THERMO-MECHANICAL DESIGN CONSIDERATIONS AT THE SERVER AND RACK LEVEL TO ACHIEVE MAXIMUM DATA CENTER ENERGY EFFICIENCY

by

RICHARD MARK EILAND

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

May 2015

Acknowledgements

I would like to thank my advisor, Dr. Dereje Agonafer, for providing me with abundant opportunities to learn and develop my skills as an engineer in the electronic packaging and data center field. His support and guidance have been instrumental to my success and reaching this milestone.

I am grateful to Dr. Haji-Sheikh, Dr. Tong, Dr. Çelik-Butler, and Dr. Mulay for serving on my dissertation committee. Thank you for your support and guidance throughout the process. A special thanks to Dr. Veerendra Mulay, whose support of the EMNSPC lab has been tremendous. He has also provided me with technical mentorship and served as a role model for transitioning from graduate student to successful professional.

The day-to-day requirements of graduate life were made simpler by the generosity and assistance of Ms. Sally Thomson. Her support in all matters administrative and morally were essential in me making it through these past several years. My colleagues in the EMNSPC lab have also been a great source of support in pursuing this degree. Special thanks to John Fernandes, whom I worked closely with on many projects and helped me grow through our thought-provoking discussions. Also, thanks to Marianna Vallejo for sharing technical discussions and unwavering support throughout this journey.

Finally, I must thank my parents, Kathy Garrett and Mark Eiland, and family for their endless love and support in all that I do. They have provided me with the firm foundation in life to take on any challenge and accomplish my goals.

April 21, 2015

Abstract

THERMO-MECHANICAL DESIGN CONSIDERATIONS AT THE SERVER AND RACK LEVEL TO ACHIEVE MAXIMUM DATA CENTER ENERGY EFFICNEICY

Richard Mark Eiland, PhD

The University of Texas at Arlington, 2015

Supervising Professor: Dereje Agonafer

Continually increasing demand for information technology (IT) applications and services has provided sustained growth and interest in data centers. The large amounts of energy consumed by data center facilities have placed a significant emphasis on the energy efficiency of their overall operation. One area of particular importance is the cooling energy required. Heat generation within a data center starts at the server level, specifically within the microelectronic devices that process digital information. Convective heat transfer is the primary driver for the removal of heat from an individual server. As such, cooling efficiency at the server level will be dictated by the pumping power required to move a cooling fluid through the system. Many methods are available for removing heat from the server, either with air or liquids as the cooling medium. This work evaluates new, efficient approaches for removing that heat and the pertinent design considerations that must be taken into account for successful implementation.

In general, smaller fans operate at lower efficiencies than larger fans of proportional linear dimensions. The applicability of replacing smaller, 60mm fans from within the chassis of web servers with an array of either 80mm or 120mm fans consolidated to the back of a rack is experimentally tested. Initial characterization of the

selected fans showed the larger 80mm and 120mm fans operate at double peak total efficiency of the smaller 60mm fans. A stack of four servers was used in a laboratory setting to represent a rack of servers. When all four servers were stressed at uniform computational loadings, the 80mm fan array resulted in between 50.1% to 52.6% reduction in total rack fan power compared to the baseline 60mm fans. The 120mm fan array showed similar reduction in rack fan power of 47.6% to 54.0% over the baseline 60mm fan configuration. Since actual data centers rarely operate at uniform computational loading across servers in a rack, a worst case scenario test was conceived. In this test, the arrays of larger fans were controlled by a single server operating at peak computational workload while the other three in the rack remained idle. Despite significant overcooling in the three idle servers, the 80mm and 120mm fan configurations still showed 35.3% and 33.8% reduction in total rack fan power compared to the best possible operation of the 60mm fans. The findings in this study strongly suggest that a rack-level fan scheme in which servers share airflow is more efficient alternative to fans contained within the server.

Air flow management is a critical tool to maintain efficient operation of a data center cooling scheme. Provisioning of airflow from CRAC units and containment systems often lead to changes in the static pressure at the inlet to server racks. Through experimental testing on an Air Flow Bench it is observer that static pressure at the inlet to servers has a significant influence on the thermal performance and fan cooling energy consumption within the server itself. Reduction in server fan power or component temperatures can be achieved by increasing the static pressure at the server inlet. Complementary design and control at the room level with this information at the server level can lead to reduction in overall system fan power and more energy efficient data center operation.

Complete immersion of servers in dielectric mineral oil has recently become a promising technique for minimizing cooling energy consumption in data centers. However, a lack of sufficient published data and long term documentation of oil immersion cooling performance makes most data center operators hesitant to apply these approaches to their mission critical facilities. In this study, a single server was fully submerged horizontally in mineral oil. Experiments were conducted to observe the effects of varying the volumetric flow rate and oil inlet temperature on the thermal performance and power consumption of the server. Specifically, temperature measurements of the CPUs, motherboard components, and bulk fluid were recorded at steady state conditions. These results provide an initial bounding envelope of environmental conditions suitable for an oil immersion data center. Comparing the results from baseline tests performed with traditional air cooing, the technology shows a 34.4% reduction in the thermal resistance of the system. The cooling loop was able to achieve partial power usage effectiveness ($pPUE_{Cooling}$) values as low as 1.03. This server-level study provides a preview of possible facility energy savings by utilizing high temperature, low flow rate oil for cooling. Following this, visual observations, microscopic measurements, and testing of mechanical properties were taken. Evaluation of the technology's impact on the mechanical reliability of components and operability of data centers is made.

Table of Contents

List of Illustrations

List of Tables

Chapter 1

Introduction

Behind the scenes of the internet services (search, email, social networks, etc.) that have become so pervasive in everyday life are data centers. Data centers are facilities that house the computing, storage, and networking equipment that serve the Information Technology (IT) needs of modern society. Almost all industries, ranging from health care, finance, government, entertainment, to enterprise businesses rely on these services. Data center facilities range in size from a few hundred square feet to over one million square feet. The environment inside data centers differs from other facilities which are built for human comfort or industrial use. Rows of racks filled with IT equipment (ITE) occupy the floor space, as seen in Figure 1-1, with much higher power density than traditional spaces. Compared to typical office buildings, schools, and commercial buildings, the power (Watts) consumed per square foot of floor space in a data center can be as much as an order of magnitude higher. These high power densities present several challenges from power delivery, cooling infrastructure, and financial perspectives. Industry trends of increasing power density to maximize productivity per square foot of floor space only exacerbate these problems. With the continued increase in demand for information technology services, all of these topics are areas of active research. The focus of this work is on evaluating more energy efficient approaches for the cooling of ITE and represents a possible paradigm shift in how data centers, especially those with high power densities, are designed and operated. The first portion of this work seeks to modify and improve traditional air cooling methods of servers at the chassis and rack level. The second part involves complete immersion of ITE into an electrically nonconductive mineral oil which circulates at a low flow rate to remove heat. An initial evaluation on the

possible impact of this fluid on the mechanical reliability and operability of ITE is also considered.

Figure 1-1: Typical interior floor space of a large data center

1.1 Data Center Power and Cooling Trends

Rapidly increasing demand for IT services led to a 56% increase in total electricity used by data centers globally over the five year period from 2005 to 2010 [1]. As an industry, data centers accounted for an estimated 1.7% to 2.2% of the total electricity used in the US and 1.1% to 1.5% globally in 2010 [1]. These numbers are expected to rise with the persistent growth of the industry. Such large proportions of energy consumed by a single industry have raised concerns over efficiencies and the sources from which they receive electrical energy.

The standard metric for gauging the overall efficiency of a data center is power usage effectiveness (PUE) which was developed by the Green Grid in 2008 [2] [3]. PUE determines efficiency by dividing total power consumed by the data center by the power consumed by IT load. IT load is considered the useful work of the data center facility. Mathematically, this metric is expressed as:

$$PUE = \frac{Total\ Facility\ Power}{IT\ Equipmet\ Power} \quad (1\text{-}1)$$

The numerator of this expression can be broken down into its component parts: power delivery systems, cooling infrastructure, IT equipment power, lighting loads, and miscellaneous power consumers throughout the facility. A 2013 survey by Digital Realty of North America data center managers found an average PUE value of 2.9 [4]. A more recent survey in 2014 by the Uptime Institute of data center owners and operator found an average self-reported PUE value of 1.7 [5]. Although there is large variation between these two values, the range indicates moderate to very inefficient operation of data center facilities. It should be noted that not all data centers operate with such inefficiencies. State of the art data center designs have the opportunity to reduce their electricity use by nearly half [6] [7]. Hyper-scale cloud computing and mega data centers built by industry leading companies typically operate with the best-in-class designs. However, they only represent a very small portion (<5%) of the overall energy consumed by the data center market [8].

Typical distribution of power consumption within a traditional air-cooled data center facility is shown in Figure 1-2 [9]. As can be seen, apart from the IT equipment (the useful work), the largest energy consumer of the data center facility is the cooling infrastructure. This is the focus of the present work and presents a great opportunity to improve overall building efficiency.

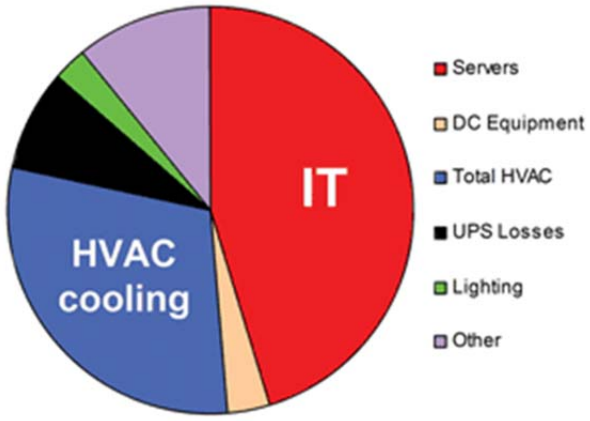

Figure 1-2: Typical energy breakdown in a traditional air cooled data center facility [9]

The leading professional organization responsible for maintaining guidelines and standards for the data center industry is ASHRAE's Technical Committee (TC) 9.9. In addition to establishing operating guidelines for cooling of facilities, TC 9.9 also tracks trends in IT equipment power densities. Figure 1-3, taken from projections made in 2005, indicates that over a twenty year period the power densities of ITE increased significantly, although they began to level off over the last ten [10]. As discussed above, these trends place more strain on the cooling infrastructure of the facility.

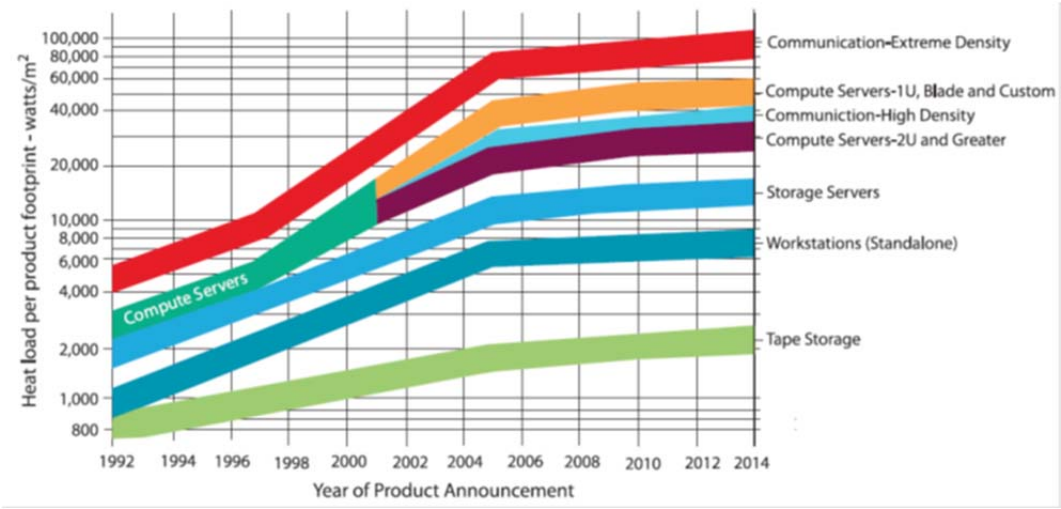

Figure 1-3: Power density trends for various types of information technology equipment (ITE) as published by ASHRAE TC 9.9 [10]

1.2 Cooling of Electronic Packages

At the heart of IT equipment are the microelectronic devices which are typically the primary sources of power consumption and heat generation. As such, these devices are the primary motivator for cooling needs within a data center. The functional limit for safe operation of silicon devices in typical commercial servers range between 85 to 105°C. At temperatures 15 to 25°C higher than this, damage to the device begins to occur [11]. This information forms the starting point for the design of thermal management solutions for ITE and around which the data center cooling system is formed.

A typical packaging architecture for a flip chip package common to today's modern microprocessors (CPUs) is shown in Figure 1-4. This simplified view shows key points in the thermal circuit which removes heat from the silicon die [12] [13]. For the

sake of simplicity, it is assumed that all heat generated from the die travels in the direction opposite of the substrate. The thermal circuit consists of two main parts:

1) Conduction resistance from the silicon die (T_j) to the base of heat sink (T_b) and
2) Convective resistance from the extended surface of the heat sink (T_b) to the ambient fluid used for cooling (T_a).

For the context of this work, the conduction resistance is a fixed parameter dictated by the design and construction of the CPU by the manufacturer. The conduction resistance will be influenced by the pumping power and hence the flow rate of the cooling fluid.

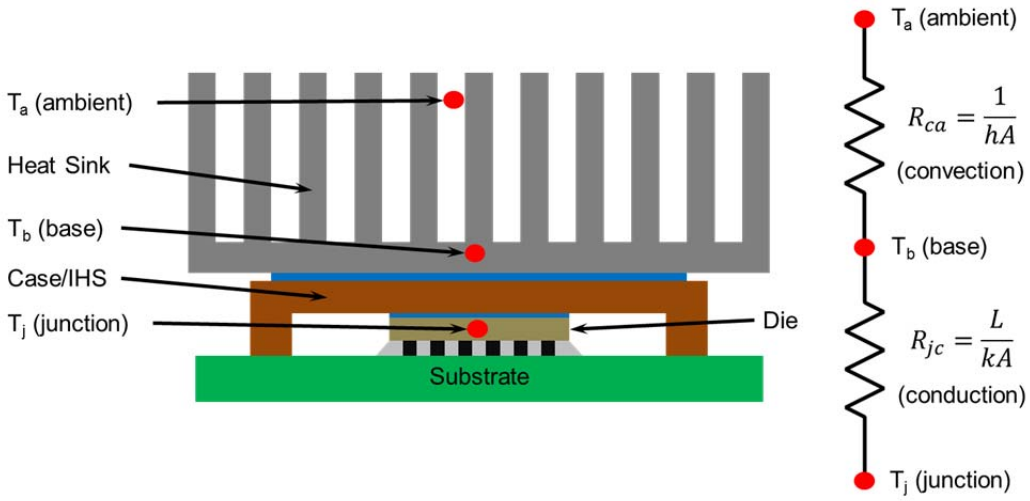

Figure 1-4: Simplified view of typical packaging architecture and thermal circuit for modern flip-chip microprocessors

1.3 Air Cooling for Data Centers

Traditional approaches to cooling data centers use air as the primary cooling medium. Heat rejected from IT hardware is absorbed by the air and either rejected to the

outside ambient, mixed with incoming fresh air, or cooled through refrigeration processes. These techniques are matured fields and well documented with safe environmental conditions established by ASHRAE TC 9.9 [14]. Energy modeling of the heat transfer path, starting at the chip level then to airflow within the room and subsequent heat rejection to building HVAC equipment, provides good understanding of the overall efficiency of the facility [15] [16] [17]. The most common floor layout orients racks in a cold aisle/hot aisle arrangement to isolate cold air entering the front of ITE from the hot air exhaust out the back as shown in Figure 1-5. Cold air is supplied through an underfloor plenum from a computer room air conditioning (CRAC) unit and the hot exhaust circulates back to the return side of the CRAC. Various alternative methods are available for distributing cold air to the inlet of racks [18] [19]. Significant work has been done to understand the physics and dynamics of airflow distribution within a data center [20] [21] [22] [23]. Best practices dictate strict separation of the cold and hot air streams to achieve the highest cooling efficiencies possible [24] [25] [26]. Any leakage, bypass, or recirculation around ITE represents wasted effort by the cooling resources. Extensive efforts to model and understand the benefits of aisle containment strategies highlight the impact of simple geometric modifications to a room on the overall thermal performance and energy efficiency of a room [27] [28] [29] [30].

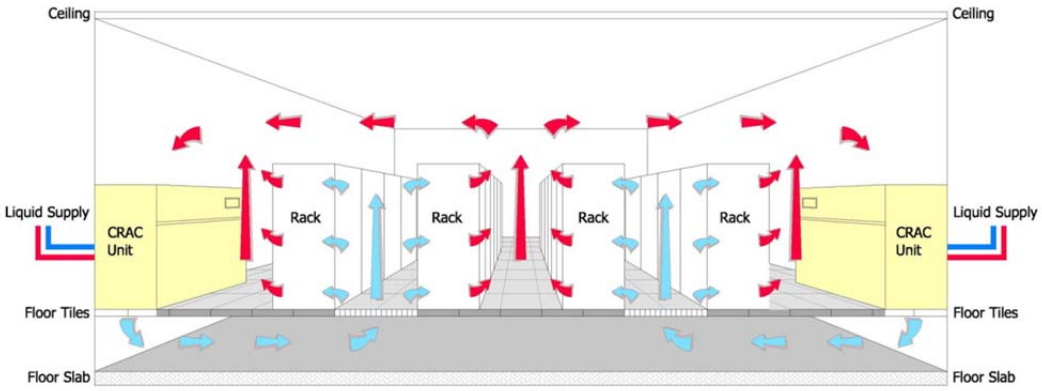

Figure 1-5: Traditional raised floor data center layout in cold aisle/hot aisle rack arrangement [10]

1.4 Limitations of Air Cooling

Data center operators have an opportunity to reduce their total cost of ownership (TCO) by increasing the density of ITE within a space [31]. However, up to a certain limit, continuing to increase heat densities of electronic components and densities at the rack level surpass the capabilities of air for efficient cooling and alternative methods should be sought [32] [33] [34]. ASHRAE TC 9.9 reports that typical air cooled rack power densities are in the range of 6 to 30 kW per rack [35]. Rack densities above this range will typically be beyond the cooling capabilities of legacy air cooling systems. Ellsworth and Iyengar showed that it is possible to cool very high density equipment (60+ kW/rack) with air cooling methods; however, the use of water cooling techniques provided anywhere from 50.1% to 92.2% operational energy savings [36]. Additional benefits of higher computational performance at lower component operating temperatures were also seen in the study [36]. An analytical study by Saini and Webb determined the upper limit heat load for “desktop” chip packages using a parallel plate fin heat sink is roughly 95 to 100W

for a 16 by 16mm heat source [37]. A white paper published by Intel shows that they are able to successfully cool heat densities of 30kW/rack with higher densities possible, but theoretical and practical limits do exist and are specific to each data center's needs [38]. Although the precise point at which air cooling will no longer be feasible is still an area of healthy debate, it is undeniable that the research and development of alternative cooling methods will be beneficial to the industry as a whole.

1.5 Liquid Cooling for Data Centers

Both direct and indirect forms of liquid cooling offer many advantages over conventional air cooling such as higher heat capacities and lower transport energy requirements. Several liquid coolants are available and have been used in the cooling of electronic equipment with varying thermal and hydraulic properties to choose from [39] [40] [41]. Indirect methods, using water as a cooling medium through cold plates or rear door heat exchangers, demonstrate the benefits of a liquid cooling strategy [36] [42]. Direct immersion of electronic equipment offers a singular cooling solution in which the entirety of a server or other ITE may be cooled by a single fluid medium. This may provide simplicity and ease in planning and implementation of a total solution [32]. Some of the most common forms of liquid cooling for data centers are discussed below.

1.5.1 Water cooling

Water cooling allows increased efficiency through use of higher temperature fluids and possible use of waste heat for other applications [43]. Cold plates have been a long standing method of bringing water cooling to high powered devices as demonstrated by the Thermal Conduction Module (TCM) of the 80s [44]. Even today, there is continued

interest in such applications and making this old approach more dynamic for the constantly changing requirements of next generation components [45]. However, the use of cold plates requires air cooling for a portion of the components within servers. Significant developments have been made to bring water cooling to a complete cooling solution (i.e. remove 100% of the ITE heat) for even the most powerful super computers [46]. Water, however, is not a dielectric fluid and as such must be implemented in an indirect way. The additional infrastructure required within the server may influence cost considerations for this technology, although the cost per performance may still be equal or better than alternatives. Ample data and guidelines for implementing water cooled data center environments are available from sources such as ASHRAE TC 9.9 [35].

1.5.2 Immersion cooling

1.5.2.1 Fluorocarbons

A legacy approach to full liquid immersion cooling is the use of dielectric fluorocarbon refrigerants in pool boiling application. These fluids are extremely good electrical insulators and have moderate thermal characteristics [47]. Additionally, their low boiling points make them suitable for two-phase flow applications capable of removing large heat densities [48] [49] [50]. These high heat transfer rates can be further enhanced though special coatings that increase boiling rates [51] [52]. Current research in this area seeks to improve upon already available fluids by enhancing the critical heat flux for two-phase applications [53]. Full scale data center implementations of these techniques show much promise as an energy saving technique [54] [55]. However, there are additional challenges to this technology that are still to be overcome before it will see wide application.

1.5.2.2 Mineral oil

Compared to air, many mineral oils have a heat capacity roughly 1200 times greater. The increased thermal properties, along with their dielectric nature, make mineral oils a possible alternative for data center applications. Mineral oils have long been used as heat transfer fluids with especially large adoption in power delivery applications such as high voltage transformers [56]. The performance of transformer oils as heat transfer fluids may be significantly improved with the addition of nanoparticles, as recently shown by Taha-Tijerina et. al. [57]. A small fraction (<0.100% by weight) of diamond nanoparticles added to mineral oil were shown to increase the thermal conductivity of the base fluid by 40 – 70% while having minimal detrimental impact to flow properties such as viscosity.

However, when looking at the specific application of mineral oil immersion cooling for data centers, limited literature is available. Recent media attention has provided general PUE values [58] [59], but the details of operation are absent. Pruncal provides a helpful overview of general operating benefits of oil-based data centers, mentioning the possibility to use facility chilled water up to 30°C compared to 7 – 13°C for traditional air-based systems [60]. The most thorough account is provided by Patterson and Best which showed a 36% improvement in thermal resistance of an oil immersion system compared to an air cooled counterpart with no adverse mechanical effects [61]. This worked showed successful cooling with oil temperatures up to 43°C and cooling PUEs in the 1.02 – 1.03 range; however, no discussion regarding the volume flow rates used was provided. With only limited data available, a large knowledge gap remains in the industry regarding environmental requirements to operate an oil immersion cooling facility.

A key concern regarding this technology by many industry professionals is the potential adverse impact related to the interaction of the fluid with electronic components and devices over time. Anecdotal discussions suggest there is no negative impact of mineral oil on system components; however, detailed documentation is lacking at this time [62].

Commercially available oil immersion systems are on the market and report very desirable performance of such a cooling system. One vendor showed that their system was capable of operating similarly configured servers at component operating temperatures 20 to 30°C lower than an air-cooled counterpart. Additionally, these results showed reduction in cooling costs up to 98% for their liquid cooled system [63]. Similar numbers are reported by another vendor who claims their system can reach a density of 100kW/rack and the opportunity to reduce cooling costs by 95% with mineral oil submersion cooling [64] [65]. However, the data presented by these sources is limited and the extent to which the system have been designed and optimized is unknown to the industry at large. This present work is supported by key technologists in the industry and the need for more widely available data and information regarding such systems.

1.6 Scope of Dissertation

This work examines the influence of different flow designs within server and rack level cooling strategies. Chapters 2 and 3 explore design strategies which influence or improve cooling efficiency when using an air cooling approach to data centers. In Chapter 2, efficiency is improved by consolidating smaller, chassis-enclosed fans with an array of more efficient, larger fans at the back of a rack. In Chapter 3, the impact of data center pressurization on the cooling performance of servers is discussed and general relations

between cold aisle pressure and fan performance are presented. From this, the impact of pressurization on the thermal design and control of a server in actual operation is discussed.

Chapters 4 and 5 look at the opportunities to improve overall data center efficiency when the using the emerging technology of direct immersion of ITE in mineral oil. The initial considerations for environmental operating conditions, in terms of volume flow rate and fluid inlet temperatures are established in Chapter 4. Chapter 5 examines some of the mechanical concerns that arise with the uncertainty of transitioning to a disruptive technology and how oil cooling may impact mechanical reliability of a server system.

Chapter 6 summarizes the work above and discusses opportunities and future work moving forward that can further enhance the improvements presented here.

Chapter 2

Rack-Level Fans: Savings through Consolidation

Significant improvements have been made to the operation of air movers within IT equipment over the last decade. In typical server systems built prior to 2005, internal server fans frequently consumed 10 to 20% of the total server power draw. Advances in fan designs, fan speed control with the use of pulse width modulation (PWM) signaling, and improved overall server thermal designs have since brought this number down to 2 to 4% in typical operation [11]. However, as with the data center facility in general, cooling power represents a parasitic load that does not contribute to the useful computational work output of a server and must be minimized or eliminated if possible. Typical servers will utilize axial, dual counter rotating, or centrifugal blower type fans as the primary air moving devices within their chassis.

Larger fans of geometrically similar proportions usually operate with higher efficiencies than their smaller counterparts. Various fan manufacturers are diligent in educating customers on this general principle [66] [67]. The typical peak operating efficiency will occur at roughly one-third of the maximum static pressure a fan can deliver. Proper system design and the use of impedance matching techniques can help optimize a system's performance to realize the best possible efficiencies. Work by Holahan and Elison outline the governing the relations between sound, flow, and pressure ratios for fans of homologous dimensions and discuss the principles of impedance matching [68]. They also performed a survey of peak total operating efficiency as a function of frame size for fans typically found in rack systems, which illustrates the trend of increased efficiency with size. Kodama, et. al. mapped the temperature response of computing nodes to the fan speeds in a fan array of a blade chassis [69]. This work highlighted

possible imbalances in CPU temperatures and system efficiency that may exist when a consolidated fan array scheme is not optimized.

A previous study performed by Nagendran and colleagues showed computational evidence (CFD) based on empirical data that significant savings in fan energy are possible through consolidation of smaller, internal server fans to a larger rack-based fan configuration [70]. Additional details of this previous study included findings that a failure of a single fan out of the four in the server under study has the same impact on die temperatures, irrespective of the location of the failed fan. These initial results provide a foundation and motivation for the present work.

The purpose of this work is to establish the feasibility and possible power saving achievable by moving from smaller, chassis enclosed fans, to larger, more efficient fans at the rack-level. To achieve this end, a simulated rack of four servers is setup in a laboratory environment. Three different axial fans sizes are experimentally tested to determine their performance over a range of computational operating conditions. This includes baseline 60mm fans original to the servers and two larger size fans, 80mm or 120mm, moved outside of the server chassis and fixed to a rack-level frame. In this manner, when a consolidated fan array is used, fans will be responsible for providing cooling airflow to multiple servers in the system as opposed to the baseline 60mm fan case in which each server has its own independent air stream. Testing over a non-uniform server computational workload (heat loads) further highlight the feasibility of the proposed modification in a more representative data center scenario.

2.1 Experimental Setup and Procedures

2.1.1 Server under study

For the present study, first generation Intel-based Open Compute servers are used as the platform for testing as seen in Figure 2-1. These particular servers have a highly efficient design in terms of their thermal performance, as well as in their mechanical design, power efficiency, and acquisition costs. Detailed description of the server design can be found in [71] [72]. The key elements of the thermal design include wider spacing of the primary heat generating components (CPUs), an air duct that directs flow over these key components, and a larger chassis height (1.5 rack units). The 1.5U height allows for use of four 60mm x 25mm fans, which are of much higher efficiency than the typical 40mm x 25mm fans found in commodity 1U servers. Under normal operation, the four 60mm fans adjust their speed based on a pulse width modulation (PWM) signal that is dictated by CPU die temperatures. A fan speed control algorithm (FSC) is set in the motherboard's BIOS settings to control CPU die temperatures within a fixed temperature range as shown in Figure 2-2. As the temperature of the CPUs goes above the target range, the FSC increases the PWM signal to the fans at a predetermined rate. When temperature drops below the target range, the PWM signal to the fans is lowered at the predetermined rate.

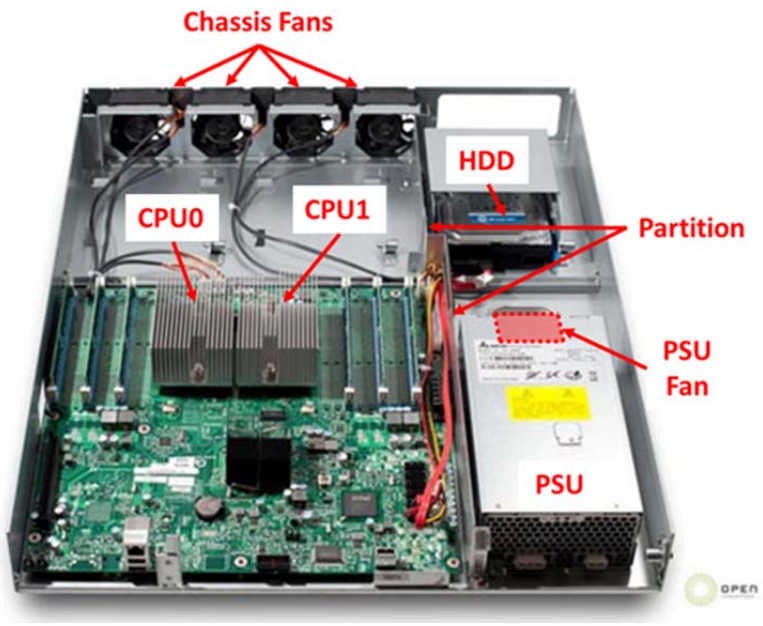

Figure 2-1: Intel-based Open Compute server used for this study

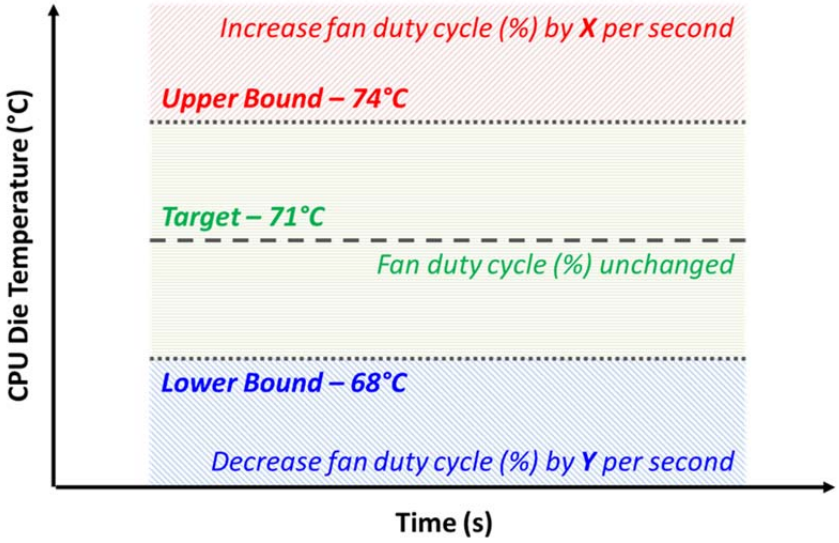

Figure 2-2: Deadband control limits used in the fan speed control algorithm for the server under study

A key feature of these particular servers that enabled simplification in the present analysis is the partitioning of the primary motherboard portion from the power supply (PSU) and hard drive (HDD) section of the server as seen in Figure 2-1. This partition allows for complete compartmentalization of airflow streams between the two sections. For the analysis of this work, air flows provided by the fan configurations under study are the only to impact temperature of the CPUs. Additionally, the physical dimensions of the motherboard section allow for a square geometry to locate a fan array. This area is roughly 330mm x 330mm if four servers are stacked on top of each other. When analyzing total rack fan power in the study, only the fans drawing air across the motherboard are considered. It is assumed that PSU airflow and power consumption will remain constant, independent from the proposed rack-level fan schemes.

2.1.2 Modified, rack-level fan configuration

In order to evaluate the possible energy savings of implementing a larger, rack-level fan system, simple modifications to the baseline servers were made. A stack of four servers were used to simulate a “rack” scenario in the laboratory setting. A simple wooden frame and acrylic panel was constructed to hold the larger fan array at the back of the servers. A preliminary computational study to this work indicated that a minimum distance of 20mm was needed between the back of the servers and the fan array to ensure uniform distribution of air flow between all four server [70]. Following this minimum requirement, the fan array was placed 25.4mm (1in) from the rear of the servers.

The simulated rack of four servers includes a total of sixteen 60mm fans in the baseline scenario. Each of these sixteen fans is enclosed within the chassis of their respective server. In the modified case of 80mm fans, nine total fans are located at the

back of the simulated rack. In the modified case of 120mm fans, four total fans are located at the back of the simulated rack. Figure 2-3 depicts these three different scenarios with both a side view and back view of the simulated rack with the PSU section of the server omitted for simplification.

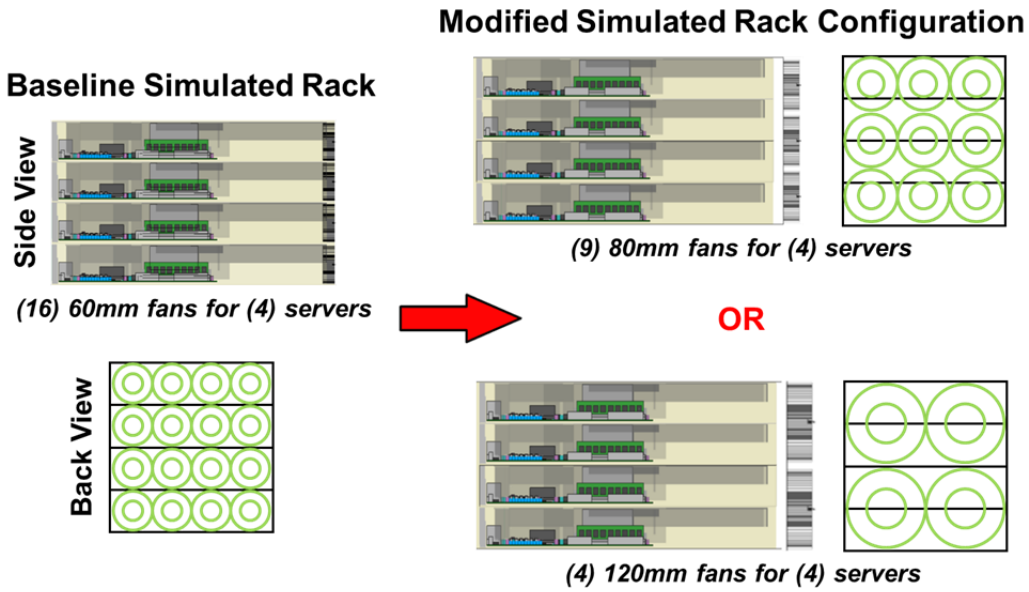

Figure 2-3: Consolidation of fans to a shared array at the back of the rack

Table 2-1 outlines the specifications for each fan size selected for this work. Selection of the larger fans (80mm and 120mm) was based on a requirement that they meet or exceed the air flow rate and static pressure delivered by the smaller 60mm fans for a given rack. An in-depth discussion of the selection process is provided in [70]. Detailed characterization of the fans was performed on an Air Flow Bench test chamber to generate the fan curves and confirm the expected efficiency increase with larger fans. Details of this testing procedure can be found in [70] and [73]. Using the Air Flow Bench test chamber, at each operating point on the fan curve, the fan's total electrical power

draw by a VDC power supply was recorded. Fan efficiency is given as the ratio of the output hydraulic power to electrical input power by Equation 2-1:

$$\eta = \frac{\dot{Q} * \Delta P}{V * I} \tag{2-1}$$

Where hydraulic power is the product of the volume flow rate and differential pressure and electrical power is the product of input voltage by current drawn. Figure 2-4 shows the efficiency curves from each of the three fan sizes when operated at their maximum respective speed. The flow rate is given in terms of the flow rate for the simulated rack which is a parallel configuration of sixteen 60mm fans, nine 80mm fans, or four 120mm fans. It can be seen that the peak efficiency of the 60mm, 80mm, and 120mm fans are 14.8%, 29.3%, and 29.3% respectively. The roughly double efficiency of the larger fans should manifest in significant reduction in cooling power consumption for a given airflow volume.

Table 2-1: Specifications for the selected fans used in this study

Setup	Frame Size (mm)	Maximum Air Flow (cfm)	Maximum Static Pressure (inH_2O)	Rated Speed (rpm)
60mm	60x60x25	31.7	0.62	7600
80mm	80x80x38	100.1	1.98	9500
120mm	120x120x25	171.0	0.90	5100

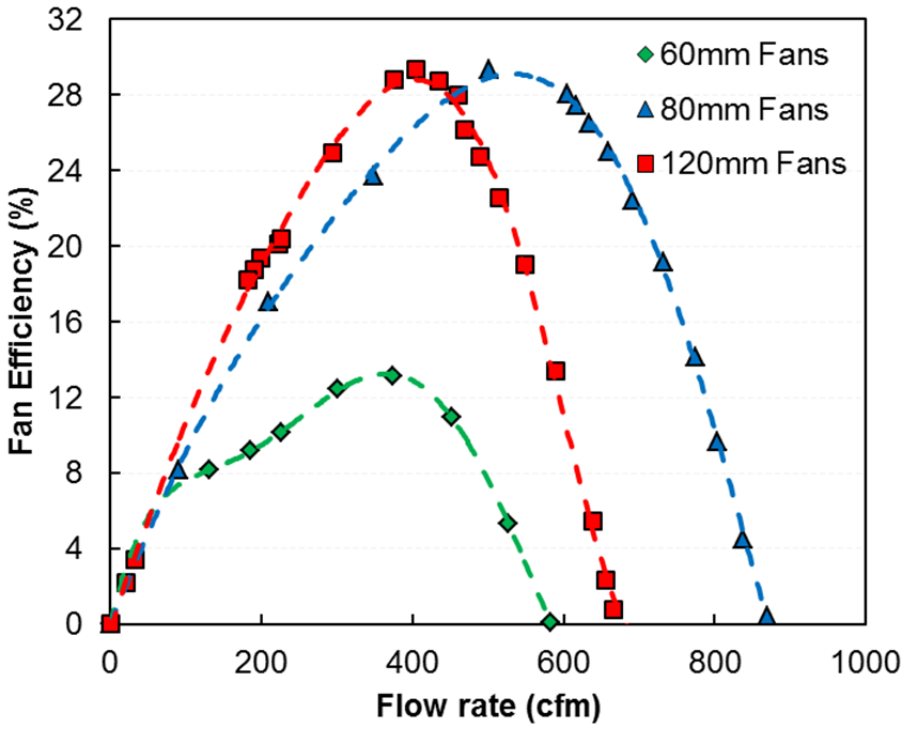

Figure 2-4: Peak total efficiencies for the 60mm, 80mm, and 120mm fans at their maximum fan speeds in the rack configuration

2.1.3 Simulated rack setup

Figure 2-5 shows the laboratory test setup of the simulated rack of four servers in the baseline configuration with sixteen 60mm internal server fans in place. The modified experimental rack setup when an array of 80mm fans is installed is shown in Figure 2-6. The servers are labeled A through D with A being the bottom server and D on top. In order to accurately evaluate the cooling energy savings of the proposed fan modifications, the power delivered to the fans is decoupled from the rest of the server. This is done by powering the fans with an external 12VDC power supply source (Agilent E3633A). A control circuit breadboard with four pin headers for each fan mimics the

connections fans would make to the motherboard. The four pins serve the functions of ground, power, tachometer sensing, and PWM control. The tachometer output signal is sent to an Agilent 34972A data acquisition unit (DAQ) to record fan speeds in RPM. The control signal is directed by either an external function generator (Arduino microcontroller) which provides a fixed PWM signal or via the servers' internal fan speed control algorithm which is relayed from the motherboard. It is critical that all these signals (power, sense, and control) maintain a common ground source for proper operation. To assess the overall efficiency of the rack, the remaining server power consumption (IT load) is recorded by a Yokogawa CW121 power meter which measures the 277VAC and current delivered to the whole rack as well as the power of two of individual servers B and D. The ambient temperature of the laboratory environment in monitored throughout testing with Omega OM-EL-USB-1-LCD temperature loggers. It is observed that the rack air inlet temperature is maintained at 25°C ± 1.0°C during all tests. A desktop workstation is used to communicate to each of the four servers and execute a command script, as well as log all data with common timestamps for effective data reduction.

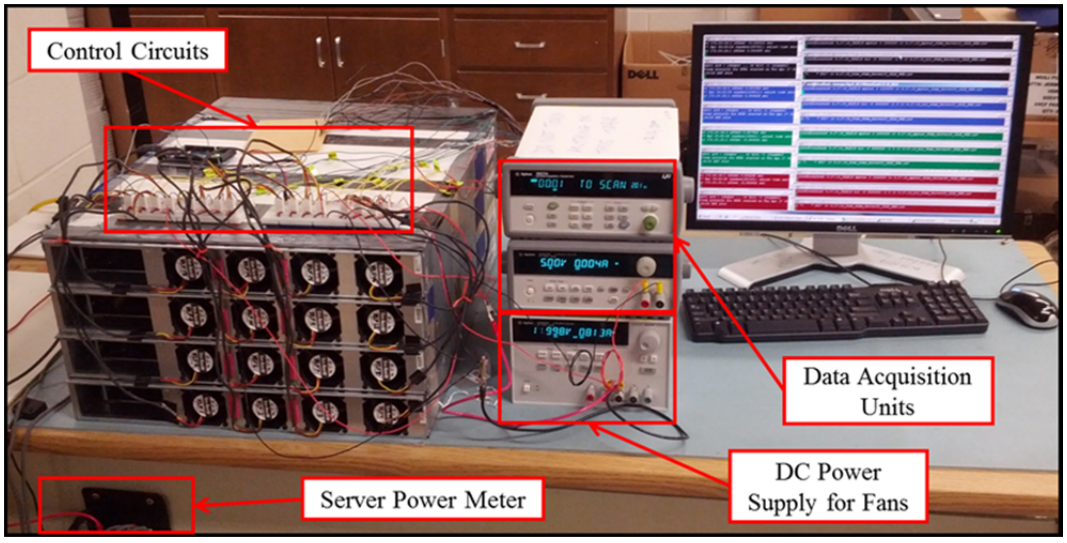

Figure 2-5: Laboratory test setup of four servers stacked to represent a rack. Pertinent testing equipment is identified accordingly

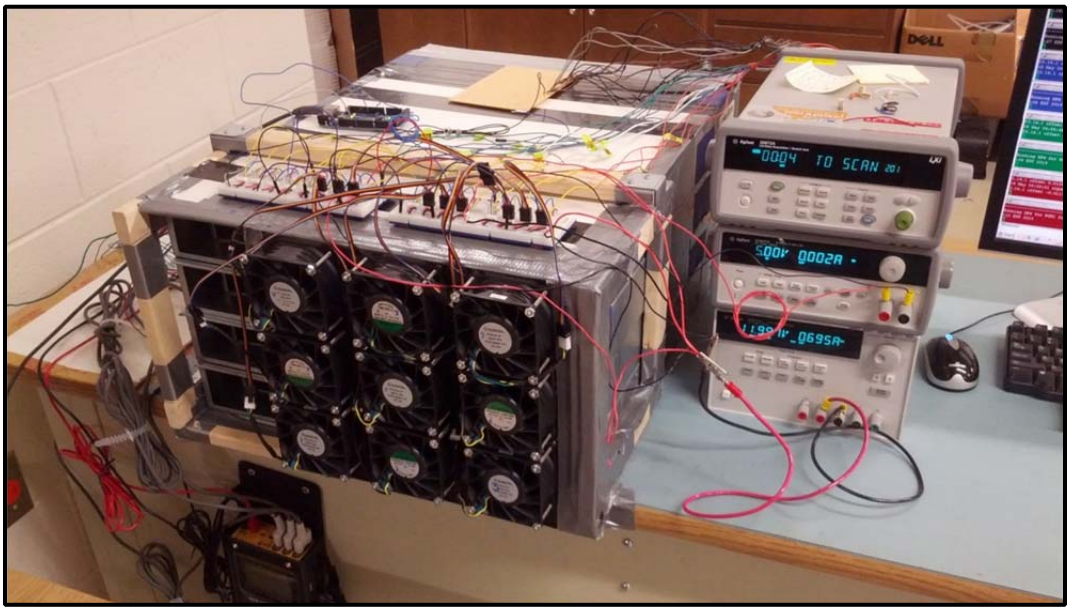

Figure 2-6: Modified rack setup with 80mm fan array installed

2.1.4 Procedures and computational loadings

For the initial testing, a synthetic computational load is applied to all four servers using a free software package, lookbusy [74]. This program allows for stressing of individual computational subsystems: CPU, memory, and network I/O. The particular servers under test typically operate computationally intensive workloads and as such, focus is given to stressing the CPUs, with only moderate memory usage allocated. The native Linux commands mpstat [75] and free [76] are used to monitor CPU utilization (U_{CPU}) and memory usage, respectively. Table 2-2 lists the various computational loadings that are ran on the servers. These loads represent the range of possible computational operating conditions a server may experience in actual service, from idle to medium to maximum utilizations. A complete thermal solution should be efficient across all these operating conditions. An inbuilt software program from the motherboard manufacturer reports the critical temperatures of the CPUs by accessing the on-die digital temperature sensors (DTS). For the tests conducted, each computational workload is applied for a thirty minute period with thirty minutes of idling between each load. Steady state temperatures are achieved within twenty minutes from the start of each applied load and the last ten minutes are averaged for reporting here. Each cycle of computational loadings is applied two more times to ensure repeatability of the measurements.

Table 2-2: Computational workload stressing conditions applied to the servers

Load	U_{CPU} (%)	Memory Usage (MB)
Idle	0 – 0.2	-
10%	10	2000
30%	30	2000
50%	50	2000
70%	70	2000
98%	98	2000

The following key terms will be used through this work when referring to the manner in which fans are powered and controlled and are defined as follows:

Fan Power:

- Internally Powered – Fans receive their 12VDC power signal directly from the motherboard fan header.
- Externally Powered – Fans receive their 12VDC power signal from a DC power supply external to the servers.

Fan Control:

- Internally Controlled – Fans receive their speed control signal directly from the motherboard fan header with the server's native fan speed control (FSC) algorithm in effect.
- Externally Controlled – Fan speeds are dictated from an outside source such as an Arduino microcontroller circuit. This may be either a fixed PWM signal or a more detailed control scheme driven by the internal CPU temperatures.

2.3 Results and Discussion

2.3.1 Comparison under uniform computational load

In order to assess the fan energy savings possible by consolidating the internal server fans to a rack-level array, the cooling performance of each configuration is established. This is done by recording the temperature of the servers' critical components, the CPUs, as a function of fan power. Fan power is varied by operating the fans over their range of speeds by adjusting the fixed PWM signal delivered externally from the microcontroller board.

A steady state temperature is established at fixed fan speeds over the range of operating speeds and computational loadings. In this externally powered, externally controlled test, precise operation of fans at discrete power levels can be achieved. For these tests, all four servers were stressed with uniform computational loading across the rack. Figure 2-7 shows the critical relation between CPU die temperature and fan power when the 60mm fans were employed at the idle, 30%, and 98% CPU utilization loadings. Fan power is given as the cumulative fan power of the rack (ignoring the PSU fan power as discussed above) and an average CPU temperature of all four servers is shown by a best curve fit line.

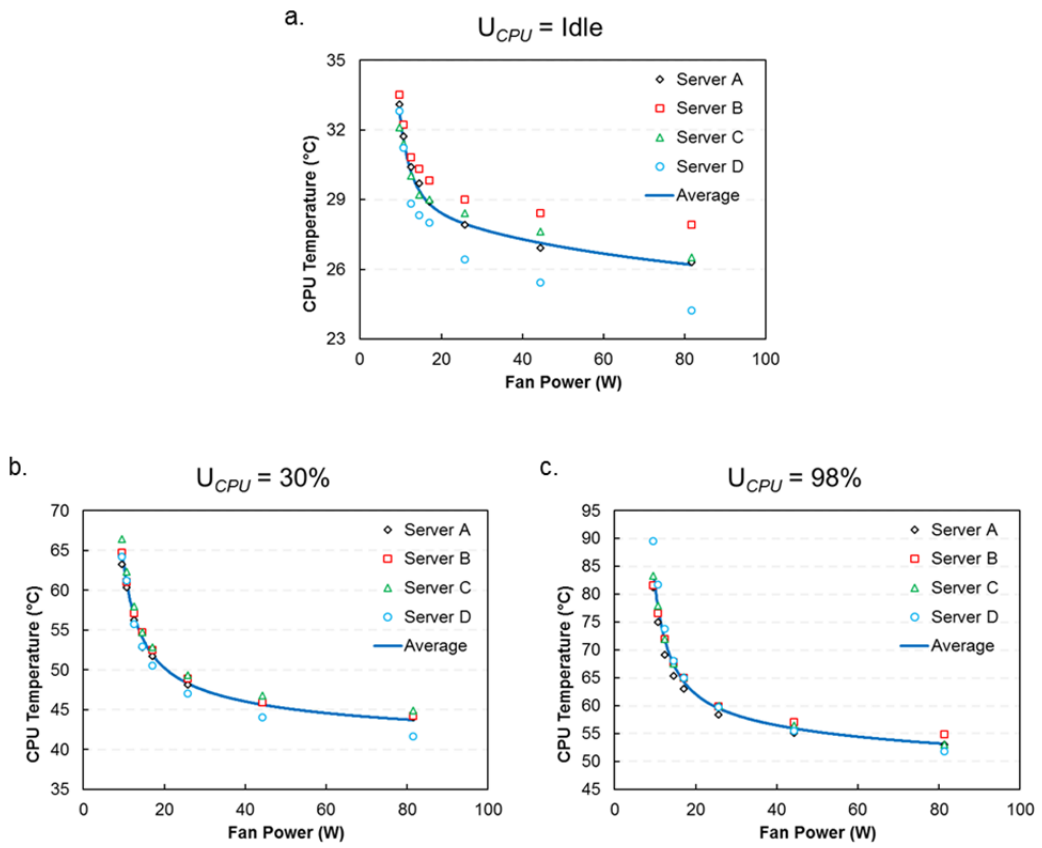

Figure 2-7: Relation between total rack fan power and CPU die temperatures for each of the four servers when operated at uniform computational workloads of (a) idle, (b) 30%, and (c) 98% CPU utilizations in the baseline 60mm fan configuration

As should be expected, when the fan speeds are increased there is a corresponding increase in fan power and subsequent reduction in CPU die temperature. As is typically seen in conventional chip packaging architectures, there is a sharp decrease in temperatures at the left leading portion of the curve due to significant reduction in the convection resistance between the heat sink and the air stream. The temperatures begin to level off towards the end of the curve as conduction resistance in

the heat sink begins to dominate the heat flow path when convective resistance is minimized.

As a check for the performance that the servers may experience during normal operation, the server fans are operated in an externally powered, internally controlled configuration. In this configuration, the native fan speed control (FSC) signal from the servers is sent from the motherboard to the control circuit powering the fans. Each row of four fans receives the individual PWM control signal from their respective server's motherboard. When the native FSC is in effect, the fans fluctuate in speed to achieve a target CPU temperature which is within the deadband range of Figure 2-2. The average of this fan speed and power over the last ten minutes of each computational load is reported here. It can be seen from Figure 2-8 that these discrete operating points lie directly on top of the previously generated curves at each computational loading conditions.

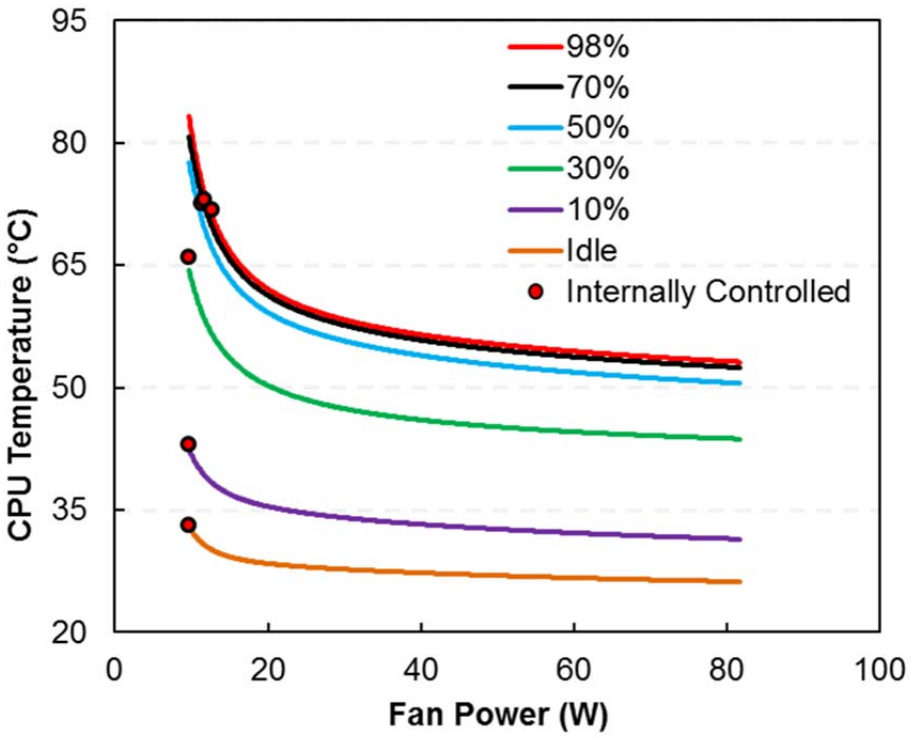

Figure 2-8: Location of the discrete system operating temperatures when the 60mm fans are internally controlled lie directly on the curve fit lines of the externally controlled tests across all computational loadings

To assess the possible fan energy savings in the larger 80mm and 120mm fan configurations, similar externally powered, externally controlled tests were conducted when the larger fan arrays were fitted to the back of the rack. Again, the fans were operated over a range of fixed power levels by adjusting the PWM signal and hence speed. A comparison of the average CPU temperature as a function of fan power for the larger 80mm and 120mm fans to the baseline 60mm fans at idle, 30%, and 98% CPU utilizations is shown in Figure 2-9..

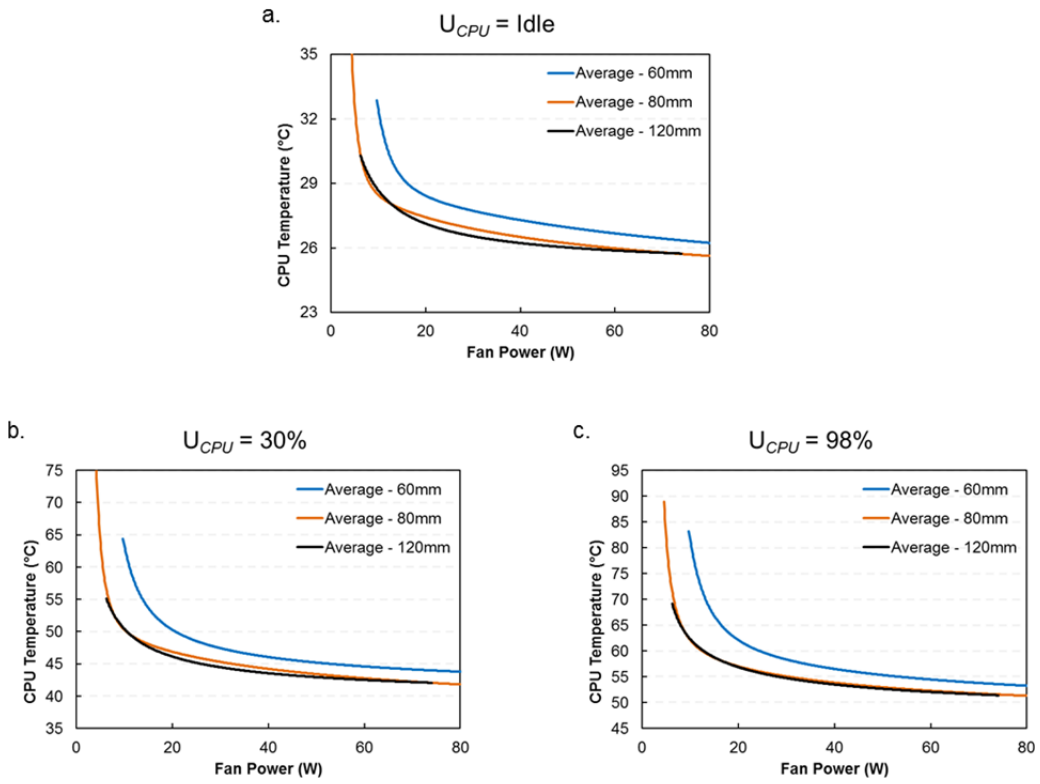

Figure 2-9:Comparison between cooling performance of the 60mm, 80mm, and 120mm fans at (a) idle, (b) 30%, and (c) 98% CPU utilizations

Across all computational loadings, to achieve a given CPU temperature requires significantly less rack fan power when larger 80mm or 120mm fans are employed. This indicates that more airflow is delivered to the servers for a fixed rack fan power. There is no distinct superiority between the 80mm and 120mm fan performance as both average curve fit lines intersect and cross paths depending on computational loading and fan power. To obtain a more precise understanding of the fan energy savings possible with the rack-level fan approach, comparison at specific operating temperatures can be made by looking at the fan power required to reach the CPU temperatures achieved during the 60mm externally powered, internally controlled test. Using the curve fit models for the

80mm fans, fan power at the discrete operating temperatures was extracted and the savings shown in Table 2-3 for all computational loadings tested. Savings from 50.1% to 52.6% were achieved with the array of nine 80mm fans when compared to the baseline 60mm fan case.

Table 2-3: Comparison between required rack fan power for 60mm and 80mm fans to achieve target operating temperatures obtained during the externally powered, internally controlled 60mm fan case

U_{CPU} (%)	CPU DT (°C)	60mm (W)	80mm (W)	% Savings
Idle	33.2	9.59	4.78	50.2%
10%	43.1	9.58	4.78	50.1%
30%	66.0	9.56	4.76	50.2%
50%	72.7	11.19	5.31	52.6%
70%	73.2	11.67	5.70	51.2%
98%	71.8	12.66	6.08	52.0%

The same comparison between the 120mm fan array and 60mm fans could not be made because even at idle speeds the 120mm fans cooled the CPUs to below the discrete operating temperatures of the 60mm externally powered, internally controlled test. To accommodate this, a comparison is made by extracting the fan power from the 60mm fan curve fit models at the highest CPU temperatures achieved with the 120mm fans. Table 2-4 shows this comparison across all computational loadings tested. A wider range of savings is seen with the 120mm fan array, spanning 47.6% up to 54.0%. In either case of the 80mm or 120mm fan arrays, savings are in line with the improved peak total efficiencies and predicted savings of Figure 2-4.

Table 2-4: Comparison between required rack fan power for 60mm and 120mm fans to achieve target operating temperatures obtained when the 120mm fans are operated at idle speed

U_{CPU} (%)	CPU DT (°C)	60mm (W)	120mm (W)	% Savings
Idle	30.3	12.42	6.51	47.6%
10%	38.1	13.08	6.49	50.4%
30%	54.9	13.86	6.40	53.8%
50%	65.4	13.62	6.36	53.3%
70%	67.6	13.78	6.34	54.0%
98%	69.0	13.66	6.33	53.7%

2.3.2 Comparison under non-uniform computational load

Actual data centers rarely, if ever, operate with uniform computational loading across all servers in a rack. The results from the previous section represent an idealized condition and best case scenario for possible savings that may be achieved with a rack-level fan configuration. In order to test the limits of the validity of a rack-level fan configuration, a simple test was conducted to evaluate the savings achievable in a worst possible condition. In this situation, the worst possible situation for the larger, rack-level fan array would be if all fans in the system are controlled by a single server operating at maximum computational load (and hence temperature) while all other servers remain at idle loading, as seen in Figure 2-10b. This would constitute the most inefficient use of fan energy possible because all fans in the array would operate at higher speeds. The shared airflow of the servers would lead to overcooling (and unnecessarily increased rack fan power) of the three idle servers. This can also be thought of as the simplest possible control scheme that may be developed for a rack-level fan configuration in which the

maximum temperature of the rack dictates the speed of all the fans in the array. A comparison can be made to the best possible case for the 60mm fans, in which all servers in the rack are operating at idle computational loadings. This represents the lowest possible energy state for the 60mm fans because, as seen in the externally powered, internally controlled baseline test, the fans will operate at their idle speeds, shown in Figure 2-10a.

To test this notion, server D, which showed consistently higher temperatures (as seen in Figure 2-7c), was operated at 98% CPU utilization. The fan control signal from the motherboard in server D was sent to all nine and four fans in the 80mm and 120mm arrays, respectively. A comparison between rack fan powers for these three cases is shown in Table 2-5. Even in the worst possible situation in which three servers are overcooled, the 80mm fans still provide 35.3% savings in rack fan power compared to the lowest possible energy state of the 60mm fans. Even at idle speeds, the 120mm fans overcool to the point that the native fan speed control (FSC) algorithm does not engage and rack fan power savings of 33.8% are achieved. These extreme situations indicate that a rack-level fan configuration provides an opportunity for significant fan energy savings when compared to smaller, chassis fans. The savings presented here also do not take into account any savings in IT energy that may be realized by the overcooling of components and hence reduced leakage power in the silicon devices.

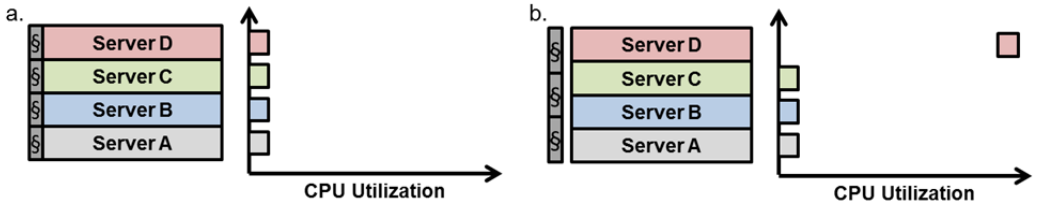

Figure 2-10: (a) “Best Case Scenario” for the 60mm fans in which all servers are operating at idle computational load and (b) “Worst Case Scenario” for the rack-level fans in which a single, high computationally loaded server dictates fan speeds

Table 2-5: Comparison of rack fan power between the “Worst Case Scenario” of the rack-level fans and the lowest energy state possible for the baseline 60m fans

Test Case	Total Rack Fan Power (W)
60mm “Best Case Scenario”	9.59
80mm “Worst Case Scenario”	6.20
120mm “Worst Case Scenario”	6.35

2.4 Further Discussion

Along with reduction in operating rack fan power, consolidation of smaller internal server fans to a rack-level array may provide additional benefits to data center owners and operators as discussed here.

Clearly, the number of fans in the system may be substantially reduced, from either sixteen to nine or four in the particular hardware used for this study. Although larger fans may have a slightly higher first cost per unit, the overall system cost may be reduced. Table 2-6 compares the relative individual unit costs and relative cost per “rack” for the 60mm, 80mm, and 120mm fans used in this study. Prices were taken from an

online electronics component distributor. The results here are not exhaustive of all possible fans that may be selected for this application and do not include consideration for price breaks for larger quantity orders. However, as a first estimate, this may provide an indication of the capital savings that may be achieved. Additional physical infrastructure may need to be put in place at the rack level to house an array of fans as well as supplemental control circuity which may have additional associated costs. A follow-on study to this work will explore how the control scheme of rack-level fans may be further designed and optimized to realize maximum fan power savings from such a configuration.

Table 2-6: Relative cost associated with the fans used in this study

Fan Size	Relative Cost per Fan Unit	Quantity per "Rack"	Relative Cost per "Rack"
60mm	0.37	16	1.00
80mm	0.49	9	0.76
120mm	1.00	4	0.68

An additional benefit of the reduced number of fans in the system is fewer points of failure. Future work on rack-level fans will investigate the impact of failure and redundancy in a rack-level fan configuration.

An emerging paradigm in the data center industry is "rack disaggregation". The concept involves separating a traditional server's compute, memory, storage, and I/O into discrete modules. The benefit being that the individual subsystems of a server may be changed and upgraded at different frequencies without having to replace all parts of the system, resulting in less material waste [77]. In this manner, a rack then becomes the

fundamental building block of a data center. The concept of rack-level fans aligns with this idea well. As could be seen in the case of the particular 80mm and 120mm fans studied here, there is a significant amount of additional cooling capacity available beyond what is required to maintain safe operating temperatures for the servers. This additional capacity may last for multiple refresh cycles of IT equipment as component heat loads and power densities increase. Prudent selection of fans during initial design stages and realistic projections of future server hardware requirements will ensure rack-level fans have their greatest impact.

Chapter 3

Impact of Rack Inlet Static Pressure on Server Level Performance

Much has been done to understand the impact of data center room pressurization on temperature gradients within the room and across the face of racks. Previous studies have worked to understand the interaction between underfloor pressure distributions on air flow rates through perforated floor tiles. Common goals for underfloor pressure management are to reduce the velocity of air and hence increase the static pressure of the underfloor. This in turn increases the volume flow rate of air through the perforated floor tiles. Ideally, every cfm of airflow provided by a CRAC unit would travel through a server. Bypass and leakage of cold air around the server and into the hot aisle are common issues which decrease the efficiency of the room level cooling system [78] [79]. Recirculation of hot exhaust air back to the cold aisle may occur, raising the inlet temperature of air supplied to the servers and diminishing cooling efficiency [80].

Most of the previous studies related to data hall pressurization stop at the rack level [81] [82]. This present work seeks to understand the impact of pressurization on the thermal performance and energy consumption within a server. This information can be helpful in optimizing the desired static pressure set points within the room to achieve the most energy efficient use of cooling resources.

3.1 Experimental Setup and Procedures

The same set of four server discussed in the previous chapter are again experimentally characterized. Similar procedures pertaining to the controlling of server fans were used (internally controlled and externally powered, etc.). An Air Flow Bench

test chamber was used to act as a pseudo-cold-aisle-containment system shown schematically in Figure 3-1. An additional view of the servers attached to the Air Flow Bench is shown in Figure 3-2. This is an idealized scenario because it is assumes uniform pressure across the face of the servers and flow straighteners within the chamber orient streamlines parallel to the servers. Previous studies have shown that actual data center racks have non-uniform pressure gradients across the face of a rack and the velocity streamlines exiting perforated floor tiles are not straight into the servers [23].

To understand the impact of inlet static pressurization on the thermal performance of the servers, the server fans were first fixed at a constant duty cycle. In this externally controlled and externally powered fan configuration, the effect of pressure on CPU temperatures could be isolated from other fan interactions. While operating at fixed fan speeds, the blower to the Air Flow Bench was adjusted to achieve a desired inlet static pressure. At each server fan speed and blower setting, the server was provided a synthetic computational workload of idle, 30%, and 98% CPU utilization with the lookbusy software tool. Each workload was run for thirty minutes and repeated three times in total for repeatability. The results gathered here are taken from the average values over the last ten minutes of each computational workload as this is when steady state CPU temperatures were achieved.

The server fans were then controlled internally by the motherboard's native fan speed control algorithm while still being powered externally. The Air Flow Bench was set to fixed blower speeds; however, the inlet static pressure became a function of the server fan speeds as they modulated to maintain the control set point temperature of the deadband operation described in Figure 2-2. Tests were conducted with the servers operating at 98% CPU utilization only to generate CPU temperatures high enough to

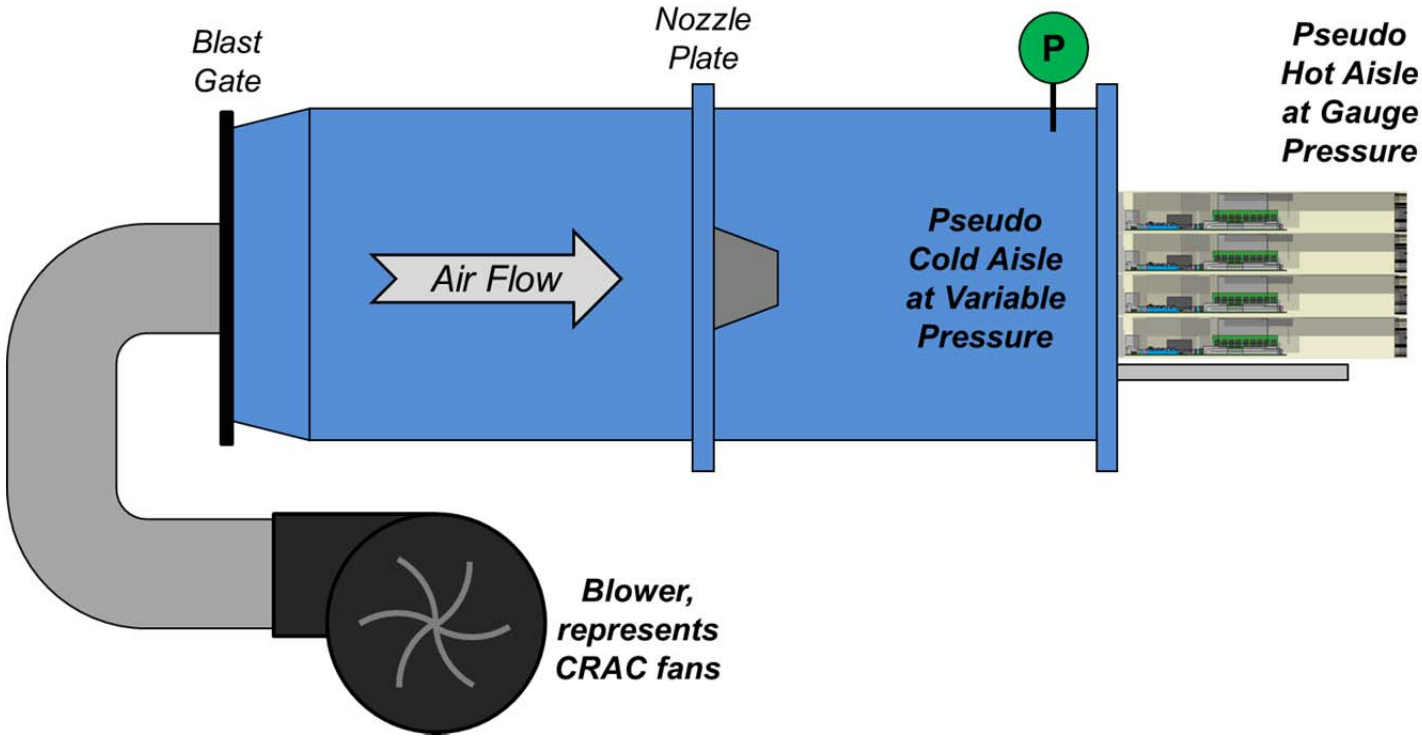

Figure 3-1: Schematic of experimental setup used to control inlet static pressure to servers

trigger the fan speed control. The workload was operated for sixty minutes in this case to allow the fluctuations in fan speed and static pressure to reach steady state conditions. Average values over the last ten minutes of steady state conditions are reported here.

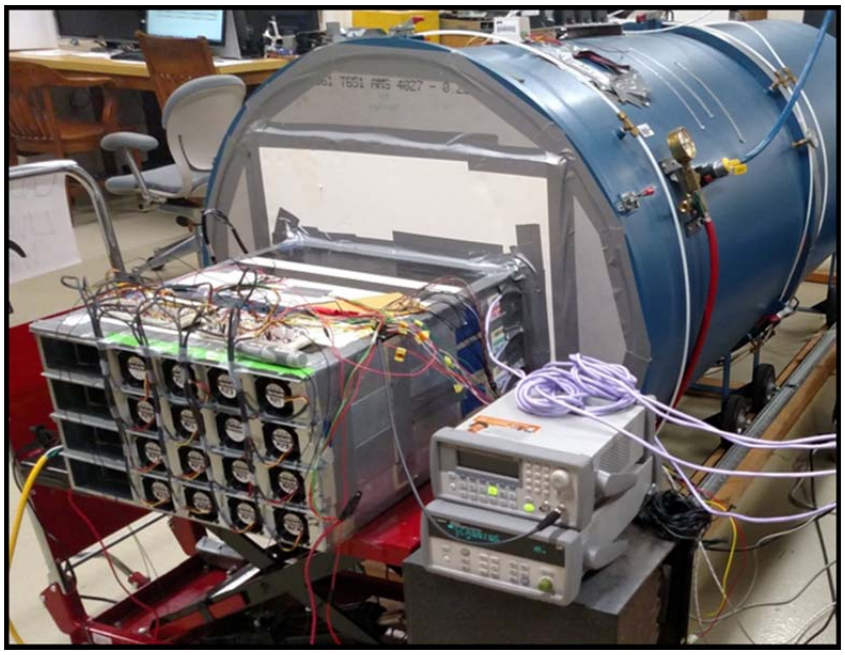

Figure 3-2: Simulated rack of servers attached to Air Flow Bench for static pressure testing

3.2 Results and Discussion

3.2.1 Externally Powered, Externally Controlled Server Fans

Figure 3-3 shows the results for the externally powered, externally controlled tests across the ranges of inlet static pressures studied when the servers were operated at maximum computational load. Note in the figure that the y-axis crosses at a negative static pressure value. A best curve fit taken from all CPU temperatures across each of

the four servers shows the average trend in the data. At the lower fan speeds of 0% and 5%PWM duty cycle, CPU temperatures begin to reach their upper functional limits when the static pressure becomes negative.

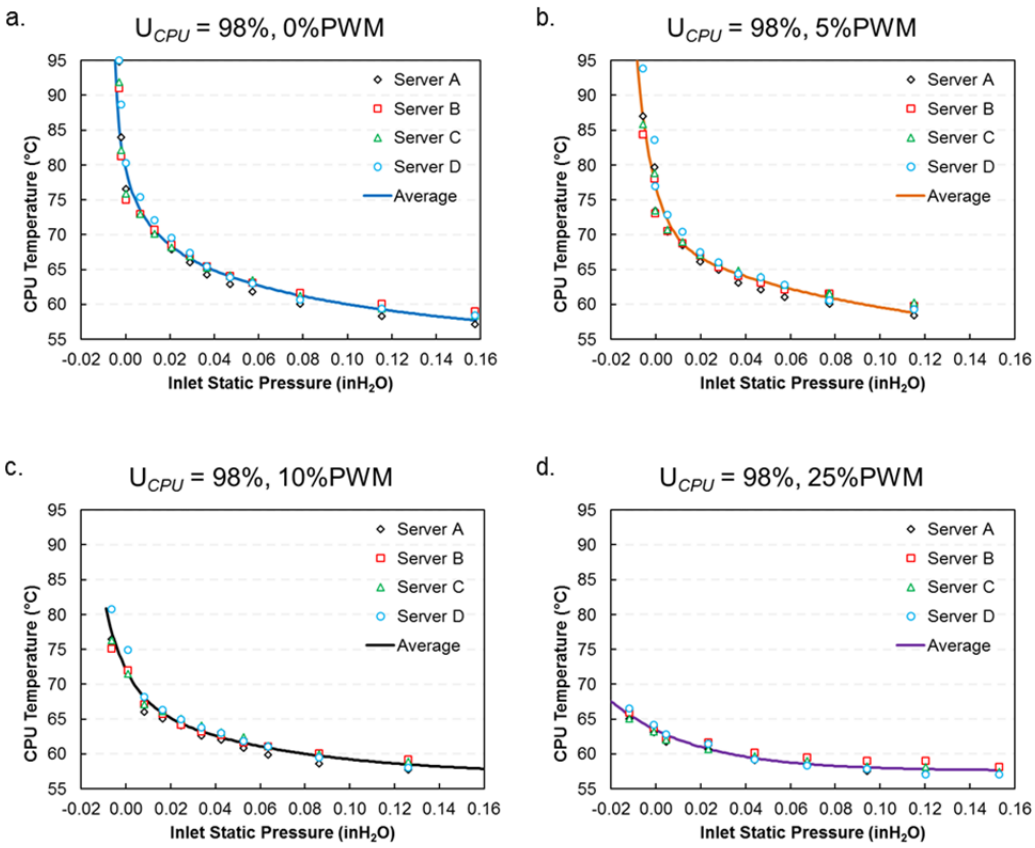

Figure 3-3: Impact of inlet static pressure on server CPU temperature at fixed fan duty cycles of (a) 0%PWM, (b) 5%PWM, (c) 10%PWM, and (d) 25%PWM

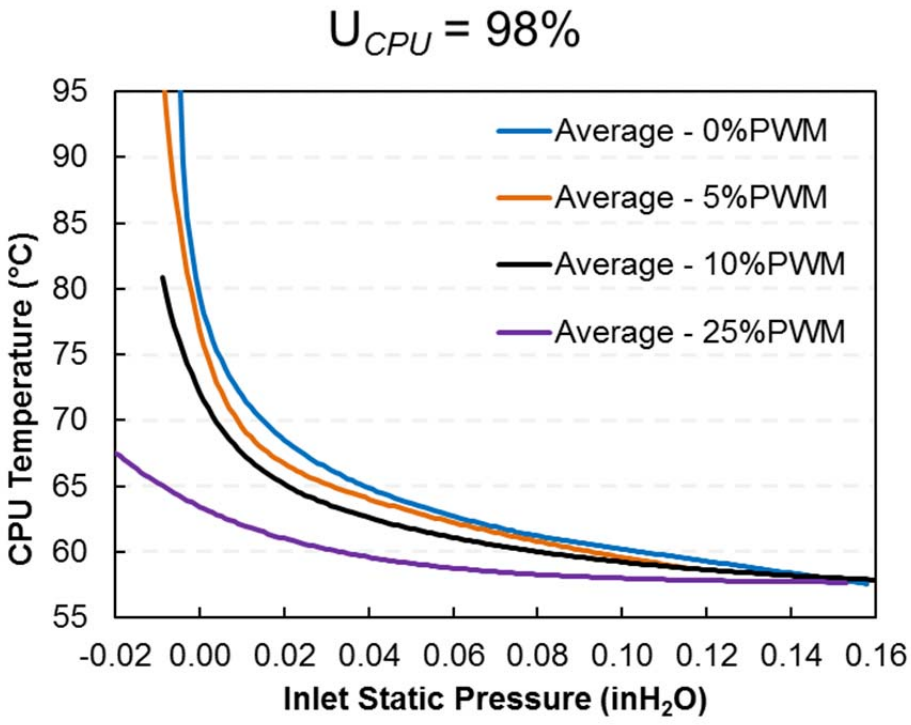

Figure 3-4: Convergence of cooling performance of 0%, 5%, 10%, and 25%PWM fan duty cycles across range of inlet static pressure

From Figure 3-4 it can be seen that these line of constant fan duty cycles eventually converge. This convergence takes place at higher static pressures for higher fan speeds. As an approximation from the curve fits shown in Figure 3-4, the die temperatures at 0% and 5%PWM signals are equal at 0.07 inH_2O inlet static pressure. CPU temperatures for 0%, 5%, and 10%PWM all converge at 0.11 inH_2O inlet static pressure. All of these curves converge with 25%PWM fan speed at 0.15 inH_2O inlet static pressure. This may be expected because, beyond a certain static pressure, the airflow provided by the blower begins to overtake the airflow provided by the internal server fans.

An approach developed by Alissa, Nemati, and colleagues, illustrates the phenomena through the application of a flow curve [83] [84]. By taking the algebraic subtraction of a server's system resistance curve from the fan curve, a cumulative flow curve is derived which accounts for the differential pressure across the server. This is illustrated in Figure 3-5 for the server under study. The system impedance curve for the motherboard portion of the Open Compute server was experimentally characterized on the Air Flow Bench. The fan curve when the fans are operated at 0% PWM was also taken. The difference of these two results in the theoretical flow curve for the server. The point where the flow curve crosses the x-axis represents the operating point of the server and fan curves in a system with neutral or zero pressure difference across the server. All points above the x-axis can be considered when the inlet static pressure to the server is a negative value during the above experiments. All points below the x-axis are indicative of the positive pressure conditions of the above experimental testing. As seen from the experimental data, positive inlet static pressure results in more airflow through the system and lower component temperatures. Beyond 38cfm the flow curve is simply an inverse of the server's system resistance. It is assumed that at pressures beyond free flow condition the fan will not be generating any additional airflow. The fan's contribution to flow and pressure is limited to the region of its fan curve.

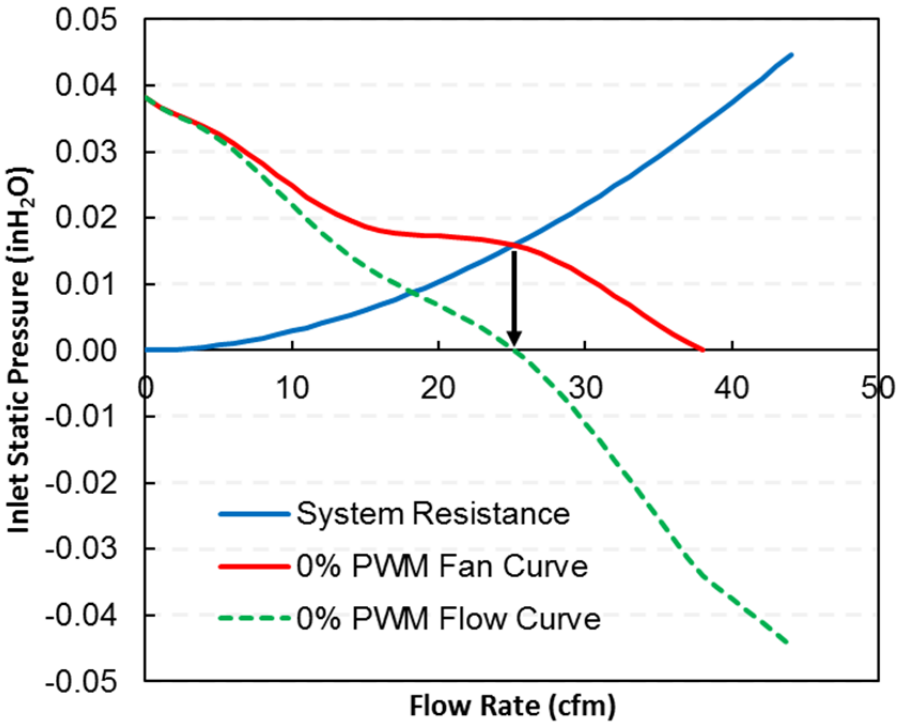

Figure 3-5: Development of server flow curve from experimental system resistance and fan curves

The resultant flow curve shown in Figure 3-5 is unique to the operation of the server when the fans are running at 0%PWM duty cycle. A new flow curve must be created for each fan operating duty cycle as the respective fan curves will change. A comparison of the flow curve for the Open Compute server when the fans operate at 0% and 25%PWM duty cycle is shown in Figure 3-6. Note again that the actual inlet static pressure is an inverse of the y-axis values presented in the figure. As can be seen, the two flow curves converge and are equal beyond 0.10 inH_2O of positive inlet static pressure. At flow rates beyond 70cfm the 25%PWM flow curve is also an inverse of the

server's system resistance and both flow curves are identical at any higher inlet static pressures.

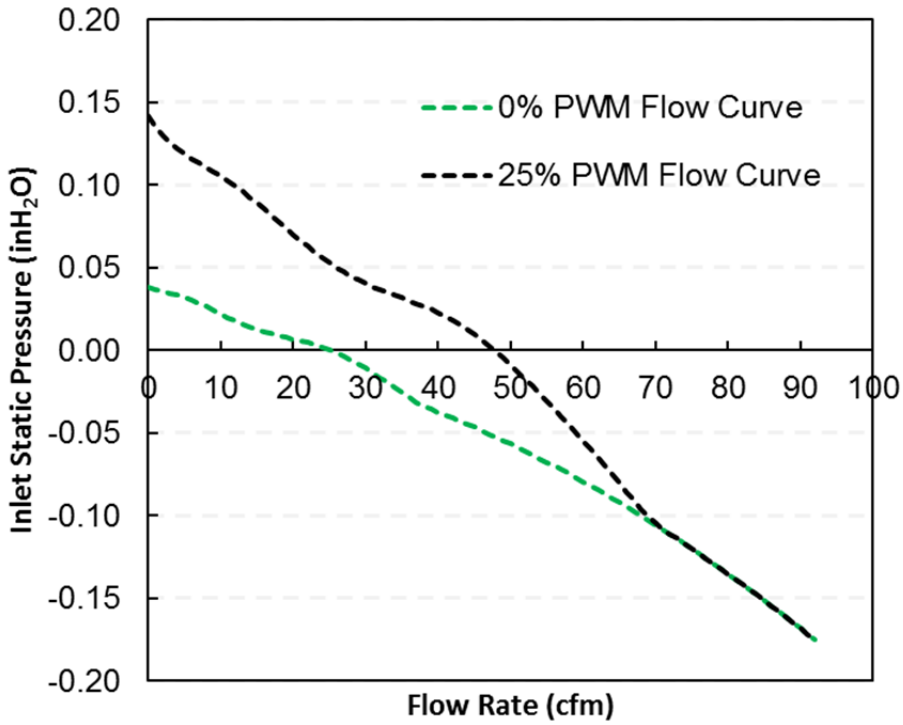

Figure 3-6: Comparison of theoretical flow curves at 0% and 25%PWM duty cycle for the motherboard potion of a single server

An additional point of note is that the convergence of the theoretical flow curves in Figure 3-6 is at approximately 0.10 inH_2O, well below the 0.14 inH_2O point seen in the experimental data of Figure 3-4. This can be explained by the fact that these flow curves are limited to the theoretical operation of just the motherboard portion of the Open Compute servers. During testing on the Air Flow Bench additional airflow passed through the power supply (PSU) potion of the server. The PSU airflow can be considered leakage

airflow past the motherboard region of the server and hence a high inlet static pressure would need to be maintained to obtain the desired airflow across the motherboard.

3.2.2 Internally Controlled, Externally Powered Server Fans

When the server fans are controlled internally, it is expected that the CPU temperatures will not exceed their target values specified in the deadband control. This is seen in Figure 3-7 across the range of inlet static pressures studied. At lower inlet static pressure values (particularly in the negative region) CPU temperatures across all the servers are observed to fluctuate around the control range of 68 to 74°C. At higher inlet static pressures, the air flow provided by the blower over powers the internal server fans and more consistent temperatures are observed.

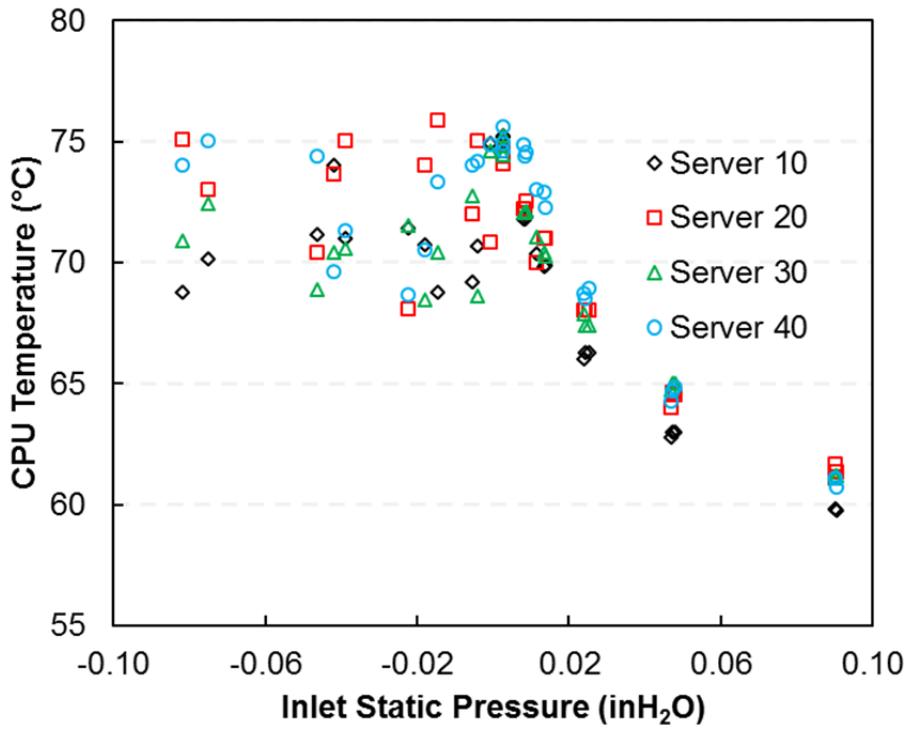

Figure 3-7: CPU temperatures maintained at or below control limit threshold across inlet static pressures range

In terms of actual fan power consumed by the rack, a piecewise linear trend is observed as seen in Figure 3-8. Below a threshold of about 0.003 inH_2O, fan power increases steadily with negative static pressure. This demand is caused by the need to maintain temperatures within the specified target range. Above 0.003 inH_2O, the rack fan power flattens out because the fan control algorithm no longer needs to engage, and the fans remain stable at idle speeds.

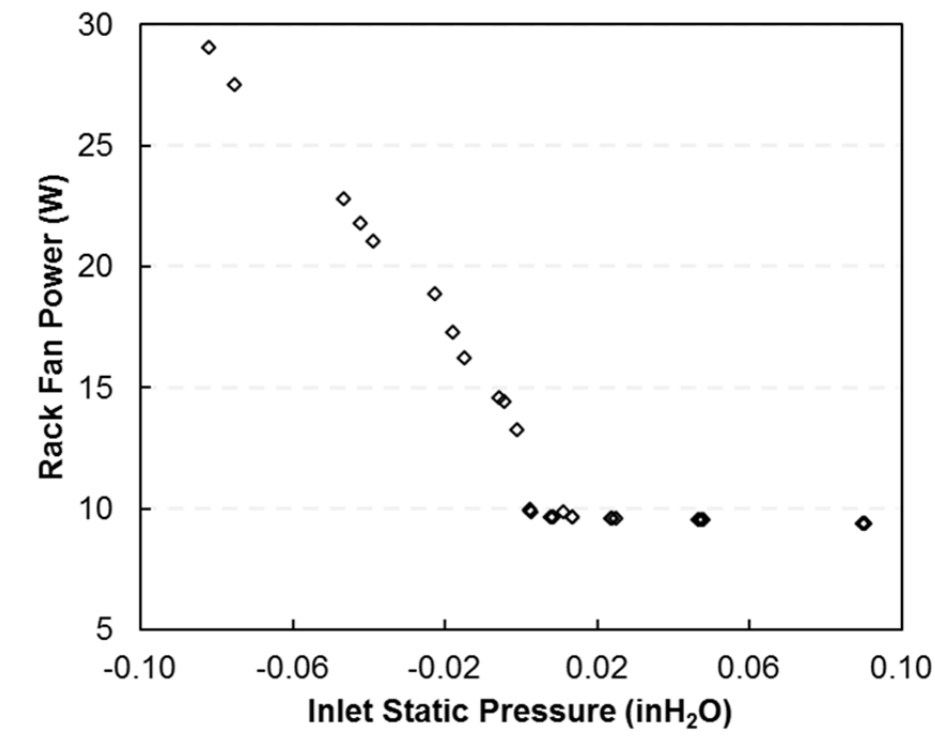

Figure 3-8: Total rack fan power as a function of inlet static pressure

3.2.3 Discussion

A truly optimized system will require dynamic balancing between the fan power within a server and the fan power from blowers within the data center's CRAC units. The overall goal should be to minimize total system fan power. This will likely require various pressure measurements throughout the data hall, either as static pressure at the front of racks in the cold aisle, or ideally, as a differential pressure measurement across a rack. The work of [85] proposed a control system utilizing dampers within either the cold or hot aisle to physically restrict airflow and maintain static pressure at specified levels. It is possible in some cases that a large enough static pressure could be built up across a

server rack allowing for complete removal of server fans. All airflow through a server may be provided solely though room level airflow from CRAC units.

Other alternatives may include additional, larger fans placed either within the cold aisle or hot aisle. As seen in the previous chapter, larger fans will typical operate more efficiently their smaller counterparts. Existing products are available on the market that have active floor tiles which include fans coupled the perforated floor tiles of the raised floor. Alternatively, an exhaust fan placed within the return air plenum of the hot aisle may serve similar function.

Chapter 4

Flow Rate and Inlet Temperature Considerations in Mineral Oil Immersion Cooling

The primary goal of this study is to establish general operating conditions, in terms of volumetric flow rate per server and oil inlet temperature, which can be expected for safe operation of servers in an oil immersion cooling configuration. For this work, an Intel-based Open Compute server was experimentally tested and characterized. Initially, the server was operated in the standard air cooled configuration with internal server fans to establish baseline operating performance and component temperatures. Next, the server was removed from its standard chassis and placed in an insulated acrylic container and submerged in white mineral oil. Data was recorded for the duration of testing, however, only steady state data is reported here with some comments made regarding transient performance. Although the complete experimental set-up is not entirely reflective of an actual data center facility implementation in fully built-out conditions, the data provides strong evidence as to the effectiveness of oil cooling for data centers.

4.1 Experimental Setup and Procedures

4.1.1 Air testing setup

To establish a baseline for comparison, the server is initially tested in the standard air cooled configuration as shown in Figure 4-1. Detailed descriptions of the Intel-based Open Compute server design can be found in [71] [72]. The server motherboard contains two CPUs each with a rated thermal design power (TDP) of 95W. These components represent the primary heat sources in the system and are cooled by

two extruded aluminum heat sinks. The key features which enable efficient air cooling are four 60 x 60 x 25.4mm fans and an air duct that directs air flow over the temperature critical components (i.e. processors and memory). The internal server fans are controlled by a BIOS fan speed control algorithm native to the server. This fan speed control operates by adjusting the fan speeds using pulse width modulation (PWM) to achieve a target CPU die temperature. In this manner, the fan speeds and die temperatures typically oscillate with some over- and undershoot of a targeted value. An average value over the duration of the test cycle is reported. To test the server in the standard configuration, it is allowed to draw air from the ambient laboratory environment for cooling. A synthetic computing workload is applied and internal monitoring tools are used for data collection, as discussed below.

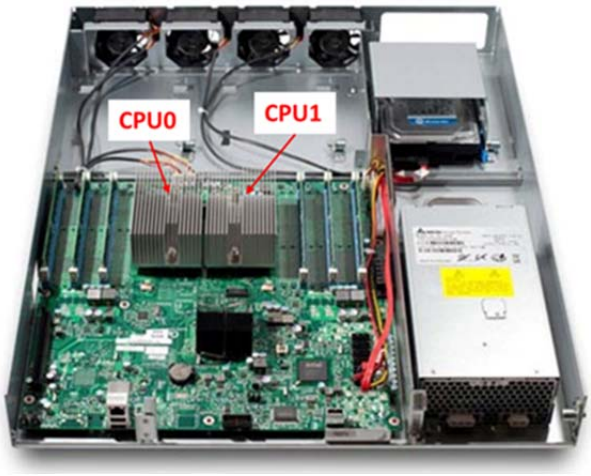

Figure 4-1: Intel-based Open Compute server in standard air cooled configuration with air duct removed for visual purposes

4.1.2 Oil testing setup

The oil immersion test setup used in the present investigation consists of the following major components as shown in Figure 4-2 and discussed in the following. A single Intel-based Open Compute server motherboard (MB) and power supply unit (PSU) are placed horizontally in an immersion tank. Some modifications are made to the server to enable testing in the container available and to enable operation in oil. The server motherboard and PSU PCBs are removed from their respective metal chasses to reduce their size to fit in the test container available as seen in Figure 4-3. In addition, the air duct is removed and the internal server fans of both the server and PSU are disconnected and removed. The hard disk drive (HDD), incapable of operating when submerged, is placed outside the tank although other options exist to leave the HDD intact. The HDD is cooled by natural convection only and represents a small portion of the total IT heat load that is not removed by oil (approximately 3.5%). Of additional note, the grease-based thermal interface materials (TIM) applied on the CPUs and chipset are kept in place for the duration of testing. Some sources suggest that oils and greases conflict and may cause the thermal grease to dissolve leading to contamination of the oil and possible fouling in the system. However, in the authors' background study, this information was merely anecdotal and was not followed for the present experiments.

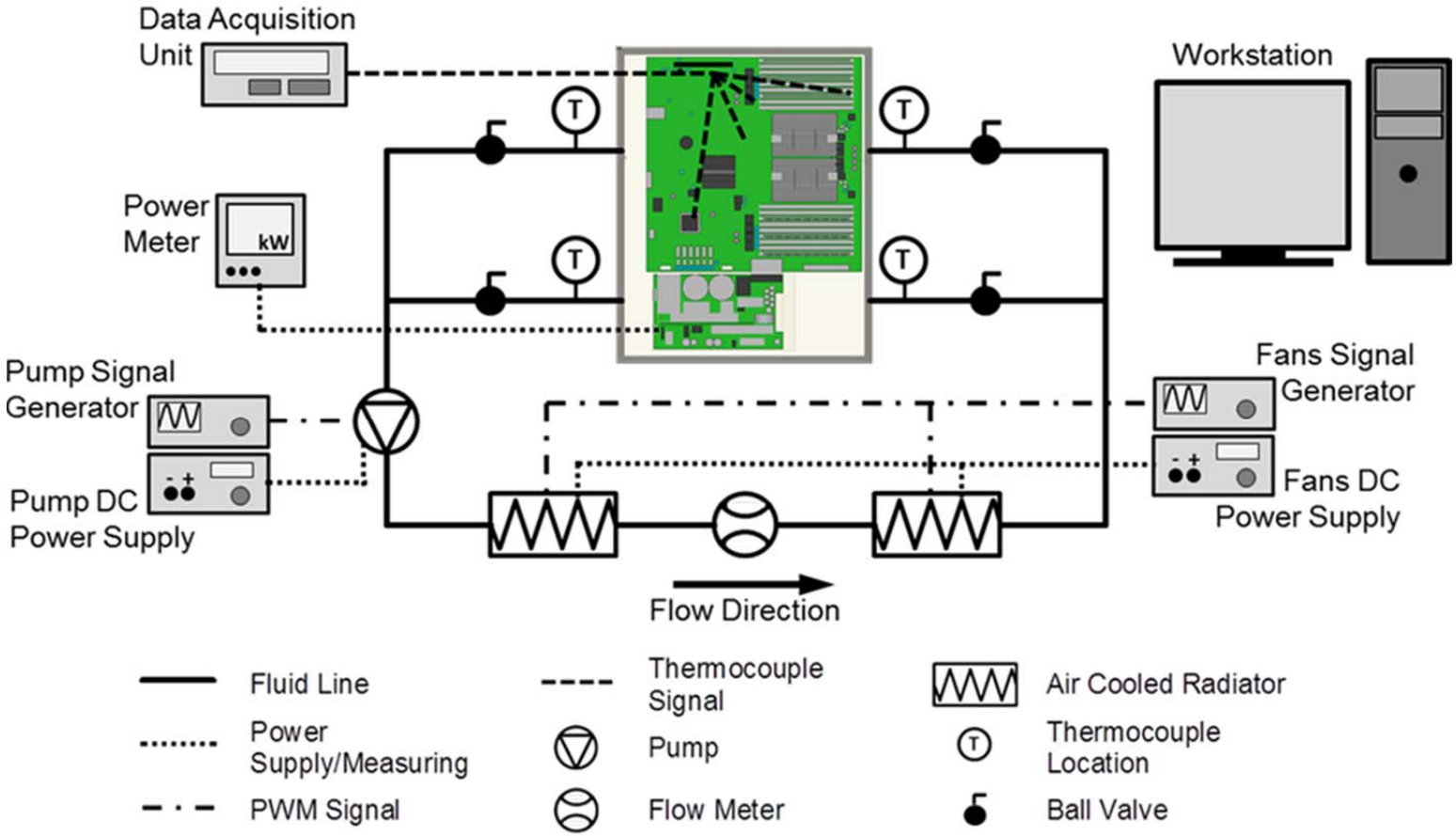

Figure 4-2: Schematic of test setup and data collection equipment

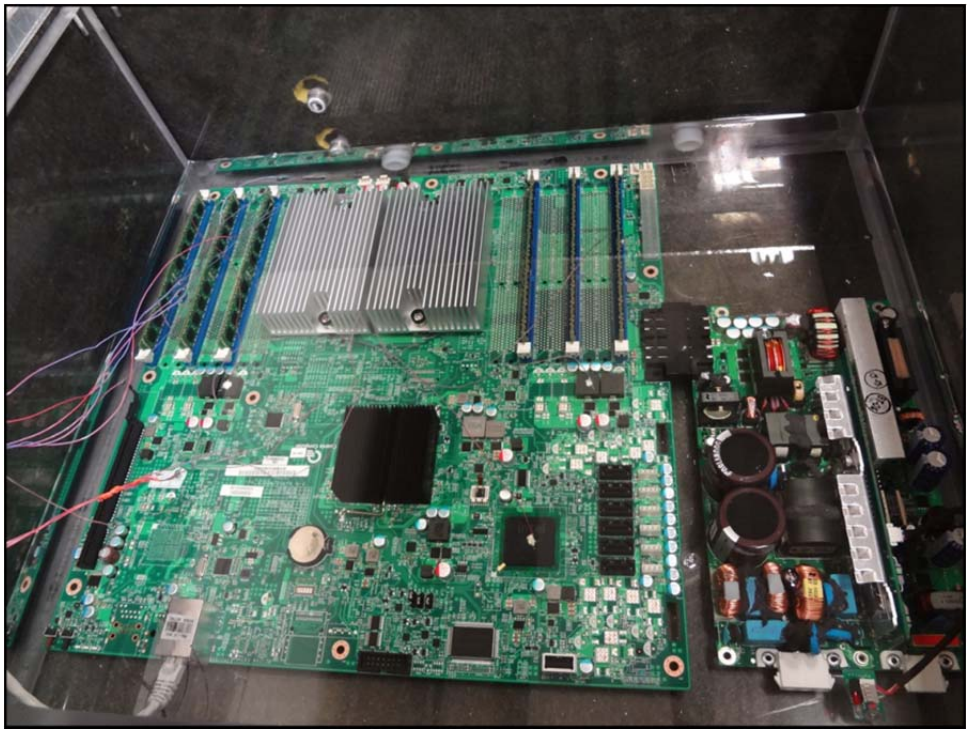

Figure 4-3: Server motherboard and power supply removed from their metal chassis and submerged in mineral oil tank

The tank is made of one half inch thick acrylic of inner dimensions 45.7 x 36.8 x 19.1cm (18.0 x 14.5 x 7.5in) and is wrapped in 2.54cm (1.0in) thick insulation. A total of four 1.27cm (0.5in) diameter ports are tapped into the container, two serving as inlets to the tank and two as outlets. The inlet and outlet ports are strategically placed with ball valves to allow for the impact of flow patterns through the system to be studied. One set of inlet and outlet are directly in-line with the center of the base of the CPU heat sinks. When only these two ports are open, it will likely represent the best case cooling scenario for the system since the highest velocity flow is directed straight across the CPU heat sinks. The second inlet/outlet port pair is in-line with the power supply unit, removed from the main motherboard components. It is predicted that this flow configuration will prompt

higher component temperatures (especially CPUs) since their cooling will be primarily induced by bulk fluid motion. These two flow configurations are used independently during testing and are termed “MB Flow Path” and “PSU Flow Path” respectively, as shown in Figure 4-4.

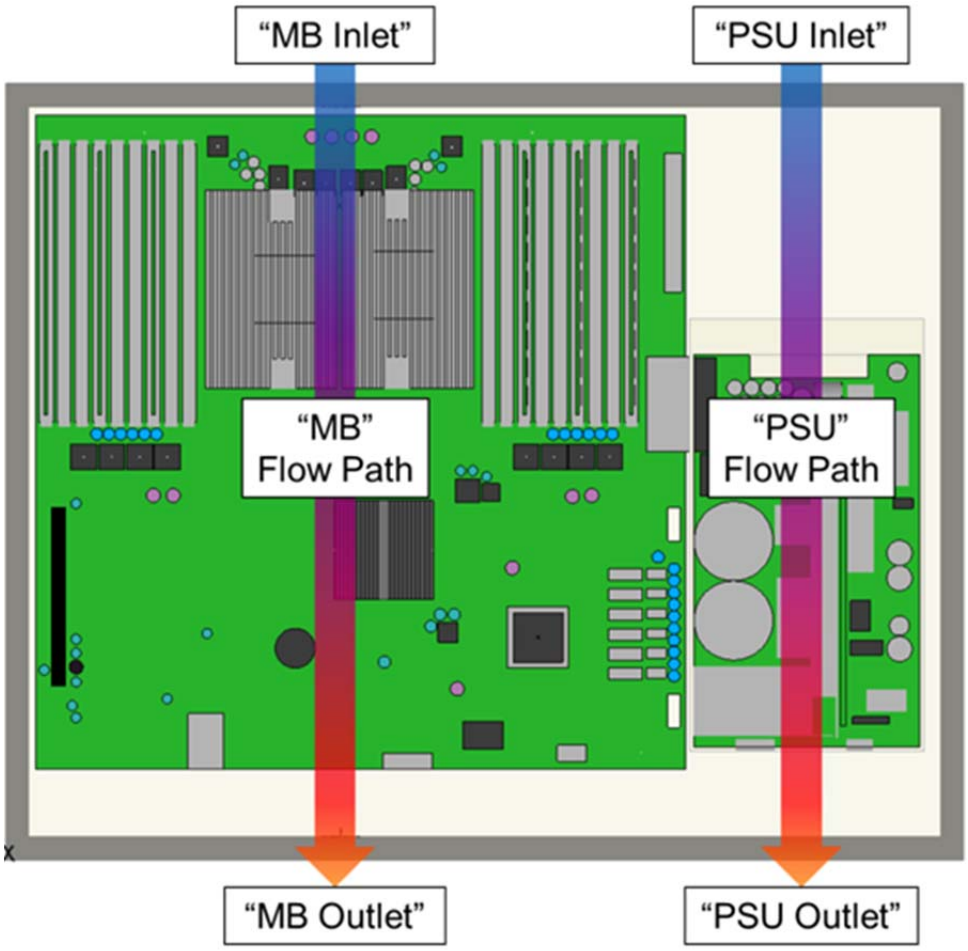

Figure 4-4: Two flow configurations through the immersion tank. “MB Flow Path” occurs when only the MB inlet and outlet valves are open. “PSU Flow Path” occurs when only the PSU inlet and outlet valves are open

A small magnetically driven centrifugal pump with a 12V DC brushless motor located on the outlet side of the tank circulates fluid through the system. A DC power supply (Agilent E3642A) delivers a constant voltage signal to the pump. The voltage delivered and current drawn by the pump are logged by a workstation in two second intervals. The pump is equipped with a 4-pin connector which enables pulse width modulation (PWM) for speed control from 1300 to 4500 rpm. A function generator (Agilent 33210A) is used to control the speed of the pump to achieve the desired volumetric flow rate to the inlet of the tank.

Heat is rejected from the oil to the laboratory environment via two 240mm radiators constructed of two-pass, single row brass tubes with louvered copper fins. Each radiator is equipped with two 120mm 12V DC brushless motor fans. A DC power supply (Agilent 3633A) delivers a constant voltage signal to the fans. The voltage delivered and current drawn by the fans are logged by the workstation. The fans are equipped with 4-pin connectors which enable PWM speed control from 1650 to 5100 rpm. A function generator (Agilent 33210A) is used to control the speed of the fans to achieve the desired inlet temperature to the tank containing the server.

An Omega FLMH-1402AL in-line flow meter is used to record the volumetric flow rate of the oil. The flow meter has a scale accuracy of ±4%. The meter is rated for oils with a specific gravity of 0.873. This leads to a correction factor of 1.0% for the oil used in the current system. The flow meter is placed midway between the outlet and inlet sections of the immersion tank. As this is an analogue device, it is visually monitored throughout a given test, with the values at the conclusion of the test reported. Although mineral oils have relatively high volumetric expansion rates (β=0.0007 $1/K$), it is assumed that the volumetric flow rate is conserved from the point of flow measurement to inlet of the immersion tank. In actual testing, the fans on the second radiator in the loop were not

needed to maintain inlet temperatures and it is assumed that temperature change between flow meter and tank inlet is minimal and any resulting change in density is negligible.

An Omega USB data logger is used to record the laboratory ambient air conditions in one minute intervals with an accuracy of ±1.0°C.

The oil inlet temperature of the tank is measured using T-type thermocouples placed in the stream of the flow, 25mm (1in) upstream of the inlet ports to the tank. These thermocouples have an accuracy of ±0.5°C and are used to maintain the tank inlet temperature to within ±0.5°C of the desired value. Similar thermocouples are 25mm (1in) downstream of the outlet ports to measure the temperature difference of the oil across the server.

The material specifications of the technical grade white mineral oil are given in Table 4-1. In total, 11.4L (3gal) of oil are used to fill the system (tank, radiators, and tubing) [86]. This allows the tank to be filled to a height of 6.99cm (2.75in), completely submerging the server to just above the top of the extruded heat sinks. This leaves an air column height of 12.1cm (4.75in) between the top free surface of the oil and the lid of the container. The system piping consists of 1.27cm (0.5in) inner diameter vinyl tubing covered with 1.27cm (0.5in) foam insulation. This size tubing allows the system to maintain low velocity, laminar flow, reducing system pressure losses.

Table 4-1: Mineral oil physical properties [86]

Specific Gravity	**0.8555**
Density	0.8493 g/cm^3
Kinematic Viscosity	<16.02mm^2/s ASTM D445
Kinematic Viscosity Temperature	40°C

T-type thermocouples are attached with epoxy to the surface of 11 components across the motherboard. The components monitored represent a range of component types (chipsets, voltage regulators, DIMM chips) to provide a survey of the thermal performance of the two cooling methods being tested. Of these components, the voltage regulators (VRDs) have the most sensitive thermal requirements, with a maximum safe operating temperature of 85°C. If VRDs exceed this limit, they begin to throttle and degrade the overall compute performance of the system. Thermocouples are also placed in three bulk fluid locations, as well as in two locations in the air gap above the oil to help establish when steady state conditions are reached. These thermocouples have error limits of ±1.0°C and are connected to a data acquisition system which records their values in five second intervals.

The total server power consumption is measured using a Yokogawa CW121 power meter by connecting voltage and current clamps to the incoming power feed to the server. Power consumption data is recorded in five second intervals and logged on the workstation.

A complete image of the laboratory setup is shown in Figure 4-5with pertinent components of the flow loop identified.

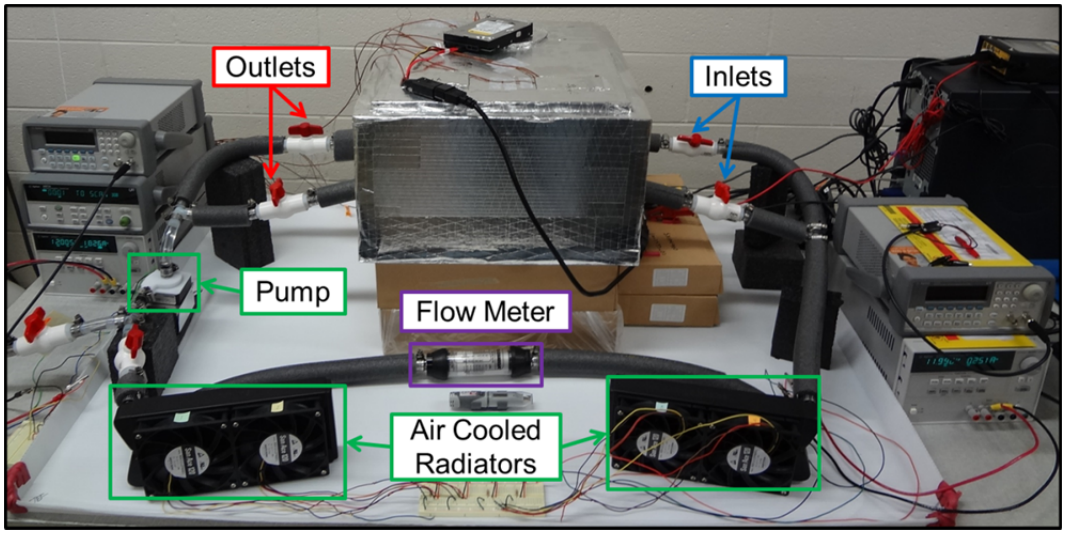

Figure 4-5: Complete experimental setup and identified flow loop elements

4.1.3 Compute load generation and data collection

To generate a computational workload on the server, a synthetic load generation program is employed. The *lookbusy* software tool allows users to set predefined CPU, memory, I/O, and networking utilization targets [74]. For this study, a workload of 75% CPU utilization with 20% memory (RAM) allocation is used as design conditions. This level of workload represents high activity that would be desired in operational service and generates near maximum heat output for the server. The native Linux operating system monitoring tools *mpstat* and *free* are used to record CPU and memory utilization levels, respectively. An internal diagnostic tool provided by the motherboard manufacturer reports data from DTS sensors in the processors, as well as, rpm readings from Hall sensors in the internal server fans.

The primary heat generating components and main driver for optimizing thermal management in the particular server under study are the CPUs. As such, CPU die temperature is the main metric of concern for this study; however, understanding and

monitoring of other motherboard components is important to the overall health of the server system.

4.1.4 Test procedure

The system oil was set to a desired inlet volumetric flow rate via speed control of the pump. By adjusting the speed of the radiator fans, a desired oil inlet temperature could be achieved. This process was manual and required several iterations of adjustments to achieve steady state conditions. The values reported here are averages over the course of at least one hour of steady state conditions, which is defined as within ±0.1lpm and ±0.5°C of the targeted volumetric flow rate and oil inlet temperature, respectively. Since typical real time workloads in data centers are not constant over long durations, the values presented here may represent worst-case conditions given the transient times required to achieve steady state. Figure 4-6 shows the temperature of several variables over the course of a test. These temperatures are the oil MB inlet temperature, which was a defining parameter of the test case, MB outlet temperature, temperature in the bulk fluid at a location in the vicinity of the center of the motherboard at a depth of about 2.54cm (1in) from the free surface of the oil, and temperature of the air gap between the top surface of the oil and the lid of the test container. It can be seen that the time to reach steady state values can take 30 minutes or more depending on the initial condition. Average steady state values over a period of at least one hour are reported here forth.

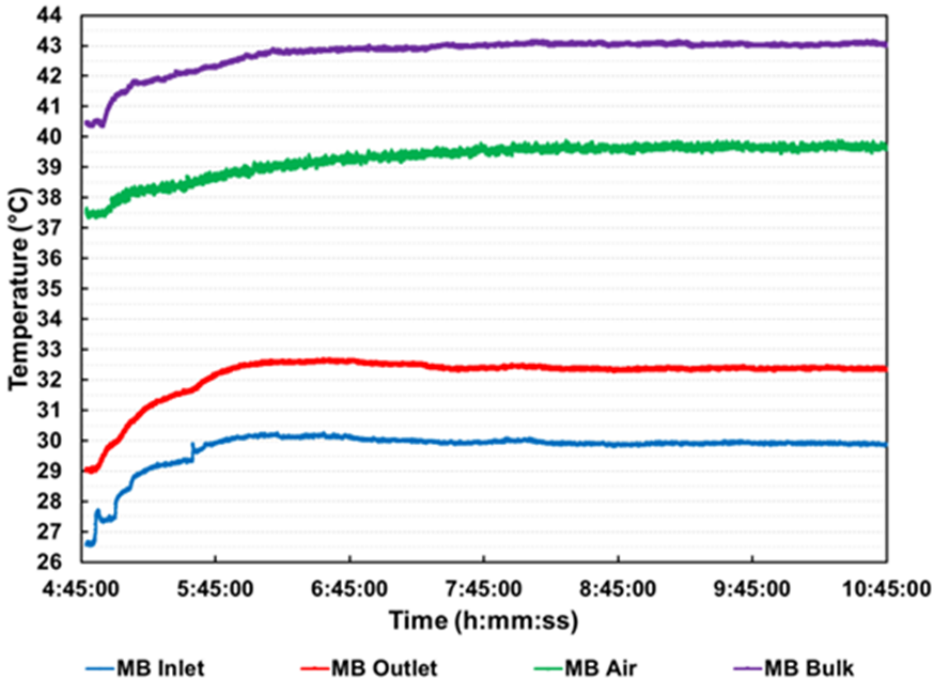

Figure 4-6: Typical test duration in establishing steady state conditions for the server under test

The test range studied included oil inlet temperatures from 30°C to 50°C, in increments of 5°C. The volumetric flow rate is varied from 0.5lpm to 2.5lpm in increments of 0.5lpm. However, at 30°C and 35°C inlet temperature, a flow rate of 2.5lpm was not achievable because the oil viscosity caused the pressure drop through the system to be prohibitively high, as will be discussed.

4.2 Results and Discussion

4.2.1 Air cooled baseline results

The results of the preliminary air cooled testing are used to establish typical operating temperatures of the server. Table 4-2 lists the average temperature of four components during steady state conditions after three repeated trials at the design compute conditions. Data recorded showed that all three of these runs occurred with an ambient inlet temperature of 25°C ±1.0°C. These components are the CPU0 die temperature, the input/output hub (IOH) chip, a memory chip on the dual in line memory module (DIMM), and a voltage regulator device (VRD) located directly behind CPU1. These four components represent a variety of component types and spaced around the motherboard. The CPU0 die temperature of 74.0°C can be used as a basis for comparison in the oil cooled results. This serves to represent an upper limit of desirable operating temperature for the CPUs. As a point of reference, the average power consumption of the server (IT components + internal server fans) during this testing was 222.4W. Roughly 4W, or less than 1% of this power, can be attributed to the fans required for cooling.

Table 4-2: Baseline temperature results of air cooled testing

Component	Run #1 (°C)	Run #2 (°C)	Run #3 (°C)	Average (°C)
CPU0	74.1	73.9	73.9	74.0
IOH	50.8	50.6	50.8	50.8
DIMM 1-B	31.6	31.7	31.6	31.6
CPU1 VRD2	70.2	70.6	70.8	70.7

4.2.2 Varying oil flow rate and inlet temperature

The resulting CPU die temperatures over the range of test conditions when using the MB Flow Path of the oil immersion setup are shown in Figure 4-7. Ambient laboratory conditions were observed to be 25°C ±1.0°C throughout the study. Quadratic curve fits are used to show the trends for increasing volumetric flow rates at given oil inlet temperatures. As should be expected, increasing the inlet flow rate at a given inlet temperature results in decreasing die temperatures. Increasing the flow rate from 0.5lpm to 1.0lpm produced the most significant impact on die temperature, resulting in a roughly 4.9°C reduction when the inlet temperature was 30°C. Figure 4-7 shows that increasing flow rate beyond 1.5lpm begins to have diminishing returns for die temperature reduction at the inlet temperatures studied. Using the 74.0°C die temperature from the air cooling results as a benchmark, it seems feasible to use a range of these operating conditions to safely cool the server. All studied flow rates at inlet temperatures of 30°C, 35°C, and 40°C provide suitable operating conditions. When the inlet temperature reaches 45°C, flow rates beyond 1.5lpm are needed to maintain CPU temperatures below the 74.0°C threshold.

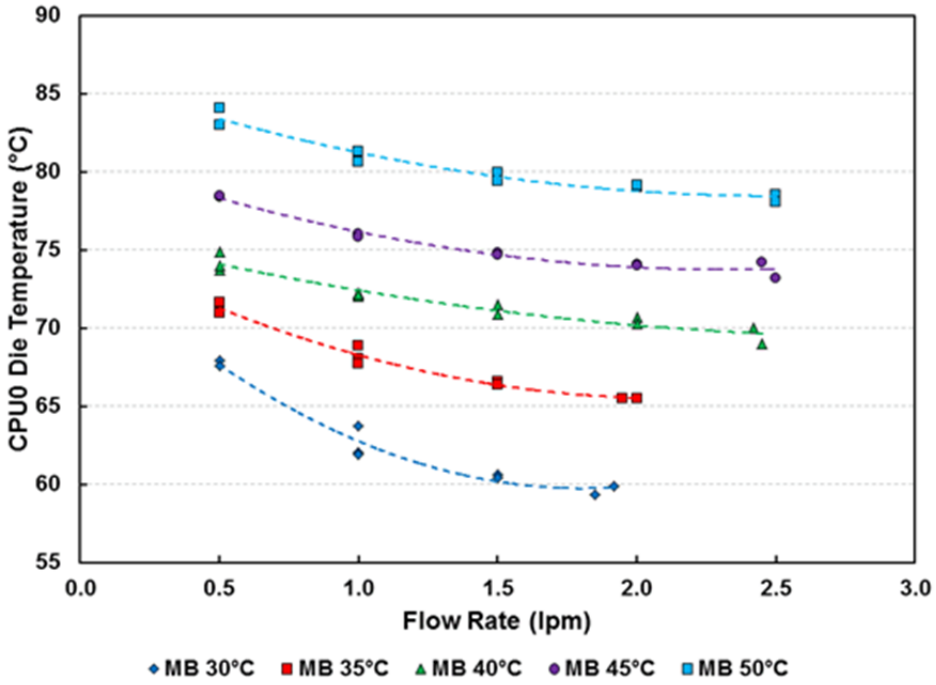

Figure 4-7: Impact of increasing oil volume flow rate on CPU die temperature along lines of constant oil inlet temperature for eh MB Flow Path case

When the PSU Flow Path was utilized, in which only the PSU inlet and outlet ports were open, resulting CPU temperatures were higher for given flow rates and inlet temperatures as seen Figure 4-8. These higher temperatures were to be expected, since the flow of oil across the CPU heat sinks is not directly focused from the container inlet as in the MB Flow Path case. Figure 4-8 also uses quadratic curve fits to highlight the trends with increasing flow rate; however, some of these trends are fairly linear. For example, at an inlet temperature of 35°C, each 0.5lpm increase in flow rate from 0.5lpm to 2.0lpm results in a roughly 2.3°C reduction in CPU die temperature.

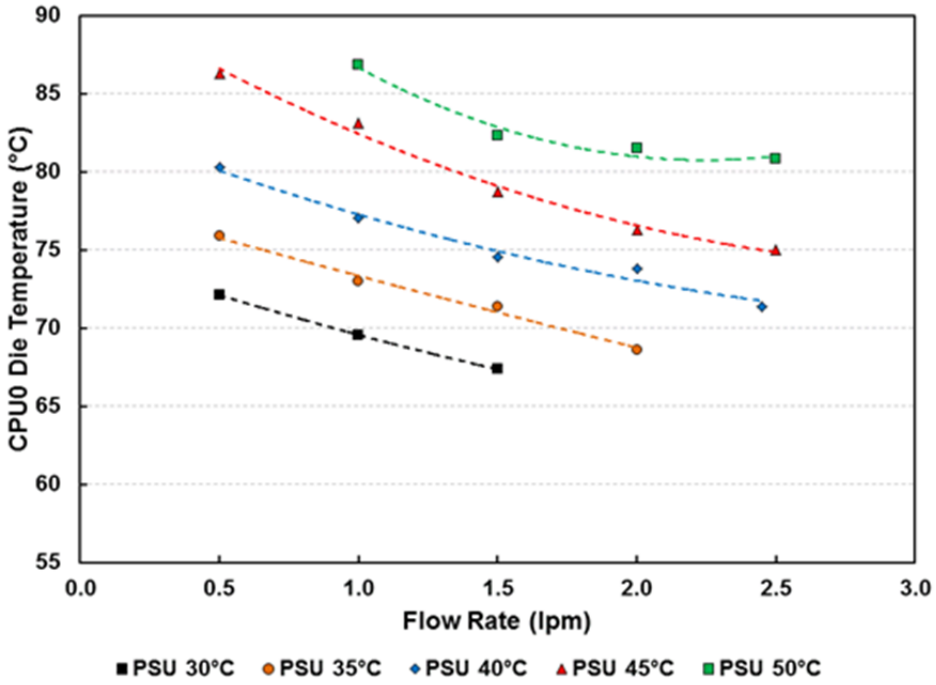

Figure 4-8: Impact of increasing volume flow rate on CPU die temperature along lines of constant oil inlet temperature for the PSU Flow Path case

Using the PSU Flow Path, inlet flow rate and inlet temperature conditions that result in 74.0°C die temperature or less are greatly reduced. Here, flow rates of 1.0lpm and above are needed at 35°C inlet temperatures and flow rates of 2.0lpm and above are needed at 40°C inlet temperatures. These results indicate that ducting of flow over key components is of importance in an oil cooling system. Although the fluid velocities are low through heat sinks, bulk fluid motion of the system does not provide the same cooling performance as directed flow, particularly at lower flow rates.

4.2.3 Partial PUE and system power consumption

Power Usage Effectiveness (PUE) has been widely adopted and used throughout as the standard efficiency metric for data centers. PUE, originated by Belady and others, determines data center efficiency by taking the total facility power consumed divided by IT load (useful work of the data center) [2]. Detailed description and development of the term can be found in [3]. Partial PUE (pPUE) metrics can be developed to understand efficiency of specific subsystems and subsets of the data center. pPUE of cooling systems can be expressed as follows:

$$pPUE_{Cooling} = \frac{Cooling\ Power + IT\ Load}{IT\ Load} \quad (3\text{-}1)$$

Equation 3-1 represents the efficiency of just the cooling system. For the present test setup, the entire cooling loop may be thought of as representing a “complete” data center loop in which the heat rejected by the IT equipment is eventually rejected to an ambient air of 25°C (77°F) from the air cooled radiators to the laboratory environment. The caveats to this assumption are discussed in a following section. The cooling energy is the sum of the power drawn by the centrifugal pump and radiator fans. Table 3-3 compiles the $pPUE_{Cooling}$ values obtained at the various conditions studied for the MB Flow Path case. Over the range of conditions studied, $pPUE_{Cooling}$ values as low as 1.027 are achievable, however, this does not necessarily coincide with the minimum total system power operating point. In the range of desired CPU temperatures (i.e. CPU0 DT < 74°C), $pPUE_{Cooling}$ values range from 1.036 to 1.170. Similar values can be seen in Table 3-4 for the PSU Flow Path case.

Table 4-3: $pPUE_{Cooling}$ at given tested operating conditions with the MB Flow Path

MB Flow Path		Oil Inlet Temperature (°C)				
		30	35	40	45	50
Flow Rate (lpm)	0.5	1.055	1.036	1.041	1.030	1.027
	1.0	1.086	1.051	1.058	1.038	1.035
	1.5	1.124	1.068	1.079	1.053	1.046
	2.0	1.170	1.088	1.102	1.072	1.059
	2.5	-	-	1.129	1.095	1.075

Table 4-4: $pPUE_{Cooling}$ at given tested operating conditions with the PSU Flow Path

PSU Flow Path		Oil Inlet Temperature (°C)				
		30	35	40	45	50
Flow Rate (lpm)	0.5	1.056	1.040	1.037	1.027	-
	1.0	1.075	1.058	1.051	1.038	1.035
	1.5	1.097	1.080	1.068	1.053	1.047
	2.0	-	1.106	1.087	1.075	1.059
	2.5	-	-	1.109	1.101	1.070

Although server power consumption was recorded for the air-cooled baseline test, this data does not lend itself to $pPUE_{Cooling}$ calculations. Rough extrapolations as to the energy required by a CRAC-based or economizer-based system to achieve the 25°C inlet air temperature would need to be made and are beyond the scope of the present work.

These results can be expanded to begin to understand efficiency of an oil immersion system at different ambient conditions. A maximum system approach temperature can be developed by taking the difference between the oil outlet temperature and ambient laboratory temperature. In this case, the oil outlet temperature is calculated using the measured inlet temperature plus the temperature rise, ΔT, across the server using the standard steady-flow thermal energy equation:

$$q = \dot{m} c_p \Delta T \tag{3-2}$$

where q is the measured IT power, c_p the oil's specific heat, and $\dot{m}$ the mass flow rate, taken as the product of the oil's density and measured system volume flow rate. The relation between the maximum system approach and system efficiency, shown in Figure 4-9, is fit with an exponential curve as can be expected in heat exchanger analysis. These results are helpful to system designers to begin estimating the system efficiency over expanded ambient conditions. This cannot replace a detailed energy flow model of a facility but serves as a starting point for reference designs.

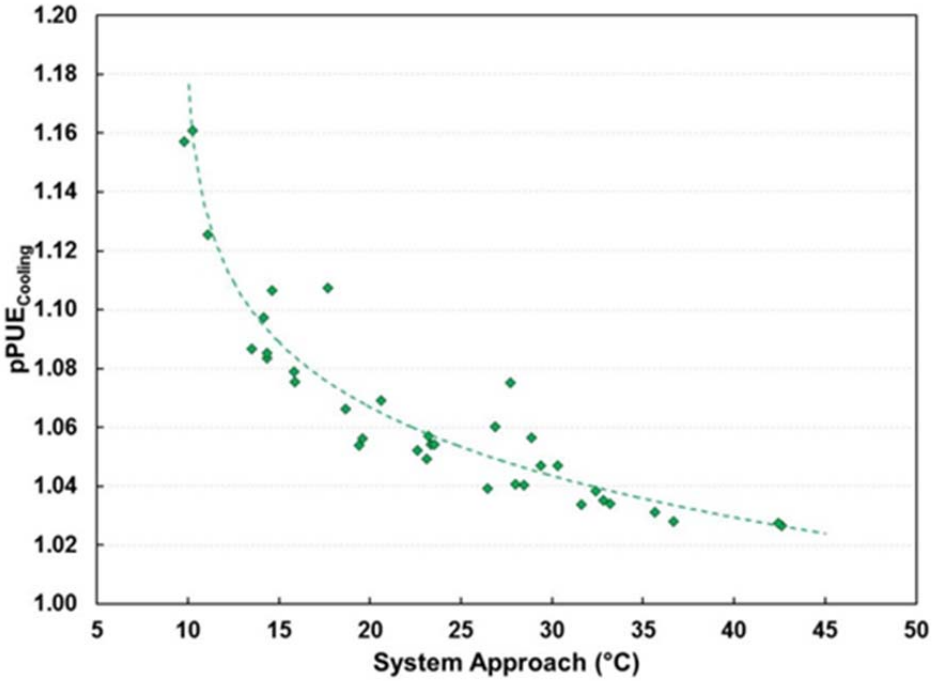

Figure 4-9: Relation between system approach temperature and efficiency

PUE and $pPUE_{Cooling}$ alone do not provide a complete picture of the energy efficiency of a data center cooling scheme. With the goal of minimizing total facility power for a data center, all components' energy consumption must be considered holistically. Figure 4-10 shows a surface plot of the minimization of total power consumption for the system for the MB Flow Path case. This includes the power to the server (IT load), the power consumed by the pump, and the power drawn by the radiator fans. It can be seen that the minimum power draw for the system occurs at 40°C (104°F) inlet temperature and 0.5lpm flow rate. At this point, the CPU die temperature is 74.1°C, just above the targeted upper limit. Although the best $pPUE_{Cooling}$ value occurred at an inlet flow rate of 0.5lpm and 50°C inlet temperature, the total system power at this point was 2% higher and had a CPU0 die temperature of 83.4°C.

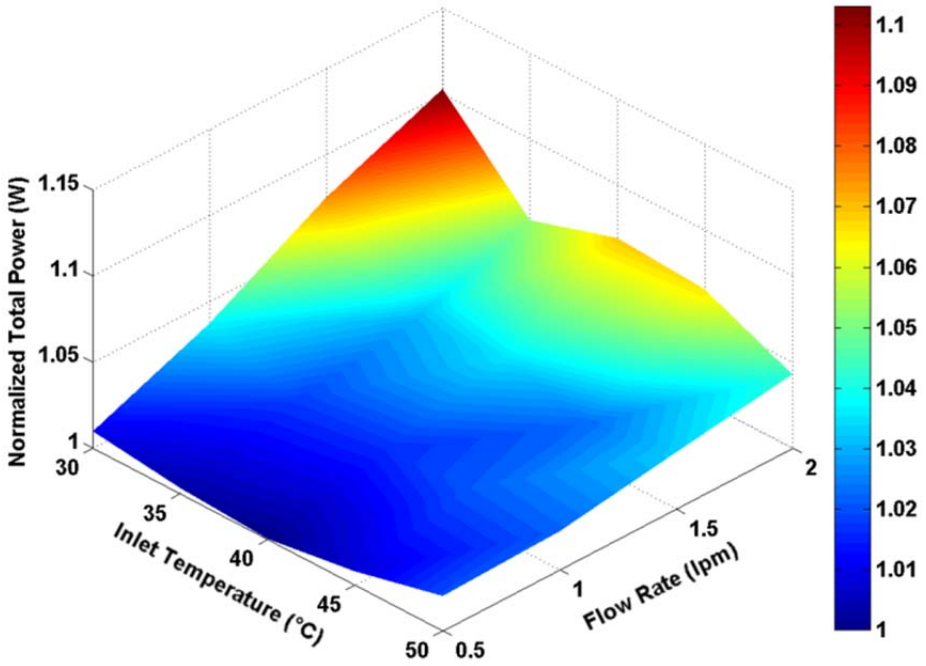

Figure 4-10: Surface contour of the normalized total system power consumption for the MB Flow Path case

The somewhat conflicting results of $pPUE_{Cooling}$ and total system power can best be understood by analyzing the individual system components (IT, pump, and fans). The general trends are as follows:

- IT power increases non-linearly with CPU die temperature (and all component temperatures in general) due to leakage current effects in silicon devices. This trend is roughly quadratic as shown in Figure 4-11, which shows the relation between CPU0 die temperature, the input/output hub (IOH), and one of CPU1's VRDs with the total server power from the data collected in the MB Flow Path cases. All three of these components have similar trends with total server power.

This is not a complete relation since the total server power depends on power consumption and leakage of all components on the MB and PSU, but it does provide a rough estimate of the general behavior. For example, increasing the CPU0 die temperature from 60 to 70°C results in a roughly 2.3W increase in total server power. Increasing the die temperature from 70 to 80°C results in a 4.7W increase in power consumption. There is certainly incentive to operate the server at lower component temperatures for this reason; however, it must be complimented with the rest of the system.

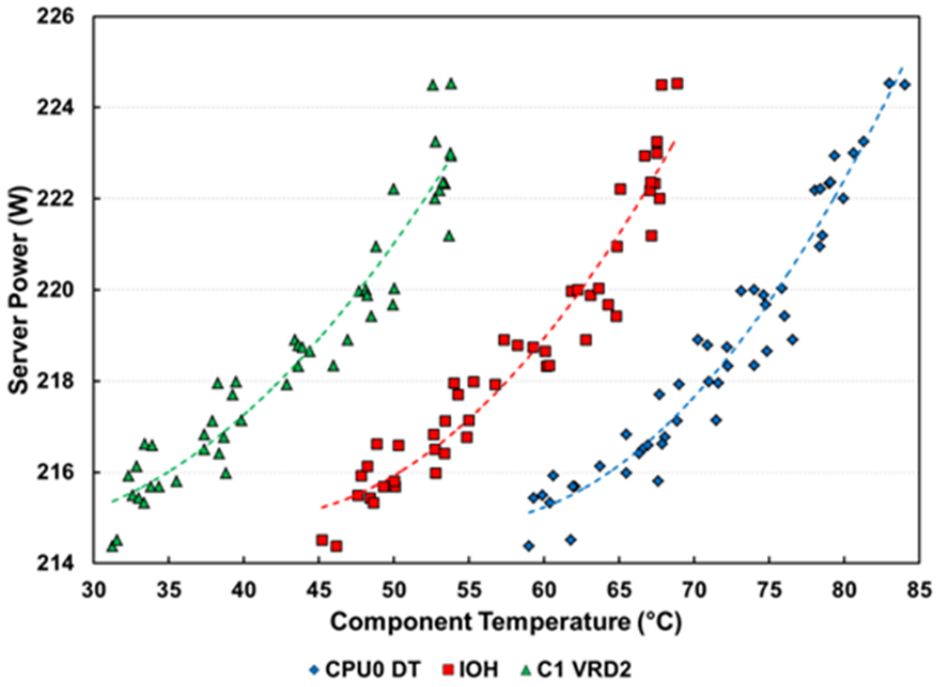

Figure 4-11: General relationship between component temperatures and total system power based on the data collected in the MB Flow Path case

- It is known that the viscosity of oils vary with temperature. This has a direct impact on the system pressure and hence, pumping power required of the

system. A standard correlation relating viscosity and temperature for transformer oils (similar to the light mineral oil used in this study) is given in [56] as:

$$\mu = C_1 * Exp\left[\frac{2797.3}{T + 273.2}\right] \tag{3-3}$$

where μ is the dynamic viscosity in centipoise, T is the temperature in °C, and C_1 is a coefficient for scaling (a value of 0.0013573 is provided in the reference for transformer oil). The interdependence of viscosity with Reynolds Number (Re), Reynolds number with friction factor (*f*), and friction factor with pressure drop (Δp) for laminar flow,

$$\mu \propto \frac{1}{Re} \propto f \propto \Delta p \tag{3-4}$$

should eventually manifest itself in the pumping power of the system by the relation:

$$P_{pump} = \Delta p * \dot{V} \tag{3-5}$$

where $\dot{V}$ is the volumetric flow rate. Figure 4-12 shows suitable agreement with the pumping power required to maintain 1.0lpm of flow rate in the immersion system for the MB Flow Path and the viscosity trend predicted by Equation 3-3. Over the range of inlet temperatures studied, a 43.5% reduction in viscosity is predicted and an average reduction in pumping power of 42.6% observed for a given flow rate.

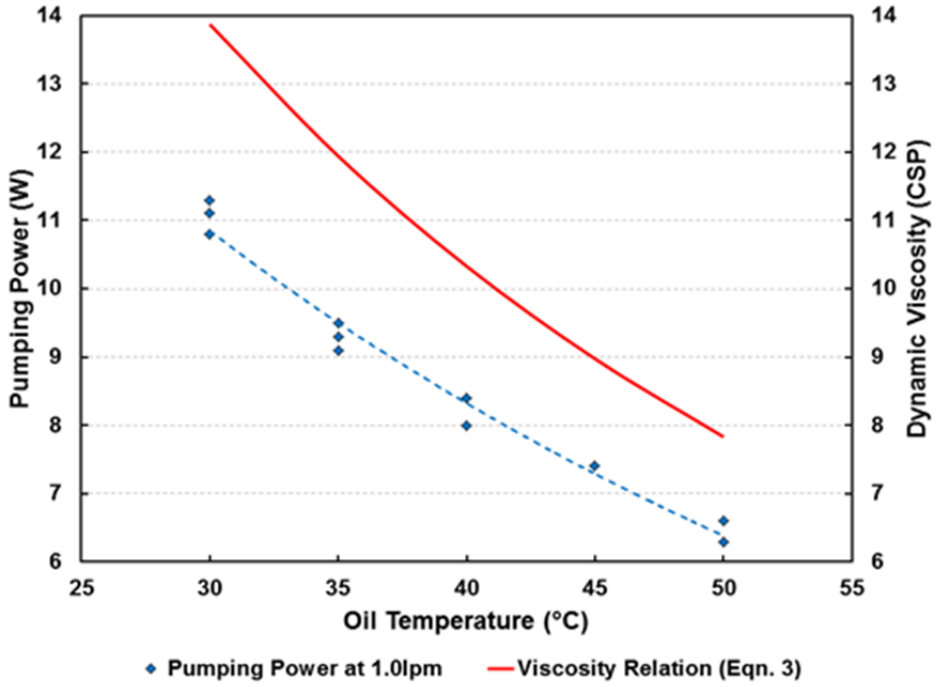

Figure 4-12: Comparison of trends for temperature dependent pumping power for the current system and temperature dependent viscosity of transformer oil as predicted by Equation 3-3

Looking at the pumping power over the range of temperatures and flow rates studied, as in Figure 4-13, it is easy to see there is ample incentive to operate at higher temperatures. The curves in Figure 4-13 are fit with cubic trend lines which should be expected from the pump affinity laws which state that pump power is proportional to the cube of the impeller or shaft speed.

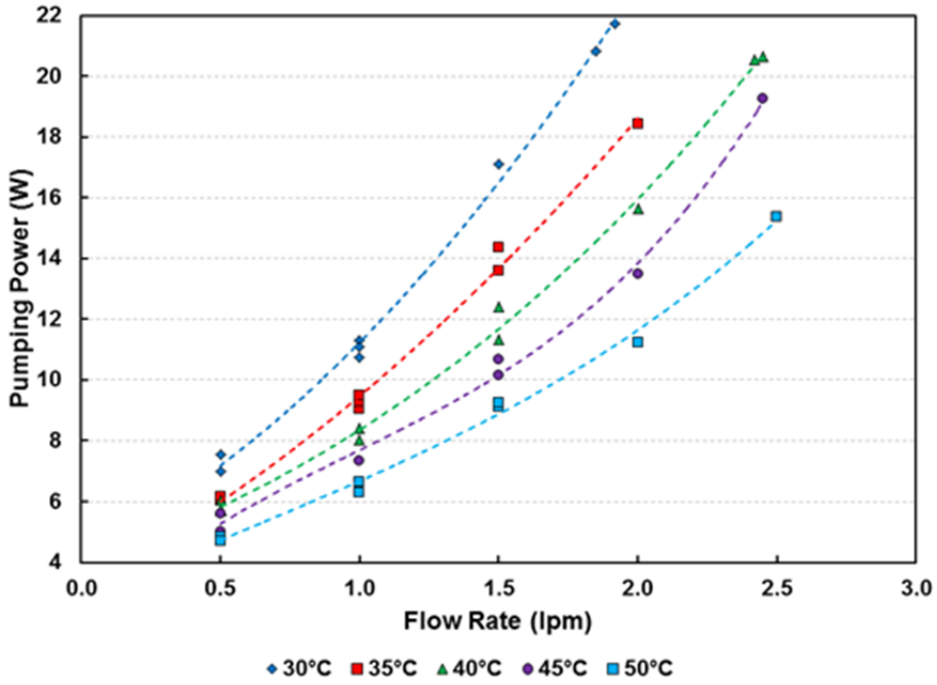

Figure 4-13: Temperature dependent oil flow rates and cubic relation of pumping power to flow rate

For all cases studied, the flow is distinctly laminar, well below the transition regime. At the lower end, with a flow rate of 0.5lpm and oil temperature of 30°C, the Reynolds number within the system piping is ~50. At the other extreme, with a flow rate of 2.5lpm and oil temperature of 50°C, the Reynolds number is ~440. These values will become substantially lower as the flow enters the larger volume of the tank and flows across the server.

- The radiator fan power contributes minimally (< 1.2%) to the total system power at oil inlet temperatures of 35°C and above because they are run at idle speeds. Their impact on $pPUE_{Cooling}$ values over the operating conditions is relatively constant

Understanding these interrelated trends of system components is central to selecting the optimal operating conditions for an oil immersion cooled system. The results presented here for the MB Flow Path and PSU Flow Path cases provide some initial bounding operating conditions for an oil cooled system. Since the MB Flow Path can be considered a “better” case scenario for the cooling, it can utilize higher temperature fluids and lower flow rates.

4.2.4 Comparison to air cooled baseline results

Although a direct comparison of PUE is not possible between the air cooled baseline tests and the oil immersion results, other performance parameters may be used. Thermal resistance to heat transfer between a device and coolant is given by

$$R_{th} = \frac{T_d - T_c}{q} \tag{3-6}$$

where T_d is the device temperature, T_c the incoming coolant temperature, and q is the heat dissipated by the device. In this way a comparison between oil and air can be made by focusing on the critical device temperature, the CPU, and the server’s inlet coolant temperature divided by the total IT power of the system. From the air cooled baseline tests, air provides a thermal resistance of 0.224°C/W. Over all oil cases studied, oil provided a thermal resistance ranging from 0.128 to 0.175°C/W, with an average resistance of 0.147°C/W. This average value is a 34.4% improvement over the baseline air cooling case. This improved performance is consistent with the results reported in [61].

Comparisons of surface temperature measurements show improvements in thermal performance across all components on the motherboard. Table 4-5 compares the rise in surface temperature of components, T_d, over the inlet coolant temperature, T_c, with

a lower value being better. The air results are taken from baseline tests and oil results are the average temperature rise over all the MB Flow Path test conditions. The most significant improvements are seen across all voltage regulating devices (VRDs) for the CPUs which are critical for power delivery to the CPUs. In all conditions, including 50°C oil inlet temperature at a flow rate of 0.5lpm, the VRDs exhibited lower surface temperatures than in the air cooled baseline case. From a thermal reliability perspective, this is significant.

Table 4-5: Thermal performance comparison of motherboard component surface temperatures

Component	$\Delta T = T_d - T_c$		Percent Reduction
	Air (Baseline)	Oil (Average)	
IOH	26.9	18.7	30.5%
DIMM1 VRD	8.1	4.6	43.2%
DIMM 1-A	20.9	16.4	21.5%
DIMM 1-B	8.2	4.1	50.0%
CPU1 VRD1	43.7	6.3	85.6%
CPU1 VRD2	43.8	3.8	91.3%
CPU0 VRD3	43.6	4.1	90.6%
CPU0 VRD4	40.1	7.3	81.8%
DIMM 0-A	6.5	4.2	35.4%
DIMM 0-B	16.6	15.1	9.0%
DIMM0 VRD	7.2	4.4	38.9%

4.3 Additional Discussion

The results of this study are helpful in establishing a range of flow requirements that can be expected on per server or per Watt basis for a data center oil cooling strategy. These figures may be helpful in design of a larger system and provide relative sizing requirements for pumps, heat exchangers, and system pressure drop. Other than using temperature difference across the server (ΔT) for the design criteria, which may be difficult to determine, these results establish flow requirements based on component temperatures, namely CPU, which are the critical elements of the system.

The efficiency values ($pPUE_{Cooling}$) presented above are unique to the system being tested. Certain simplifications and modifications were required for laboratory testing that may not be in place in an actual data center implementation. A fully built oil immersion data center may contain additional heat exchanges (i.e. oil-to-water, water-to-air) before final rejection of heat to the environment. The additional exchanges and piping will require increased cooling energy consumption. However, larger components (pumps, fans, etc.) are generally more efficiency than geometrically similar, smaller components. Because of these counteracting efficiency factors, the results here are expected to be indicative of what may be seen at larger scales.

As discussed in [71], this Open Compute server was strategically designed and optimized for air cooling. Many of the design aspects are beneficial in oil cooling, but it is expected that further optimization for oil cooling can be achieved. For example, the heat sinks may be better designed for oil flow conditions as fin efficiencies in liquid and air are markedly different. This modification may be especially helpful in the PSU Flow Path case in which heat transfer is mainly from bulk fluid motion. Modifications to the current

system, such as the removal of the air duct and PSU chassis present material savings that may not be achievable in standard air cooled server designs.

Further enhancements to the system's performance may be gained by more prudent selection of the oil used. Additional oil types such as vegetable oils and other mineral oils may provide better heat transfer and fluid characteristics compared to the white mineral oil used for this study. In general, there are a wide range of topics that can be explored to better understand this promising cooling technique.

4.3.1 Mixed convection in oil immersion applications

As mentioned, current servers are designed and optimized for performance using traditional air cooling methods. This typically consists of fans and blowers moving large volumes of air though servers in turbulent forced convection. However, the fluid performance of mineral oil is starkly different, depending heavily on two nonlinear influences, namely, the temperature dependent nature of oil viscosity and mixed convection. The first of these was discussed in the findings of section 3.2.3. The second nonlinear influence, mixed convection, may greatly change the design standards and operating conditions for a server system.

As a quick proof, assuming the current dimensions of a typical air-cooled heat sink as shown in Figure 4-14, a sample calculation to determine the convection regime of an oil cooled system is performed. The natural convection component is described by the Grashof number which is defined as:

$$Gr_L = \frac{g\beta\Delta T L^3}{\nu^2} \tag{3-7}$$

where g is the component of gravity, β is the expansion coefficient, ∆T is the driving temperature difference, L is the length scale physically relevant to the problem, and v is

the kinematic viscosity of the fluid. The forced convection component is described in terms of the Reynold's numbers which is given as:

$$Re = \frac{VD_h}{\nu} \tag{3-8}$$

where V is the fluid velocity and D_h is the hydraulic diameter when considering internal channel flow. The convection regime is determined by the ratio of the Grashof number and square of the Reynold's number for the flow.

$$\frac{Gr_L}{Re^2} \tag{3-9}$$

When this ratio is much greater than unity (>> 1), buoyancy forces dominate and the flow corresponds to natural convection. Buoyancy forces can be neglected when the ration is much less than unity (<< 1) and conditions correspond to forced convection. As the ratio approaches unity (~ 1) both buoyancy and inertial forces must be considered and the flow corresponds to mixed convection.

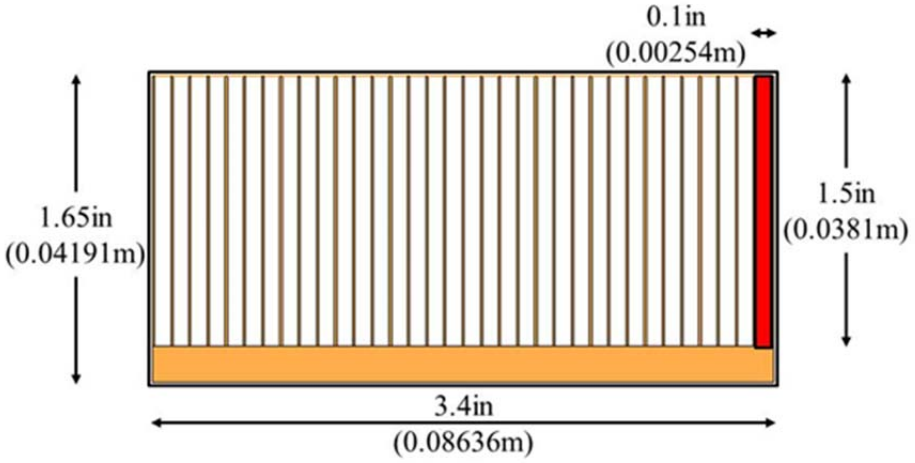

Figure 4-14: Sample heat sink dimensions for use in the calculation of nondimensional parameters. The highlighted red region is used for calculation purposes

Table 4-6 shows the calculated values of this ratio for the simplified case, assuming various flow rates though the heat sink of Figure 4-14. It is apparent that depending on the particular flow conditions, the combined effects of natural and forced convection will influence the system.

Table 4-6: Calculated mixed convection ratio values over a range of flow rates for the heat sink considered

Flow Rate	**0.5**	**1.0**	**1.5**	**2.0**	**2.5**	**3.0**
$\frac{Gr}{Re^2}$	43.1	5.4	1.6	0.7	0.3	0.2

A well designed system should be able to take advantage of enhanced thermal performance due to buoyancy induced flow at a reduced pumping power requirement. These conditions help drive the motivation for this work. An example of the improved heat transfer performance in mixed convection was shown by Maughan and Incropera in Figure 4-15 [87]. Here it is seen that as secondary flows are initiated due to buoyancy forces, improved heat transfer over pure forced convection is achieved.

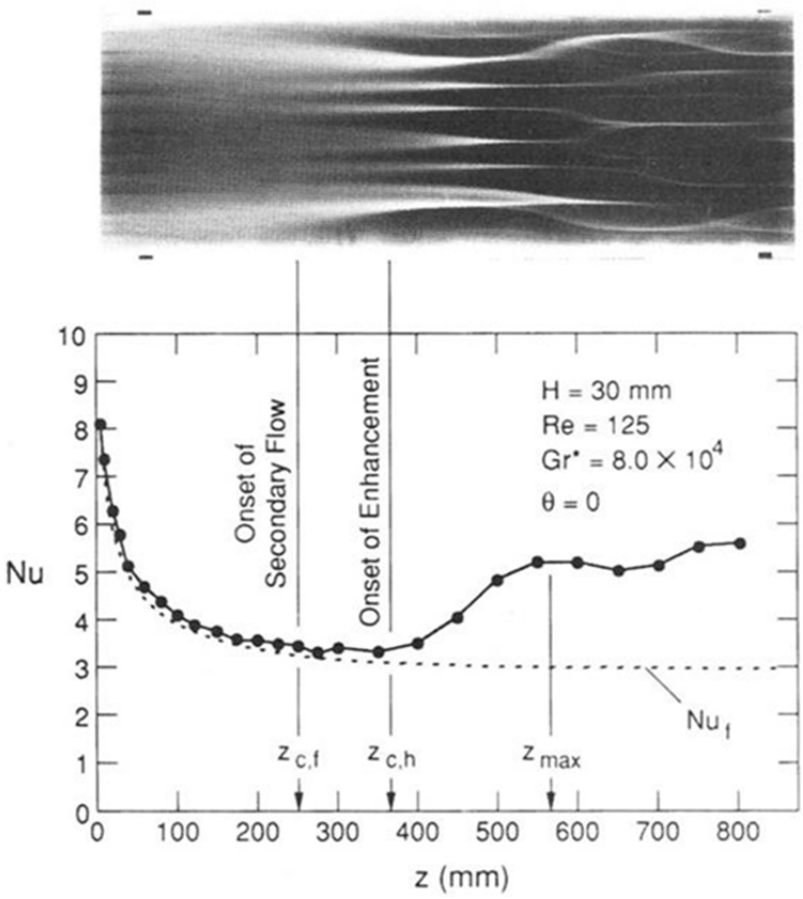

Figure 4-15: Improved heat transfer performance with the onset of secondary flows (shown in solid line) over forced convection alone (dashed line) [87]

4.3.2 Future work

To maximize the advantages shown and make oil immersion cooling more practical for adoption by industry, it is imperative to begin designing systems specifically for this use. The goals of such optimization should be to increase the functional density of IT equipment within a given footprint (kW/m^2) by taking advantage of the improved thermal properties of mineral oil. Additionally, opportunities to further reduce the operating costs associated with cooling and if possible, minimize capital costs should be

explored. An experimental fluid test circuit has already been constructed and in place to expand this work and is described below.

Three third generation Open Compute servers, shown in Figure 4-16, are to be used for baseline study and optimization. Each server contains two CPUs, each with a thermal design power (TDP) of 115W and four banks of DIMMs which can produce up to 20W per bank. This area in which a total of six CPUs and twelve banks of DIMMs reside can generate roughly 960W of heat load in a 12” by 28” planar footprint. This is considered high density in the data center industry, and in total, the servers can produce up to 1kW of IT within a 2U (height) form factor.

Figure 4-16: Three third-generation Open Compute servers in prepared experimental testing tank

A polycarbonate tank has been constructed to house the three servers and allows for circulation of mineral oil though the servers for heat removal. The test container, shown in Figure 4-17, consists of an inlet plenum where cool oil enters. It is designed such that the servers are oriented vertically as is most commonly done in commercial systems and to take greatest advantage of the mixed convection properties discussed in the previous section. Temperature is monitored within the plenum for system control purposes. From here, oil flows through two orifices and is distributed to the three servers. As oil flows over the components of the servers it is heated and temperature measurements are taken at the outlet of each server. Once oil reaches the top of the tank it spills over into a return reservoir where it will exit the tank to an external fluid circuit.

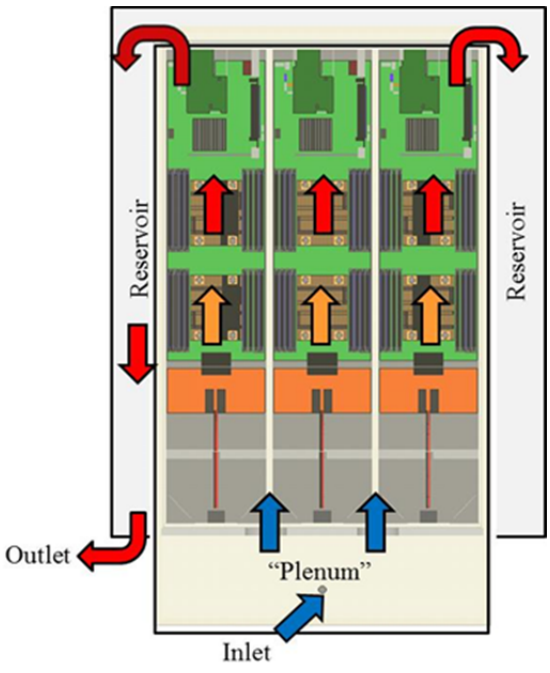

Figure 4-17: Schematic of the fluid flow path and heating through the server test container

The external fluid loop is shown schematically in Figure 4-18. The loop consists of several regulating valves, a 177 micron particulate filter, two fluid pumps, an air cooled radiator, and a positive displacement flow meter. As an improvement upon the cooling loop used in the previous study of just a single server, this fluid oop takes advantage of monitoring and control circuitry to automate much of the testing. Of the two pumps, one (Laing D5-38/710B Vario) operates with a five-speed variable control setting to establish a minimum desired flow setting for a given test condition. The second, smaller pump (Swiftech MCP-35X) operates across continuous variable speeds with the aid of PWM functionality and is used as fine tuning to regulate the flow rate of oil through the system to precise values. The air cooled radiator, used for heat rejection from the oil to the ambient environment, has four 120mm DC fans attached. These fans operate at variable speed and, with the aid of PWM control, are used to maintain the oil inlet temperature based on measured values within the plenum of the polycarbonate tank. The temperature of the oil at the inlet and outlet of the radiator are also measured for calculating total heat rejection from the fluid. The flow meter (Omega FPD2003) is a positive displacement type, specifically selected for its ability to maintain accuracy across the range of viscosity values experienced by mineral oil. It was selected with a pulse output, recorded as a frequency response by the data acquisition unit. The reported value of volumetric flow rate is sent to a microcontroller circuit board (Arduino Mega 2560) that relays a PWM signal to the smaller pump to maintain system flow rate within a desired range.

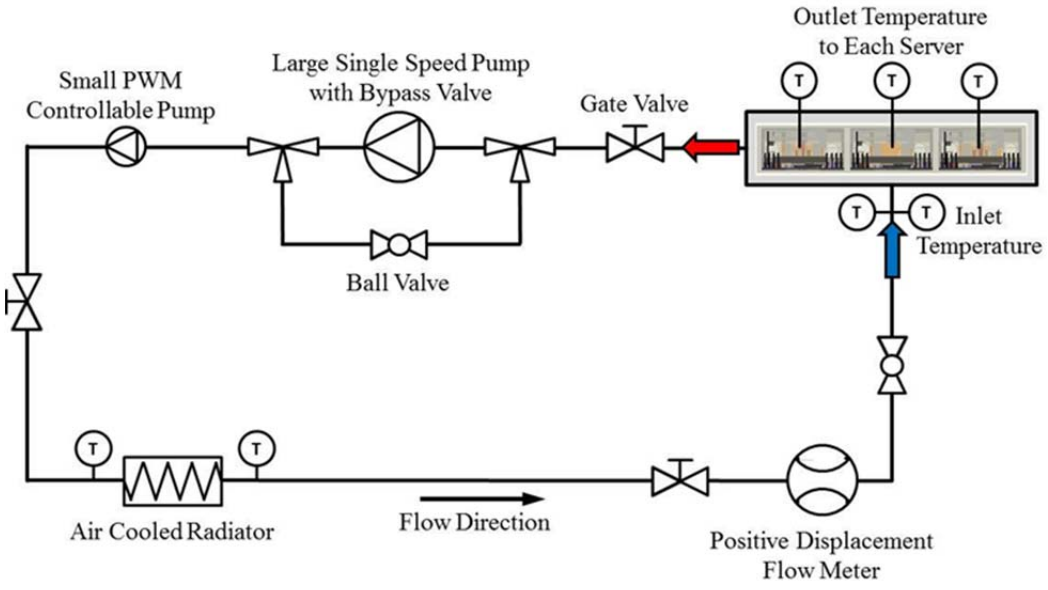

Figure 4-18: Schematic of fluid circuit for use in future testing

4.3.3 Test methodology

Using the above experimental setup, the servers will be characterized for their performance in mineral oil in the current configuration. Using synthetically generated computational workloads, the amount of heat generated by the primary server components (CPUs and DIMMs) can be well controlled. Various heating loads and patterns will be placed on the server to stress the system. Following baseline testing, modifications shall be made to the server hardware to test new design goals. The results and outcomes of these experiments with baseline and modified server hardware should provide detailed information for advancing the performance of IT equipment immersed in mineral oil. Design guidelines for future servers and rack systems for use in oil immersion applications may be established based on these finding.

Chapter 5

Mechanical Observations and Material Properties in Mineral Oil

A primary concern by data center industry professionals regarding mineral oil immersion techniques is the impact of the fluids on the long term reliability of components and systems. By fully immersing a server in oil, a company may be voiding the warranty on their equipment and expose themselves to potential failure costs. Current industry data regarding the reliability of server systems after immersion in mineral oil suggest that there is no detrimental impact to components [38] [88]. However, the remarks made in literature are anecdotal, not providing detail information or data, limiting their utility to the industry at large.

The study undertaken and discussed below presents a first look at the impact of mineral oil on server components. This includes high level visual observations, microscopic observations made by sectioning server components, and a more detailed study of the change in material properties that results from exposure of printed circuit boards (PCBs) to mineral oil. These results provide general design guidelines that help improve server performance in oil immersed systems.

5.1 Visual/Cosmetic Observations

A sample of three third-generation Open Compute servers immersed in mineral oil for a six month period for thermal testing were taken apart, photographed, and sectioned for imaging to document the effects of oil on server components. General observations include oil's impact on adhesives such as that found under labels, inks components and bar code identifiers, and the retention of particulate contaminates on

residual surface oils after removal of equipment from a bulk oil fluid. Figure 5-1 show the fading of screen printed component markings on the memory chips of the DIMMs. Similar issues with the bar code identifier labeling ink of the server chassis itself is seen in Figure 5-2. In Figure 5-3 the label associated with a small peripheral component board easily falls off due to pooling of oil under its adhesive surface. After removal of servers from a bulk fluid bath of oil, a significant amount to of residual oil remains on the surface of surfers due to the surface tension properties of the fluid. In Figure 5-4, this residual surface fluid has captured airborne dust particles to the surface of the server.

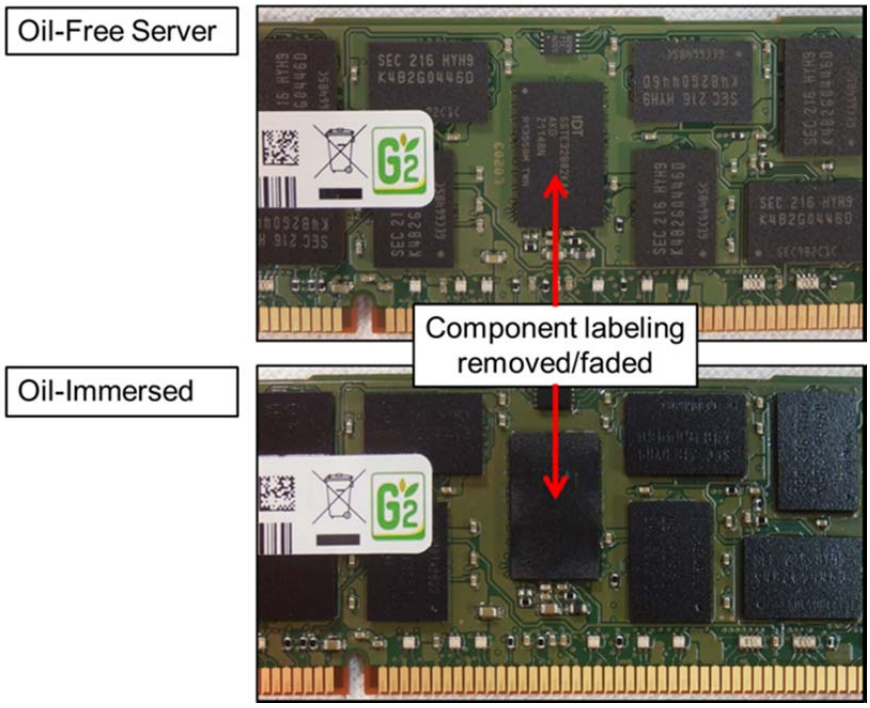

Figure 5-1: Fading component identifies as a result of oil exposure

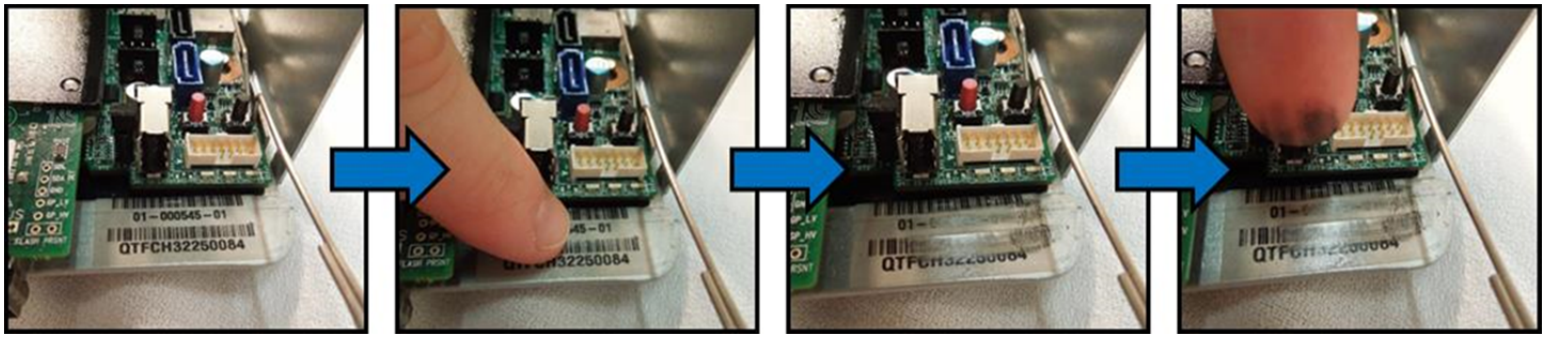

Figure 5-2: Smearing of labeling inks by light touch after immersion in mineral oil

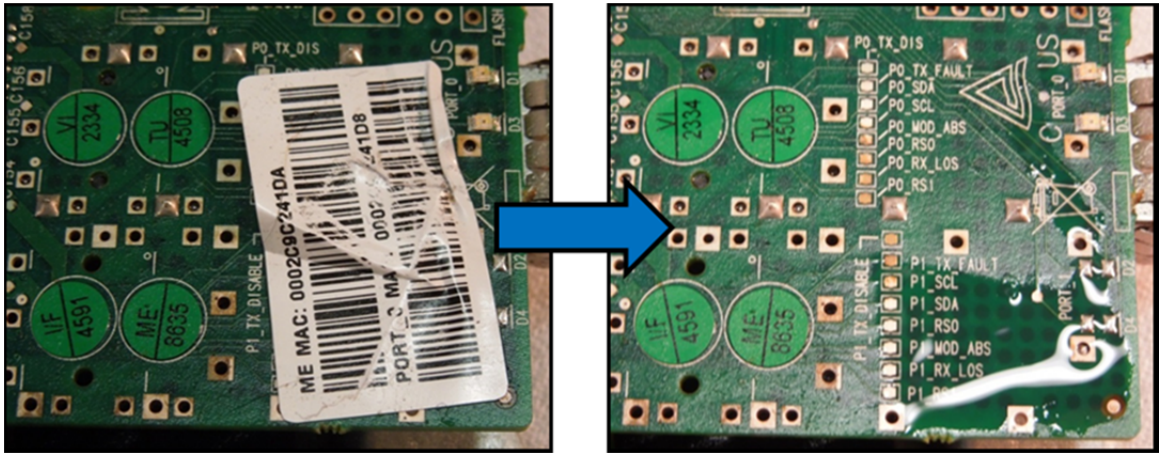

Figure 5-3: Removal of part labeling due to pooling of oil under adhesives

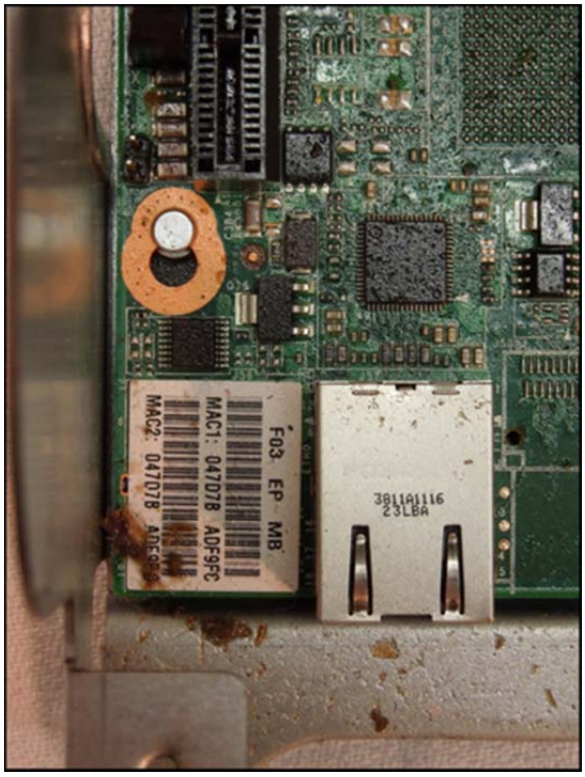

Figure 5-4: Dust and debris accumulation due to residual surface oil after removal of servers from oil bath

Based on the observations made during this study, it is expected that submerging servers in mineral oil will have an impact on the operational performance of data centers. Inventory tracking of individual components or whole servers may be compromised if labels were to fall off or should their ink and barcodes become unreadable. Dislodged labels may also pose a clogging or contamination concern should they fall into the bulk fluid or server tank. Additional impact of the residual oils on the servers will impact serviceability of equipment. The propensity to attract particulate contaminates may be an issue should servers be placed back into the bulk fluid with dust

and dirt still on the surface. Careful process procedures may need to be developed to accommodate these new operational challenges associate with oil cooling.

5.2 Microscopy Observations

A more detailed visual study was carried out by taking cross sections of various components to determine the microstructure of electronic packages. The same third-generation Open Compute servers as in the previous section were placed in molding compound, sectioned, and polished. Control samples of servers that were not exposed to oil and used in traditional air-cooled based testing underwent the same treatment. The details of the package structure were observed under microscopes.

Figure 5-5 offers a comparison of solder balls from the backside of the memory module attached to the DIMMs. As can be seen, there are no noticeable deformations, change in size, or cracking of solder balls. In addition, the intermetallic compound (IMC) layers which provide the mechanical and electrical connection between PCB-solder ball and solder ball-substrate interfaces showed no change in thickness between air cooled and oil cooled samples. The chip underfill material, which strengthens the mechanical connection between a flip chip package and substrate, also showed no detectable variation between air and oil cooled samples. In Figure 5-6 it is seen that there are no size variations in the metal layers of the packaging substrate. The trace thickness does not change or alter after exposure of the server in an immersive environment.

Additional samples of PCB boards showed no delamination, swelling, or warpage of layers after extended periods of submersion in mineral oil. In Figure 5-7, a cross section at a plated through hole location in the motherboard has maintained its structural

integrity. Similar observations can be made from Figure 5-8 at an edge location on the PCB.

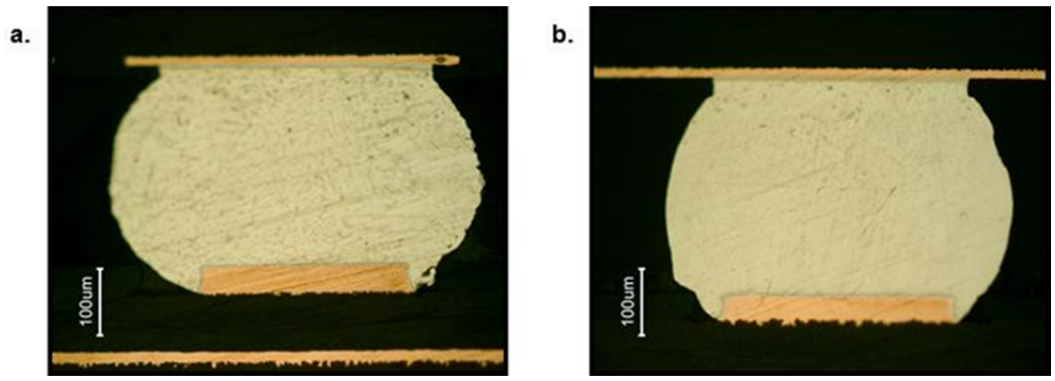

Figure 5-5: Comparison of microstructure of solder balls taken from (a) an air cooled server and (b) an oil immersed server

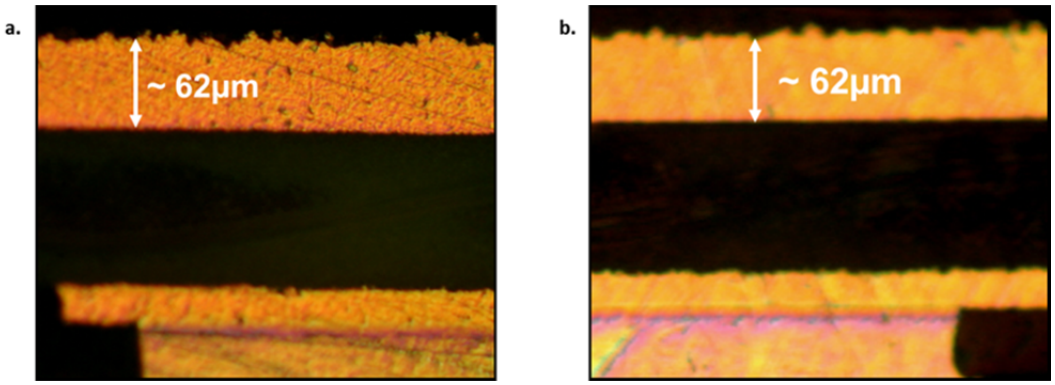

Figure 5-6: Comparison of substrate layer of BGA package taken from (a) an air cooled server and (b) an oil immersed server

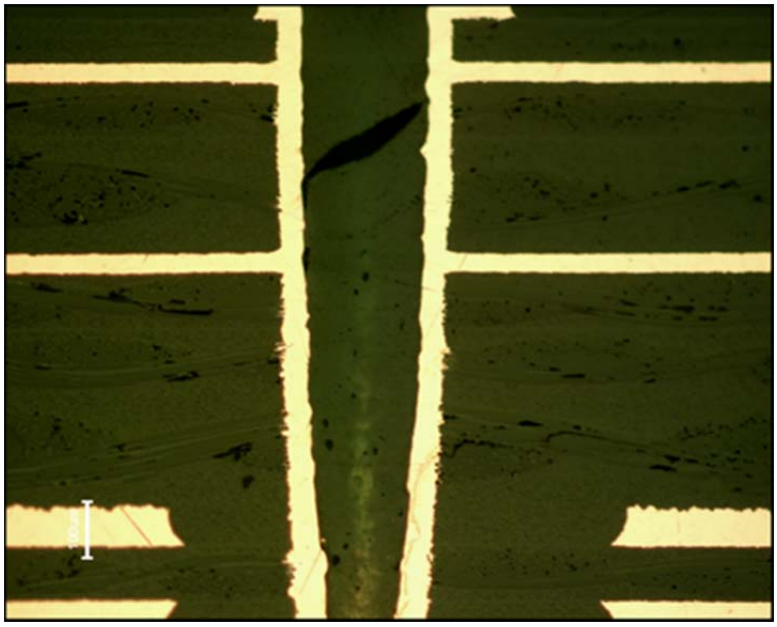

Figure 5-7: Cross section of PCB plated through-hole on oil exposed server

Figure 5-8: Edge of oil exposed PCBs maintain structural integrity and show no indication of delamination

The images and results gathered here provide a more detailed account to support the anecdotal claims made in the literature. In terms of component reliability, when submerging servers over a six month period, there is not any indicated reason for concern. However, typical servers operate in a data center for longer durations, anywhere from three years up to 10 years. A larger sample size of components and materials tested over extended periods or with the aid of accelerated thermal cycling can help strengthen the conclusions made here.

5.3 Material Mechanical Properties

An observation made when handling servers after submersion in mineral oil for extended periods is that some materials become noticeably stiffer. This includes the primary PCBs, as well as, the plastic and insulating materials used for connection cords for power, networking, and hard disk drives (HDD). A concern amongst industry professionals is that this hardening may lead to cracking of insulators, exposing wiring or full failure of connectors. A study has been initiated to determine the extent to which oil exposure alters the material properties of PCBs.

Samples of motherboard PCBs from air cooled and oil exposed servers were taken and prepared for mechanical testing using an Instron machine, shown in Figure 5-9. Preliminary stain measurements showed a significant increase in the Young's modulus of PCB material from 27.2GPa to 38.1GPa for servers that had been immersed in oil for an eight month period. An increase of this type may severely limit the reliable life a motherboard, based on the trend presented by Cheng et. al. as shown in Figure 5-10 [89]. The results from this study may be input to Finite Element Models (FEA) to further

simulate the impact of changes in material properties on component and solder ball fatigue life.

Figure 5-9: Typical setup for strain measurements on Instron microtester machine

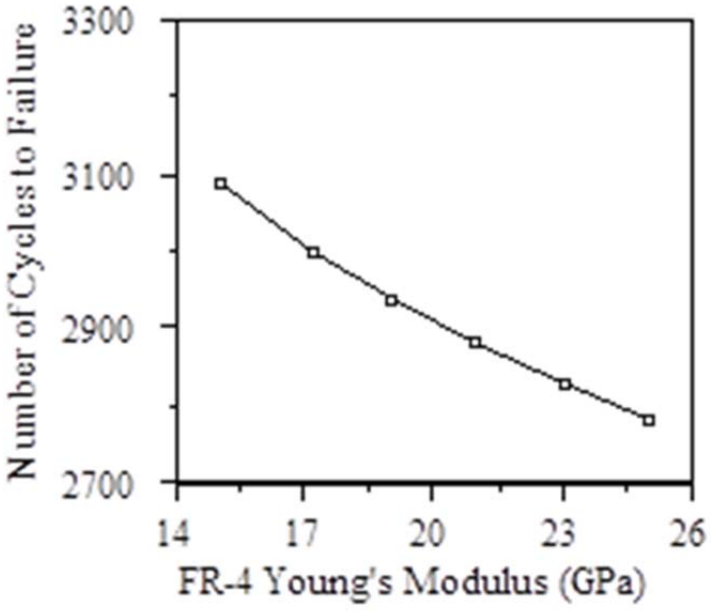

Figure 5-10: Relation between PCB stiffness and cycles to failure as reported by [89]

Chapter 6

Conclusions

With the continued increase in demand for information technology (IT) applications and services there are growing energy requirements to sustain their operation. The studies reported here offer insight into new, energy efficient methods for cooling data centers starting at the server and rack level. The principle results and conclusions of these findings are outlined below.

6.1 Rack-Level Fans: Savings through Consolidations

Consolidation of smaller fans from within servers to a rack-level array of larger fans offers an opportunity to significantly reduce total rack fan power. Initial fan performance characterization showed that the 80mm and 120mm fans chosen for this study have peak total efficiencies two times greater than that of the baseline 60mm fans. When configured into an array at the back of a rack and at uniform computational loading across four servers, the 80mm fans offer 50.1% to 52.3% reduction in total rack fan power at a fixed target temperature. Similar savings of 47.6% to 54.0% in rack fan power were observed when the 120mm fans were employed. When a highly non-uniform computational workload was applied across all four servers, savings in rack fan power were found to be 35.3% and 33.8% for the 80mm and 120mm fans respectively when compared to the lowest possible fan operating energy of the 60mm fans. Part II of this work will study additional operational issues for a rack-level fan array such as developing a control scheme, as well as redundancy in failure situations.

6.2 Impact of Rack Inlet Static Pressure on Server Level Performance

Static pressure built up at the inlet of a rack, either due to provisioning of air flow from CRAC units, aisle containment systems, or other room geometrical features can have a significant impact on the thermal and cooling energy performance within a server. Testing done on an Air Flow Bench to simulate this condition showed that at higher static pressure levels, the bulk airflow from a data center can overpower the contribution of higher fan speeds within a server. This may have a significant impact on the fan energy consumption within a server, especially in scenarios where a negative pressure exists at the server inlet. A reduction in fan power for four servers from 29.1 to 9.8W was observed as inlet static pressure was increased from -0.082 to 0.003 inH_2O. A tradeoff exists between the fan energy consumed at the CRAC level and within servers to minimize total fan power of the system. Additional studies into sensing and controls options to manage airflow though the room can further enhance the benefit of proper pressurization.

6.3 Flow Rate and Inlet Temperature Considerations in Mineral Oil Cooling

The purpose of this work was to establish the general operating trends that may be seen in an oil immersion cooled data center setup. By operating a single server fully immersed in mineral oil and varying volumetric flow rate and oil inlet temperature, bounding environmental operating conditions were established. From these results, it is possible to utilize oil inlet temperatures up to 45°C for cooling. $pPUE_{Cooling}$ values ranging from 1.03 to 1.17 were achieved in the current experimental setup. Comparison with baseline air cooling tests showed 34.4% reduction in the thermal resistance of the system and significant reduction in the temperature difference between component surfaces and

the inlet coolant. Improvements to system and hardware design may increase the efficiency and potential for use of oil cooling in data centers. Future work in this area could include optimizing heat sinks for oil cooling, constructing a larger experimental setup to understand how efficiency scales with size, understanding dynamic loading effects of IT equipment in mineral oil, exhaustive reliability studies, and understanding servicing challenges present in an oil immersed data center.

6.4 Mechanical Observations and Material Properties in Mineral Oil

The issues associated with the mechanical reliability and operational service impact of adopting an oil immersion cooing strategy must be addressed before the technology will see widespread adoption. The visual studies presented here indicate that certain compromises or adjustments in data center operation, such as inventory tracking and servicing procedures, will need to be made for this technology. From a structural point of view, no damage or alterations to components were observed from both the macro and microscopic views. However, changes in mechanical properties due to exposure of oil may alter the length of the useful life of components in oil. Studies with an increased sample size and over multiple time durations will help strengthen the findings presented above. Understanding the rate of change of material properties over extended service times may help build industry confidence in this technology and promote future adoption.

References

[1] J. Koomey, "Growth in Data Center Electricity Use 2005 to 2010," Analytics Press, Burlingame, CA, 2011.

[2] The Green Grid, "Green Grid Data Center Power Efficiency Metrics: PUE and DCIE, White Paper #6," The Green Grid, Beaverton, OR, 2008.

[3] The Green Grid, "PUE™: A Comprehensive Examination of the Metric, White Paper #49," The Green Grid, Beaverton, OR, 2012.

[4] "North American Campos Survey Results," Digital Realty, San Francisco, CA, 2013.

[5] M. Stansberry, "2014 Data Center Industry Survey Results," Uptime Institute, New York, NY, 2014.

[6] U.S. Environmental Protection Agency: ENERGY STAR Program, "Report to Congress on Server and Data Center Energy Efficiency Public Law 109-431," 2007.

[7] R. Brown, "Report to Congress on Server and Data Center Energy Efficiency: Public Law 109-431," eScholarship, 2008.

[8] Natural Resources Defense Council (NRDC), "Data Center Efficiency Assessment: Scalng Up Energy Effiiency Across the Data Center Industry: Evaluating Key Drivers and Barriers," New York, NY, 2014.

[9] M. Iyengard and R. Schmidt, "Energy Consumption of Information Technology Data Centers," *Electronics Cooling Magazine (Online),* 6 December 2010.

[10] ASHRAE TC 9.9, "Datacom Equipment Power Trends and Cooling Applications," ASHRAE, Atlanta, GA, 2005.

[11] ASHRAE TC 9.9, "IT Equipment Thermal Management and Controls White Paper," ASHRAE, Atlanta, GA, 2012.

[12] R. Tummala, Fundamentals of Microsystems Packaging, New York, NY: McGraw-Hill, 2000.

[13] G. Xu, B. Guenin and M. Vogel, "Extension of Air Cooling for High Power Processors," in *IEEE ITherm*, Las Vegas, NV, 2004.

[14] ASHRAE TC 9.9, "2011 Thermal Guidelines for Data Processing Environments - Expanded Data Center Classes and Usage Guidance," ASHRAE, Atlanta, GA, 2011.

[15] M. Iyengar and R. Schmidt, "Analytical Modeling for Thermodynamic Characterization of Data Center Cooling Systems," *Journal of Electronic Packaging,* vol. 131, no. 2, pp. 021009-1-9, 2009.

[16] M. Beitelmal and C. Patel, "Model-Based Approach for Optimizing a Data Center Centralized Cooling System," Hewlett-Packard Laboratories, Palo Alto, CA, 2006.

[17] T. Breen, E. Walsh, J. Punch, A. Shah and C. Bash, "From Chip to Cooling Tower Data Center Modeling: Influence of Server Inlet Temperature and Temperature Rise Across Cabinet," *Journal of Electronic Packaging,* vol. 133, no. 1, pp. 011044-1-8, 2011.

[18] S. Shrivastava, B. Sammakia, R. Schmidt and M. Iyengar, "Comparative Analysis of Different Data Center Airflow Management Configurations," in *ASME InterPACK*, San Francisco, CA, 2005.

[19] V. Mulay, "Analysis of Data Center Cooling Strategies and the Impact of the Dynamic Thermal Management on the Data Center Energy Efficiency," PhD Thesis, Arlington, TX, 2009.

[20] S. Patankar, "Airflow and Cooling in a Data Center," *Journal of Heat Transfer,* vol. 132, no. 7, pp. 073001-1-17, 2010.

[21] J. Rambo and Y. Joshi, "Modeling of Data Center Airflow and Heat Transfer: State of the Art and Future Trends," *Distributed Parallel Databases,* vol. 21, pp. 193-225, 2007.

[22] A. Beitelmal and C. Patel, "Thermo-Fluids Provisioning of a High Performance High Density Data Center," *Distributed and Parallel Databases,* vol. 21, pp. 227-238, 2007.

[23] P. Kumar, Y. Joshi, M. Patterson, R. Steinbrecher and M. Mena, "Cold Aisle Distribution in a Raised Floor Data Center with Heterogeneous Opposing Orientation Racks," in *ASME InterPACK*, Portland, OR, 2011.

[24] The Green Grid, "Guidelines for Energy-Efficient Datacenters," The Green Grid, Beaverton, OR, 2007.

[25] O. VanGeet, "Best Practices Guide for Energy-Efficient Data Center Design," US Department of Energy, National Renewable Energy Laboratory (NREL), Golden, CO, 2011.

[26] The Green Grid, "Case Study: The ROI of Cooling System Energy Efficiency Upgrades (White Paper #39)," The Green Grid, Beaverton, OR, 2011.

[27] V. Arghode, V. Sundaralingam, Y. Joshi and W. Phelps, "Thermal Characteristics of Open and Contained Data Center Cold Aisle," *Journal of Heat Transfer,* vol. 135, no. 6, pp. 061901-1-11, 2013.

[28] S. Alkharabsheh, S. Shrivastava and B. Sammakia, "Effect of Cold Aisle Containment Leakage on Flow Rates and Temperatures in a Data Center," in *ASME InterPACK*, Burlingame, CA, 2013.

[29] B. Muralidharan, S. Shrivastava, M. Ibrahim, S. Alkharabsheh and B. Sammakia, "Impact of Cold Aisle Containment on Thermal Performance of Data Center," in *ASME InterPACK*, Burlingame, CA, 2013.

[30] S. Shrivastava and M. Ibrahim, "Benefit of Cold Aisle Containment During Cooling Failure," in *ASME InterPACK*, Burlingame, CA, 2013.

[31] M. K. Patterson, D. G. Costello, P. F. Grimm and M. Loeffler, "Data Center TCO; A Comparison of High-Density and Low-Density Spaces," Intel Corporation, 2005.

[32] L. Stahl, "System Architectures and Fluids for High Heat Density Cooling Solutions," *ASHRAE Transactions,* vol. 116, no. 1, 2010.

[33] P. Hughes and W. Tschudi, "Energy Efficient and Lower Capital Cost - an Alternative Data Center Cooling Strategy," in *ASHRAE Annual Conference*, Montreal, 2011.

[34] H. Villa, "Liquid Cooling for Extreme Heat Densities," *ASHRAE Transactions,* vol. 113, no. 1, 2007.

[35] ASHRAE TC 9.9, "2011 Thermal Guidelines for Liquid Cooled Data Processing Environments," ASHRAE, Atlanta, GA, 2011.

[36] M. Ellsworth and M. Iyengar, "Energy Efficiency Analysis and Comparison of Air and Water Cooled High Performance Server," in *ASME InterPACK*, San Francisco, CA, 2009.

[37] M. Saini and R. Webb, "Heat Rejection Limits of Air Cooled Plane Fin Heat Sinks for Computer Cooling," *IEEE Trans. on Components and Packaging Technologies,* vol. 26, no. 1, pp. 71 - 79, 2003.

[38] M. Patterson and D. Fenwick, "The State of Data Center Cooling: A review of current air and liquid cooling solutions," Intel White Paper, 2008.

[39] S. Mohapatra, "An Overview of Liquid Coolants for Electronics Cooling," *Electronics Cooling Magazine (Online),* 1 May 2006.

[40] M. Ellsworth, "Comparing Liquid Coolants from Both a Thermal and Hydrualic Perspective," *Electronics Cooling Magazine (Online),* 1 August 2006.

[41] S. Mohapatra and D. Loikits, "Advances in Liquid Coolant Technologies for Electronics Cooling," in *IEEE SEMI-THERM*, San Jose, CA, 2005.

[42] A. Almoli, A. Thompson, N. Kapur, J. Summers, H. Thompson and G. Hannah, "Computational Fluid Dynamic Investigation of Liquid Cooling in Data Centers," *Applied Energy,* vol. 89, pp. 150-155, 2012.

[43] S. Zimmermann, I. Meijer, M. K. Tiwari, S. Parades and B. Michel, "Aquasar: A Hot Water Cooled Data Center with Direct Energy Reuse," *Energy,* vol. 43, pp. 237-245, 2012.

[44] R. C. Chu, U. P. Hwang and R. E. Simons, "Conduction Cooling for an LSI Package: A One-Dimensional Approach," *IBM Journal of Research and Development,* vol. 26, no. 1, pp. 45-54, 1982.

[45] J. Fernandes, S. Ghalambor, A. Docca, C. Aldham, D. Agonafer, E. Chenelly, B. Chan and M. Ellsworth Jr., "Combining Computational Fluid Dynamics (CFD) and Flow Network Modeling (FNM) Design of a Multi-Chip Module (MCM) Cold Plate," in *ASME InterPACK Conference*, San Diego, CA, 2013.

[46] M. J. Ellsworth Jr., G. F. Goth, R. J. Zoodsma, A. Arvelo, L. A. Campbell and W. J. Anderl, "An Overview of the IBM Power 775 Supercomputer Water Cooling System," *Journal of Electronic Packaging,* vol. 134, 2012.

[47] F. P. Incropera, Liquid Cooling of Electronic Devices by Single-Phase Convection, New York: A Wiley-Interscience Publication, 1999.

[48] T. M. Anderson and I. Mudawar, "Microelectronic Cooling by Enhanced Pool Boiling of a Dielectric Fluorocarbon Liquid," *Journal of Heat Transfer,* vol. 111, pp. 752-759, 1989.

[49] P. Tuma, "The Merits of Open Bath Immersion Cooling of Datacom Equipment," in *IEEE SEMI-THERM Symposium*, Santa Clara, 2010.

[50] B. Chan, H. Lin, S. Polzer, N. Harff, E. Daniel, B. Gilbert and P. Tuma, "Performance of Passive 2-Phase Immersion Cooling of Server Hardware," in *IMAPs ATW on Thermal Management*, Palo Alto, CA, 2010.

[51] M. Arik, "Enhancement of Pool Boiling Critical Heat Flux in Dielectric Liquids by Microporous Coatings," *International Journal of Heat and Mass Transfer,* vol. 50, pp. 997-1009, 2006.

[52] M. S. El-Genk, "Immersion Cooling Nucleate Boiling of High Power Computer Chips," *Energy Conversion and Management,* vol. 53, pp. 205-128, 2012.

[53] P. Warrier, A. Sathyanarayana, D. V. Patil, S. France, Y. Joshi and A. S. Teja, "Novel Heat Transfer Fluids for Direct Immersion Phase Change Cooling of Electronic Systems," *International Journal of Heat and Mass Transfer,* vol. 55, pp. 3379-3385, 2012.

[54] T. Cader, V. Sorel, L. Westra and A. Marquez, "Liquid Cooling in Data Centers," *ASHRAE Transactions,* vol. 115, no. 1, 2009.

[55] V. Sorell, P. Tuma and L. Newcombe, "Liquid Cooling in Data Centers, Part 2," in *ASHRAE Annual Conference*, Denver, CO, 2013.

[56] L. W. Pierce, "An Investigation of the Thermal Performance of an Oil Filled Transformer Winding," *IEEE Transactions on Power Delivery,* vol. 7, no. 3, pp. 1347-1358, 1992.

[57] J. J. Taha-Tijerina, T. Narayana, C. S. Tiwary, K. Lozano, M. Chipara and P. M. Ajayan, "Nanodiamond-Based Thermal Fluids," *ACS Applied Materials and Interfaces,* vol. 6, pp. 4778 - 4785, 2014.

[58] P. Jones, "Data Center Dynamics," 5 September 2012. [Online]. Available: http://www.datacenterdynamics.com/focus/archive/2012/09/intel-says-liquid-and-servers-do-mix. [Accessed 15 November 2013].

[59] J. Rath, "Data Center Knowledge," 21 November 2013. [Online]. Available: http://www.datacenterknowledge.com/archives/2013/11/21/submerged-supercomputer-named-worlds-efficient-system-green-500/. [Accessed 22 November 2013].

[60] D. Prucnal, "Doing More With Less: Cooling Computers with Oil Pays Off," *The Next Wave,* vol. 20, no. 2, pp. 20 - 29, 2013.

[61] M. Patterson and C. Best, "Oil Immersion Cooling and Reductions In Data Center Energy Use," in *ASHRAE Summer Conference*, San Antonio, TX, 2012.

[62] R. Miller, "Data Center Knowledge," 4 September 2012. [Online]. Available: http://www.datacenterknowledge.com/archives/2012/09/04/intel-explores-mineral-oil-cooling/2/. [Accessed 22 November 2013].

[63] A. Regimbal and R. Tufty, "Power-to-Cool: Liquid Submersion Cooled Server vs. Air-Cooled Servers," Liquid Cooled Solutions White Paper, Rochester, MN, 2013.

[64] Green Revolution Cooling, "CarnotJet Submersion Cooling System Product Data Sheet," Green Revolution Cooling, Austin, TX, 2013.

[65] Green Revolution Cooling, "Submerged Servers in the Data Center: Not as Scary as We All Think," Green Revolution Cooling, Austin, TX, 2014.

[66] N. D. Strike, "Fan Efficiency, An Increasingly Important Selection Criteria," [Online]. Available: http://www.nmbtc.com/fans/white-papers/fan_efficiency_important_selection_criteria/. [Accessed 21 January 2015].

[67] Greenheck, "Understanding Fan Efficiency Grades (FEG)," Greenheck Fan Corp., Schofield, WI, 2013.

[68] M. F. Holohan and B. P. Elison, "Fan Laws for Rack Systems," in *ASME InterPACK*, Vancouver, 2007.

[69] Y. Kodama, S. Itoh, T. Shimizu, S. Sekiguchi, H. Nakamura and N. Mori, "Imbalance of CPU Temperatures in a Blade System and its Impact for Power Consumption of Fans," *Cluster Computer,* vol. 16, no. 1, pp. 27-37, 2013.

[70] B. Nagendran, S. Nagaraj, J. Fernandes, R. Eiland, D. Agonafer and V. Mulay, "Improving Cooling Efficiency of Servers by Replacing Smaller Chassis Enclosed Fans with Larger," in *IEEE ITherm Conference*, Lake Buena Vista, FL, 2014.

[71] E. Frachtenberg, A. Heydari, H. Li, A. Michael, J. Na, A. Nisbet and P. Sarti, "High-Efficiency Server Design," in *International Conference for High Performance Computing, Networking, Storage, and Analysis*, November 2011.

[72] H. Li and A. Michael, "Intel Motherboard Hardware v1.0," 2011. [Online]. Available: http://www.opencompute.org/rojects/intel-motherboard/. [Accessed 22 November 2013].

[73] Airflow Measurements Systems, *Instruction Manual for AMCA 210-99 Airflow Test,* Chula Vista, CA.

[74] D. Carraway, "Lookbusy - A Synthetic Load Generator," [Online]. Available: http://www.devin.com/lookbusy.

[75] die.net, "mpstat(1): Report processors related statistics - Linux man page," [Online]. Available: http://linux.die.net/man/1/mpstat.

[76] die.net, "free(1): amount of free/used memory in system - Linux man page," [Online]. Available: http://linux.die.net/man/1/free.

[77] R. Skillern, "Data Center Knowledge," 20 February 2013. [Online]. Available: http://www.datacenterknowledge.com/archives/2013/02/20/meet-the-future-of-data-center-rack-technologies/. [Accessed 21 January 2015].

[78] T. Breen, E. Walsh, J. Punch, A. Shah, C. Bash, B. Rubenstein, S. Heath and N. Kumari, "From Chip to Cooling Tower Data Center Modeling: Influence of Air-Stream Containment on Operating Efficiency," *Journal of Electronic Packaging,* vol. 134, no. 4, pp. 041006-1-9, 2012.

[79] J. VanGilder and S. Shrivastava, "Real-Time Prediction of Rack-Cooling Performance," *ASHRAE Transactions,* vol. 112, no. 2, pp. 151-162, 2006.

[80] D. Demetriou and H. Ezzat Khalifa, "Optimization of Enclosed Aisle Data Centers Using Bypass Recirculation," *Journal of Electronic Packaging,* vol. 134, no. 2, pp. 020904-1-8, 2012.

[81] B. Fakhim, N. Srinarayana, M. Behina and S. Armfield, "Thermal Performance of Data Centers - Rack Level Analysis," *IEEE Transactions on Components, Packaging, and Manufacturing Technology,* vol. 3, no. 5, pp. 792-799, 2013.

[82] B. Fakhim, M. Behnia, S. W. Armfield and N. Srinarayana, "Cooling Solutions in an Operational Data Centre: A Case Study," *Applied Thermal Engineering,* vol. 31, no. 14, pp. 2279-2291, 2011.

[83] H. Alissa, K. Nemati, B. Sammakia, M. Seymour, K. Schneebli and R. Schmidt, "Experimental and Numerical Characterization of a Raised Floor Data Center Using Rapid Operational Flow Curves Model," in *ASME InterPACK*, San Francisco, CA, 2015.

[84] K. Nemati, H. Alissa, B. Murray, B. Sammakia and M. Seymour, "Experimentally Validated Numerical Model of a Fully-Enclosed Hybrid Cooled Server Cabinet," in *ASME InterPACK*, San Francisco, CA, 2015.

[85] D. Kennedy, "Ramification of Server Airflow Leakage in Data Centers with Aisle Containment," Tate Inc., 2012.

[86] STE Oil Company, Inc., *MSDS Crystal Plus 70T,* San Marcos, TX, 2014.

[87] J. R. Maughan and I. F. P., "Regions of Heat Transfer Enhancement for Laminar Mixed Convection in a Parallel Plate Channel," *Int. Journal of Heat & Mass Transfer,* vol. 33, no. 3, pp. 555-570, 1990.

[88] C. Tulkoff and C. Boyd, "Improved Efficiency & Reliability for Data Center Servers Using Immersion Oil Cooling," in *Electronic System Technology Conference & Exhibition*, Las Vegas, NV, 2013.

[89] H. C. Cheng, K. N. Chiang and C. K. Chen, "Parametric Analysis of Thermally Enhanced BGA Reliability Using a Finite-Volume-Weighted Averaging Technique," in *ASME IMECE*, Anaheim, CA, 1998.

Biographical Information

Richard Mark Eiland received his bachelor's degree (BS) in Physics from Southwestern University in Georgetown, Texas in 2008. After a brief time working in the aircraft industry with Hawker Beechcraft in Wichita, KS, he returned to school in 2009 to pursue a graduate degree at the University of Texas at Arlington. In August 2010 Rick entered the BS to PhD track in Mechanical Engineering after receiving an Enhanced Graduate Teaching Assistantship and Graduate Doctoral Scholarship. As a member of the EMNSPC research lab, he has worked on research projects related to electronic packaging and data center thermal management. Over the course of his graduate degree, Rick worked on internships with Terremark Worldwide, Intel Corp., and Baird, Hampton, & Brown Engineering. Rick received his Doctorate of Philosophy in Mechanical Engineering from the University of Texas in Arlington in May 2015.