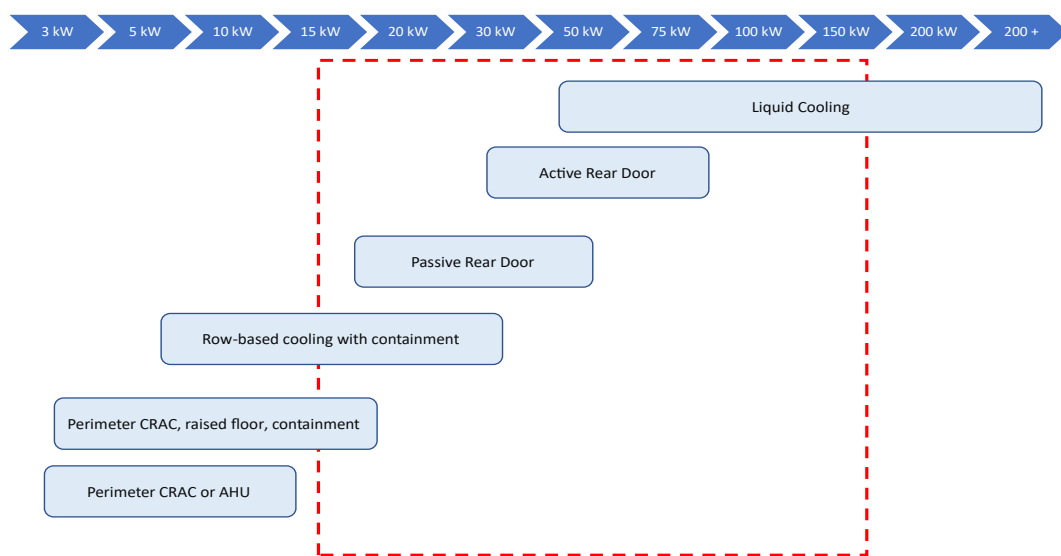


Figure 5-1 Air cooling limits: rack power density [3]

Rear Door Heat eXchangers (RDHx) have emerged as an effective solution for addressing the cooling challenges in data centers. RDHx is a type of liquid cooling system that is installed on the rear doors of server racks to remove heat directly from the IT equipment. This approach offers several advantages, including improved cooling efficiency, reduced energy consumption, and enhanced heat dissipation capabilities. Several studies have investigated the performance and benefits of RDHx in data centers. Conventional methods of air cooling have significant disadvantages in heat dissipation due to high energy consumption at the CRAH units, humidity excursions, recirculation into the cold aisle, extensive site survey effort etc. [4]. Even though there are optimization strategies for air cooling provisioning, there

are limitations due to increased demand in high power consuming devices [5]. The growing rate of world-wide servers is 2.5×10^6 annually based on the industry data, and this growth is expected to continue [6]. Also, as per ASHRAE, heat load per 42U rack shows that rack density is continuously increasing. This makes it difficult for conventional air-cooling methodologies to provision additional capacity [7]. Therefore, there are potential advancements in methods of air cooling; one such method is retrofitting existing data centers with RDHx. Simon et al. [8] studied the feasibility of deploying Rear door heat exchangers in high-capacity data centers, where it is to be noted that passive RDHx will offer significant advantage in cooling, provided, the server fans overcome the Hx pressure drop also in active mode the benefit is substantially more than the CRAH based cooling. Hybrid cooling systems are typically self-contained components like the side car, which is an air-liquid heat exchanger installed on the side of a server cabinet. IBM introduced this enclosed heat exchanger, which is designed to remove up to 35 kW [9]. A fully enclosed server cabinet, which commonly uses a V-shaped heat exchanger on the bottom of the cabinet, is another example of a hybrid cooling system [10]. The RDHx is a heat exchanger located at the back of the cabinet, where hot exhaust air leaves the servers. It was demonstrated in [11] that using an RDHx in various data center configurations can have several advantages, including 1) reducing the number of CRAH units required, 2) eliminate hot spots by removing heat closer to the source, and 3) allowing higher chilled water temperatures while still maintaining IT inlet temperatures within recommended ranges. RDHXs were used in tandem with room level air conditioning in a hybrid manner [12]. The CRAH fan speed was optimized using a control mechanism based on the heat load of the servers. They were able to save 47 kWh of energy, or around 6% of the total energy consumed by the CRAH.

OpEx (Operational Expenditure) analysis and TCO (Total Cost of Ownership) evaluation play crucial roles in understanding the financial aspects of data centers. OpEx

analysis involves assessing the ongoing operational expenses associated with running and maintaining a data center, while TCO considers the overall cost incurred for owning the equipment throughout its lifespan.

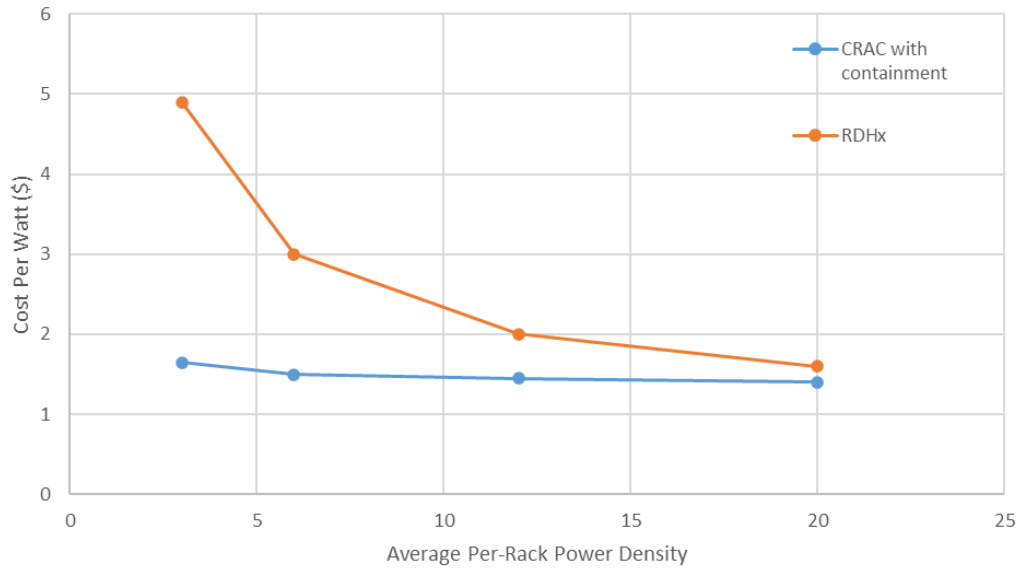


Figure 5-2 Cost of Equipment vs rack power density [13]

OpEx analysis encompasses various cost components, including energy consumption, cooling, maintenance, personnel, and equipment upgrades. It enables data center operators to identify areas of high expenditure and implement strategies for cost reduction and operational efficiency. Additionally, OpEx analysis provides insights into the financial implications of different operational decisions, such as equipment selection, cooling system optimization, and energy management strategies. Fig. 5-2 from a whitepaper published by Schneider Electric shows cost per watt consumed by various air-cooling methodologies based on the rack power density. RDHx is cost effective when the rack density increases, and it reduces further beyond 20kW as room-based cooling units become ineffective.

5.3 Computational Model

6SigmaRoom is a powerful computational fluid dynamics (CFD) software designed specifically for data centers. It offers numerous advantages. Firstly, it provides modeling ability

to precisely define control parameters and response curves for fans, pumps, control valves etc. By monitoring airflow, temperature, and pressure distribution within data centers, the CFD tool enables data center operators to optimize energy efficiency by identifying hotspots, temperature gradients, and areas of inefficiency.

The data center model encompasses a room size of 2,447 sq ft with a height of 14.76 ft and a total power capacity of 1.28 MW. The height of the raised floor is 2 ft. It accommodates a total of 8 rows with 10 racks each, in a hot-aisle containment configuration for the CRAH-based DC model. Each rack houses 10 4U servers. A compact model of 4U servers was populated in all the racks with maximum rack power of 16 kW. All the ITEs are provided with power and temperature rise specifications of 10°C ΔT . The CFD model includes 8 peripheral Computer Room Air Handling Units (CRAHs) with a net cooling capacity of 200 kW/CRAH. The data center layout is shown in Fig 5-3.

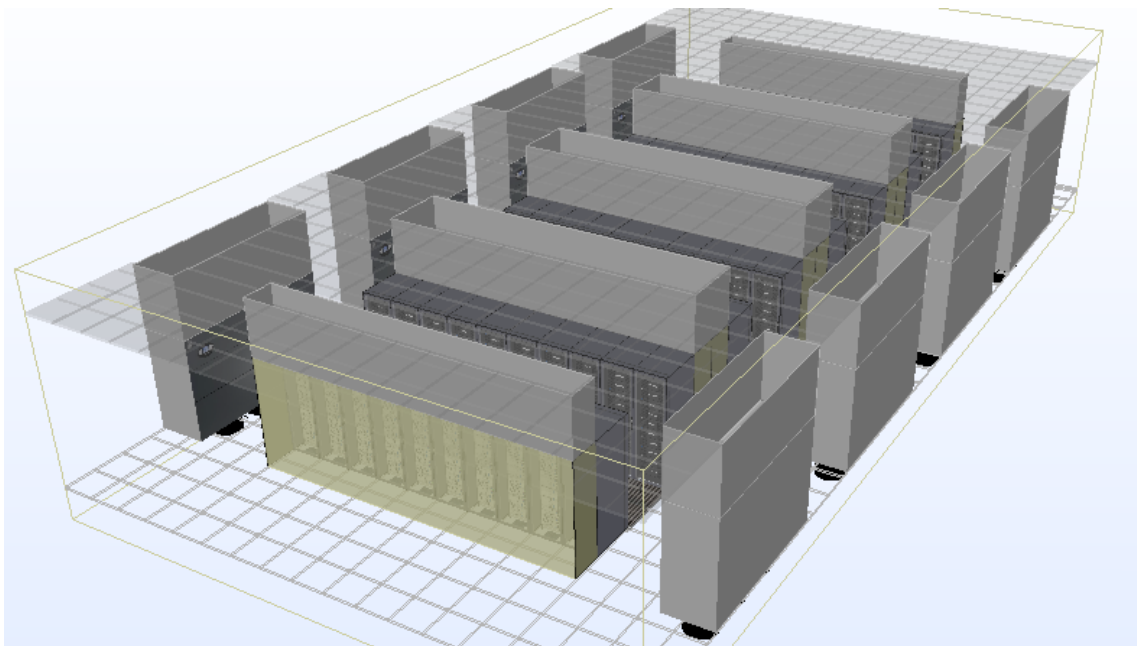


Figure 5-3 Raised floor data center with CRAH units.

The CRAH based DC model has hot aisle containment as it minimizes recirculation which is typical in most of the hyperscale data centers. The CRAH units supply cold air through

the underfloor plenum to the cold aisle via slotted floor grills in a raised floor configuration. After cooling the IT Equipment, the hot air is contained and ducted to the ceiling where the return air mixes and directed to the return side of the CRAH unit through a ceiling duct as shown in figure 5-3. Another CFD model has the same data center IT layout and specifications with RDHx attached to all the racks, and no peripheral CRAH units, raised floor or the false ceiling as shown in Fig. 5-4.

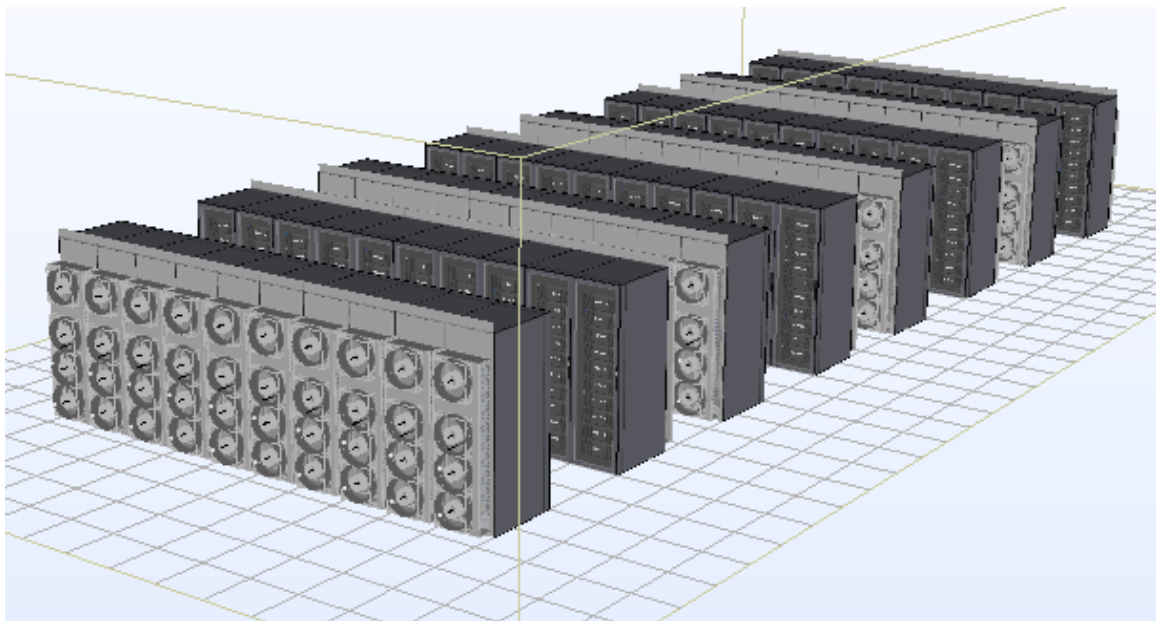


Figure 5-4 Data center with RDHx

5.4 Rear Door Heat Exchanger

Rear door heat exchangers are typically built to fit at the rear end of the rack, which is 2 ft wide. A commercially available rear door heat exchanger (68 x 21.3 x 7 in) was chosen to fit at the back of the ORV2 rack (84.8 x 23.6 x 42 in) from the Open Compute Project (OCP). It is typical for an active RDHx to have a specification of 600W of total fan power to move 3,000 cfm of air across the assembly. For our use case, based on the required flow rate across the rack, fans usually run at approximately maximum speed of 80% of the rated speed which results in a much lower power consumption when calculated using fan laws. A typical 5%

leakage is assumed to account for gaps in between the servers, mounting rails etc. In idle scenario, the server inlet air temperature would be the same as exit air temperature of the RDHx.

The RDHx typically consists of a heat exchanger unit with fans mounted on the rear side, accessible from the aisle. A control valve modulates the flow of coolant (facility water) through the coil typically depending on the leave air temperature setpoint. The position of the valves depends on the data center layout and the placement of chilled water pipes. They are either located at the top (for overhead piping) or at the bottom (raised floor configuration). The RDHx is connected to the main loop via an isolation valve, which allows for disconnecting the rack from the chilled water loop without disrupting facility operations at large. Fig. 5-5 shows the schematic of a traditional rack and a rack with RDHx.

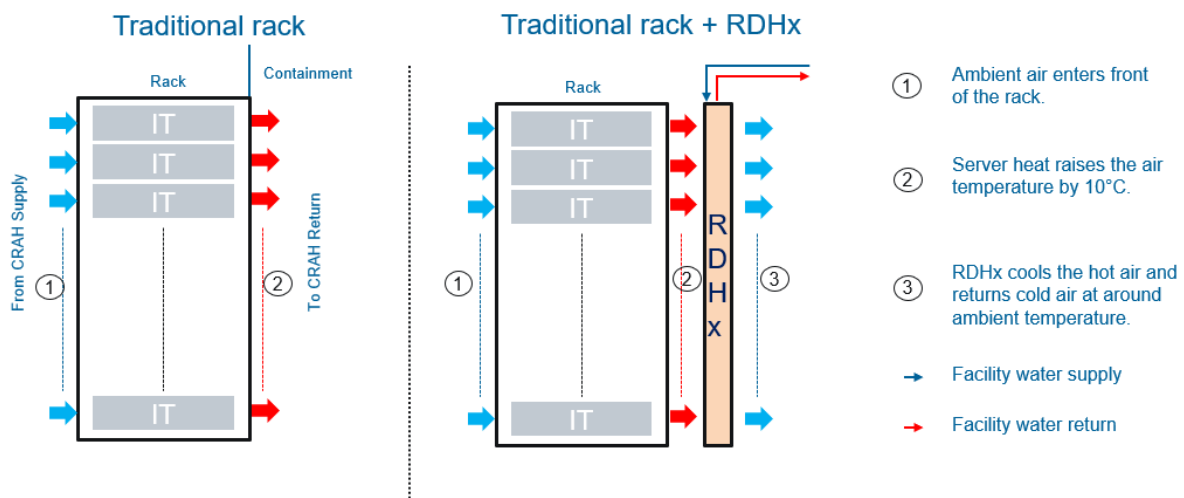


Figure 5-5 Traditional rack vs RDHx Rack Schematic (Side View)

The coolant flow through each of the heat exchangers is modulated by a flow control valve attached to the RDHx unit, based on the downstream air temperature at the heat exchanger. Depending on the design of the coil, chilled water inlet direction, and variation in face velocity at the inlet of the RDHx, air may be cooled non-uniformly. To counter this, axial fans are placed at the rear of the heat exchanger unit to uniformly pull air from the cabinet.

These fans can be controlled based on sensors measuring the inlet temperature at the rack or based on approach air temperature of the heat exchanger.

Although the addition of fans consumes more power, it helps to maintain a pressure-neutral data center environment for the server fans. Apart from distributing the flow uniformly over the coil, it also helps overcome the air-side pressure drop of the heat exchanger. The boundary conditions for the CRAH and RDHx units are shown in Table 5-1 and 5-2, respectively.

Table 5-1 CRAH unit specification and BC

CRAH Specification and Boundary conditions	
Supply CHW Temp	22 °C
Supply Air Temp	29 °C
Max. air flowrate	24,000 cfm
Max. CHW flowrate/CRAH	108 gpm
No. of EC fans	3
Fan power	3.7 kW/fan

Table 5-2 CRAH unit specification and BC

RDHx Specification and Boundary conditions	
Coolant flow direction	Top to Bottom
Nominal Cooling Capacity	35 kW
Reference coolant flowrate (max.)	23 gpm
Hx Effectiveness	0.8
Reference Air flowrate (max.)	3,000 cfm
Coolant Temperature	22°C
Coolant flowrate	Controlled: Hx exit air temperature (30°C)
No. of fans	4
Fan power (Full speed)	500 W

The CRAH unit air flow rate control was set to pressure-based condition so that the cold aisle is slightly over pressurized compared to hot aisle. It is a common practice to maintain a minimum ΔP of 0.05 in.H₂O across the racks to avoid hot air recirculation. CRAH supply air temperature control was set to 29°C based on the feedback from the sensors at the cold aisle.

The coolant (chilled water) temperature was set to 22°C as mentioned in Table 5-1. CRAH units have individual flow control valves to modulate the coolant flow rate based on the return air temperature. The rest of the coolant bypasses the cooling coil and mixes with the return water from the coil. Coolant flowrate was set to modulate (using the valve) based on the cooling energy demand and maintain the set supply air temperature.

5.5 Uniform workload distribution

In this case, we considered a typical scenario of uniform workload distribution with similar-performance servers populated throughout the data center. Two different CFD models were created: one for CRAH-based cooling units and another for RDHx-based cooling units. All IT equipments are set to maintain a ΔT of 10-15°C across the racks, based on the rack inlet temperature. The ITE specifications are given in Table 5-3.

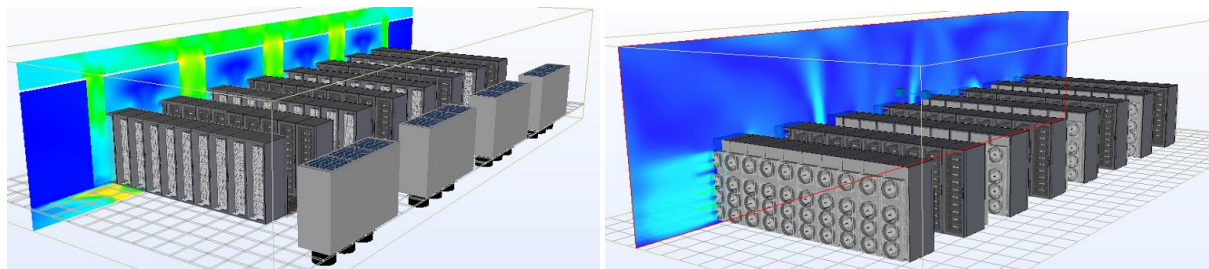
All the IT Equipment power utilization was set to a certain % utilization uniformly throughout the data center space and simulations were performed to observe the ΔT across the cooling coils in the CRAH unit as well as the CRAH unit with the bypass. This would let us know the ΔT across the chiller.

Table 5-3 ITE Specifications

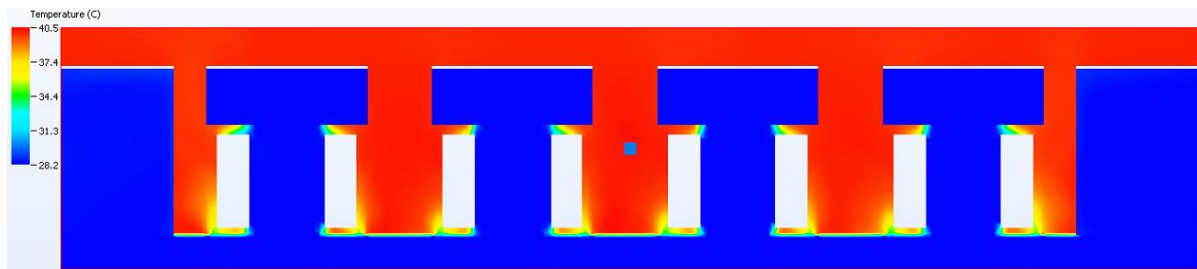
ITE Utilization	Rack power (W)	Rack air flow required (cfm)
30%	4800	696
65%	10400	1508
100%	16000	2320

Fig.5- 6 (a) shows the temperature result plane for both configurations. The figure on the left shows the cold and hot aisle with raised floor and hot air return plenum. Based on the temperature contours for RDHx (right figure), it can be seen that RDHx provides a room neutral solution. The difference between the two set-ups can be prominently noted in Fig. 5-6 (b) and (c). It shows the temperature profile across the racks for the baseline case The temperature

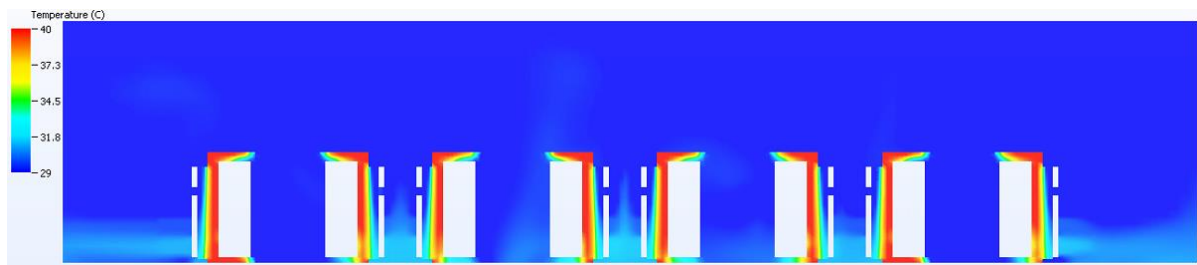
profile looks similar in all three cases mentioned above since the ΔT across the racks are similar.



(a) Result plane location for CRAH based DC (left) and RDHx based DC (right)



(b) CRAH-based DC Temperature contour



(c) RDHx-based DC Temperature contour

Figure 5-6 Temperature Contours for uniform workload distribution

5.6 Spatial workload distribution

In this case, we considered a typical scenario of workload distribution with similar-performance servers populated throughout the data center. Two different CFD models were created as mentioned in the previous case. These models were used to compare and observe the effects of spatial workload distribution on thermal profile of the room and operating energy consumption.

Out of 80 racks in the room, 20% of the fully populated racks were randomly selected and set at maximum utilization (100%) with a capacity of 16 kW per rack, while the remaining set of racks were set at idle power (30% of the total rack power). Fig. 5-7 shows the highlighted

racks at 100% utilization with CRAH units. Similarly, another CFD model with the same racks at 100% utilization was modeled with RDHx and no peripheral CRAH units. Simulations were run based on the previously mentioned boundary conditions.

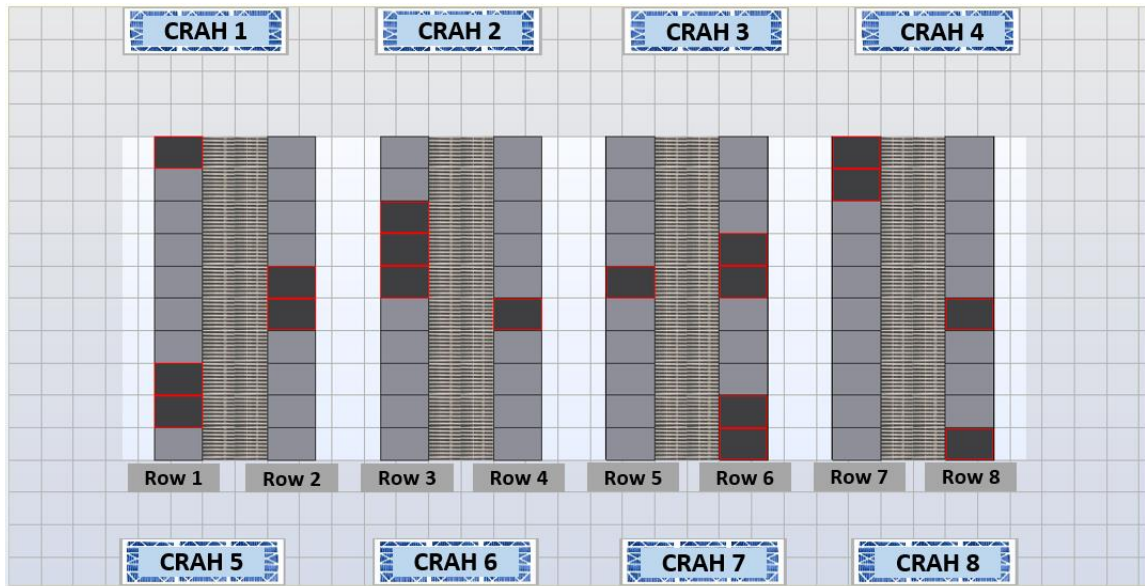
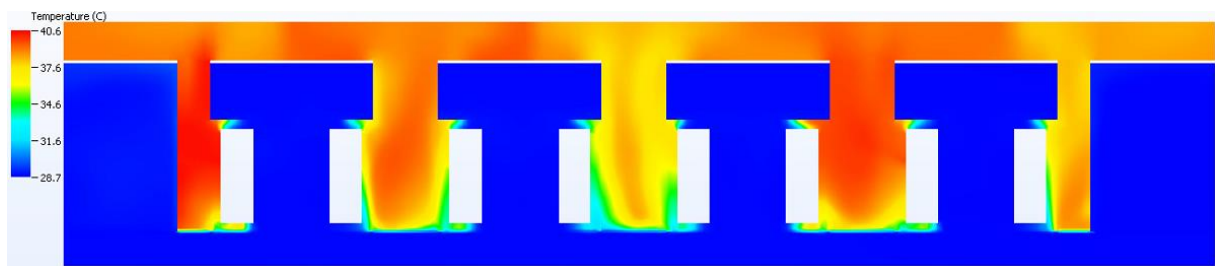


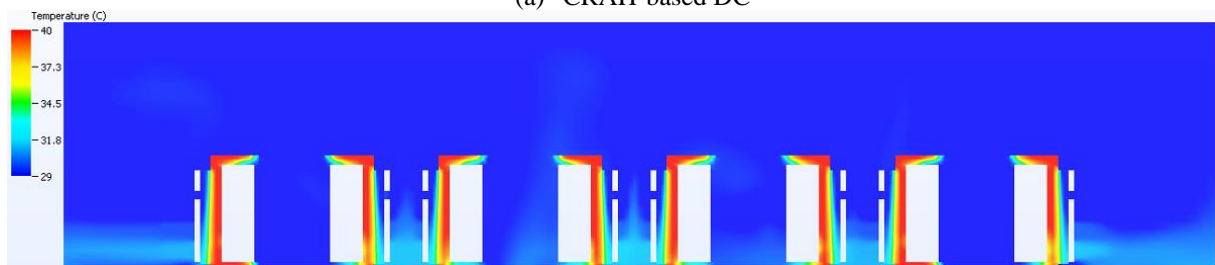
Figure 5-7 Randomly selected racks to mimic spatial workload distribution in data center

The result plane was taken at the same location as shown in previous case (Fig. 5-6 (a)). It was observed that RDHx delivered a uniform temperature distribution throughout the data center, irrespective of the rack utilization. The CRAH-based data center model had varying CRAH return air temperatures since the hot air from various heat sources at varying velocities mixes in the ceiling before returning to the cooling coils. This might be because of the uneven air flowrate provided by each CRAH unit that works based on sensor-based feedback while the chilled water flowrate remains constant at the facility side. Therefore, chilled water bypasses the cooling coils when there is lower energy demand leading to lower ΔT across the cooling coil and the CRAH unit while maintaining the set supply air temperature. It is a known fact that chiller units work efficiently when the ΔT across the chilled water is higher. When the difference between supply and return chilled water temperature goes below the design value (in this case equivalent to maximum data center capacity of 1.2MW) it results in cooling facility to operate at lower efficiency for majority of the time since data centers rarely operate at design

capacity. For most of the data centers, maximum capacity utilized is between 65-70% of the design capacity. During off peak hours, it can lower to 30% of the design capacity. Low ΔT results in chiller consuming more energy resulting in higher Capital Expenditure (CapEx) cost. CRAH based units are observed to have much lower ΔT across the cooling coils compared to rear door heat exchangers. Under low load conditions, the coils in CRAH units can be considered “oversized” which ultimately results in lower ΔT across facility water. In contrast, the localized cooling control provided by the RDHx results in higher ΔT . It also provides uniform temperature distribution across the data center keeping the inlet temperature of all the racks within a range of 1°C difference. Hotspots caused by high heat rejection and recirculation at the random high utilization racks were eliminated using RDHx.

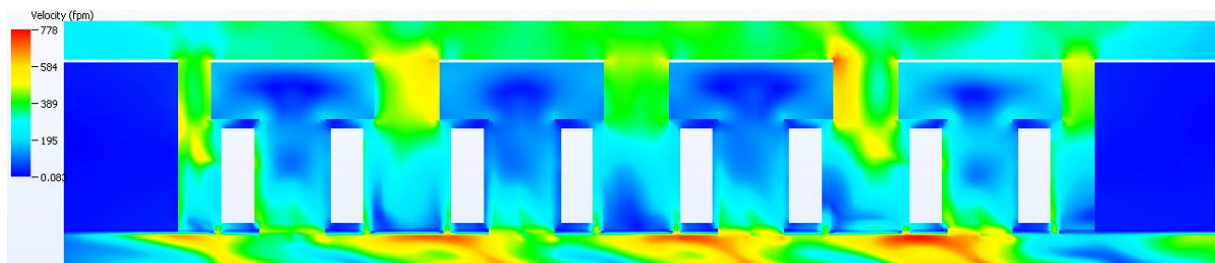


(a) CRAH-based DC

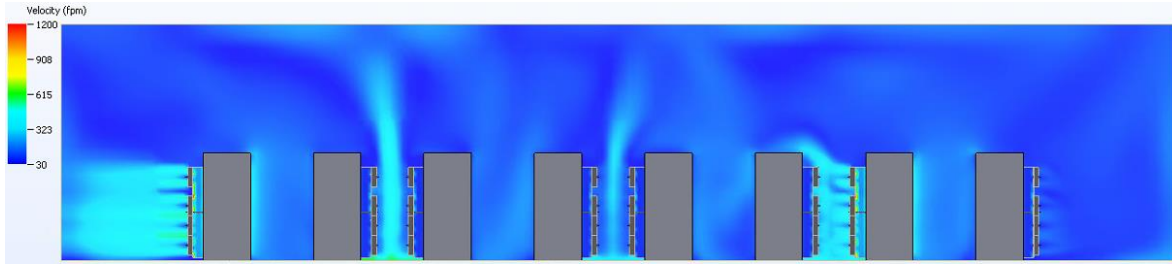


(b) RDHx-based DC

Figure 5-8 Temperature contours for spatial workload distribution



(a) CRAH-based DC



(b) RDHx-based DC

Figure 5-9 Velocity contours for spatial workload distribution

Based on the data from table 5-4, the total energy consumed by the RDHx would be lower compared to traditional style of cooling due to several reasons. The chilled water supply temperature for the RDHx is higher compared to the CRAH units, resulting in significant energy savings on the primary side. Additionally, RDHx has rack level flow control device which modulates based on the downstream temperature of the air. This level of sophistication provides localized control of coolant flow rate through the RDHx. The difference in pumping power when all the servers were at idle load compared to Case A was negligible relative to the power consumed by the CRAH units that were provisioned by chilled water pumps from the facility. Future work would include calculating the annualized energy consumption for a specific location.

Table 5-4 CRAH and RDHx-based DC chilled water results

% IT Utilization		CRAH coil ΔT ($^{\circ}C$)	Net Facility water ΔT ($^{\circ}C$)	Total CRAH coolant flowrate (gpm)	RDHx water ΔT ($^{\circ}C$)	RDHx flow rate (gpm)
Uniform workload distribution	30%	5.6	1.8	268.8	8.1	2.0
	65%	5.9	3.4	536.8	8.4	4.7
	100%	7.2	5.7	844.7	8.9	7.5
Spatial workload distribution	18 racks at 100% 62 racks at 30%	5.76	2.84	385.8	8.5	265.2 (Total coolant flow rate)

Again, similar results in terms of temperature distribution across the data center were observed. RDHx offered an even temperature distribution, and the average room temperature

was maintained at or 30°C with an average difference of 1°C tolerance at the rack inlet temperature.

5.7 Conclusion

In this paper, two data center configurations were analyzed using CFD to estimate the operational efficiency of data centers under different scenarios. ΔT between facility return and supply water temperature was compared for a legacy CRAH based cooling system and RDHx based cooling system. RDHx irrespective of operating load conditions can provide consistent higher ΔT resulting in higher operational efficiency. When the ΔT on the facility side is lower than the design, chillers operate less efficiently for more than the 90 percent of the year since the peak load conditions are primarily observed during hot weather conditions (summer). Thus, it would result in chillers consuming more energy or less efficient.

In addition, for typical raised floor-based data centers with peripheral CRAH/CRAC units, it is difficult to manage airflow. As a result, CRAH units will oversupply cold air by about 10-20% of what is required by IT racks. In contrast, RDHx can provide localized thereby modulating airflow and coolant flow rates as per rack demand. In future, authors plan to calculate the annualized energy consumption for chillers based on both technologies for a particular location.

5.8 References

- [1] ASHRAE, 2018. IT Equipment Power Trends, 3rd Edition. Atlanta: ASHRAE
- [2] Shehabi, A., Smith, S.J., Horner, N., Azevedo, I., Brown, R., Koomey, J., Masanet, E., Sartor, D., Herrlin, M., Lintner, W. 2016. United States Data Center Energy Usage Report. Lawrence Berkeley National Laboratory, Berkeley, California. LBNL-1005775

- [3] Understanding Liquid Cooling Options and Infrastructure Requirements for Your Data Center (vertiv.com)
- [4] S. Singh, K. Nemati, V. Simon, A. Siddarth, M. Seymour, and D. Agonafer, "Sensitivity Analysis of a Calibrated Data Center Model to Minimize the Site Survey Effort," 2021 37th Semiconductor Thermal Measurement, Modeling & Management Symposium (SEMI-THERM), 2021, pp. 50-57.
- [5] V. S. Simon, A. Siddarth, and D. Agonafer, "Artificial Neural Network Based Prediction of Control Strategies for Multiple Air-Cooling Units in a Raised-floor Data Center," 2020 19th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), 2020, pp. 334-340, Doi: 10.1109/ITherm45881.2020.9190431.
- [6] Scaramella, J., 2008, "Next-Generation Power and Cooling for Blade Environments," IDC, Technical Report No. 215675.
- [7] IT Equipment Power Trends is authored by ASHRAE Technical Committee (TC) 9.9, Mission Critical Facilities, Data Centers, Technology Spaces and Electronic Equipment.
- [8] Shalom Simon, V., Modi, H., Sivaraju, K. B., Bansode, P., Saini, S., Shahi, P., ... & Agonafer, D. (2022, October). Feasibility Study of Rear Door Heat Exchanger for a High Capacity Data Center. In International Electronic Packaging Technical Conference and Exhibition (Vol. 86557, p. V001T01A018). American Society of Mechanical Engineers.
- [9] R. Schmidt, M. Iyengar, D. Porter, G. Weber, D. Graybill, and J. Steffes, "Open side car heat exchanger that removes entire server heat load without any added fan power," in Proc. 12th IEEE Intersoc. Conf. Thermal Thermomech. Phenomena Electron. Syst. (ITherm), Jun. 2010, pp. 1–6.

- [10] K. Nemati, H. A. Alissa, B. T. Murray, B. Sammakia, and M. Seymour, “Experimentally validated numerical model of a fully-enclosed hybrid cooled server cabinet,” in Proc. ASME Int. Tech. Conf. Exhibit. Packag. Integr. Electron. Photon. Microsystem. Collocated, ASME 13th Int. Conf. Nanochannels, Microchannels, Minichannels, 2015, p. V001T09A041.
- [11] F. Douchet, D. Nortershauser, S. Le Masson, and P. Glouannec, “Experimental and numerical study of water-cooled datacom equipment,” *Appl. Thermal Eng.*, vol. 84, pp. 350–359, Jun. 2015.
- [12] R. Schmidt and M. Iyengar, “Server rack rear door heat exchanger and the new ASHRAE recommended environmental guidelines,” in Proc. ASME InterPACK Conf. Collocated ASME Summer Heat Transf. Conf. ASME 3rd Int. Conf. Energy Sustainability, 2009, pp. 851–862.
- [13] “Choosing Between Room, Row, and Rack-based Cooling for Data Centers”, White paper 130, Schneider Electric

Chapter 6 Experimental and CFD Analysis of Heat Capture Ratio in a Hybrid Cooled Server

Reprinted with permission © 2022 ASME

6.1 Abstract

Rising power densities at the server level due to increasing performance demands are being met by using efficient thermal management methods such as direct-to-chip liquid cooling. The use of cold plates that are directly installed yields a lower thermal resistance path from the chip to the ambient. In a hybrid cooled server arrangement, high-heat-generating components are cooled with water or a water-based fluid, while the rest of the components are cooled with air using server-level fans. It is imperative to characterize the heat capture ratio for various server boundary conditions to ascertain the best possible liquid and airflow rates and temperatures. These parameters serve as inputs in defining the Total Cost of Ownership (TCO). The present investigation numerically evaluates the heat capture ratio in a hybrid cooled server for peak server load and varying inlet temperature for air and liquid. The CFD model of a Cisco Series C220 server with direct-to-chip liquid-cooled CPUs was developed. The cold plate for the CPU was experimentally characterized for pressure drop and thermal resistance characteristics and a black-box model was used for CFD simulations using 25% propylene glycol as the coolant. The heat capture ratio value was obtained under the varied temperature and flow rate boundary conditions of air and liquid. Based on the heat capture ratio values obtained, optimum values of inlet temperatures and flow rates are recommended for air and liquid for the server being investigated.

6.2 Introduction

Conventional methods of air cooling have significant disadvantages in heat dissipation due to lower heat transfer coefficient, high energy consumption at the CRAC/CRAH units, humidity excursions, recirculation into the cold aisle, etc. Even though there are optimization strategies for air cooling provisioning, there are limitations due to increased demand for high power-consuming devices [1-3]. Direct on-chip liquid cooling overcomes most of the issues compared to air cooling at the system level. The combination of both air and liquid cooling at the server level must be studied critically since there are multiple electronic chips in a server other than high heat-dissipating electronic devices such as CPUs and GPUs. These low-power consuming devices usually have a small form factor as well as low heat flux and therefore it is impractical to attach cold plates for liquid cooling. Liquid-based cooling technologies offer higher heat transfer coefficients and are being increasingly used to dissipate high power densities of the new generation central processing units (CPUs) and graphics processing (GPUs) [4]. Direct on-chip Liquid cooling is usually implemented only on CPUs, GPUs, and sometimes on high-power-consuming DIMMs. Immersion cooling is one such liquid cooling method to address the concerns of rising powers in chipsets used in high-performance servers. However, the cost of maintaining and serviceability is a major concern when deploying at a large scale. Hybrid cooling addresses most of the issues in terms of cooling high-performance servers.



Figure 6-1 Cisco M220 series M3 server

6.3 Experimental Setup and Methodology

Typically, water-based coolants are used in cold plates since the heat-carrying capacity is much higher. Since contamination is due to waterborne bacteria, water is mixed with a proportion of propylene glycol or ethylene glycol. Even though the heat capacity of PG and EG is slightly lower, it minimizes the risk of contamination over a long period of usage. In our study, we considered using 25% propylene glycol in water. Liquid cooling using cold plates is modeled in CFD as a black box model using thermal resistance and flow resistance values from experimental data and validated. The heat capture ratio is defined as the ratio of heat carried by the coolant using cold plates to the total power consumption of the server. This is an important metric to calculate the availability of coolant as well as airflow. It is important to analyze the trade-off in energy consumption when both air and liquid cooling systems work efficiently. The purpose of this study is to quantify and calculate the thermal and flow

parameters of the hybrid cooled server under maximum utilization and at extreme inlet temperatures of both the heat transfer medium.

A commercially available CFD tool, Future facilities 6SigmaET was used to model the spread core Cisco series M220 M3 server using an inbuilt black-box model of cold plates for the CPU while the rest of the system was air-cooled. Two mediums of fluids were defined in the model to simulate the hybrid environment. The baseline version of the model was solved for only air cooling with heat sinks for CPUs. This model was compared with experimental data and validated. The heat sink was retrofitted with commercially available cold plates from Asetec for direct on-chip liquid cooling.

Cisco server M220 series M3 (Fig. 1) has two CPUs each with a TDP of 130 Watts (W) supporting 16 DIMMs (DDR3) each consuming 4 W. Platform Controller Hub (PCH) is a chip that controls the I/O and peripherals consumes 7 W. It is attached to a pin fin heat sink since the heat flux is high for a component with such a small form factor. The server supports 8 hard drives consuming 4 W each. The above-mentioned heat-dissipating electronic components are modeled in 6SigmaET with exact dimensions using the built-in modeling features in the computational software.

Table 6-1 Server specification

Components	Quantity	Values
CPU	2	115 W
DIMMs	16	2.5 W
Hard Disk	8	1.5 W
Platform Controller Hub (PCH)	1	5 W
Fans	5	15 CFM each (max.)

We noted the power consumed by DIMMs, Hard disk, and PCH as 2.5 W, 1.5 W and 5 W each respectively based on signals from internal sensors monitored using IPMI tool at peak power input. The baseline model has heat sinks attached to the CPUs for air cooling. The server

is tested for various air inlet temperatures and flow rates. The data collected from the CFD serves as a baseline for comparing with the server simulated with cold plates attached to the CPUs instead of heat sinks.

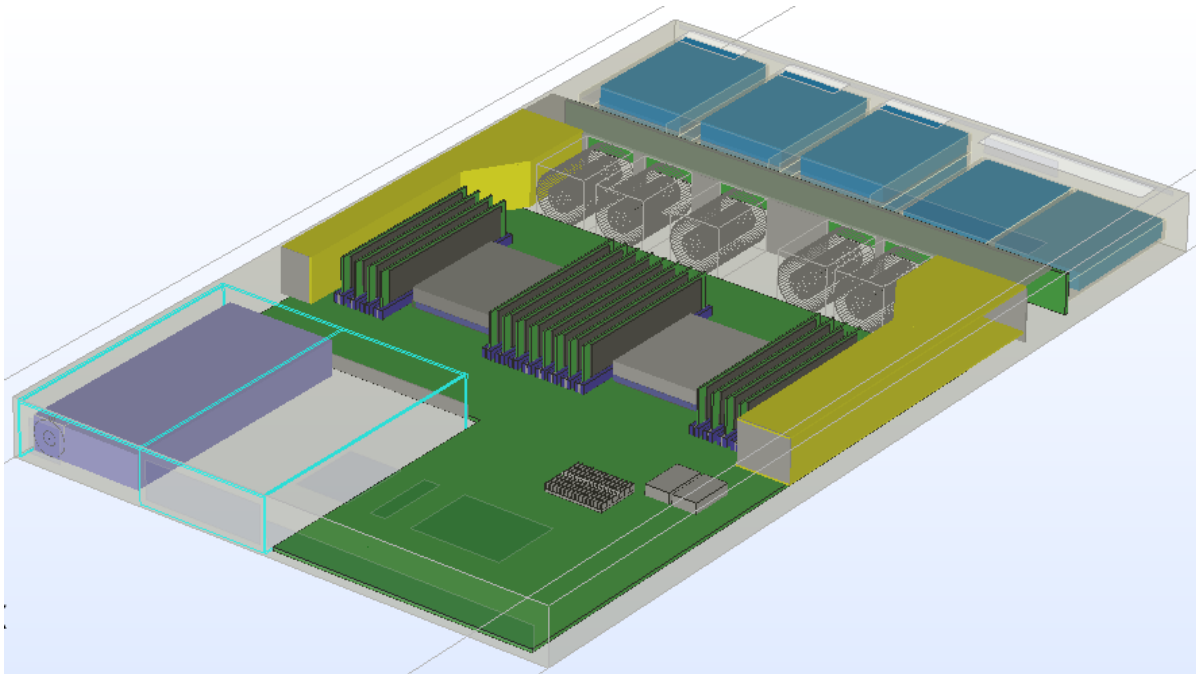


Figure 6-2 CFD model of Cisco M220 series M3 server

Fig 6-2 shows the CFD model of the Cisco M220 series M3 server containing all the targeted components that are critical to measuring thermal characteristics.

Table 6-2 CFD model validation using experimental result.

		Experimental	CFD
		Air & coolant inlet temp 25°C	
Coolant flowrate	Air flowrate (CFM)	80	80
0.4 lpm	CPU1 exit Temp (°C)	28.37	29.1
	CPU2 exit Temp(°C)	30.77	31.4
	Avg. air exit Temp (°C)	27.6	27.87
0.6 lpm	CPU1 exit Temp (°C)	28.16	27.8
	CPU2 exit Temp (°C)	29.84	30.5
	Avg. air exit Temp (°C)	28.4	27.81

The hybrid CFD model was compared with experimental results to validate the CFD model as shown in Table 6-2 so that further investigations for higher power dissipation can be simulated. The model was validated with $\pm 1^\circ\text{C}$ since there are experimental measurement accuracy errors. The model was validated for two different flowrates by comparing all the air and coolant temperature values at the exit.

To model cold plates in CFD, we need to calculate the thermal resistance and flow resistance of the cold plates at various inlet temperatures and flow rates. Thermal resistance across the cold plate is measured using the following equation for various flow rates and coolant inlet temperature in the system (Fig. 6-3). Since specific heat is a function of temperature, coolant inlet temperature is varied to calculate R_{th} .

$$\text{Thermal Resistance } (R_{th}) = \frac{(\text{CPU surface temperature} - \text{Coolant Inlet temperature})}{\text{CPU Power}}$$

Flow resistance curve (Fig. 6-5) is calculated using the measured values of pressure using pressure transducers placed at the inlet and outlet of the cold plate as shown in Fig 6-4. Cold plates in the server were modeled in a 1D flow network. Cold plates are in series connection in this particular server where the coolant enters the first cold plate attached to CPU 1 and the coolant exiting the first cold plate is supplied to the second cold plate attached to CPU 2. For simplicity, the cold plate attached to CPU 1 is referred to as CP 1 and the cold plate attached to CPU 2 is referred to as CP 2. The coolant entering CP 1 picks the heat from CPU 1 and becomes warm then it flows through CP 2 and picks heat from CPU 2 and returns to the reservoir. The 1D flow network is modeled as shown in Fig 6-6.

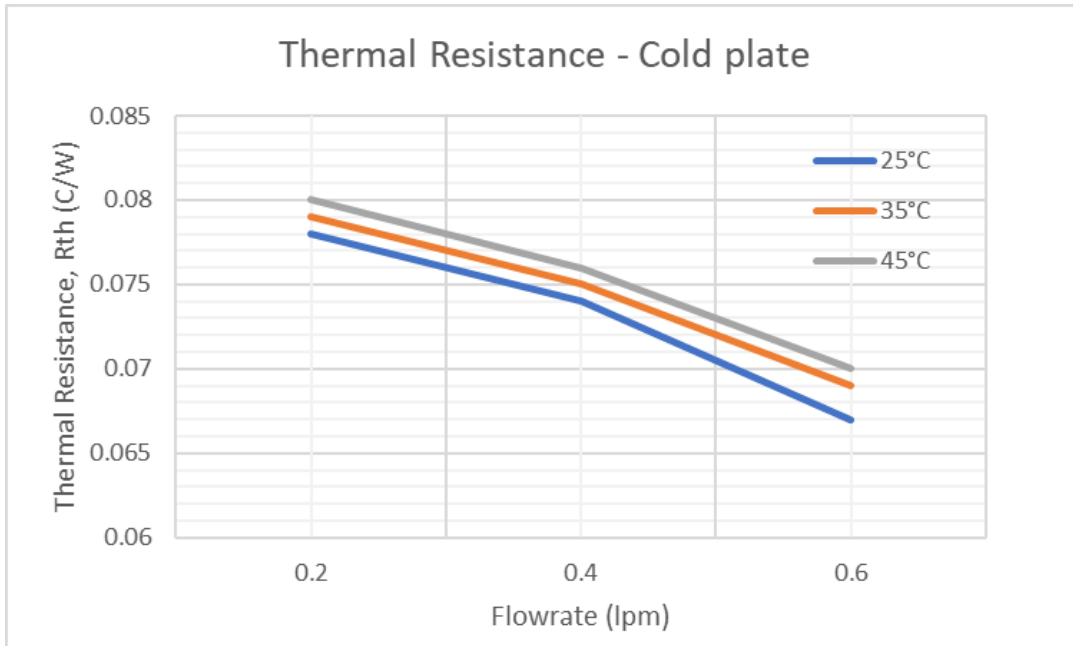


Figure 6-3 Thermal Resistance (Rth) Curve



Figure 6-4 Flow resistance measurement setup

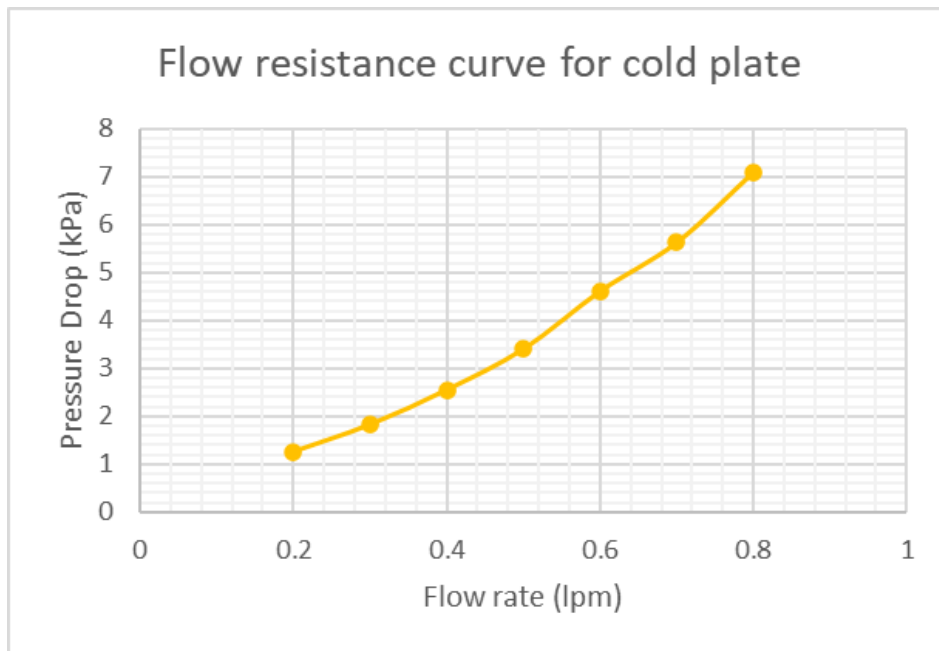


Figure 6-5 Flow resistance (PQ) curve

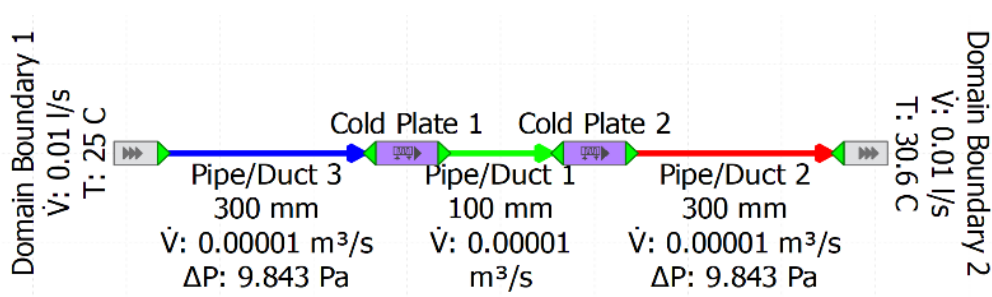


Figure 6-6 1-D Flow network model

Domain boundary 1 is the inlet condition of the coolant for CP 1 and Domain boundary 2 is the outlet condition of the coolant. We keep the Domain Boundary 2 at 0 Pa simulating it as if the return coolant is supplied to the reservoir. To manipulate the inlet conditions of the coolant the thermal and flow parameters are changed at Domain boundary 1. Domain boundary 2 is provided with the same flow rate as the inlet and there is no control over the thermal boundary since it is dependent on the thermal resistance of the cold plates.

The 3D model of the server consists of cold plates and a heat sink modeled with exact dimensions and material properties given accordingly. Platform Control Hub (PCH) is attached with a small heat sink to dissipate the heat. Since the heat flux from the PCH is high, it needs

an appropriate heat sink. The cold plates attached to the CPUs are made of copper and the heat sink attached to the PCH was made of aluminum.

Table 6-3 Boundary Conditions

Parameters	Values
Air Inlet Temperature	25°C, 35°C, 45°C
Air Flowrate	80 CFM
Coolant Inlet Temperature	25°C, 35°C, 45°C
Coolant Flowrate	0.2 l/m, 0.4 l/m, 0.6 l/m

Table 6-3 shows the input parameter space and constraints of the parameters that were simulated. The inlet temperature of the air varies from 25 C to 45 C and the air flowrate is fixed at 80 CFM. The minimum air flowrate that is provided by all the 5 fans in the server is based on the minimum rated voltage (5 V) to run each fan. The maximum flowrate that all fans in parallel can supply is 80 CFM and is calculated based on the maximum rated voltage (12 V). To compare the experimental results with CFD, we need to match the exact conditions of the experiments with the CFD model. Therefore, power supplied to the fans were not considered when simulating in CFD since the experimental setup was designed in such a way that the fans are powered separately using a power supply. Also, all perforations in the CFD are modeled exactly as in actual server.

The focus of this study is to observe the heat transfer phenomenon due to the interaction of the cold plate material with the air passing through the server. Air entering the server passes through various electronic components such as hard drives, DIMMs and PCH also comes in contact with the cold plate. From experimental studies, it is observed that when the air is colder than the inlet temperature of the coolant, the cold plate rejects heat to the air thereby reducing the heat captured by the coolant and vice versa.

6.4 Results and Discussion

The heat capture ratio is calculated using the following equation. This parameter is useful to compare the amount of heat captured by the coolant in a hybrid cooled server with cold plates attached to the CPUs.

$$\text{Heat Capture Ratio (HCR)} = \frac{\dot{m} * C_p * (T_{e, cp2} - T_{i, cp1}) (W)}{\text{Total power given to server (W)}}$$

The HCR equation shows the heat captured by the coolant in the numerator as \dot{m} is the total mass flowrate of the coolant, C_p is the specific heat of the coolant, $T_{e, cp2}$ is the exit temperature at cold plate 2 and $T_{i, cp1}$ is the inlet temperature at the cold plate 1. Table 6-4 shows the heat capture ratio of all the simulations done using the CFD model. We could observe that heat captured by the coolant increases for a constant coolant inlet temperature and increasing air inlet temperature in the system. When the air inlet temperature is more than the coolant inlet temperature, the heat is transferred to the cold plates/tubes that supply the coolant, thereby, increasing the temperature of the coolant. This phenomenon leads to misconception of understanding heat captured by the coolant in a hybrid cooling environment.

Table 6-4 Heat Capture Ratio (HCR)

		Liquid Heat Capture Ratio			
		Air Temp \ Coolant Temp	25°C	35°C	45°C
0.2 lpm	25°C	0.7605	0.7177	0.6702	
	35°C	0.808	0.7652	0.7177	
	45°C	0.8508	0.808	0.7652	
		Air Temp \ Coolant Temp	25°C	35°C	45°C
0.4 lpm	25°C	0.77	0.732	0.6844	
	35°C	0.8175	0.77	0.732	
	45°C	0.865	0.8175	0.77	

		Liquid Heat Capture Ratio			
		Coolant Temp	25°C	35°C	45°C
		Air Temp			
0.6 lpm	25°C		0.7842	0.7415	0.6844
	35°C		0.827	0.7842	0.7272
	45°C		0.8698	0.827	0.7842

To quantify the heat captured by the coolant, the CFD model is simulated for various inlet temperatures of the coolant and air. Keeping air flowrate constant at 80 CFM for the entire server and varying inlet air temperatures, we produced the results tabulated in Table 6-4.

Case 1, we found that at constant air inlet temperature the heat captured by the coolant decreases as the coolant inlet temperature increases. This trend follows for all coolant flowrates. Fig. 6-7 shows the trend lines for heat capture ratio at various coolant flowrates at 25°C air inlet temperature. This trend follows 35°C and 45°C air inlet temperatures.

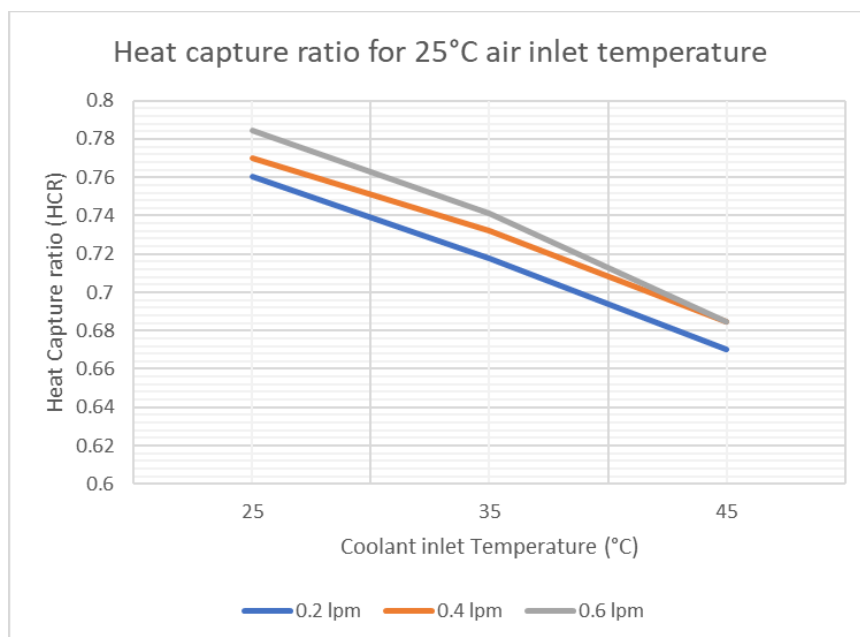


Figure 6-7 Heat capture ratio trend lines for varying flowrates

Similarly, we monitored HCR keeping the coolant inlet temperature constant at 25°C and varying the air inlet temperature. We observed, when air inlet temperature increases, the HCR increases proportionally. Therefore, from the observations, we could quantify that heat

is exchanged between the fluids when air comes in contact with the cold plates which is causing changes in HCR. Generally, we assume that heat dissipated by the CPU is carried away by the coolant through the cold plates but, from the study we could observe that the cold plate acts as a heat exchanging medium between the fluids. We also observed the max. temperatures at the surface of the CPU and PCH heat spreader and found that the cold plates could carry away maximum heat dissipated from the CPU and the surface temperatures are much lower than the max. allowable junction temperature considering a nominal internal resistance from the die to the surface of the heat spreader. Table 6-5 shows one such example comparing the experimentally observed junction temperature and CFD model-based results at 25°C air and coolant inlet temperature.

Table 6-5 Experimental vs CFD results

		Experimental	CFD
		Air & coolant inlet temp 25 C	
Coolant flowrate	Air flowrate (CFM)	80	80
0.4 lpm	CPU1 Temp (°C)	33.5	37.4
	CPU2 Temp (°C)	36.5	41.5
	PCH Temp (°C)	48	54
0.6 lpm	CPU1 Temp (°C)	33.5	36.7
	CPU2 Temp (°C)	36	39.5
	PCH Temp (°C)	49	53.8

6.5 Conclusion

The present investigation numerically evaluates the heat capture ratio in a hybrid cooled server for peak server load and varying inlet temperature for air and liquid. The CFD model of a Cisco Series C220 server with direct-to-chip liquid-cooled CPUs was developed. The cold plate for the CPU was experimentally characterized for pressure drop and thermal resistance characteristics and a black-box model was used for CFD simulations using 25% propylene glycol as the coolant. It can be concluded from the results shown above that heat capture ratio

of hybrid cooled servers depends mainly on the air and coolant supply temperature. It was observed that true heat capture ratio can be obtained only when the air and coolant supply temperature is same. In all the other cases it is transfer of heat between two fluids which does not represent the true heat captured by liquid. Similarly, it was observed that at any constant air supply temperature when the coolant supply temperature was varied, the PCH temperature increase with the increase in coolant supply temperature, this is due to heat transfer from the cold plate to air before reaching the PCH. Cooling liquid is more cost efficient than cooling air at higher temperatures. From this study we could also suggest keeping coolant inlet temperature low compared to air inlet temperature allowing heat transfer from air to the coolant. This would help in saving energy at CDU or CRAH unit deployed in hybrid cooling environment.

6.6 References

- [1] Li, Z., & Satish G. Kandlikar, S.G., 2015, "Current Status and Future Trends in Data-Center Cooling Technologies," *Heat Transfer Engineering*, 36(6), 523-538, DOI: 10.1080/01457632.2014.939032
- [2] V. S. Simon, A. Siddarth, and D. Agonafer, "Artificial Neural Network Based Prediction of Control Strategies for Multiple Air-Cooling Units in a Raised-floor Data Center," 2020 19th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), 2020, pp. 334-340, Doi: 10.1109/ITherm45881.2020.9190431.
- [3] S. Singh, K. Nemati, V. Simon, A. Siddarth, M. Seymour, and D. Agonafer, "Sensitivity Analysis of a Calibrated Data Center Model to Minimize the Site Survey Effort," 2021 37th Semiconductor Thermal Measurement, Modeling & Management Symposium (SEMI-THERM), 2021, pp. 50-57.

- [4] P. Shahi, S. Saini, P. Bansode and D. Agonafer, "A Comparative Study of Energy Savings in a Liquid-Cooled Server by Dynamic Control of Coolant Flow Rate at Server Level," in IEEE Transactions on Components, Packaging and Manufacturing Technology, vol. 11, no. 4, pp. 616-624, April 2021, doi: 10.1109/TCPMT.2021.3067045.
- [5] Shah, J. M., Padmanaban, K., Singh, H., Duraisamy Asokan, S., Saini, S., and Agonafer, D. (October 14, 2021). "Evaluating the Reliability of Passive Server Components for Single-Phase Immersion Cooling." ASME. J. Electron. Packag. June 2022; 144(2): 021109.
- [6] Mohapatra, S. and Loikits, D., "Advances in Liquid Coolant Technologies for Electronics Cooling," Proceedings of the 21st IEEE Semiconductor Thermal Measurement and Management Symposium, San Jose, CA, 2005, pp. 354-360.
- [7] ASHRAE 2021, Emergence and expansion of liquid cooling.

Chapter 7 Compact Modeling and Thermal Analysis of Immersion Based Hybrid Cooled Server

Reprinted with permission © 2023 ASME

7.1 Abstract

In recent years there has been a phenomenal development in cloud computing, networking, virtualization, and storage, which has increased the demand for high performance data centers. The demand for higher CPU (Central Processing Unit) performance and increasing Thermal Design Power (TDP) trends in the industry needs advanced methods of cooling systems that offer high heat transfer capabilities. Maintaining the CPU temperature within the specified limitation with air-cooled servers becomes a challenge after a certain TDP threshold. Among the equipments used in data centers, energy consumption of a cooling system is significantly large and is typically estimated to be over 40% of the total energy consumed. Advancements in Dual In-line Memory Modules (DIMMs) and the CPU compatibility led to overall higher server power consumption. Recent trends show DIMMs consume up to or above 20W each and each CPU can support up to 12 DIMM channels. Therefore, in a data center where high-power dense compute systems are packed together, it demands efficient cooling for the overall server components. In single-phase immersion cooling technology, electronic components or servers are typically submerged in a thermally conductive dielectric fluid allowing it to dissipate heat from all the electronics. The broader focus of this research is to investigate the heat transfer and flow behavior in a 1U air cooled spread core configuration server with heat sinks compared to cold plates attached in series in an immersion environment. Cold plates have extremely low thermal resistance compared to standard air cooled heatsinks. Generally, immersion fluids are dielectric, and fluids used in cold plates are electrically conductive which exposes several problems. In this study, we focus only on understanding the

thermal and flow behavior, but it is important to address the challenges associated with it. The coolant used for cold plate is 25% Propylene Glycol water mixture and the fluid used in the tank is a commercially available synthetic dielectric fluid EC-100. A Computational Fluid Dynamics (CFD) model is built in such a way that only the CPUs are cooled using cold plates and the auxiliary electronic components are cooled by the immersion fluid. A baseline CFD model using an air-cooled server with heat sinks is compared to the immersion cold server with cold plates attached to the CPU. The server model has a compact model for cold plate representing thermal resistance and pressure drop. Results of the study discuss the impact on CPU temperatures for various fluid inlet conditions and predict the cooling capability of the integrated cold plate in immersion environment.

7.2 Introduction

A data center is a facility that centralizes an organization's shared Information Technology operations and equipment for the purposes of storing, processing, and disseminating data and applications. The Switches, routers, firewalls, storage systems, servers, and controllers make up most of the data center's hardware [1]. As per 2018 study on total energy consumption, it is observed that there is a total energy consumption of 205 terawatt of energy consumed by the data centers which is 1% of energy consumption worldwide which is a total of 6% increase in energy consumption for data centers worldwide [2] and in the meantime the computing capacity has increased more than 550% over the time. Out of the total electricity consumed by the data center 30-50% is consumed by the IT and cooling equipment [3].

There has been extensive research on improving the cooling capabilities and adopting emerging cooling technologies at different levels of data centers. Researchers have explored different cooling technologies to address the thermal challenges faced due to the limits of air

cooling, the most common approach of cooling the components in a server has been via air cooling but with rise in chip power density, air cooling is becoming challenging with increase in chip power, the server airflow requirement increases and the power consumption via fans and acoustic noise also increases which decreases the cooling efficiency and increases the cost of cooling. Alternative cooling technologies are introduced and widely accepted by the industry is liquid cooling, in liquid cooling approaches, the components are either cooled via direct or indirect liquid cooling. In direct liquid cooling the server is completely immersed in a dielectric coolant and the coolant carries the heat, and in indirect liquid cooling approach, it includes cold plates placed on the top of the heat generating components [4]. In either type of liquid cooling approach, the fan power loads are replaced with lower pumping power [5]. Though being an effective solution, the deployment of liquid cooling is challenging and thus should be comprehensively considered. The ASHRAE provides guidelines [6] W1-W5 classes of liquid cooling based on facility water conditions. As being one of the efficient options to cool high heat generating components, hybrid cooling approach where utilizing the cold plates to cool the high heat generating components and dielectric coolant for cooling the secondary components inside the server can help in lowering the cooling power demands and increase the efficiency as the heat transfer coolants will have a higher heat carrying capacity then compared to air.

Several studies have contributed to the understanding and optimization of cooling performance in different configurations. Analytical models were developed to investigate split flow microchannel liquid-cooled cold plates with flow impingement, aiming to enhance heat dissipation [7][8]. Another study focused on determining the thermal performance limits of single-phase liquid cooling, using an improved effectiveness cold plate model [9]. Heat sink design optimization was explored through the implementation of guided vanes to target hotspots in liquid cooling systems [10][11]. A comprehensive CFD analysis examined the

impact of various parameters on the heat capture ratio of liquid cooling in a hybrid cooled server [12]. Experimental investigations were conducted to assess the effects of design modifications on the chassis and ducting of a server [13][14]. Transient studies were performed to optimize cooling performance in direct-to-chip liquid cooling at the rack level by implementing control strategies [15-17]. These research efforts collectively aimed to enhance cooling performance in hybrid cooled servers, combining both air and liquid heat transfer fluids. The current study specifically focuses on assessing the impact of using both liquid heat transfer coolants in a hybrid cooled server.

This research focuses on investigating heat transfer and flow behavior in a 1U air-cooled spread core configuration server with heat sinks compared to cold plates attached in series in an immersion environment and aims to understand the capability of cooling high power dense components while maintaining the 1U form factor. It also explores the effects of targeted coolant delivery in an immersion environment on CPU temperature and this research examines the cooling capability of associated electronic components present in the server, such as PCH and DIMMs. We also compare the cooling capability of cold plates in an immersion environment to that of heat sinks, and aim to predict the effectiveness of adapting cold plates over heat sinks. The results provide insights into the impact on CPU case temperatures under various fluid inlet conditions, as well as the cooling capability of the integrated cold plates in an immersion environment.

7.3 Server Description

The server used for the experiment was Cisco M220 M3, designed for performance and compute density over a wide range of business workloads from web serving to distributed database. The data server has a 1U form factor, 1.75in in height, width 16.92in and depth of 28.5in. The server consists of two CPUs in spread core configuration, Intel Xeon E5-M2600

and ME5-2600 processor with a Thermal Design Power (TDP) of 115 Watts. The server supports 16 DDR4 DIMMs, up to 8 drives and 2 x 1 GbE LAN-on-motherboard (LOM) ports delivering outstanding levels of density and performance in a compact 1U package. Figure 7-1 shows the server retrofitted with cold plates.



Figure 7-1 Server retrofitted with cold plates.

The server has five 40 mm counter rotating fans, which were removed to adapt immersion cooled environment. The server is equipped with a copper heatsink for air cooling solution. It was later retrofitted with Asetek cold plates for the study. In immersion cooling the entire server is submerged completely in a thermally conductive dielectric fluid and the heat is removed through natural or forced convection from the electronic components in the server.

Generally, immersion cooled systems use optimized heat sinks based on the fluid's thermo-mechanical properties to dissipate heat from high power consuming components such as CPUs and GPUs. With the rise in power density over the years, heat sinks do not satisfy the relative power dissipation from the electronics for the desired form factor of the servers. Associated components such as DIMMs, HDDs and other chipsets like Platform Control Hub (PCH) consume proportionally higher power which poses a challenge for air cooling systems to work efficiently.

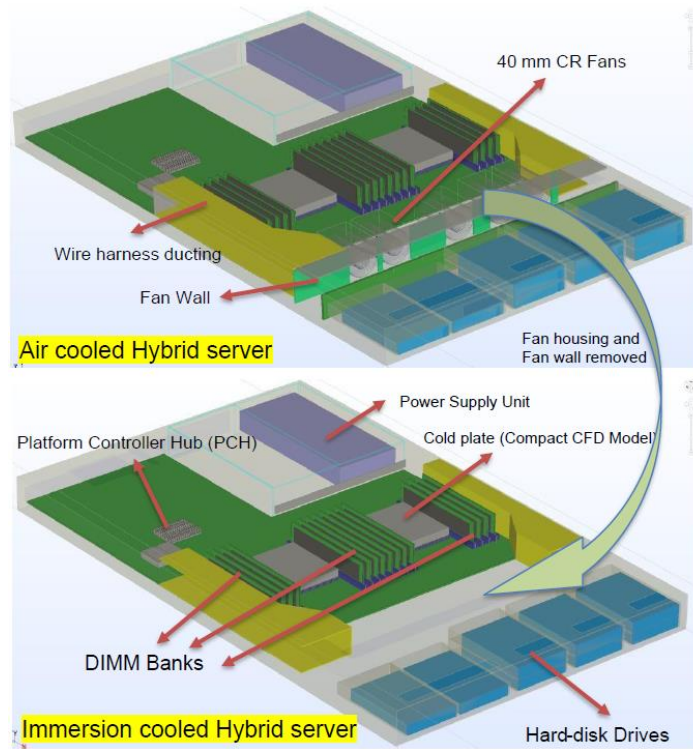


Figure 7-2 CFD model of a 1U server

7.4 CFD Model

A detailed CFD model of the server is shown in Figure 7-2. The CFD model has all the heat generating and flow impeding components including CPUs, DIMMs, Chipsets, cold plates, power supply unit, PCH, HDDs and baffles for cable routing. The CFD model was created using commercially available software, 6SigmaET, Future Facilities. For the baseline study and validation, an air-cooled hybrid server was modeled. A commercially available synthetic dielectric fluid EC-100 compatible for immersion cooling was used as the immersion coolant; it has good heat transfer properties with no/minimal risk of corrosion. A water-based coolant, PG25 (25% Propylene Glycol) was used for the cold plate as the heat transfer medium.

Table 7-1 Critical heat generating components.

Heat dissipating components	Qty	Power per Qty (W)
CPU	2	115
DIMMs	16	4
HDD	8	1.5
PCH	1	7

The specifications of critical heat generating components in the server are shown in Table 7-1. Figure 7-3 shows the flow network diagram attached to the two, 3D compact cold plate in series connection using the same CFD software.

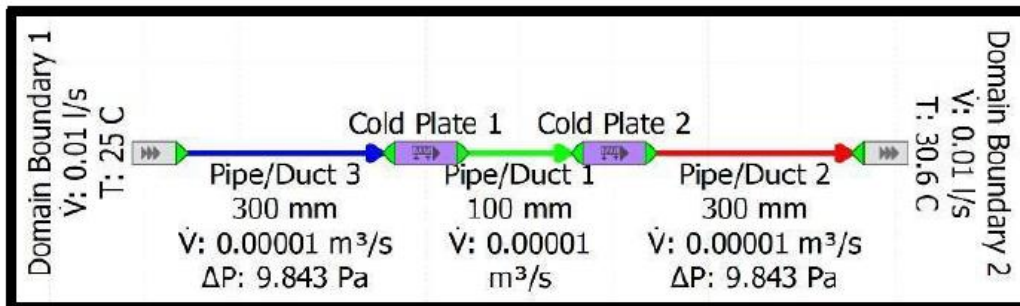


Figure 7-3 Flow network model for cold plates

Mesh Sensitivity analysis is performed to ensure that the model is independent of the grid count. The grid count considered is 2-7 million grid counts with the inlet flow rate at 0.6 lpm at the cold plate and 80 CFM at the server, inlet temperature at 25°C, and the front and rear end open to the environment. The optimum grid was decided at approximately 4.36 million cells for the baseline model to have an acceptable result (less than 5% variation).

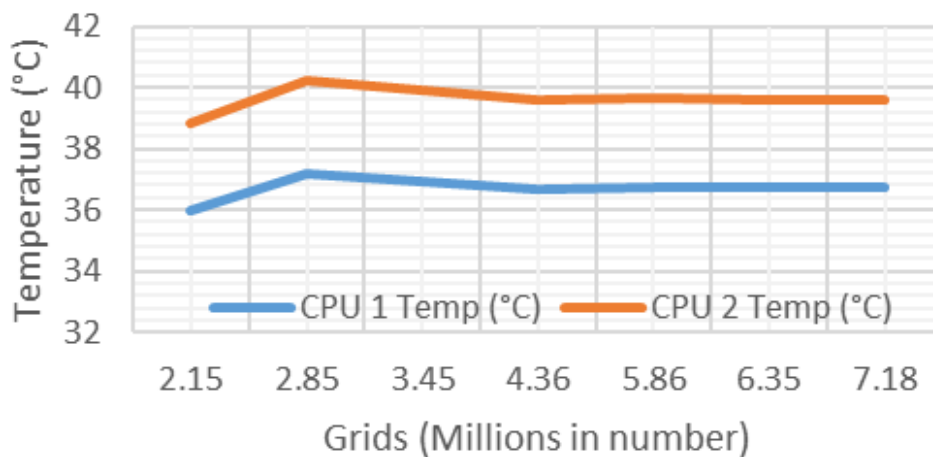


Figure 7-4 Mesh Sensitivity Analysis

The cold plate was experimentally characterized using PG-25 as the coolant. The pressure drop across the cold plate was measured using calibrated pressure sensors and thermal resistance from chip to coolant was calculated using the below equation (1).

$$R_{th} = \frac{(T_j - T_{inlet})}{Q} \quad (^\circ\text{C}/\text{W}) \quad (1)$$

Where Q is the heat dissipated by the CPU under max. utilization, T_j is the junction temperature of the component and T_{inlet} is the coolant inlet temperature. Figure 7-5 shows the thermal resistance for the cold plate at different coolant flow rates respectively. The cold plate was over-designed for water-based fluids for this specific CPU model which is why we observe very low thermal resistance at low flowrates.

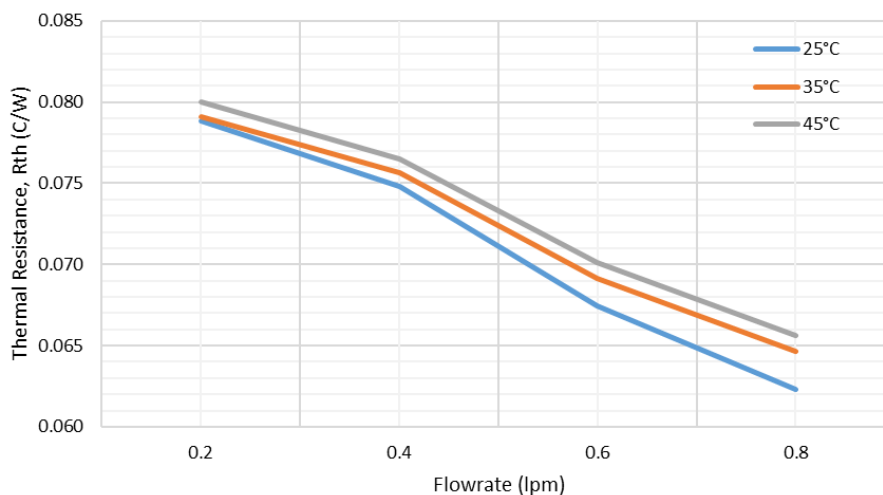


Figure 7-5 Thermal resistance for the cold plate at various coolant inlet temperature and flowrate

Flow resistance is the difference between the measured values of pressure using pressure transducers placed at the inlet and outlet of the cold plate.

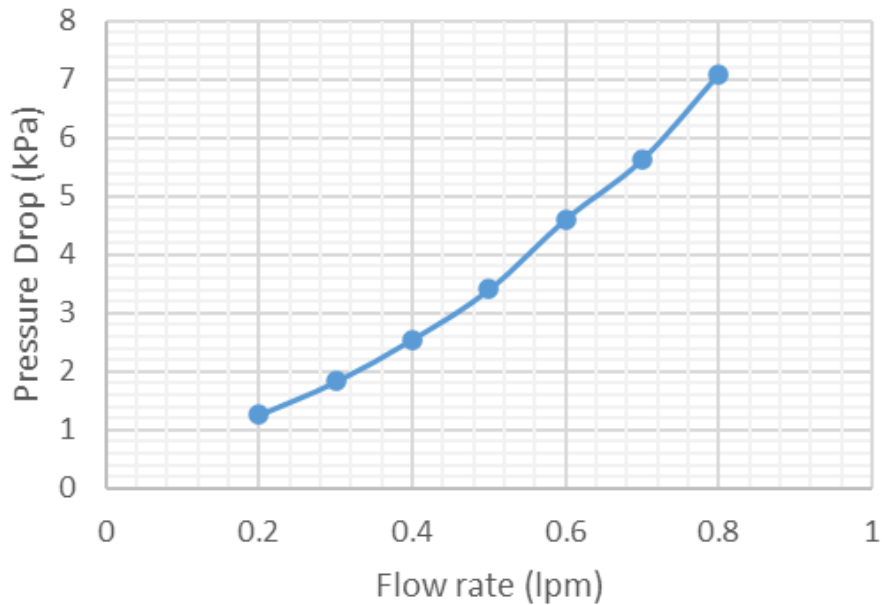


Figure 7-6 Pressure drop for the cold plate.

Figure 7-6 shows the pressure drop for the cold plate at different coolant flow rates respectively. Thermal resistance and pressure drop curves are set as compact cold plate characteristics in the flow network model in CFD.

An experimental study of server at various coolant flow rates was performed at constant air inlet temperature and flow rate and compared with the CFD results. The computational model was simulated for fluid inlet temperatures of 25 °C 35 °C and 45 °C at 115 W of TDP at various flow rates. The CFD model was validated with less than 3% error comparing the average junction temperatures and average air exit temperature from CFD and experiment. Table 2 shows the CFD model validation.

7.5 Test Cases

Simulations were run with various air inlet temperatures at fixed flow rates as well as various immersion fluid temperatures. Later, cooling capability for CPUs, PCH and DIMMs were calculated and compared.

7.5.1 Immersion Hybrid Cooling

The model was set up in such a way that the server was submerged vertically in the tank. Simulations were run with EC100 at various temperatures 25°C, 35°C, 45°C in natural convection environment. The cold plate coolant inlet temperature was also varied 25°C, 35°C, 45°C at fixed flow rate at 0.6 LPM. Convective heat transfer from the cold plate to immersion cooled environment is quantified and compared with air cooled server to evaluate the cooling capability at CPU, PCH and DIMMs.

7.5.2 Boundary Conditions

Simulations were run in sets of 3 keeping the air inlet temperature constant and varying cold plate coolant inlet temperature, air flow rate across the servers were kept constant (maximum fan speed speed) throughout all tests. The coolant flow rate for cold plates was also fixed at 0.6 lpm for all the test cases. Table 7-2 shows the boundary conditions for all the test cases. Similarly, for immersion hybrid scenario, the immersion fluid was varied while keeping the inlet and outlet of the server in a tank configuration.

Table 7-2 Boundary Conditions for all Test Cases

Boundary conditions for Air/Immersion hybrid server	
Air Inlet Temperature	25°C, 35°C, 45°C
Air Flow Rate	80 cfm (max.)
PG-25 Inlet Temperature	25°C, 35°C, 45°C
Cold Plate liquid Flow Rate	0.6 lpm (fixed)
EC100 Immersion Tank Temperature	25°C, 35°C, 45°C

7.5.3 Flow Pattern

Flow patterns were observed for Air hybrid and Immersion hybrid cases when set to different inlet temperatures and fixed boundary conditions. As discussed in test cases the Air flow is set to max fan speed which is 16 W per fan (total 5 fans), the server is placed

horizontally hence the gravity is in negative Y direction. Figure 7-7 shows the simulated flow pattern for an Air hybrid cooled server and immersion hybrid server. The velocity of air flow throughout the server is uniform unlike the immersion fluid velocity which changes based on the natural convection from the heat dissipating components.

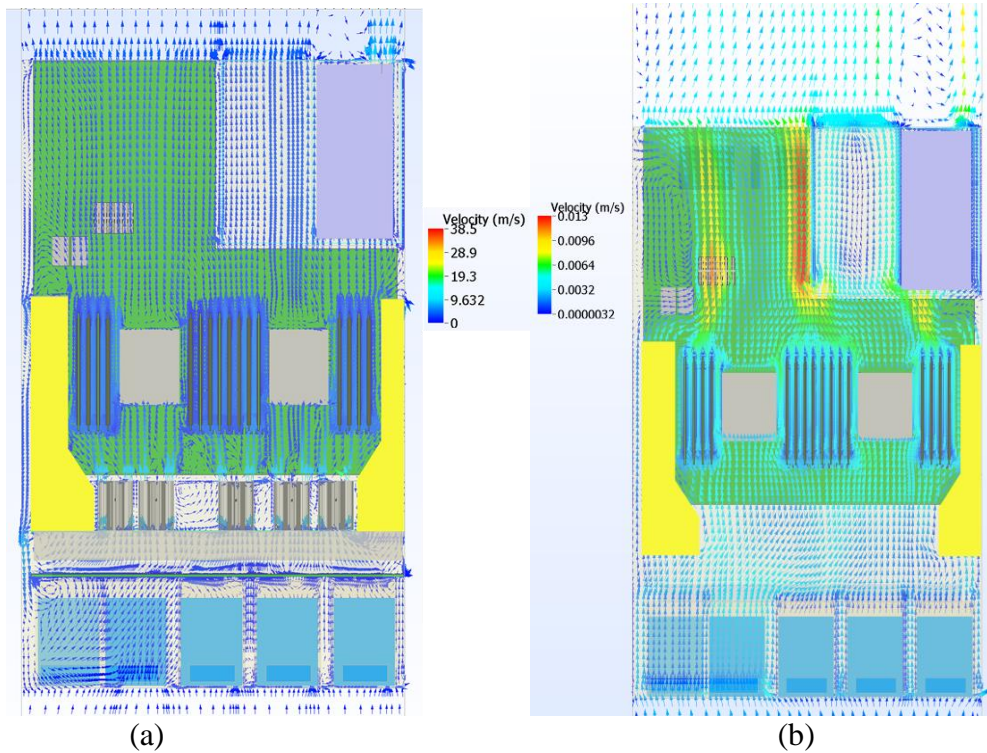


Figure 7-7 (a) Air hybrid server (b) Immersion hybrid server

7.6 Results and Observations

The simulation results showed that in air environment the PCH temperatures were high when compared to immersion fluid environment even at high fluid temperatures which is obvious because the heat transfer capacity of the immersion fluid is much higher than air. Even when there is a change in cold plate inlet temperature, the PCH temperature remains consistent in the immersion fluid environment, whereas in the air environment we observed at about 2°C rise in temperature at increasing cold plate inlet temperature. The approach air temperature at the PCH kept increasing since the PCH is at the downstream of the DIMMs and there was

additional heat dissipation from the server fans. Whereas the approach temperature at the PCH was constant in immersion environment. Therefore, PCH cooling capability in an immersion cooled environment increases compared to air, provided heat sink geometry is optimized accordingly.

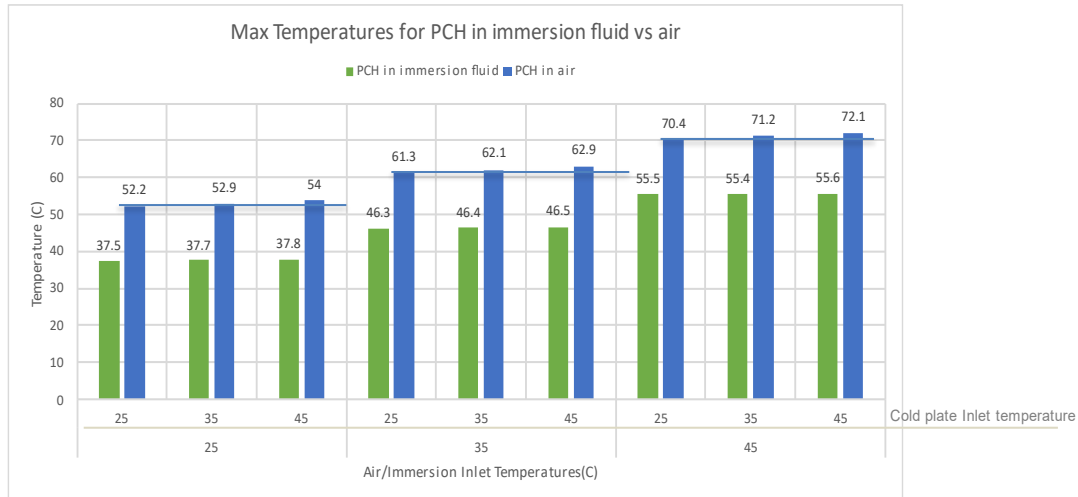


Figure 7-8 PCH temperature vs Air and Immersion server inlet temperatures

DIMM temperatures in both air and immersion environment were compared and the results showed that in air environment, the DIMM temperatures were high when compared to immersion fluid environment which was again an obvious observation as mentioned earlier. Even when there was a change in cold plate inlet temperature the DIMM temperature remains consistent in both air and immersion fluid environment.

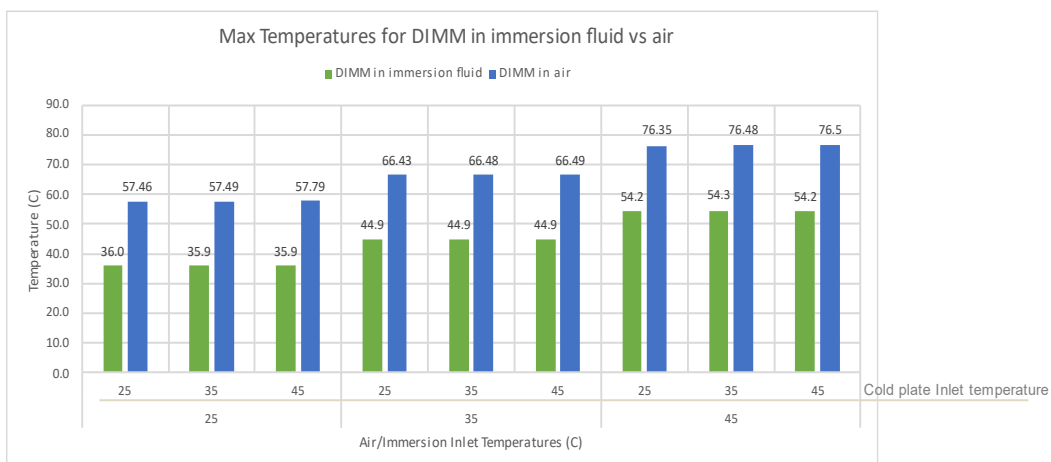


Figure 7-9 DIMM temperature vs Air and Immersion server inlet temperatures

Moreover, the max. DIMM temperature occurs at a localized area in the server due to the insufficient air flow between the DIMMS. This is attributed to the server design and fan placement. The air flow through the middle DIMM bank with 16 DIMMs (in between the CPUs) is comparatively lower since it is provisioned by one 40 mm counter rotating (CR) fan while the rest of the DIMM banks with 8 DIMMs each are provisioned with two fans each and baffling for the CPU.

CPU Temperature for both air and immersion environment were compared for various cold plate and server inlet temperatures. Since we have cold plates in both the environments, we see increase in CPU temperatures in both the cases as cold plate inlet temperature increases (proportionally), but there is a slight temperature drop of about 3°C to 4°C at the CPUs in case of immersion environment since percentage of convection at the cold plate surface is higher in immersion environment compared to air. This gives extra room to increase the TDP of the CPU in the immersion environment.

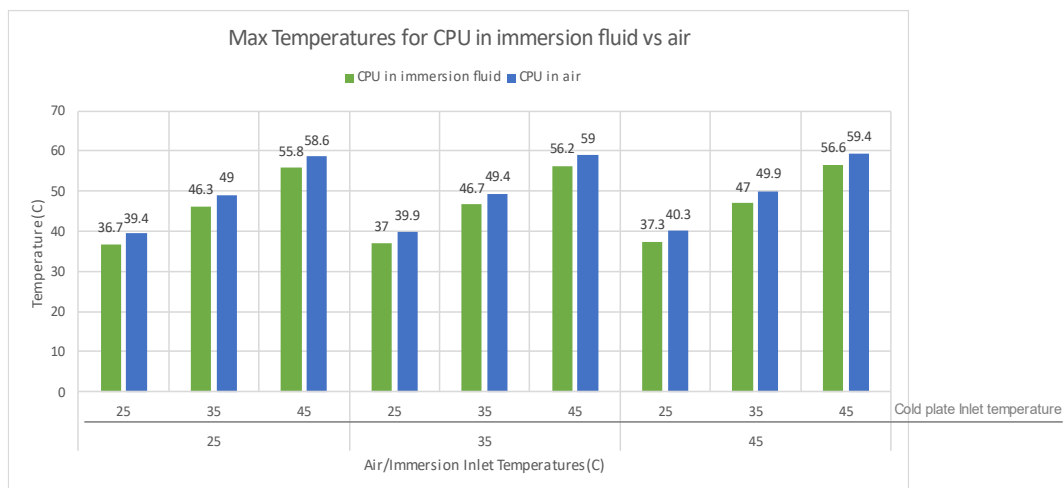


Figure 7-10 CPU temperature vs Air and Immersion server inlet temperatures

7.7 Cooling Capability Predictions

The cooling capability of individual electronic components in various environments is calculated using the equation (2) given below. Q is the amount of cooling/power consumption that can be provided to the components for the same server form factor and fluids used. TTT

is the thermal throttling temperature for electronic chips. Usually, commercially available CPUs throttle at/above 85°C after which the CPU performance drops due to insufficient cooling provided. Similarly, the DIMMs change refresh cycles after reaching at or above 80°C. Therefore, it is necessary to design the cooling system to support the throttling limits of the electronics. In this section, we compared three cases: 1. Cold plate in air environment. 2. Cold plate in immersion environment. 3. Heatsink in immersion environment.

$$\text{Cooling Capability } (Q) = \frac{(T_{TT} - T_{in})}{\text{Thermal resistance } (R_{th})} \quad (2)$$

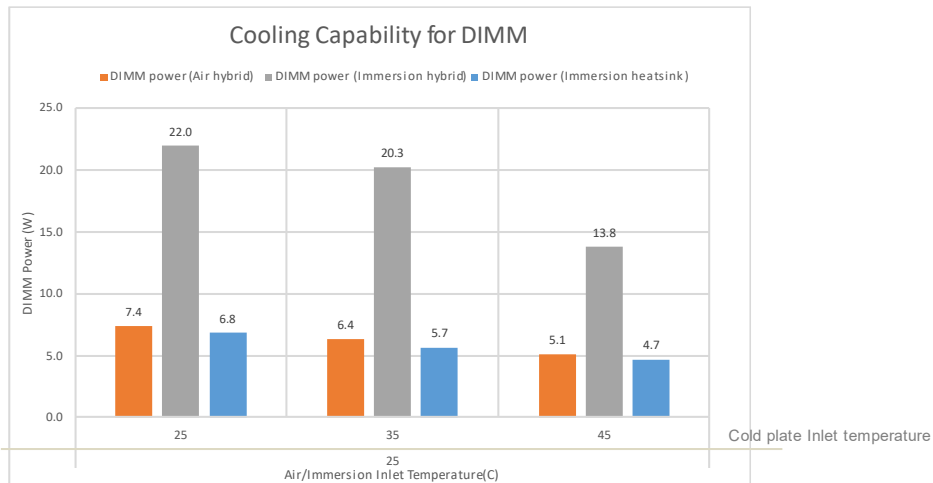


Figure 7-11 DIMM cooling capability

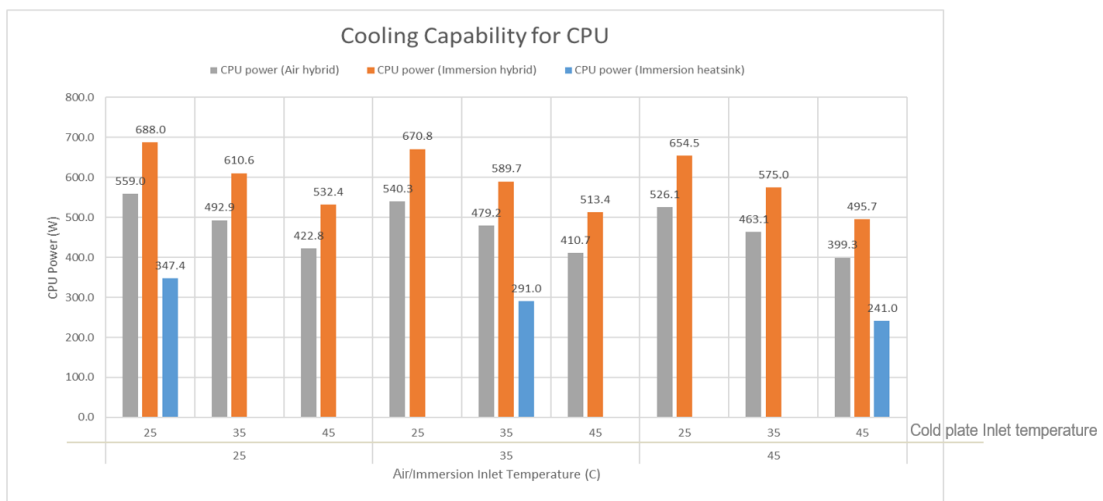


Figure 7-12 CPU cooling capability

There are certain limitations for the predictions made above:

- The form factor of the server was assumed to be same

- The cold plate thermal resistance and pressure drop were assumed to have the same characteristics. Hence these predictions can be used to follow the capability trends for further investigations on the cooling system design.

- Cooling capability analysis was done for this specific design of the server.

It was observed, for all the components the immersion-hybrid cooling environment has the best cooling capability prediction when compared to other methods of cooling. Moreover, in the case of targeted liquid delivery without temperature control, CPU temperatures are lower in cold plate-based models compared to heat sink models, even when both tank and cold plate inlet temperatures are the same. Close observation on CPU temperature shows 2°C to 9°C temperature variations when comparing heatsink based and cold plate immersion cooling model obviously due to higher heat transfer due through forced convection at the cold plates. In the case of targeted liquid delivery with temperature control CPU temperatures can individually be optimized based on coolant inlet temperature. This opens the opportunity/potential for precision control using cold plates in high performance computing systems. However, the true cooling capability (air-hybrid model) might be slightly higher than the reported values since it was calculated based on the max. DIMM and CPU temperatures which might differ based on the server design and fan placements. Moreover, the average DIMM temperature was found to be 4 ~5°C lower than the max. DIMM temperature.

7.8 Conclusion

Single-phase immersion cooling has proven to be more efficient than air cooling of data centers according to several research and case studies in the electronic cooling industry. In addition to being more efficient than air cooling, the servers are cut-off from the external

environment and air contamination as seen in air cooling is completely taken aback. Although immersion cooling has its own reliability risks, the author believes there are risk mitigation strategies being researched which help in increasing the reliability of immersion cooling. The power density of the servers can be increased by reducing/maintaining the same form factor of the server using immersion cooling techniques. In this study, it is observed that Immersion Hybrid cooling has a better impact on all the component temperatures while keeping the power consumption minimum compared to air hybrid cooling (fans consuming roughly 20% of the server power). Cooling capability predictions show the amount of cooling/power consumption that can be provided to the components for the same form factor using the same fluid properties.

7.9 References

- [1] What is a data center - cisco, website. [Online]. Available: <https://www.cisco.com/c/en/us/solutions/data-center-virtualization/what-is-a-data-center.html>
- [2] E. Masanet, A. Shehabi, N. Lei, S. Smith, and J. Koomey, “Recalibrating global data center energy-use estimates.”.
- [3] J. Ni and X. Bai, “A review of air conditioning energy performance in data centers,” *Renewable and Sustainable Energy Reviews*, vol. 67, pp. 625–640, Jan. 2017, doi: 10.1016/j.rser.2016.09.050.
- [4] “As Rack Densities Rise, Liquid Cooling Specialists Begin to See Gains,” *Data Center Frontier*, Jun. 14, 2019. <https://datacenterfrontier.com/as-rack-densities-rise-liquid-cooling-specialists-begin-to-see-gains/> (accessed Apr. 28, 2021).
- [5] A. C. Kheirabadi and D. Groulx, “Cooling of server electronics: A design review of existing technology,” *Applied Thermal Engineering*, vol. 105, pp. 622–638, Jul. 2016, doi: 10.1016/j.applthermaleng.2016.03.056.

- [6] American Society of Heating, Refrigerating and Air-Conditioning Engineers, Ed., Liquid cooling guidelines for datacom equipment centers, Second edition. Atlanta, [GA]: ASHRAE, 2014.
- [7] Kisitu, Deogratius, and Alfonso Ortega. "Thermal-Hydraulic Analytical Models of Split-Flow Microchannel Liquid-Cooled Cold Plates With Flow Impingement." In International Electronic Packaging Technical Conference and Exhibition, vol. 85505, p. V001T02A015. American Society of Mechanical Engineers, 2021.
- [8] Kisitu, Deogratius. "Approximate Compact Thermal-Hydraulic Analytical Models for Laminar Microchannel Impingement Liquid-Cooled Cold Plates for Data Center Thermal Management." PhD diss., Villanova University, 2021.
- [9] Ortega, Alfonso, Carol Caceres, Umut Uras, Deogratius Kisitu, Uschas Chowdhury, Vahideh Radmard, and Ali Heydari. "Determination of the Thermal Performance Limits for Single Phase Liquid Cooling Using an Improved Effectiveness-NTU Cold Plate Model." In International Electronic Packaging Technical Conference and Exhibition, vol. 86557, p. V001T01A008. American Society of Mechanical Engineers, 2022.
- [10] Gharaibeh, Ahmad R., Mohammad I. Tradat, Srikanth Rangarajan, Bahgat G. Sammakia, and Husam A. Alissa. "Multi-objective optimization of 3D printed liquid cooled heat sink with guide vanes for targeting hotspots in high heat flux electronics." *International Journal of Heat and Mass Transfer* 184 (2022): 122287.
- [11] Gharaibeh, Ahmad R., Yaman M. Manaserh, Mohammad I. Tradat, Firas W. AlShatnawi, Scott N. Schiffres, and Bahgat G. Sammakia. "Using a Multi-Inlet/Outlet Manifold to Improve Heat Transfer and Flow Distribution of a Pin Fin Heat Sink." *Journal of Electronic Packaging* 144, no. 3 (2022): 031017.

- [12] Shalom Simon, Vibin, Lochan Sai Reddy, Pardeep Shahi, Amrutha Valli, Satyam Saini, Himanshu Modi, Pratik Bansode, and Dereje Agonafer. "CFD Analysis of Heat Capture Ratio in a Hybrid Cooled Server." In International Electronic Packaging Technical Conference and Exhibition, vol. 86557, p. V001T01A013. American Society of Mechanical Engineers, 2022.
- [13] Modi, Himanshu, Uschas Chowdhury, and Dereje Agonafer. "Impact of Improved Ducting and Chassis Re-design for Air-Cooled Servers in a Data Center." In 2022 21st IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (iTherm), pp. 1-8. IEEE, 2022.
- [14] Modi, Himanshu, Pardeep Shahi, Lochan Sai Reddy Chinthaparthi, Gautam Gupta, Pratik Bansode, Vibin Shalom Simon, and Dereje Agonafer. "Experimental Investigation of the Impact of Improved Ducting and Chassis Re-Design of a Hybrid-Cooled Server." In International Electronic Packaging Technical Conference and Exhibition, vol. 86557, p. V001T01A019. American Society of Mechanical Engineers, 2022.
- [15] Modi, Himanshu, Pardeep Shahi, Akiilessh Sivakumar, Satyam Saini, Pratik Bansode, Vibin Shalom, Amrutha Valli Rachakonda, Gautam Gupta, and Dereje Agonafer. "Transient CFD Analysis of Dynamic Liquid-Cooling Implementation at Rack Level." In International Electronic Packaging Technical Conference and Exhibition, vol. 86557, p. V001T01A012. American Society of Mechanical Engineers, 2022.
- [16] Heydari, Ali, Pardeep Shahi, Vahideh Radmard, Bahareh Eslami, Uschas Chowdhury, Chandraprakash Hinge, Lochan Sai Reddy Cinthaparthi et al. "A Control Strategy for Minimizing Temperature Fluctuations in High Power Liquid to Liquid CDUs Operated at Very Low Heat Loads." In International Electronic Packaging Technical Conference and

Exhibition, vol. 86557, p. V001T01A011. American Society of Mechanical Engineers, 2022.

- [17] Shalom Simon, V, Modi, H, Sivaraju, KB, Bansode, P, Saini, S, Shahi, P, Karajgikar, S, Mulay, V, & Agonafer, D. "Feasibility Study of Rear Door Heat Exchanger for a High Capacity Data Center." Proceedings of the ASME 2022 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems. ASME 2022 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems. Garden Grove, California, USA. October 25–27, 2022. V001T01A018. ASME. <https://doi.org/10.1115/IPACK2022-97494>

Chapter 8 Parametric Multi-Objective Optimization of Cold Plate for Single-Phase Immersion Cooling

Reprinted with permission © 2023 ASME

8.1 Abstract

The increasing demand for high-performance computing in applications such as the Internet of Things, Deep Learning, Big data for crypto-mining, virtual reality, healthcare research on genomic sequencing, cancer treatment, etc. have led to the growth of hyperscale data centers. To meet the cooling energy demands of HPC datacenters efficient cooling technologies must be adopted. Traditional air cooling, direct-to-chip liquid cooling, and immersion are some of those methods. Among all, Liquid cooling is superior compared to various air-cooling methods in terms of energy consumption. Direct on-chip cooling using cold plate technology is one such method used in removing heat from high-power electronic components such as CPUs and GPUs in a broader sense. Over the years Thermal Design Power (TDP) is rapidly increasing and will continue to increase in the coming years for not only CPUs and GPUs but also associated electronic components like DRAMs, Platform Control Hub (PCH), and other I/O chipsets on a typical server board. Therefore, unlike air hybrid cooling which uses liquid for cold plates and air as the secondary medium of cooling the associated electronics, we foresee using immersion-based fluids to cool the rest of the electronics in the server. The broader focus of this research is to study the effects of adopting immersion cooling, with integrated cold plates for high-performance systems. Although there are several other factors involved in the study, the focus of this paper will be the optimization of cold plate microchannels for immersion-based fluids in an immersion-cooled environment. Since immersion fluids are dielectric and the fluids used in cold plates are conductive, it exposes us to a major risk of leakage into the tank and short-circuiting the electronics. Therefore, we

propose using the immersed fluid to pump into the cold plate. However, it leads to a suspicion of poor thermal performance and associated pumping power due to the difference in viscosity and other fluid properties. To address the thermal and flow performance, the objective is to optimize the cold plate microchannel fin parameters based on thermal and flow performance by evaluating thermal resistance and pressure drop across the cold plate. The detailed CFD model and optimization of the cold plate were done using Ansys Icepak and Ansys OptiSLang respectively.

8.2 Introduction

A data center is a physical facility where organizations store their critical applications and data. The design of a data center is built on a network of computing and storage resources that allow the delivery of shared applications and data. In the AI era, deep Learning, machine learning, and big data need enormous amounts of CPU power and computing resources. This implies that a large number of high-performance processors, such as high-performance CPUs, GPUs, FPGAs, and ASIC devices are required [1]. Due to the limitations of air cooling in dissipating growing power densities in servers, researchers have been driven to seek newer and more effective cooling alternatives [2-8]. The next radical change in the thermal management of data centers is to shift from conventional cooling methods like air cooling to direct liquid cooling (DLC) [2-8].

Submerging servers and IT equipment in a dielectric medium for cooling results in significant energy savings due to the high energy loads and density. Its heat capacity per volume is 1120–1400 times that of air [9]. Furthermore, the rack density as a function of Power Usage Effectiveness (PUE) shows that the rack power density for single-phase immersion cooling is nearly three times greater than for air cooling. While a conventional air-cooled system has a PUE of about 1.5, immersion cooling has a PUE of about 1.07, meaning a 36%

reduction in power usage when employing immersion cooling [10]. The single-phase immersion cooling uses a dielectric fluid that helps in dissipating much higher heat from the components of the server. Some of its advantages include a high heat transfer coefficient, stable hydrodynamic flow, and the ability to directly cool hot components using the fluid. In single-phase immersion cooling, there is no phase change phenomenon, and the server along with its electronic components is immersed completely in the dielectric fluid. The heat transferred from the server components to the tank is cooled using an external heat exchanger or a coolant distribution unit (CDU) and finally discharged to the ambient through the primary side cooling units [11].

Thus, liquid cooling can assist to boost performance per watt while cutting total energy use [12-14]. A heat exchanger dumps the heat collected from the server modules to the facility water in a typical liquid-cooled data center server rack that utilizes a warm water direct liquid cooling system (DLCS). The heat exchanger's primary and secondary loops are the building facility water loop and the coolant loop to the server manifolds. The secondary loop coolant is routed through manifolds to cold plates, which are mounted on processor chips in individual servers with the thermal interface material (TIM) at the interface between the processor chips and the cold plates. The heat transfer coefficient 'h' between the entering coolant and the channel determines the amount of heat that can be taken up by the coolant going through the cold plate [15].

This paper focuses on the application of single-phase liquid immersion cooling along with a cold plate for high heat flux components. The cold plates typically have thin fin microchannels which transfer the heat from surfaces with high heat load to the fluid used in the system. It has two connectors at the end for the inlet and outlet flow. The fluid flows inside the cold plate, through the microchannels removing heat from the source. A baseline numerical model was created in Ansys Icepak with a cold plate attached to a heater with the whole setup

immersed in a dielectric fluid. The same dielectric fluid is pumped through the cold plate as well. The baseline cold plate model was parametrically optimized for immersion cooling using ANSYS OPTISLANG with the fin spacings, fin thickness, and fin height as the variable parameters.

8.3 Computational model and experimental validation

The cold plates are designed to suit each of the CPUs or GPUs in a server arrangement. A commercially available cold plate of overall dimensions – 110 mm x 85 mm x 41 mm, and with a base of - 55 mm x 55 mm x 4.5 mm, was chosen for this study. The CFD model of the server and cold plate setup is shown in Fig. 8-1.

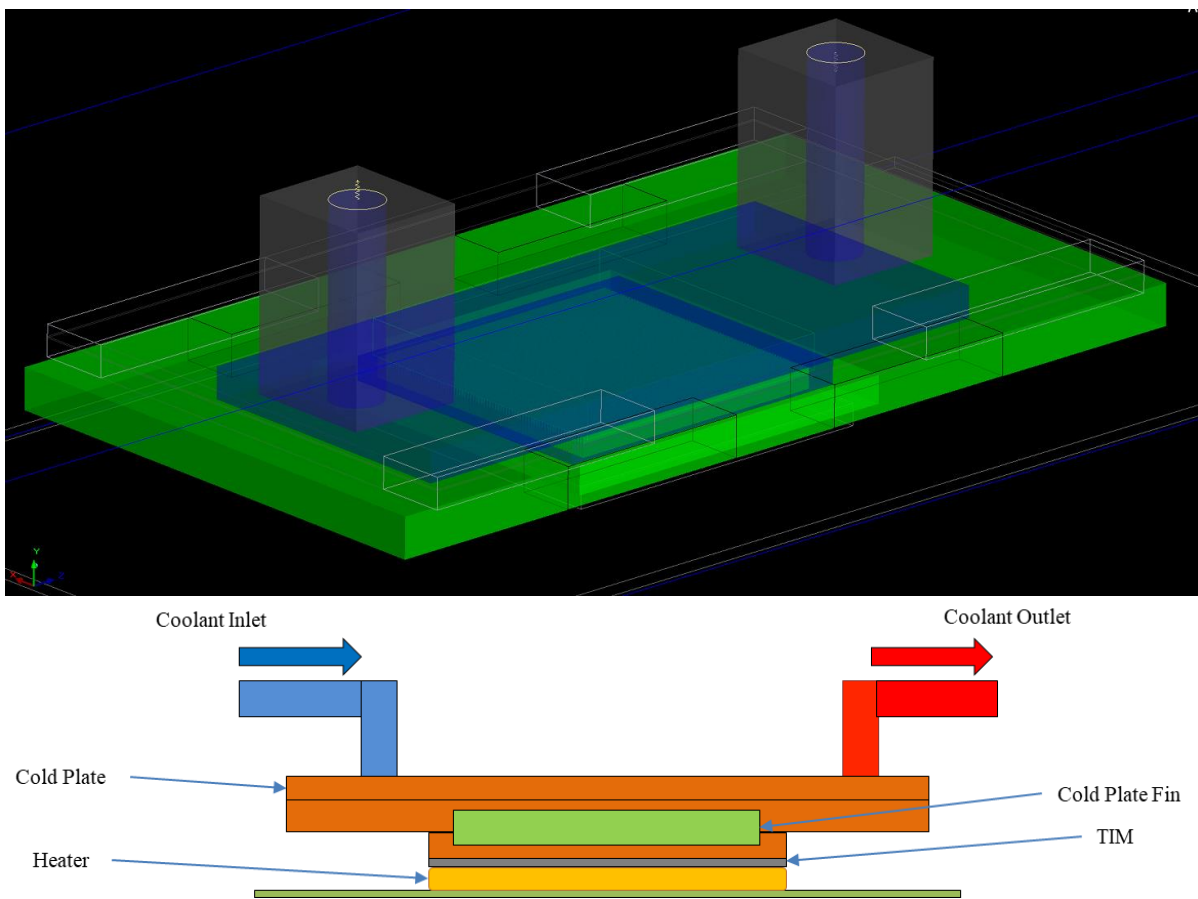


Figure 8-1 MODEL OF COLD PLATE & STACK-UP OF THE COLD PLATE ON THE HEATER.

The detailed CFD model of the cold plate was created using Ansys Icepak. The model stack-up consists of a ceramic heater (50 mm x 50 mm x 3 mm) simulating the heat load of an electronic chip CPU placed on top of a PCB. The fin thickness of the original cold plate was 0.2 mm, and the height was 3mm. A commercially available immersion cooling-specific Indium Heat-spring TIM (50 mm x 50 mm x 0.3 mm) was used as the interface material between the heater and the cold plate, and its thermal conductivity is 8 W/mK (calculated based on the data provided by vendor) which is slightly higher than the commonly used thermal grease. The cold plate was attached to the heater with a torque of 8 lbf-ft force which creates optimum contact and better heat transfer. The heater and cold plate setup are kept inside the domain containing immersion fluid. EC-100 is used as the dielectric coolant as that has high thermal conductivity and is compatible with components of the server with no or minimal risk of corrosion and has low GWP.

Table 8-1 ELECTROCOOL (EC) 100 THERMO-MECHANICAL PROPERTIES

EC100 thermo-mechanical				
Temperature	Dynamic Viscosity	Density	Thermal conductivity	Specific Heat
°C	Kg/m-sec	Kg/m ³	W/m/K	KJ/kg-K
20	0.02193	845.86	0.1389	2.1317
25	0.01763	842.56	0.13853	2.1499
30	0.01439	839.26	0.13815	2.1683
35	0.01191	835.96	0.13778	2.1868
40	0.00998	832.66	0.1374	2.209
45	0.00847	828.7	0.13703	2.2243
50	0.00725	826.06	0.13665	2.2433

The heater was assumed to represent a high-performance CPU that could reach a very high operating temperature during operation. Various sets of simulation were performed to

validate the thermal resistance and pressure drop across the cold plate and use the data to optimize the cold plate design for immersion-based fluids to integrate the cold plate in immersion cooling. The boundary conditions were set as shown in Table 8-2.

Table 8-2 BOUNDARY CONDITIONS



Boundary Conditions	
Coolant Inlet Temperature	40°C
Coolant Type	EC-100
Coolant Flowrates	3 lpm
Heater Power	400 W
Server Flow	Natural Convection






Experimental Procedure

- ✓ A benchtop setup is created which consists of a set of sensors and other equipment. The detailed list and specifications of each of the components are given in Table 3.
- ✓ The initial experiment was performed with PG-25 (Direct-on chip coolant) and later EC-100 immersion fluid was used to perform and validate the experiment.
- ✓ First add some amount of the coolant to the reservoir and run the pump for some time so that the loop is rinsed with the coolant and any external particles and contaminants are removed from the loop and the associated components.
- ✓ After a few runs, remove the liquid from the loop and use fresh coolant fluid to perform the actual experiment.
- ✓ Run the pump at a higher flow rate for a few minutes and make sure all air-bubbles are removed from the loop and there is proper flow of the fluid.
- ✓ Switch on the DAQ and make sure the temperature, pressure and flow rate readings are recorded and monitor the feedback.
- ✓ The supply voltage to the pump can be adjusted so that the flow rate can be controlled and maintained in the desired value.

- ✓ Switch on the chiller unit connected to the heat exchanger and set the desired inlet temperature.
- ✓ Gently power up the heater in a slowly increasing pace and set it to desired power (experiment set TDP)
- ✓ Frequent monitoring of data is needed to make sure that the heater does not go to a higher temperature than expected.
- ✓ Once the entire system is in a stable condition start collecting the data and store it in the local PC using DAQ
- ✓ After the desired amount of time, and reliable amount of data is collected, change the parameters, and repeat the experiment.

Table 8-3 List of Equipments

Sl. No	Details	Image	Specification
1	In-Line Strainer	 A cylindrical in-line strainer with a clear plastic housing and a black base. It features a mesh screen inside and two ports for fluid flow.	50um Mesh (SS) Material Housing – polypropylene Bowl material - Nylon
2	Gear Pump SEAFLO (SFGP1-032-003-01)	 A small, cylindrical gear pump with a white body and brass-colored ports. It has a black base and a CE mark on the side.	Power source – 12V DC Max Flow rate – 12 lpm Max Head – 3m Material – plastic
3	Thermistor 10k sensor	 A cylindrical thermistor sensor with a black cable. The sensor head is made of a metal mesh and has a black plastic cap.	

4	Pressure Sensor GP-M001		<p>Range - -14.5 to 145 PSI Medium temperature - -20 to +100°C Power voltage – 10-30V DC Material - SS304</p>
5	Clamp-on Micro Flow Sensor		<p>Flow rate – 0.0831pm to 5 lpm Temp range - -40°C to 80°C Max pressure - 145psi Accuracy - +/- 1% Material – Pa66 + GF/PPs</p>
6	Power supply E3642A		<p>Output range – 30 to 100W Low range 0 to 30V/2.2A High range 0 to 60V /1.3A</p>
7	Agilent DAQ Unit		<p>2-wired 22 channel inputs 20 voltage inputs Two current inputs</p>
8	Heat Exchanger		

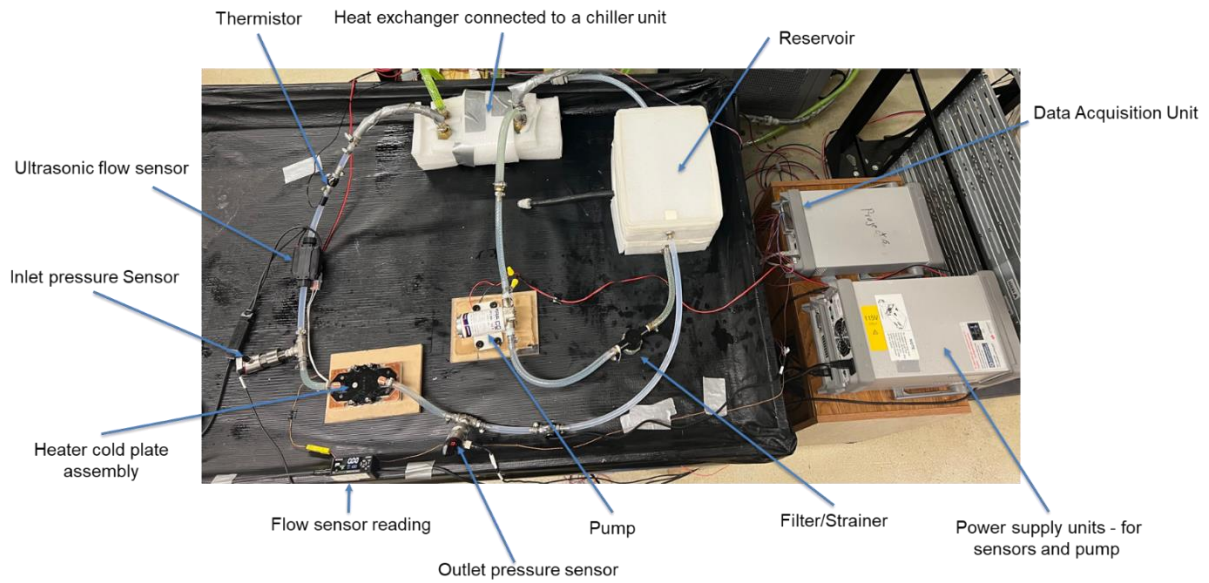


Figure 8-2 Experimental Setup

The above image shows the experimental setup in the EMNSPC Research Facility. The system relates to temperature, pressure and flow sensors as shown and data is collected. All the components are marked in the image as listed in Table 8-3.

To validate the CFD model, experimental data was collected for 400 W heater power with 3 sets of inlet temperatures as 25°C, 35°C and 45°C and a combination of different flow rates - 0.5 lpm, 1.0 lpm, 1.5 lpm, 2.0 lpm, 2.5 lpm & 3.0 lpm. EC100 coolant was used for experiments and validation. Figure 8-3 shows the schematic of the experimental setup.

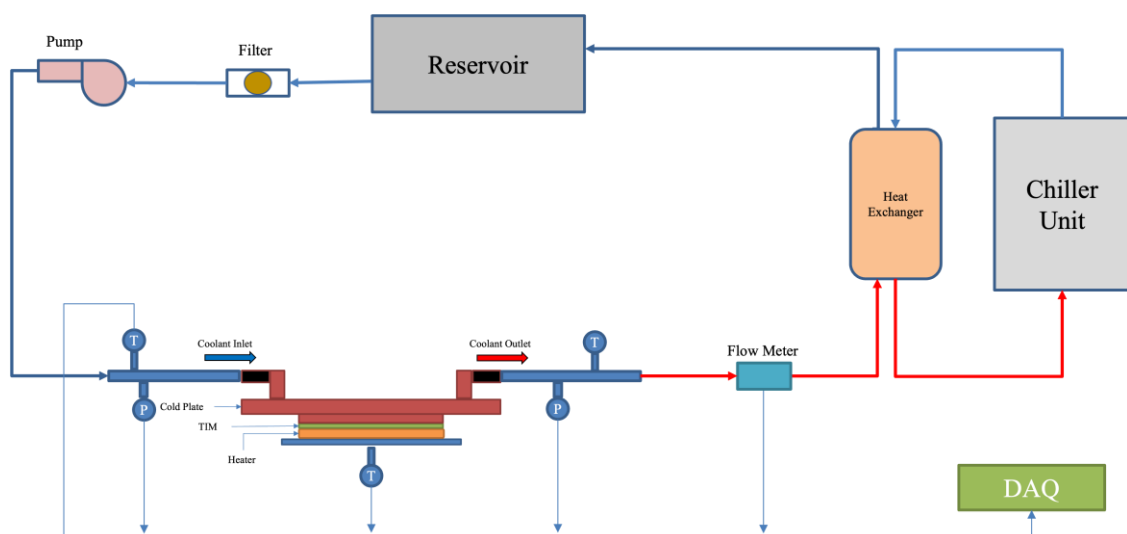


Figure 8-3 EXPERIMENTAL SETUP SCHEMATIC

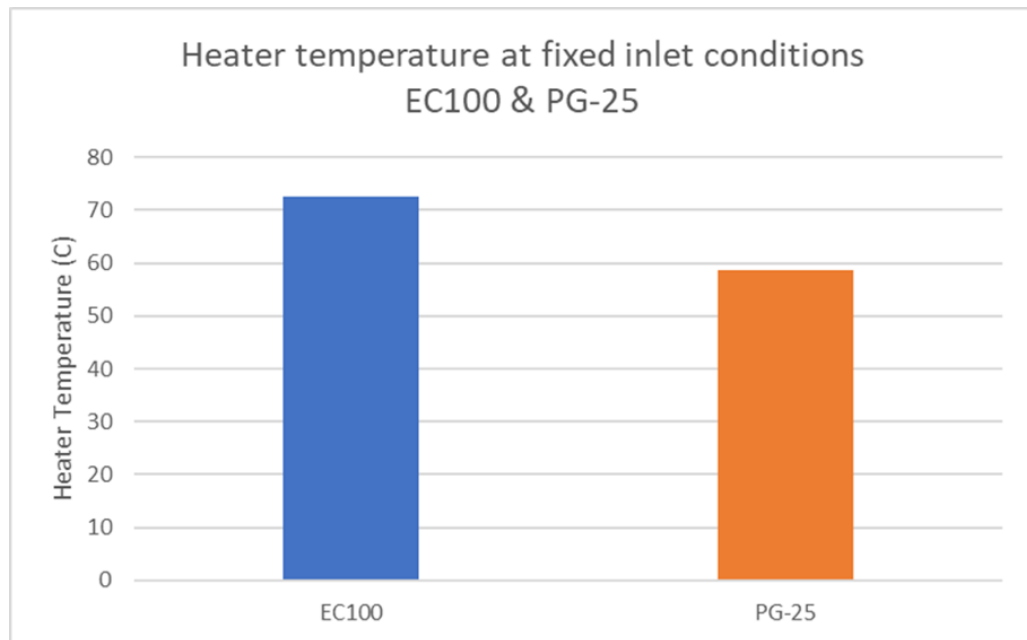


Figure 8-4 Heater temperature - EC100 vs PG25

From the experiment comparing the heater temperature at fixed inlet conditions for EC100 and PG25 (Figure 8-4) we can see that the thermal performance of EC100 is not as good as water-based coolants which are generally used in cold plates due to its lower heat transfer capacity. The properties and data values of these immersion fluids are obtained from industry, and these are used for simulations and comparison. Using these details, the experiments are done, and the physics is observed.

The model validation is shown in Fig. 8-5. Heater temperature was monitored for the same coolant inlet temperature and different flow rates. The experiments were conducted while the cold plate was in an air environment and the CFD model was validated with the same boundary conditions. The validated model was used to simulate the immersion fluid EC100 using OptiSLang to optimize the fin parameters (fin spacing, fin height, fin thickness) for the given boundary conditions.

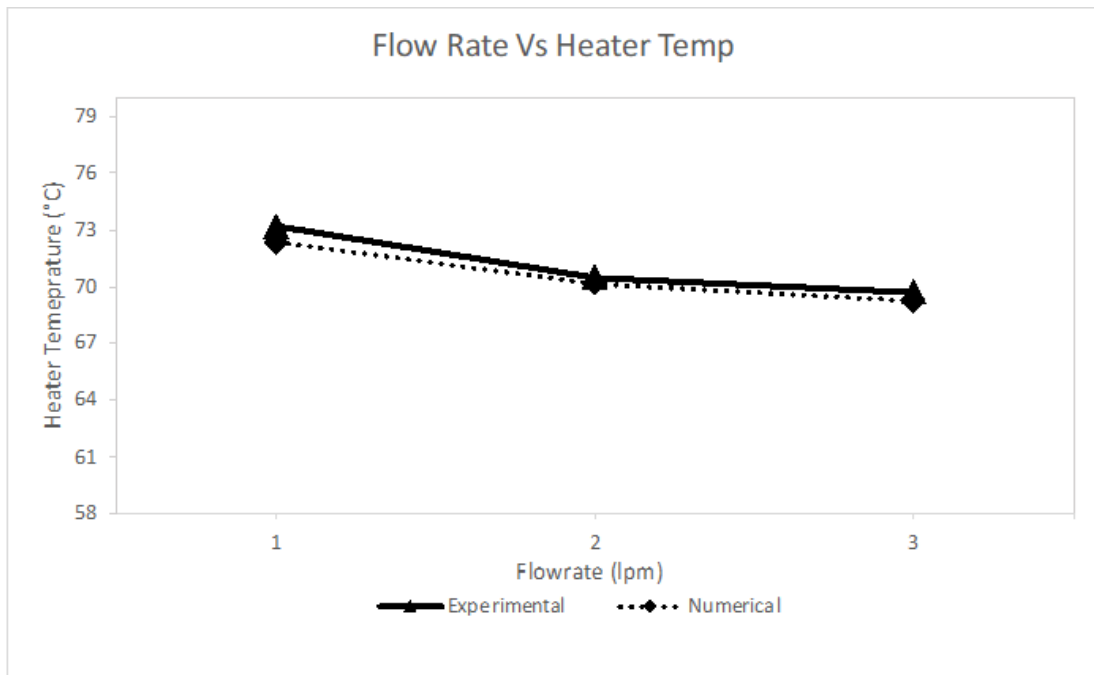


Figure 8-5 MODEL VALIDATION PLOT - HEATER TEMPERATURE VS. FLOW RATE

Table 8-4 VARIABLE INPUT PARAMETERS

Design Points for fin geometry			
No	Height	Thickness	Spacing
1	3	0.2	0.3
2	4	0.6	0.65
3	5	1.0	1
4		1.4	
Step Size	1	0.4	0.35
Total Discrete Values	3	4	3
Total No. of Design Points = $3*4*3 = 36$			

To optimize the fin parameters for the given inlet boundary conditions a set of 36 combinations were chosen as shown in table 8-4. Three different values of fin heights and three different values of fin spacing with four distinct values of fin thickness are derived for the optimization variables. All the combinations models were simulated using Ansys Icepak coupled with Ansys OptiSLang to compare and get the optimized cold plate fin parameters. The results are compared, and the best combination is derived. Fig 8-6, 8-7 show the temperature contours for the baseline model.

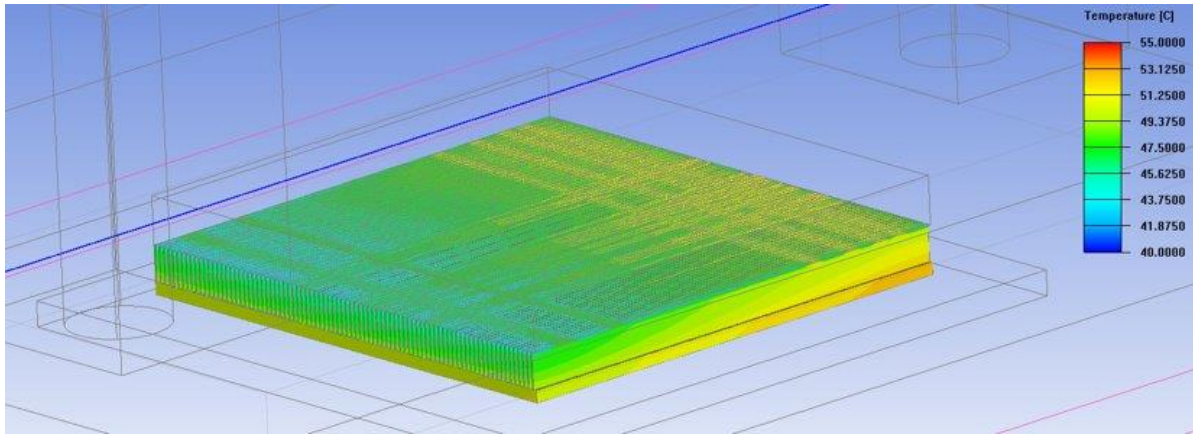


Figure 8-6: Temperature contour of cold plate fins

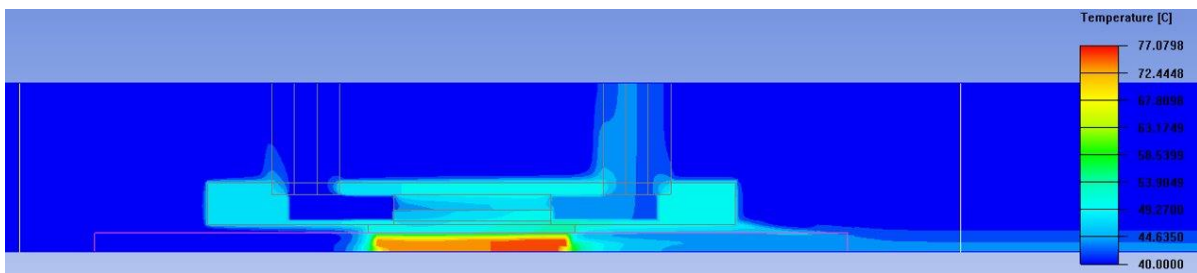


Figure 8-7 Temperature contour at X-Plane

8.4 Results and Observation

8.4.1 Thermal performance

Fig. 8-8 shows thermal resistance as a function of fin thickness for different fin spacings at a fixed fin height of 3 mm. It was observed that the thermal resistance decreases as the fin thickness increases and this trend was observed for all fin spacings. Moreover, it is to be noted that the thermal resistance decreases as the fin spacings are increased. The fact that single-phase immersion fluids have higher viscosity can be attributed to the observation of decreasing thermal resistance with increased fin spacing as higher fin spacing, allows the viscous fluid to pass through the channel with ease preventing choking of the fluid at the channels. The decrease in thermal resistance with an increasing fin thickness was because of increased conduction heat

transfer to the fins. These two reasons can be attributed to the reason why single-phase immersion heat sinks have higher pitch and fin thickness.

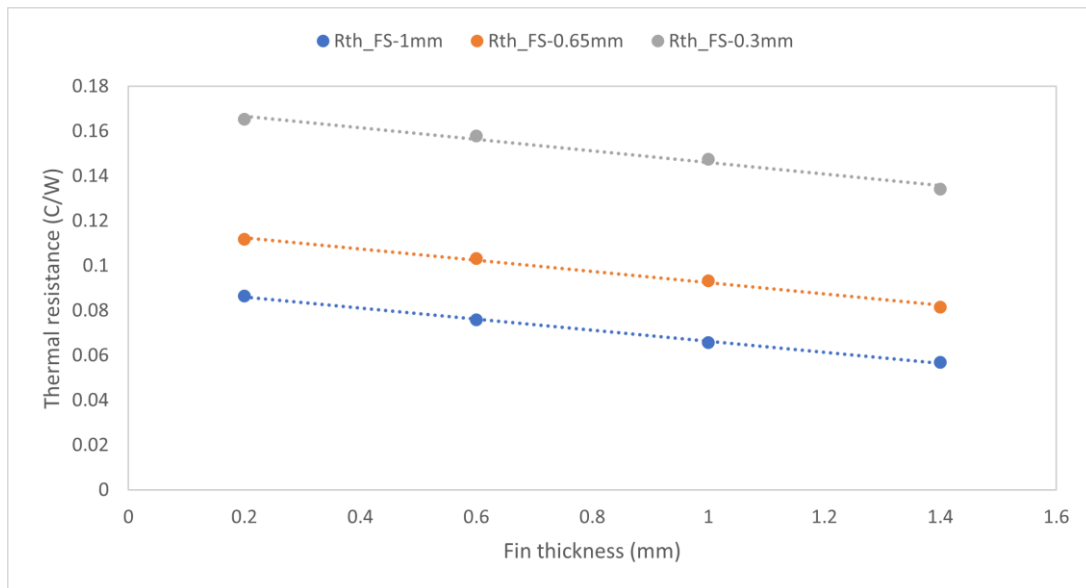


Figure 8-8 THERMAL RESISTANCE VS. FIN THICKNESS FOR DIFFERENT FIN SPACING (FS) AT A FIXED FIN HEIGHT OF 3 MM

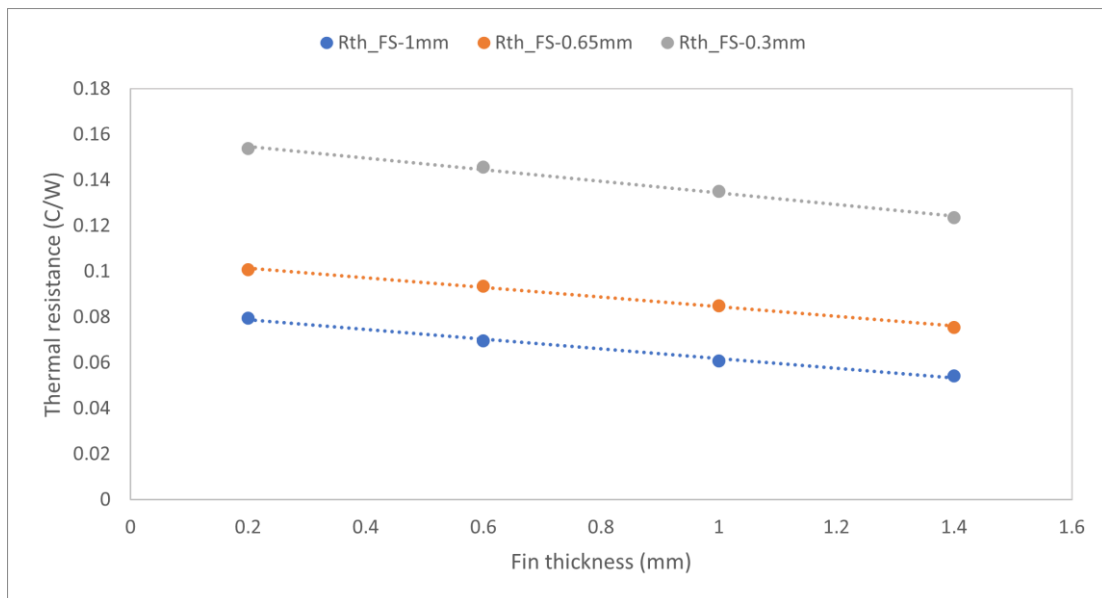


Figure 8-9 THERMAL RESISTANCE VS. FIN THICKNESS FOR DIFFERENT FIN SPACING (FS) AT A FIXED FIN HEIGHT OF 4 MM

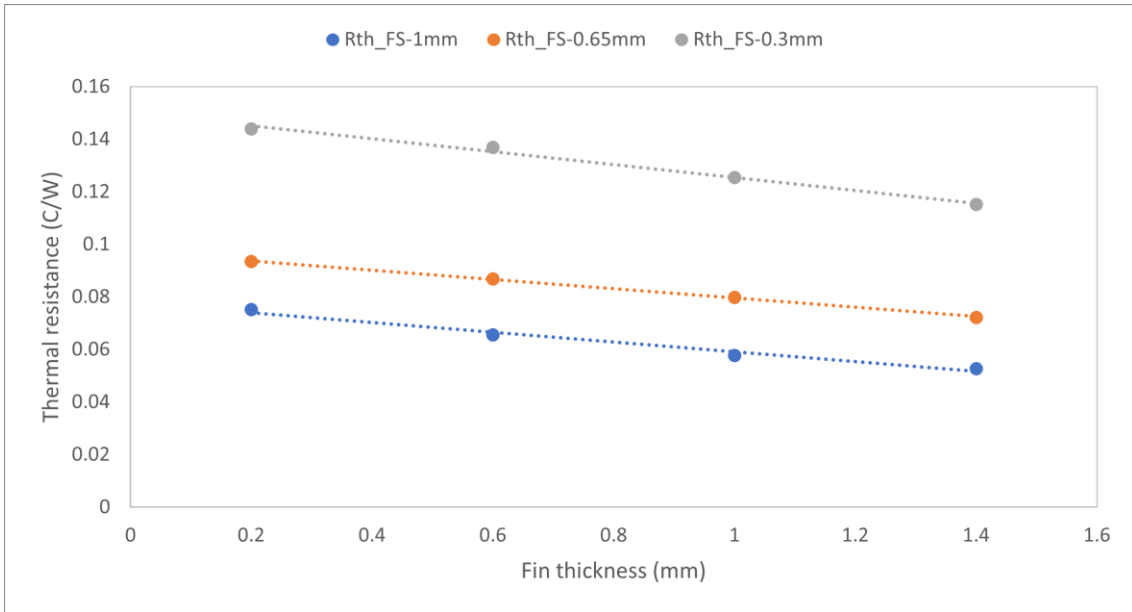


Figure 8-10 THERMAL RESISTANCE VS. FIN THICKNESS FOR DIFFERENT FIN SPACING (FS) AT A FIXED FIN HEIGHT OF 5 MM

It was observed that the trend from Fig.-8-8 can be seen in Fig.8-9 and Fig.8-10 which show the thermal resistance as a function of fin thickness and fin spacing with fin height lengths of 4 mm and 5 mm respectively.

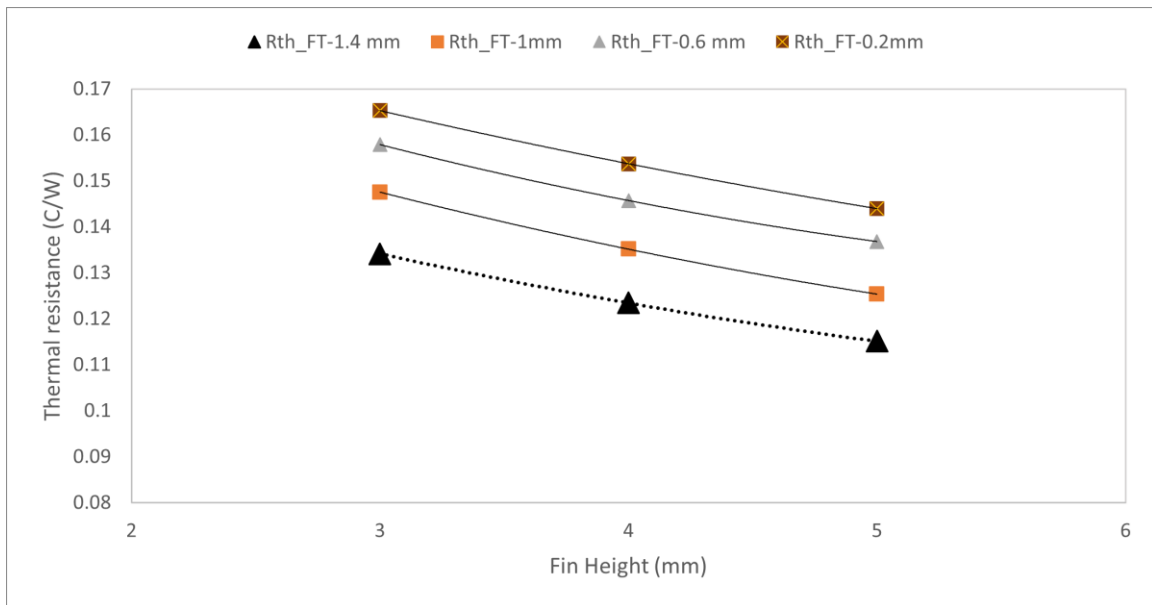


Figure 8-11 THERMAL RESISTANCE VS. FIN HEIGHT FOR DIFFERENT FIN THICKNESSES AT A FIXED FIN SPACING OF 1 MM

The above figure shows the thermal resistance vs fin height and fin thickness for a fixed fin spacing of 1mm and it was observed that the thermal resistance at the cold plate reduces when fin height increases, this is due to an increase in heat transfer surface area as the fin height increases thus enhancing the convection heat transfer mechanism.

8.4.2 Flow performance

The pressure drops across the cold plate as a function of fin thickness and fin spacing is shown in Fig.8-12. It can be seen that the pressure drop at the cold plate reduces dramatically when fin spacing increases (a maximum reduction of ~ 80% was observed). The fin thickness and fin height do not seem to affect the pressure drop significantly. The Coefficient of Prognosis (COP) matrix in Fig 8-13 forecasts the quality of the regression model for prognosis. This shows the change in pressure at the inlet and outlet of the cold plate with varying fin thickness, height, and spacing.

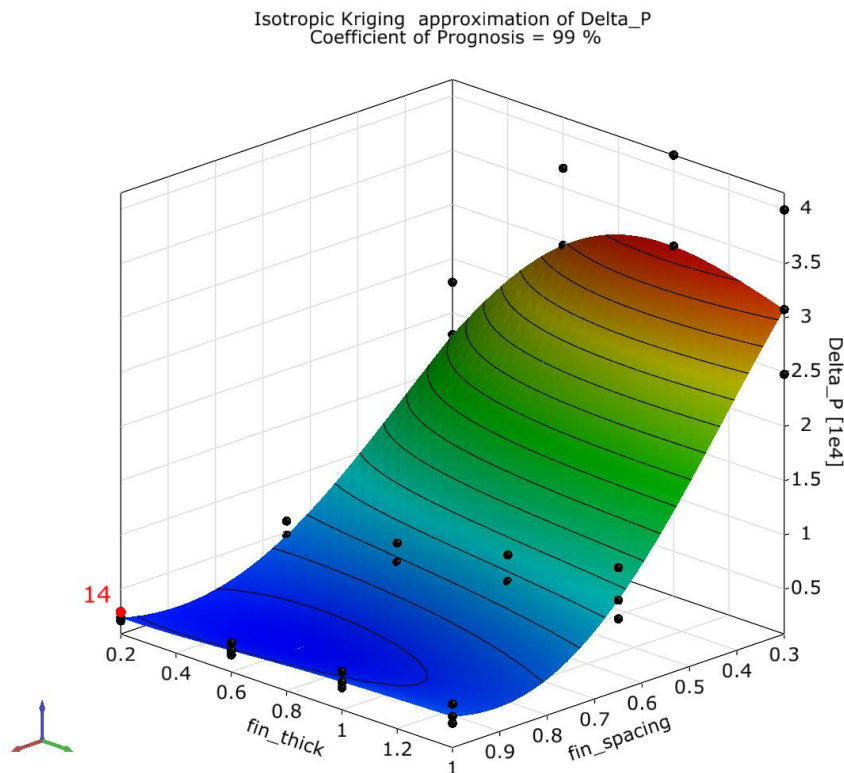


Figure 8-12 PRESSURE DROP ACROSS THE COLD PLATE AS A FUNCTION OF FIN THICKNESS AND FIN SPACING

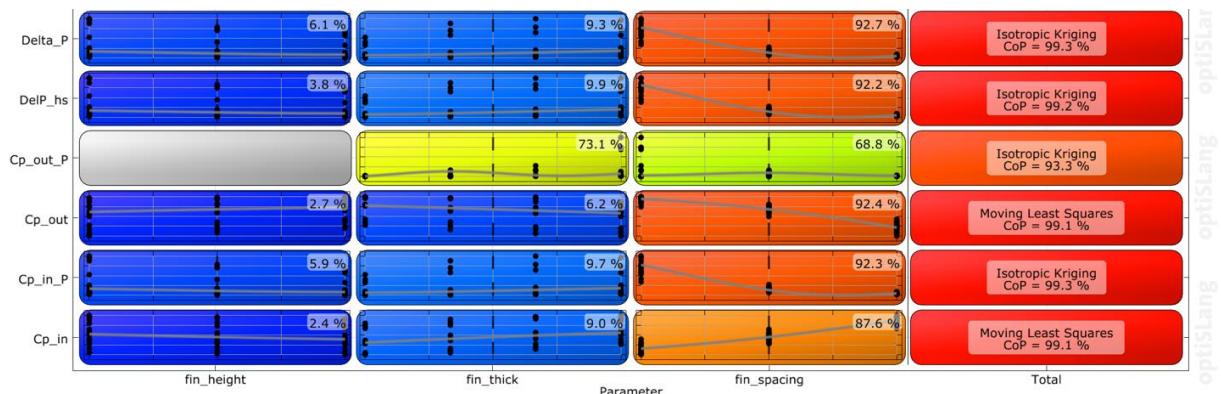


Figure 8-13 COP MATRIX OF RESULTS

Optimized Cold Plate:

From the 36 combinations of fin thickness, fin spacings, and fin heights the optimized cold plate was chosen which had the least thermal resistance. A comparison of the fin dimensions/ parameters of the optimized cold plate and the baseline cold plate used for optimization are shown in TABLE 8-5. As discussed in the thermal performance section previously it can be seen that an immersion-optimized cold plate's fins have higher spacing (owing to the viscous fluid), and higher fin thickness (to increase the conduction heat transfer to the fins).

Table 8-5 FIN PARAMETERS OF BASELINE COLD PLATE VS OPTIMIZED COLD PLATE

Variables	Baseline Cold plate	Optimized Cold plate
Fin Spacing	0.3 mm	1.0 mm
Fin Height	3.0 mm	5.0 mm
Fin Thickness	0.2 mm	1.4 mm

TABLE 8-6 shows the thermal performance of the baseline cold plate and the optimized cold plate. It was observed that the thermal resistance of the optimized cold plate was 0.025 °C/W.

Table 8-6 RTH FOR SAME BOUNDARY CONDITIONS

Parameters	Baseline Cold Plate	Optimized Cold Plate
Heater Temperature (°C)	72	62
Inlet Fluid Temp (°C)	40	40
ΔT	32	37
Q (W)	400	400
Rth (°C/W)	0.08	0.025

8.5 Conclusion

Also, with increasing edge deployments and expansion of 5G, immersion tanks operating at elevated ambient temperatures will be useful owing to simple cooling infrastructure and high heat capture capabilities. The major challenge in using cold plates in immersion tank environment is to set up the coolant delivery to the micro channels which involves internal pumps that supply the cold fluid from the tank directly to the cold plate. This adds on to additional power consumption which is a major trade-off with the heat transfer performance compared to heat sinks. Several conclusions drawn from the study are as follows:

- Using cold plates in an immersion-cooled environment, the heat transfer from the cold plate to the environment also contributes to enhanced heat transfer capability in cold plates.
- Optimization of cold plate fin geometries or heatsinks is purely based on the heat flux of the component and the thermo-mechanical properties of the fluid
- Immersion cooling supports provisioning even higher powered associated electronic components such as DIMMs

8.6 References

- [1] Patterson, M., Fenwick, D., “The State of Data Center Cooling-A review of current air and liquid cooling solutions.” Intel white paper, 2008.

- [2] S. Singh, K. Nemati, V. Simon, A. Siddarth, M. Seymour, and D. Agonafer, "Sensitivity Analysis of a Calibrated Data Center Model to Minimize the Site Survey Effort," 2021 37th Semiconductor Thermal Measurement, Modeling & Management Symposium (SEMI-THERM), 2021, pp. 50-57.
- [3] V. S. Simon, A. Siddarth, and D. Agonafer, "Artificial Neural Network Based Prediction of Control Strategies for Multiple Air-Cooling Units in a Raised-floor Data Center," 2020 19th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), 2020, pp. 334-340, Doi: 10.1109/ITherm45881.2020.9190431.
- [4] Scaramella, J., 2008, "Next-Generation Power and Cooling for Blade Environments," IDC, Technical Report No. 215675.
- [5] Sivaraju, K. B., Bansode, P., Gupta, G., Lamotte-Dawaghreh, J., Saini, S., Simon, V., Herring, J., Karajgikar, S., Mulay, V., & Agonafer, D. (2022). Comparative Study of Single-Phase Immersion Cooled Two Socket Server in Tank and Sled Configurations. Proceedings of ASME 2022 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems, InterPACK 2022. <https://doi.org/10.1115/IPACK2022-97429>
- [6] R. Schmidt, M. Iyengar, D. Porter, G. Weber, D. Graybill, and J. Steffes, "Open side car heat exchanger that removes entire server heat load without any added fan power," in Proc. 12th IEEE Intersoc. Conf. Thermal Thermomech. Phenomena Electron. Syst. (ITherm), Jun. 2010, pp. 1–6.
- [7] K. Nemati, H. A. Alissa, B. T. Murray, B. Sammakia, and M. Seymour, "Experimentally validated numerical model of a fully-enclosed hybrid cooled server cabinet," in Proc. ASME Int. Tech. Conf. Exhibit. Packag. Integr. Electron. Photon. Microsystem.

- Collocated, ASME 13th Int. Conf. Nanochannels, Microchannels, Minichannels, 2015, p. V001T09A041.
- [8] Shalom Simon, V, Modi, H, Sivaraju, KB, Bansode, P, Saini, S, Shahi, P, Karajgikar, S, Mulay, V, & Agonafer, D. "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. Garden Grove, California, USA. October 25–27, 2022. V001T01A018. ASME.
- [9] F. Douchet, D. Nortershauser, S. Le Masson, and P. Glouannec, "Experimental and numerical study of water-cooled datacom equipment," *Appl. Thermal Eng.*, vol. 84, pp. 350–359, Jun. 2015.
- [10] Modi, H, Shahi, P, Sivakumar, A, Saini, S, Bansode, P, Shalom, V, Rachakonda, AV, Gupta, G, & Agonafer, D. "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. Garden Grove, California, USA. October 25–27, 2022. V001T01A012. ASME.
- [11] Modi, H, Shahi, P, Chinthaparthi, LSR, Gupta, G, Bansode, P, Shalom Simon, V, & Agonafer, D. "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. Garden Grove, California, USA. October 25–27, 2022. V001T01A019.
- [12] 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.

- [13] 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).
- [14] 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).
- [15] Shalom Simon, V, Reddy, LS, Shahi, P, Valli, A, Saini, S, Modi, H, Bansode, P, & Agonafer, D. "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*. Garden Grove, California, USA. October 25–27, 2022. V001T01A013. ASME.

Chapter 9 Fluid based optimization scheme for heat sinks in single-phase immersion cooling under natural convection

9.1 Introduction

Data centers have swiftly evolved into a foundational pillar of contemporary economies, owing to the emergence of transformative technologies such as cloud computing, online media, social networking, as well as advancements in Artificial Intelligence (AI) and Machine Learning. Notably, prevailing statistics on active internet users reveal that approximately 5 billion users engage with online platforms at any given moment, thereby underscoring the substantial reliance placed upon network servers and data centers [1]. Although the escalation of data center energy consumption has been mitigated by the implementation of efficient IT equipment [2], the trajectory of growth persists on an upward trajectory annually. Nevertheless, the capacity of enhancements in cooling and energy efficiency to effectively counterbalance the rapid expansion of computational workloads remains uncertain [3]. Remarkably, power densities within conventional data centers can range from 15 to 100 times greater than those observed in generic commercial edifices [4]. The escalation in this demand also carries consequential environmental implications, including amplified greenhouse gas emissions and elevated utilization of water resources, both through direct and indirect avenues [5].

Over the preceding decade, the augmentation in energy requisites and processor power densities, driven by heightened processing demands, has constrained conventional air-cooling techniques to cater primarily to processors with more restrained Thermal Design Power (TDP) limits. In response to escalating energy consumption and the escalating intricacy of thermal management challenges, a spectrum of approaches has been suggested by researchers to ameliorate thermal performance [6,7]. These strategies encompass endeavors to amplify energy

conservation by employing alternative operational methodologies for extant cooling techniques [8,9], as well as the integration of liquid-based cooling technologies [10, 11].

Among the array of liquid-cooling methodologies, the Single-Phase Immersion Cooling technique distinguishes itself due to its facile implementation, economical deployment of dielectric fluids, and the simplicity of its cooling infrastructure [12]. Many of the prevailing proprietary cooling solutions have validated the efficacy of single-phase immersion cooling through attributes such as diminished Power Usage Effectiveness (PUE) values and heightened dependability of server components [13,14]. A primary, direct advantage offered by the comprehensive submersion of servers within dielectric fluids lies in the severance of server components from hostile environments, encompassing gaseous contaminants. This, in turn, mitigates failures arising from fan vibrations and obviates the necessity for cooling peripheral components, given that thermally elevated constituents are in direct interaction with the coolant. While immersion cooling presents pronounced benefits in comparison to air-cooling, it mandates meticulous contemplation of both thermal and non-thermal design aspects when applied to air-cooled hardware.

To illustrate, in the context of immersing air-cooled hardware, certain imperatives come to the forefront: the extraction of fans from the server configuration, the hermetic sealing of hard drives, the resolution of material compatibility concerns, and the optimization of heat sink design. These aforementioned considerations constitute the focal point of inquiry in the present study.

To consistently attain dependable and optimal functionality from both Central Processing Units (CPUs) and Graphics Processing Units (GPUs) within immersion cooling setups, it becomes imperative to employ an optimized heat sink, diverging from the use of conventional air-cooled heat sinks. Extensive investigations into the optimization of parallel

plate-fin heat sinks have been conducted, encompassing not only the refinement of geometric attributes [15, 16], but also the examination from both single and multi-objective optimization perspectives [17, 18]. Critical geometric parameters such as fin spacing, fin thickness, base thickness, and fin count play a pivotal role in maximizing the thermal efficacy of both the heat sink and the processors themselves. The current body of literature demonstrates an increasing proliferation of research endeavors related to heat sink optimization across both air-cooled and liquid-cooled systems, employing diverse computational fluid dynamics (CFD)-based and analytical methodologies.

Chen and Chen, for instance, employed a multi-objective, innovative direction-based algorithm to optimize plate-fin heat sinks in conjunction with an impingement fan, utilizing a commercially available Multiphysics tool [19]. The outcome of this optimization process revealed heightened heat transfer performance coupled with reduced weight in the optimized parallel plate heat sink configuration. Additionally, methodologies grounded in fuzzy logic have been harnessed to quantitatively assess the impact of heat sink design parameters on thermal efficiency [20]. Experimental investigations targeted design parameters of pin-fin heat sinks, encompassing aspects such as fin spacing, pin-fin diameter, and height. Subsequently, Analysis of Variance (ANOVA) was harnessed to explore the influence of these design parameters on critical heat sink attributes including thermal resistance, pressure drop, and average heat transfer coefficient.

In a distinct vein, Chaing and Chang leveraged the Response Surface Method (RSM) to achieve optimum design parameters for pin-fin heat sinks, thereby attaining elevated thermal performance [21]. The quest for minimizing entropy generation rate as a designated objective function was undertaken by Chen and colleagues [22]. To this end, they optimized a plate-fin heat sink for CPU usage by employing a coded genetic algorithm, thus identifying the optimal design parameters for the heat sink. Addressing diverse objective functions, Devi et al.

employed a Taguchi-based non-gradient method to minimize three distinct objective functions—namely, radiation emission, thermal resistance, and heat sink mass [23].

Nonetheless, the existing body of research exhibits a marked generality in terms of the applicability of heat sink configurations. To the authors' knowledge, no study within the current literature has addressed the optimization of heat sinks with a focus on accommodating the diverse thermophysical properties of immersion fluids within immersion-cooled server systems that rely on natural convection cooling.

Drawing from Kim et al. (2013) [24] and the base specifications elucidated by Intel within the Open Compute Project (OCP) forum [25], the governing equations underpinning the phenomena of conduction and natural convection across the heat sink fins are provided below. It is important to note that the OCP publication extensively deliberates on the assessment of diverse fluids through a figure of merit, crucial for ascertaining the natural convection attributes, grounded upon the distinctive fluid properties at specific temperatures.

In light of this, a clear demarcation can be established between the variables contingent upon fluid characteristics and the design parameters intrinsic to the heat sink configuration. This present study is primarily dedicated to forging connections between these factors, with the ultimate objective of promptly predicting the attributes of the heat sink predicated upon the designated fluid properties.

$$Nu_{fin} = \left[(0.09112El^{0.6822})^{-3.5} + (0.5170El^{0.2813})^{-3.5} \right]^{-1/3.5}$$

$$El = \text{Elenbass number} = \frac{g \beta (T_w - T_{amb}) W_c^4}{Lv^2}$$

$$h = \frac{k}{L} \left(\frac{g \beta \Delta T \cdot Pr \cdot W_c^4}{Lv^2} \right)^n$$

FOM for natural convection $\left(\text{FOM1} \right) = k \left(\frac{\beta c_p \rho^2}{\mu k} \right)^n$ where $n=0.2813$ \rightarrow Fluid Variables

$\left(\text{Heat sink design parameters} = W_c^{4n} / L^n \right) \rightarrow$ Heat Sink variables

This investigation seeks to meticulously explore various permutations of design parameters and objective functions, specifically targeting the optimization of a parallel plate-fin heat sink tailored for employment in immersion cooling scenarios. A Computational Fluid Dynamics (CFD) model of the server was developed using ANSYS Icepak, with validation against experimental data from a previously published work, thereby confirming the thermal performance accuracy of the model.

Preliminary Computational Fluid Dynamics (CFD) simulations were executed with the heat sink submerged within a range of dielectric fluids, specifically PAO6, EC-100, EC110, and EC120 [26]. These selections were deliberate, as these fluids were chosen due to the congruent manner in which most temperature-dependent properties fluctuate, with the notable exception of fluid viscosity as shown in Figure 9-1.

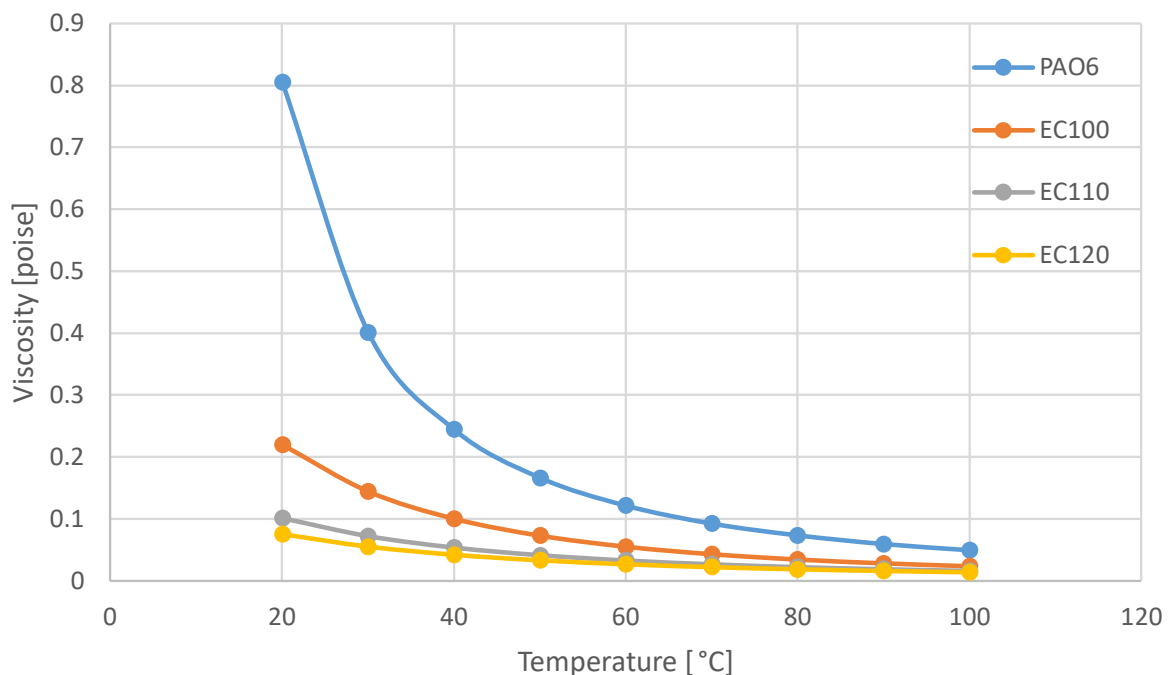


Figure 9-1 Viscosity Vs Fluid Temperature

Typically, the comprehensive dimensions of a heat sink are typically governed by the intricacies of server design, the positioning of the CPU, and the configuration of the CPU socket. In alignment with the tenets of this investigation, we have adhered to a fixed

specification for the length, width, overall height, and base height dimensions. This specification closely adheres to Intel's established reference design for Icelake CPUs. Through systematic modifications to the heat sink fin thickness, fin count, fin height (H), and base thickness, diverse combinations were examined while keeping the total thermal power (TTV) constant. By defining a spectrum of values for these heat sink design parameters within Icepak, a Design of Experiment (DOE) was established, predicated on the parameter range. Subsequently, these designated design points were iteratively solved, leading to the generation of response surfaces as well as plots illustrating the impact of fin spacing (w_c) on thermal resistance.

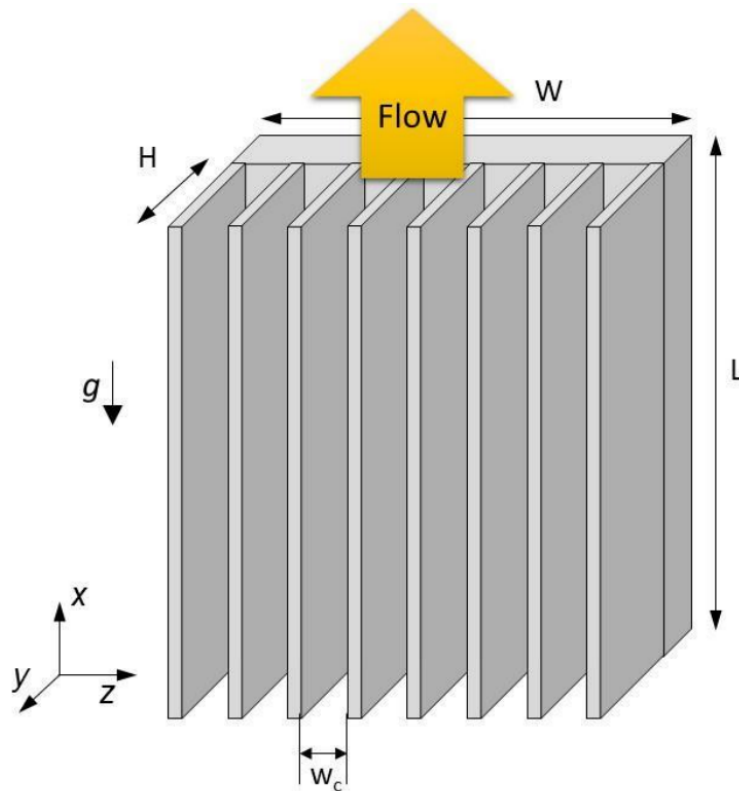


Figure 9-2 Heat sink design variables [25]

Thorough scrutiny of the influence exerted by individual design variables on the designated objective functions was executed, thereby discerning the parameters that wield a

predominant influence on the performance of aluminum heat sinks under natural convection conditions for the specific fluids selected for this study.

9.2 Computational Model

A computational model encompassing a Thermal Test Vehicle (TTV) board along with an integrated heatsink was meticulously formulated within the ANSYS Icepak software. This endeavor involved streamlining the representation of the server system, focusing exclusively on components that wield a critical influence on the heat transfer mechanism, such as the Central Processing Unit (CPU) and the associated heat sink. The configuration of the CPU/TTV unit within the Computational Fluid Dynamics (CFD) model was meticulously devised, with its foundation rooted in the heat flux values emanating from Intel Icelake CPUs.

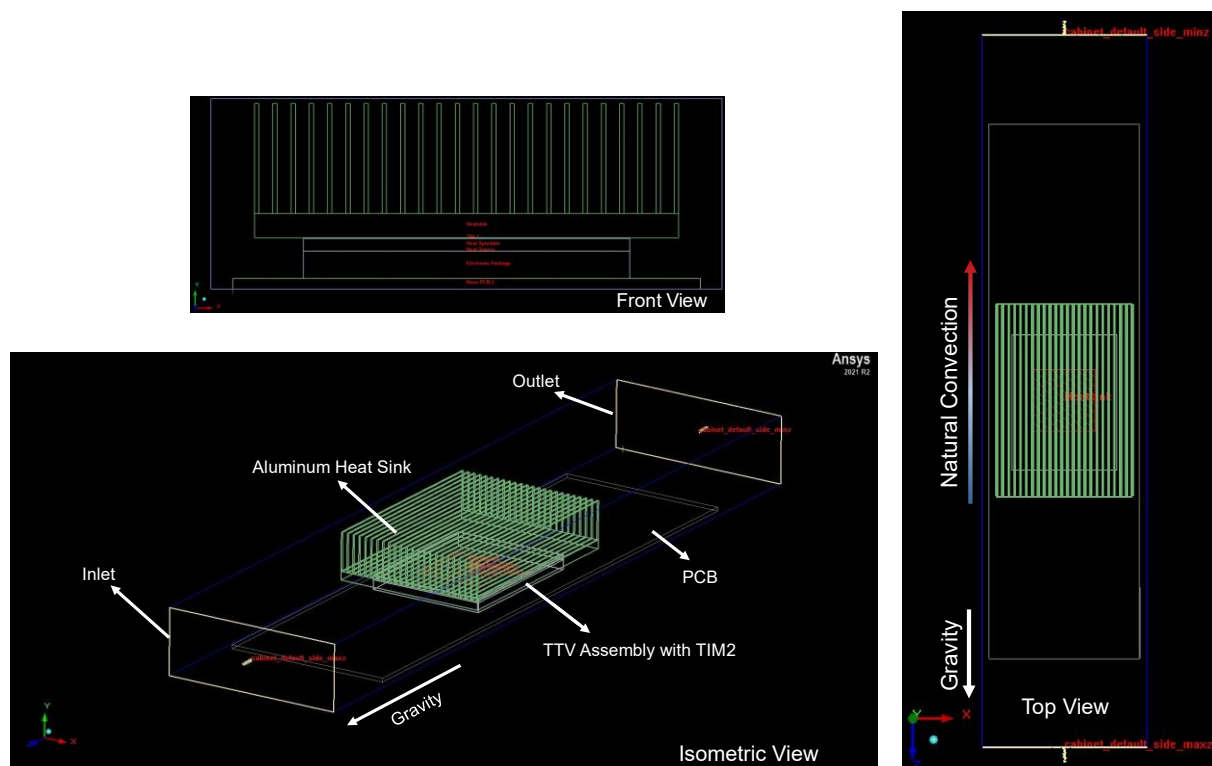


Figure 9-3 CFD model in Ansys Icepak

The thermal structure, symbolizing the CPU, was composed of a 2-Dimensional heat source situated at the lowermost stratum of a heat sink. This assembly featured an interposed

layer of Indium foil, functioning as a Thermal Interface Material (TIM), characterized by a thickness of 0.125 mm and a thermal conductivity coefficient of 8 W/mK. To establish the credibility of the CFD model, an initial validation step was undertaken. This validation procedure was conducted through a comprehensive comparison between the CFD model's predictions and outcomes from a previously published investigation executed by Intel, wherein an analogous configuration was studied [27].

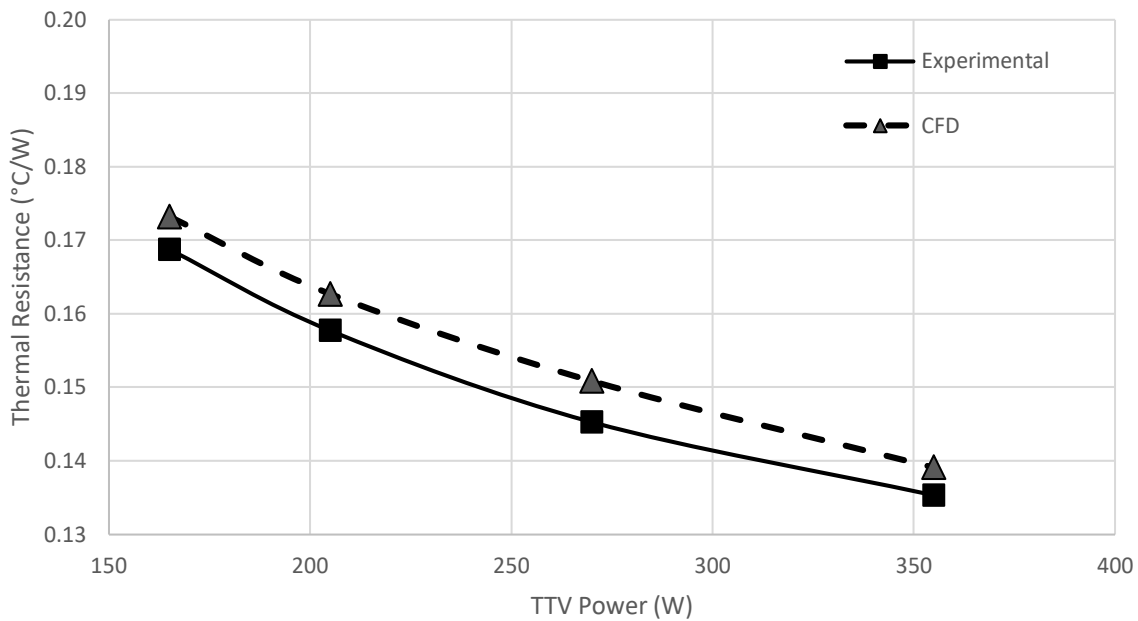


Figure 9-4 Model Validation

Displayed in Figure 9-4 are the thermal resistance measurements contingent upon the temperature at the casing under conditions of natural convection. Notably, a marginal divergence of 3.9% in thermal resistance values emerged between the CFD analysis and the empirical study. Discrepancies apparent in the outcomes could potentially be attributed to the absence of heat pipes at the heat sink's base within the model, a feature that facilitated enhanced heat spreading during the practical experiment. It should be acknowledged, however, that the integration of heat pipes within the CFD analysis was consciously omitted in order to preserve a simplified heat sink design and expedite the optimization procedure.

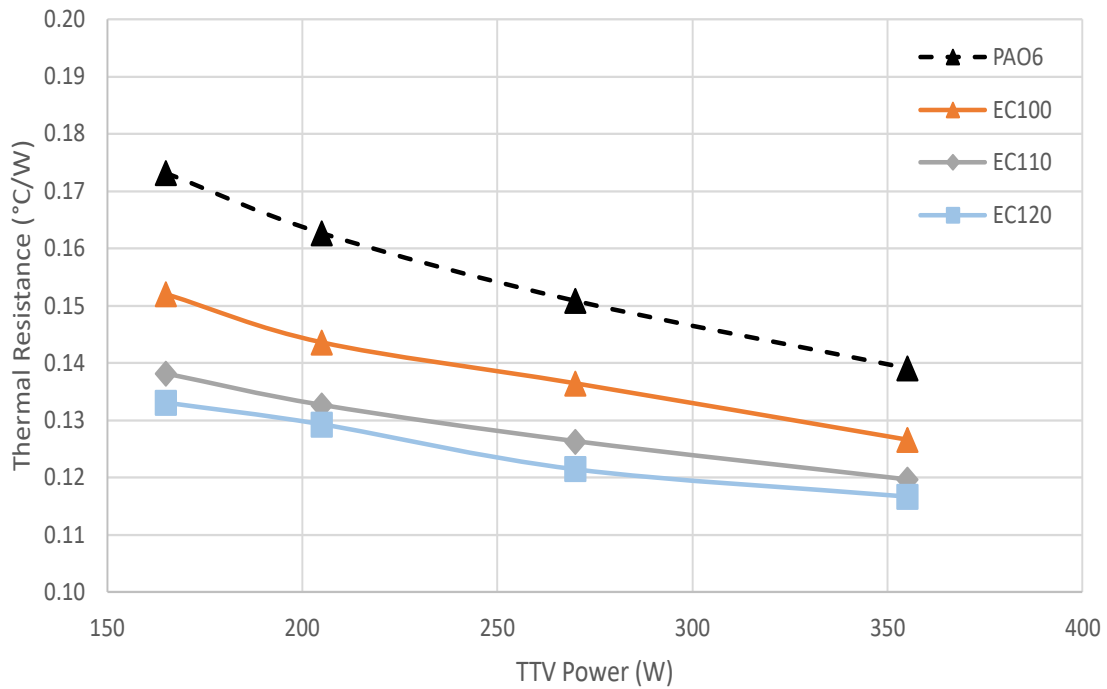


Figure 9-5 Thermal resistance for selected immersion compatible fluids

A noteworthy observation emerged from the analysis, revealing that the thermal resistance associated with the range of EC fluids consistently exhibited marked reductions in comparison to the thermal resistance of PAO 6 at a temperature of 40°C, encompassing all TTV power scenarios, as visually presented in Figure 9-5. This distinctive outcome can be attributed to the discernible disparities in the thermophysical attributes specifically pertaining to viscosity variations, characterizing the EC range of fluids when juxtaposed with PAO 6.

Table 9-1 Thermal resistance and viscosity variation for selected immersion compatible fluids

Fluids at 40°C	Viscosity (poise)	X Viscosity of EC120	(350 W) R_{th} (°C/W)	(350 W) X R_{th} (°C/W) of EC120
PAO6	0.244	5.8 x	0.1391	1.27 x
EC100	0.100	2.4 x	0.124	1.13 x
EC110	0.053	1.3 x	0.1139	1.04 x
EC120	0.042	1 x	0.1093	1 x

The divergent viscosity variation with respect to temperature of the EC fluid series manifestly contributed to the substantiated decrease in thermal resistance observed across the scenarios as shown in Table 9-1.

9.3 Methodology

This section expounds upon the fundamental methodology employed, the associated assumptions, and the range of parametric scenarios investigated in the context of optimized heat sink configurations. The optimization process targeted heat sinks designed for natural convection conditions and was executed at the pinnacle of performance demand, denoted by a power utilization of 350W for CPU/TTV.

The baseline configuration of the heat sink was characterized by a fin count of 24, a base thickness measuring 4.5 mm, fins with a thickness of 0.8 mm, and a fin height of 20.2 mm. For the Computational Fluid Dynamics (CFD) study, the selected dielectric fluids were PAO6, EC100, EC110, and EC120, all of which are commercially available synthetic fluids compatible with immersion setups. CFD results from PAO6 were used as a reference point to evaluate the impact of viscosity variation on the heat sink design variables.

Table 9-2 Fin spacing for different designs of Heat Sink

		Fin Spacing (mm)					
		22	24	26	28	30	32
Fin thickness (mm)	0.6	3.08	2.765	2.5	2.27	2.07	1.89
	0.8	2.876	2.55	2.28	2.06	1.86	1.69
	1	2.67	2.34	2.08	1.85	1.65	1.48
	1.2	2.45	2.14	1.872	1.64	1.44	1.27
	1.4	2.24	1.93	1.664	1.43	1.24	1.07

Increasing the number of fins while keeping fin thickness unchanged reduces fin spacing, and vice versa – altering the fin count and fin thickness adjusts fin spacing as shown in Table 9-2. Lower fin spacing with higher fin count increases the overall heat transfer area of the heat sink fins. However, this shift is accompanied by a reduction in the volume of fluid residing amidst the fins, which consequently restricts the fluid flow pathways within the heat sink. Conversely, selecting larger fin spacing leads to a contraction of the aggregate heat transfer surface area, ultimately resulting in suboptimal heat transfer performance. Therefore, in natural convection dominant mode an optimal fin count and fin thickness is necessary for a given overall dimension of the heat sink to achieve maximum heat transfer performance.

The same heat sink design parameters that were already optimized for PAO6 was used as the baseline or reference. The fin count and fin thickness were further increased to observe the thermal performance under different fluids. This specific approach was adopted in this study to effectively quantify the disparities between optimized heat sink designs and to elucidate the impact of the fluid properties on the designated design variables. Consequently, these outcomes will facilitate a streamlined selection process for the optimal design parameters, catering to the specific thermos-physical properties inherent to the chosen fluid medium.

9.4 Results and discussion

Upon conducting an exhaustive array of permutations involving the heat sink's design variables, a specific combination of fin thickness and fin count was discerned to yield the lowest thermal resistance. This phenomenon can be attributed to the predominant role of natural convection as the driving force behind the heat transfer capability of the heat sink. To conduct a more in-depth scrutiny of the design variables, a strategic selection was made to focus on the fin spacing. This choice was made to facilitate a comparative analysis of the thermal performance, specifically thermal resistance, for the chosen fluids.

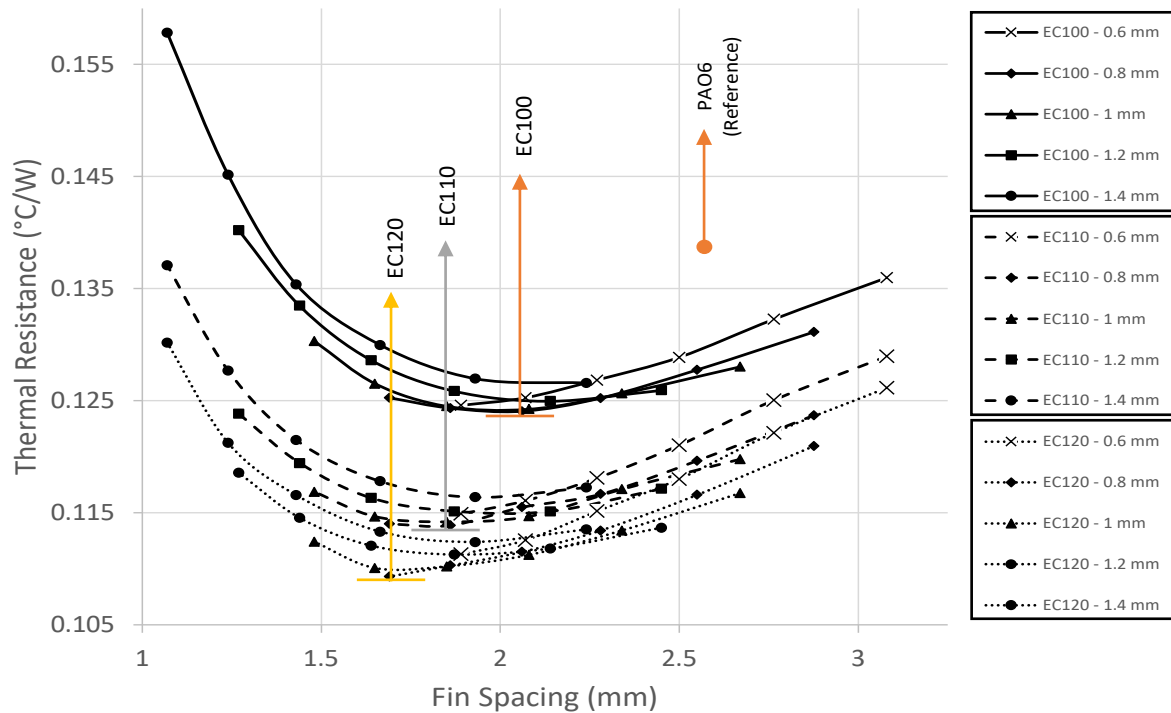


Figure 9-6 Thermal resistance for selected immersion compatible fluids

From this thorough analysis, a coherent pattern consistently surfaced, demonstrating that elevating the fin spacing—achieved by augmenting the fin count while holding the fin thickness steady—conducted to a gradual reduction in thermal resistance. This reduction led to an ultimate minimum value within a specified range of fin spacing. However, beyond a certain point of fin spacing, the thermal resistance exhibited a subsequent increase across all the considered fluids. It is crucial to note that the fluid chosen for this study exhibits a uniform variation in all its thermos-physical properties with temperature, except for fluid viscosity. This enabled us to confidently establish a correlation between the fluctuations in viscosity and the corresponding shifts in fin spacing in convection dominant regime, within the context of the pre-defined overall design configuration of the heat sink.

Further studies in this topic would include various fluids, all temperature dependent properties connected to heat sink fin design variables, impact of heat flux of the CPU/TTV, dependencies of fluid properties on heat sink in forced convection.

9.5 Conclusion

The insights derived from this study can be extrapolated to encompass a variety of single-phase immersion-compatible fluids, thereby constructing a comprehensive tool encapsulating a gamut of heat sink design possibilities, the heat flux channeled to the TTV, and the heat sink material composition. Through the assembly of a robust database, users would be facilitated in directly selecting a heat sink design tailored to the fluid property variations—an accomplishment aligned with the research objectives.

REFERENCES

- [1] <https://www.internetlivestats.com/>
- [2] Masanet, E., Shehabi, A., Lei, N., Smith, S. and Koomey, J., Recalibrating global data center energy-use estimates. *Science*, vol 367 no. 6481, pp.984-986, <https://doi.org/10.1126/science.aba3758>, 2020
- [3] Shehabi, A., Smith, S.J., Masanet, E. and Koomey, J., Data center growth in the United States: decoupling the demand for services from electricity use, *Environmental Research Letters*, vol 13 no. 12, p.124030, <https://doi.org/10.1088/1748-9326/aaec9c>, 2018
- [4] Greenberg, S., Mills, E., Tschudi, B., Rumsey, P. and Myatt, B., Best practices for data centers: Lessons learned from benchmarking 22 data centers, *Proceedings of the ACEEE Summer Study on Energy Efficiency in Buildings in Asilomar, CA*. ACEEE, August, 3, pp.76-87, 2006
- [5] Siddik, M.A.B., Shehabi, A. and Marston, L., 2021, The environmental footprint of data centers in the United States. *Environmental Research Letters*, vol 16(6), <https://doi.org/10.1088/1748-9326/abfba1>, 2021

- [6] Shahi, P., Agarwal, S., Saini, S., Niazmand, A., Bansode, P., & Agonafer, D., CFD Analysis on Liquid Cooled Cold Plate Using Copper Nanoparticles, Proceedings of the ASME 2020 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems. ASME 2020 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems. Virtual, Online. October 27–29, 2020. V001T08A007. ASME. <https://doi.org/10.1115/IPACK2020-2592>, 2020
- [7] Niazmand, A., Murthy, P., Saini, S., Shahi, P., Bansode, P., & Agonafer, D., Numerical Analysis of Oil Immersion Cooling of a Server Using Mineral Oil and Al₂O₃ Nanofluid, Proceedings of the ASME 2020 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems. ASME 2020 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems. Virtual, Online. October 27–29, 2020. V001T08A009. ASME. <https://doi.org/10.1115/IPACK2020-2662>, 2020
- [8] Shahi, P., Saini, S., Bansode, P., Agonafer, D., A Comparative Study of Energy Savings in a Liquid-Cooled Server by Dynamic Control of Coolant Flow Rate at Server Level, in IEEE Transactions on Components, Packaging and Manufacturing Technology, vol. 11, no. 4, pp. 616-624, doi: 10.1109/TCPMT.2021.3067045, 2021
- [9] Shahi, P., Deshmukh, A. P., Hurnekar, H. Y., Saini, S., Bansode, P., Kasukurthy, R., and Agonafer, D., Design, Development, and Characterization of a Flow Control Device for Dynamic Cooling of Liquid-Cooled Servers, ASME. J. Electron. Packag., vol 144(4): 041008. <https://doi.org/10.1115/1.4052324>, 2022
- [10] Niazmand, A., Chauhan, T., Saini, S., Shahi, P., Bansode, P.V., & Agonafer, D., CFD Simulation of Two-Phase Immersion Cooling Using FC-72 Dielectric Fluid, Proceedings

of the ASME 2020 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems. ASME 2020 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems. Virtual, Online. October 27–29, 2020. V001T07A009. ASME. <https://doi.org/10.1115/IPACK2020-2595>

- [11] Hoang, C.H., Hoang, C.H., Tradat, M., Manaserh, Y., Ramakrisnan, B., Rangarajan, S., Hadad, Y., Schiffres, S. and Sammakia, B., A Review of Recent Developments in Pumped Two-Phase Cooling Technologies for Electronic Devices," in IEEE Transactions on Components, Packaging and Manufacturing Technology, vol. 11, no. 10, pp. 1565-1582, doi: 10.1109/TCPMT.2021.3117572, 2021
- [12] Shah, J.M., Eiland, R., Siddarth, A. and Agonafer, D., Effects of mineral oil immersion cooling on IT equipment reliability and reliability enhancements to data center operations. In 2016 15th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), 31 May-3 June, pp. 316-325, IEEE, <https://doi.org/10.1109/ITHERM.2016.7517566>, 2016
- [13] Eiland, R., Fernandes, J., Vallejo, M., Agonafer, D. and Mulay, V., "Flow Rate and inlet temperature considerations for direct immersion of a single server in mineral oil," Fourteenth Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), 2014, pp. 706-714, doi: 10.1109/ITHERM.2014.6892350, 2014
- [14] Shah, J. M., Padmanaban, K., Singh, H., Duraisamy Asokan, S., Saini, S., and Agonafer, D., Evaluating the Reliability of Passive Server Components for Single-Phase Immersion Cooling, ASME. J. Electron. Packag., 144(2): 021109. <https://doi.org/10.1115/1.4052536>, 2022

- [15] Yazicioğlu, B. and Yüncü, H., 2007. Optimum fin spacing of rectangular fins on a vertical base in free convection heat transfer. *Heat and Mass Transfer*, 44(1), pp.11-21, <https://doi.org/10.1007/s00231-006-0207-6>, 2007
- [16] Jang, D., Yu, S.H. and Lee, K.S., 2012. Multidisciplinary optimization of a pin-fin radial heat sink for LED lighting applications. *International Journal of Heat and Mass Transfer*, 55(4), pp. 515-521, <https://doi.org/10.1016/j.ijheatmasstransfer.2011.11.016>, 2012
- [17] Kim, D.K., Thermal optimization of plate-fin heat sinks with fins of variable thickness under natural convection, *International journal of heat and mass transfer*, 55(4), pp.752-761, <https://doi.org/10.1016/j.ijheatmasstransfer.2011.10.034>, 2012
- [18] Subasi, A., Sahin, B. and Kaymaz, I., Multi-objective optimization of a honeycomb heat sink using Response Surface Method. *International Journal of Heat and Mass Transfer*, 101, pp.295-302, <https://doi.org/10.1016/j.ijheatmasstransfer.2016.05.012>, 2016
- [19] Chen, C.T. and Chen, H.I., 2013. Multi-objective optimization design of plate-fin heat sinks using a direction-based genetic algorithm. *Journal of the Taiwan Institute of chemical Engineers*, 44(2), pp.257-265, <https://doi.org/10.1016/j.jtice.2012.11.012>, 2013.
- [20] Chiang, K.T., Chang, F.P. and Tsai, T.C., 2006. Optimum design parameters of Pin-Fin heat sink using the grey-fuzzy logic based on the orthogonal arrays. *International communications in heat and mass transfer*, 33(6), pp.744-752, <https://doi.org/10.1016/j.icheatmasstransfer.2006.02.011>, 2006
- [21] Chiang, K.T. and Chang, F.P., 2006. Application of response surface methodology in the parametric optimization of a pin-fin type heat sink. *International communications in heat and mass transfer*, 33(7), pp.836-845, <https://doi.org/10.1016/j.icheatmasstransfer.2006.04.011>, 2006

- [22] Chen, C.T., Wu, C.K. and Hwang, C., 2008. Optimal design and control of CPU heat sink processes. *IEEE Transactions on components and Packaging Technologies*, 31(1), pp.184-195, <https://doi.org/10.1109/TCAPT.2008.916855>, 2008
- [23] Devi, S.P., Manivannan, S. and Rao, K.S., 2012. Comparison of nongradient methods with hybrid Taguchi-based epsilon constraint method for multiobjective optimization of cylindrical fin heat sink. *The International Journal of Advanced Manufacturing Technology*, 63(9-12), pp.1081-1094, DOI 10.1007/s00170-012-3985-7, 2012
- [24] Kim, T. H., Kim, D. K., and Do, K. H., 2013, "Correlation for the fin Nusselt number of natural convective heat sinks with vertically oriented plate-fins," *J. of Heat Mass Transfer*, Vol. 49, pp. 413–425
- [25] <https://www.opencompute.org/documents/ocp-base-specification-for-immersion-fluids-20221201-pdf>
- [26] Engineered Fluids, EC-100 Dielectric Coolant, <https://engineeredfluids.store/products/ec-100>
- [27] S. Sarangi, E. D. McAfee, D. G. Damm and J. Gullbrand, "Single-Phase Immersion Cooling Performance in Intel Servers with Immersion Influenced Heatsink Design," 2022 38th Semiconductor Thermal Measurement, Modeling & Management Symposium (SEMI-THERM), San Jose, CA, USA, 2022, pp. 1-5.

BIOGRAPHICAL INFORMATION

Vibin Shalom Simon received his Bachelor of Engineering Degree in Mechanical Engineering from Anna University (Loyola-ICAM College of Engineering and Technology) in 2016. He received his Master of Science Degree in Mechanical Engineering from The University of Texas at Arlington in 2019. Vibin served as a Graduate Research Assistant at Electronics, MEMS, and Nanoelectronics Packaging Center (EMNSPC) and worked on various industry funded projects from 2019 to 2023. His PhD work constitutes evaluation and optimization of various thermal solutions in air and liquid cooling for electronics focusing data center industry. He also had an opportunity to work at Intel Corporation as a Thermal Analysis Intern for a year (2022 – 2023).