AN ABSTRACT OF THE DISSERTATION OF

Babak Lajevardi for the degree of Doctor of Philosophy in Industrial Engineering presented on May 28, 2015.
Title: Energy Analysis of Novel Data Center Cooling Technology: Evaporative Cooling System Operation and Microchannel Heat Exchanger Manufacturing

Abstract approved:

Karl R. Haapala Brian K. Paul

Worldwide, many organizations are pursuing higher energy efficiency by reducing power consumption of their processes, systems, and supporting infrastructure. The rapid growth of the information technology (IT) industry and the miniaturization of semiconductors have resulted in substantial increases in energy consumption and power density of IT equipment, and, subsequently, heat generated by data center equipment contained within data center racks. Energy efficiency and thermal management effectiveness are two major issues facing data centers due to increases in heat dissipated from data center racks. Higher data center energy efficiency will lower total cost of ownership (TCO) and enable organizations to better manage increasing computing and network demands. To improve data center energy efficiency, efforts have been focused on novel center-level and rack-level cooling technologies to remove the heat generated by high-density servers. The research presented herein investigates the operational energy performance of a data center evaporative cooling system and the manufacturing energy requirements for a server-scale microchannel heat exchanger (MCHX). Energy monitoring and analysis was conducted to evaluate an evaporative cooling system installed at a data center located in Gresham, OR. A holistic metric and measurement approach is developed to evaluate the impact of changes for data center infrastructure and information technology (IT) equipment. It was found that the developed metric is more responsive to changes in cooling power and environmental conditions than commonly used metrics. Further, the evaporative cooling technology was shown to be more efficient and effective than conventional cooling technology. Liquid cooling has been demonstrated as an effective strategy to provide a reliable environment for servers and to reduce the load on conventional cooling systems. While microchannel process technology (MPT)-based devices offer a space-efficient approach to liquid cooling of high-density servers, MPT device manufacturing, in particular device patterning and bonding, has been shown to be energy intensive. A weld depth model for bonding of MPT devices is developed and used to understand the capabilities and limitations of the laser welding process. Energy analysis is conducted for the production of a MCHX device to liquid cool the warm exiting air from server racks. Analysis of the patterning, photochemical machining (PCM), and bonding, diffusion bonding and laser welding, processes revealed a considerable reduction in cumulative energy demand (CED) and global warming potential (GWP) when laser welding is used in place of diffusion bonding. This environmental impact reduction was due to reduced process time, reduced energy use, and improved process yield.

Energy Analysis of Novel Data Center Cooling Technology: Evaporative Cooling System Operation and Microchannel Heat Exchanger Manufacturing

by
Babak Lajevardi

A DISSERTATION

Submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented May 28, 2015
Commencement June 2015

Doctor of Philosophy dissertation of Babak Lajevardi
presented on May 28, 2015.

APPROVED:

Co-Major Professor, representing Industrial Engineering

Co-Major Professor, representing Industrial Engineering

Head of the School of Mechanical, Industrial and Manufacturing Engineering

---

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Babak Lajevardi, Author

ACKNOWLEDGEMENTS

I would first like to express my sincere gratitude to Dr. Karl. R. Haapala for his invaluable advice, support, and guidance throughout the PhD program. His supervision and inspiration greatly improved my research skills. I would also like to express my sincere appreciation to Dr. Brian K. Paul, for giving me the opportunity to pursue my studies at Oregon State University. His support and guidance are the key to this dissertation.

I would also like to express my appreciation for the advice and support given by my graduate advisory committee Dr. David Porter, Dr. Zhaohui Wu, and Dr. Jair Lizarazo-Adarme.

Special thanks should be given to Joe Junker, Director of OSU Energy Efficiency Center, for his valuable insights and advice as well as his support during this research project.

I wish to express gratitude to Oregon BEST, Portland Development Commission, City of Gresham, Energy Trust of Oregon (ETO), and the School of Mechanical, Industrial, and Manufacturing Engineering and the Energy Efficiency Center at Oregon State University for support of this research. I also gratefully acknowledge the support of this research by the U.S. Army (W15P7T08CV201).

My appreciation is extended to Ian C. Garretson for assisting in sustainability analysis of this study and Hamed Seyedmahmoudi for his assistance in setting up the energy monitoring system, as well as Michael Dexter from the OSU Energy Efficiency Center for assisting in psychrometric chart preparation.

In the end, my sincere gratitude goes to my family, especially my parents.

TABLE OF CONTENTS

Page

Chapter 1: Introduction 1
1.1 Introduction 1
1.2 Background 2
1.3 Problem Statement 5
1.4 Research Tasks 6
1.5 Dissertation Outline 7
References 7
Chapter 2: Real-time Monitoring and Evaluation of Energy Efficiency and Thermal Management of Data Centers 11
2.1 Abstract 11
2.2 Introduction 11
2.3 Energy Efficiency and Thermal Management Metrics for Data Centers 13
2.3.1 Power Usage Effectiveness (PUE) and Data Center Infrastructure Energy (DCiE) Metrics 14
2.3.2 Rack Cooling Index (RCI) 15
2.3.3 Return Temperature Index (RTI) 16
2.3.4 Supply and Return Heat Indices (SHI, RHI) 17
2.4 Experimental Setup 18
2.5 Results and Discussion 20
2.6 Conclusions and Future Work 25
Acknowledgements 26
References 26
Chapter 3: A Metric for Evaluating Energy Efficiency and Thermal Management Effectiveness in Data Centers 30
3.1 Abstract 30
3.2 Introduction 30
3.3 Current Energy Efficiency and Thermal Management Metrics 34

TABLE OF CONTENTS (Continued)

Page

3.4 Efficiency Metric Model Development 38
3.5 Experimental Setup 42
3.6 Results and Analysis 43
3.7 Conclusions 45
Acknowledgements 46
References 47
Chapter 4: Manufacturing Energy Analysis of a Microchannel Heat Exchanger for High-Density Servers 50
4.1 Abstract 50
4.2. Introduction 50
4.3. Model Development 53
4.4 Application 56
4.5 Energy Analysis Method 58
4.6 Results and Discussion 62
4.7 Conclusions (summarize primary learning) 65
Acknowledgements 66
References 66
Chapter 5: Conclusions 70
5.1. Summary 70
5.2. Conclusions 71
5.3. Contributions 72
5.4. Opportunities for Future Work 72
Appendix A: Data Center Cooling System Evaluation 75
A1.1 Abstract 75
A1.2 Introduction 76
A1.3 Background 78
A1.4 Data Center Cooling Technology 79

TABLE OF CONTENTS (Continued)

Page

A1.4.1 Direct Expansion Air Conditioning (DX-AC) Systems....79

A1.4.2 Indirect Evaporative Cooling Systems....80

A1.5 Methodology....81

A1.5.1 Power Usage Effectiveness (PUE) and Data Center Infrastructure Energy (DCiE)....82

A1.5.2 Rack Cooling Index (RCI)....83

A1.6 Experimental Setup....84

A1.7 Results and Discussion....86

A1.8 Bin Analysis....89

Acknowledgements....93

References....93

LIST OF FIGURES

Figure Page

Figure 0.1. Worldwide and U.S. data center energy use……………………………………………2

Figure 1.2. Worldwide and U.S. data center energy use growth rate……………………………..3

Figure 1.3. Average power density of data centers...4

Figure 0.1. Schematic top view of the wireless sensor network (WSN) installed at the Gresham City Hall (Gresham, OR)…………………………………………………………………..…19

Figure 2.2. Average and standard deviation of PUE over the warmest and coolest weeks of the monitoring period……………………………………………………………………………...20

Figure 2.3. The impact of outdoor temperature and cooling power on PUE…………………..…21

Figure 2.4. IT rack intake temperature compared to ASHRAE environmental guideline……….22

Figure 2.5. Psychrometric chart of IT rack intake temperatures with ASHRAE guideline envelopes shown……………………………………………………………………………....23

Figure 2.6. Average and standard deviation of RTI over the warmest and coolest weeks of the monitoring period……………………………………………………………………………...24

Figure 2.7. Daily average and standard deviation of SHI and RHI over the warmest and coolest weeks of the monitoring period…………………………………………………………………24

Figure 3.1. Schematic of indirect evaporative cooling……………………………………………..33

Figure 3.2. The hot aisle/cold aisle (HACA) strategy for data center cooling………………......38

Figure 3.3. The energy balance in data center……………………………………………………...39

Figure 3.4. Variations in metric values for 1kW change in cooling power……………………...43

Figure 3.5. Effects of cooling power on data center efficiency………………………………......44

Figure 3.6. Data center efficiency under varying weather conditions…………………………...45

Figure 4.1. Schematic diagram of a transverse cross-section of a keyhole weld………………...54

LIST OF FIGURES (Continued)

Figure Page

Figure 4.2. Schematic of a moving line energy source forming a keyhole in laser welding of two metal shims (lap weld configuration)…………………………………………………………55

Figure 4.3. Lamina designs for microchannel heat exchanger (a) before bonding and (b) after bonding………………………………………………………………………………………..57

Figure 4.4. (a) High-density data server and rear microchannel heat exchanger (MCHX) assembly and (b) MCHX installed on a high density data center server………………………...58

Figure 4.5. The manufacturing process flow for the MCHX devices. a) Laser welding flow and b) Diffusion bonding flow…………………………………………………………………….60

Figure 4.6. Global Warming Potential (kg CO2 eq.) for the selected design and manufacturing scenarios…………………………………………………………………………………….64

Figure 4.7. Cumulative Energy Demand (MJ eq.) for the selected design and manufacturing scenarios…………………………………………………………………………………......65

Figure A1.1. Worldwide and U.S. data center energy use…………………………………….....77

Figure A1.2. Schematic of indirect evaporative cooling……………………………………..….81

Figure A1.3. Schematic top view of the wireless sensor network (WSN) installed in the data center under study (City Hall, Gresham, OR)…………………………………………………85

Figure A1.4. Measured PUE over two weeks of the monitoring period (one week before and one week after installation of the IT Aire indirect evaporative cooling system)……………………..87

Figure A1.5. Rack intake temperatures compared to ASHRAE environmental guidelines……..88

Figure A1.6. Energy savings after installation of the IT Aire indirect evaporative cooling system…………………………………………………………………………………………....89

Figure A1.7. PUE as a function of dry bulb temperature for the DX-AC system……………….90

Figure A1.8. PUE as a function of wet bulb temperature for the IT Aire indirect evaporative system…………………………………………………………………………………………....90

Figure A1.9. Dry bulb temperature bin distribution over a monitoring period of one year……..91

LIST OF FIGURES (Continued)

Figure Page

Figure A1.10. Wet bulb temperature bin distribution over a monitoring period of one year…....92

Figure A1.11. Estimated annual energy consumption of the two cooling systems……………...92

LIST OF TABLES

Table Page

Table 3.1. Overview of existing data center efficiency metrics……………………………….…37

Table 4.1. MCHX device data……………………………………………………………...61

Table 4.2. Global Warming Potential (kg CO_2 eq.) for the selected design and manufacturing scenarios………………………………………………………………………………….…62

Table 4.3. Cumulative Energy Demand (MJ eq.) for the selected design and manufacturing scenarios…………………………………………………………………………………….63

Table A1.1. Rack cooling index (RCI) values at different intake temperature ranges…………..84

Chapter 1: Introduction

1.1 Introduction

Worldwide, many organizations are pursuing higher energy efficiency by reducing power consumption of their processes, systems, and supporting infrastructure. Likewise, increased energy efficiency is recognized to lower operational costs of data centers, which are common and proliferating throughout industrial, academic, and governmental organizations. More efficient data centers not only lower total cost of ownership (TCO), but also enable organizations to better manage increasing computing and network demands.

A data center is a facility housing networked computers, servers, and associated infrastructure, for the purposes of storage and management of large amounts of data and information [1]. Data center power use and cooling challenges are two major issues facing organizations and, in particular, enterprises that operate conventional data centers. The rapid growth of the information technology (IT) industry and the miniaturization of semiconductors have resulted in substantial increases in energy consumption and power density of IT equipment, and, subsequently, heat generated by data center equipment contained within data center racks [2]. Heat generated inside the racks must be evacuated from the data center to avoid damaging IT equipment (e.g., servers). Energy in data centers is consumed by two main categories of equipment: the IT equipment and the infrastructure that supports the IT facilities, including the systems that provide reliable cooling and an adequate environment for the IT equipment. Since the power utilized by the IT equipment is transformed into heat dissipated within the racks [3], a significant increase in rack level heat density and resulting thermal management challenges have been rising over the past few decades [4], [5].

1.2 Background

The United States Environmental Protection Agency (EPA) reported that the annual energy use of U.S. data centers in 2005 was about 61 billion kWh, or approximately 1.5% of the total U.S. energy consumption [6]. It was also reported that data center energy use in 2005 was approximately double that of 2000 [6]. According to Koomey [7], data center energy consumption significantly increased worldwide and in the U.S. from 2000 to 2010. As seen in Figure 1.1 and Figure 1.2, data centers consumed about 240 and 80 billion kilowatt-hours (kWh) of energy worldwide and in the U.S., respectively, in 2010, and this consumption has been on a steady increase for the decade up to 2010. This trend is expected to continue with increasing data storage and processing demands. One of the most common metrics used to evaluate data center performance, power use effectiveness (PUE), was also studied by the EPA [6]. Benchmarking of more than 100 North American data centers revealed an average PUE of 1.91 in 2007 (a PUE of 2.00 implies that for every 2 kWh of energy supplied to the data center, 1 kWh is used by the IT equipment).

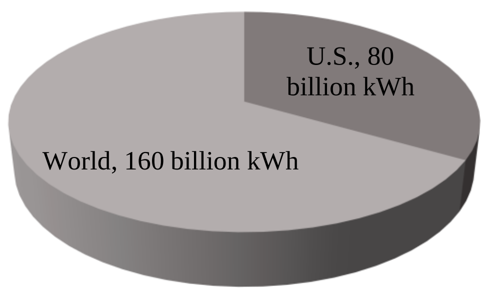

Figure 1.1. Worldwide and U.S. data center energy use [6]

Digital Realty Trust, Inc., a provider of data center solutions, surveyed more than 300 North American corporations (annual revenue of higher than $1 billion or at least 5000 employees) in 2013 [8]. An average PUE value of 2.9 was reported, while only 20% of the data centers reported a PUE of less than 2.0 and 9% stated a PUE of 4.0 or greater [9]. Despite efforts over the past decade to improve data center energy efficiency, the reported PUEs by Digital Realty Trust suggest energy efficiency of North American data centers is decreasing.

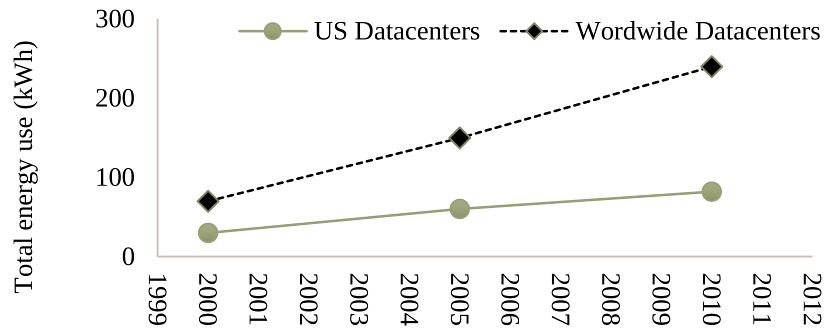

Figure 1.2. Worldwide and U.S. data center energy use growth rate [7].

The continuous growth in rack power density and the high degree of inefficiencies in existing data centers imply that additional investment in innovative data center cooling solutions is of great importance.

Due to miniaturization of IT equipment, incorporation of high density servers, and a rapid growth in rack level power densities, data center thermal management challenges have been rising over the past few decades [4], [5], [13]. Semiconductor miniaturization sustained the Moore's law prediction of microprocessor performance doubling every 18-24 months up to 2005 [14]. Since 2005, due to challenges and limitations associated with reduction in feature sizes such as heat dissipation, the performance rate has decreased. It is noted, however, that the performance of IT equipment has increased at a rate that sustains Moore's law [15]. Typical IT rack dimensions (standard height of 42U or 1.86m) have remained unchanged [16]. Thus, an average rack power density of 8.5 kW was reported in 2012 with a growth rate of about 8% over the previous year (7.9 kW) [8].

The reported rack power densities by Digital Reality Trust [8], [9], [17] during 2006 to 2012 are plotted in Figure 1.3. Considerable increases in rack power density suggest growing thermal management challenges, which lower data center industry energy efficiency. The cooling system accounts for a significant fraction of the total data center energy use, and drives the total operational cost of a typical data center [18]. It is estimated that five times the server cost is needed for cooling and supporting infrastructure to provide an adequate thermal environment [19].

Since the power consumed by servers is converted into heat dissipated through the racks [3], reliable cooling is required to provide an adequate environment for IT devices. Poor thermal management lowers energy efficiency and can lead to higher risks of server failures and lower IT equipment longevity.

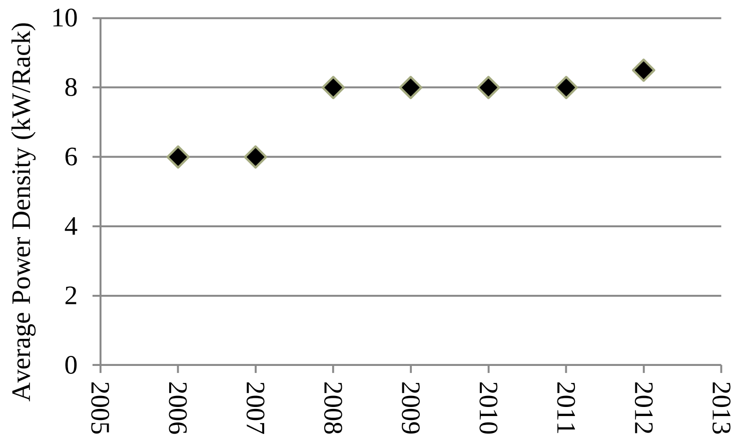

Figure 1.3. Change in average power density of data centers [8], [9], [17]

Despite efforts that have been undertaken over recent decades, energy consumption of data center cooling systems has remained a major concern, and opportunities remain for efficiency improvements. In conventional data centers, due to inefficiencies of conventional thermal management systems, which often rely on direct expansion air conditioning (DXAC), there is additional pressure to cool the ever-increasing rack power densities using novel strategies.
The demands for data center and IT solutions continue to increase, including demands within the manufacturing sector to understand process and equipment operating conditions in real time to improve energy, environmental, and quality performance. This demand is driving the installation of dedicated data centers within the existing building infrastructure as well as new infrastructure, which necessitates thermal management of those spaces – often through conventional refrigeration cooling systems. These data centers thus represent a significant energy load, which in turn must be managed through new technologies and control strategies.

The thermal management of data centers is increasingly complex due to the additional dynamic elements involved in the operation of data centers, such as variations in power dissipation, changes in the IT equipment within the racks, and data center layout. A change in power dissipated inside the racks will affect the temperature distribution within the data center on the order of minutes. Thus, as a result of significant increases in rack power density, demands for effective data center cooling strategies have significantly increased in recent years. This demand has created an opportunity for development of innovative technologies and systems for energy efficiency improvements and cost savings, as well as the need to better understand the impacts of these technologies on life cycle environmental impacts.

1.3 Problem Statement

Efforts have been made to increase the efficiency and effectiveness of cooling strategies in data centers. The research presented herein investigates the operational energy performance of a data center evaporative cooling system and the manufacturing energy requirements for a server-scale microchannel heat exchanger (MCHX). In order to assess the energy efficiency and thermal management effectiveness of a data center evaporative cooling system an energy monitoring system was established. A data center efficiency metric was developed and efficiency performance of data center was evaluated. MPT device manufacturing has been shown to be cost and energy intensive. To assess the system-level manufacturing energy requirements for a server-scale microchannel heat exchanger (MCHX) a prototype server-scale MCHX was established. A laser welding model to quantify process energy use was adapted and the process and system level models to evaluate MCHX device manufacturing energy use integrated.
This work has the joint goals of developing a more comprehensive understanding of data center efficiency challenges, suggesting innovative strategies to improve the performance of data centers, and demonstrating approaches for more integrated energy, environmental, and operational performance analysis of data centers.

1.4 Research Tasks

This study focused on evaluation of data center efficiency and measuring the impact of incorporating novel center-level and rack-level cooling technologies to remove the heat generated by high-density servers. The work was conducted under two primary tasks, as described below.

In the first primary task, energy monitoring and analysis was conducted to evaluate an evaporative cooling system installed at a data center located in Gresham, OR. To support this analysis a holistic energy efficiency and thermal management effectiveness metric was developed. A comprehensive metric was not previously reported in literature. An effective holistic metric allows better insight into the inefficiencies and enables higher level decision makers to take more effective investment strategies. Thus, a holistic metric and measurement approach was developed to evaluate the impact of changes for data center infrastructure and information technology (IT) equipment.

It was found that the developed metric is more responsive to changes in cooling power and environmental conditions than commonly used metrics. Further, the evaporative cooling technology was shown to be more efficient and effective than conventional cooling technology.

In the second primary task, environmental impact assessment for equivalent MCHX devices designs for liquid cooling of the warm exiting air from server racks is conducted. A weld depth model for investigating the application of laser keyhole welding to the microchannel lamination of stainless steel 316 and 3003 aluminum alloy is developed. MCHX Process-based life cycle assessment (LCA) used to evaluate the impacts of energy use in manufacturing of the MPT-based MCHX devices. Results were compared for photochemical machining and two joining methods: diffusion bonding and laser welding. Laser welding is shown to reduce process time, reduce energy use, and improve process yield, and thus reduce environmental impacts compared to diffusion bonding.

1.5 Dissertation Outline

This dissertation is presented in five chapters. Chapter 1 provides an introduction and motivation for the work, as well as presenting the research problem and tasks. A variety of data center efficiency metrics developed and applied for evaluation of data center performance are reviewed in Chapter 2. The energy and thermal management performance of a selected data center located in Gresham, Oregon is then evaluated by applying the chosen metrics. Supporting data was collected using a wireless monitoring network developed and installed within the Gresham City Hall, which is also described.

Due to the limitations of existing metrics to evaluate the performance and the impact of changes in data centers, Chapter 3 focuses on developing a new metric which can more effectively evaluate the energy efficiency and thermal management of data centers. Data that were collected using the wireless monitoring network installed at the data center located in the Gresham City Hall are analyzed to validate the new metric and compare its performance to existing metrics.

Manufacturing energy analysis and environmental impacts for the production of a MCHX device to liquid cool the warm exiting air from server racks is discussed in Chapter 4. A process model for laser welding is developed, which enhances energy analysis. Results compared for photochemical machining and two joining methods: diffusion bonding and laser welding.

Finally, the findings and learnings as a result of this study as well as the contributions are presented in Chapter 5 and the potential for future work are discussed. The performance of the data center (located in Gresham, Oregon) and efficiency analysis when operating with two different cooling systems (direct expansion air conditioning and evaporative cooling system) is presented in Appendix A.

References

[1] S. F. Anderson, “Improving Data Center Efficiency,” *Energy Eng.*, vol. 107, no. 5, pp. 42–63, 2010.

[2] R. K. Sharma, C. E. Bash, and C. D. Patel, “Dimensionless parameters for evaluation of thermal design and performance of large-scale data centers,” presented at the 8th ASME/AIAA Joint Thermophysics and Heat Transfer Conference, 2002, pp. 1–11.

[3] Y. Joshi and P. Kumar, *Energy efficient thermal management of data centers*. Springer, 2012.

[4] B. Kenneth, “Heat Density Trends in Data Processing, Computer Systems, and Telecommunications Equipment,” UpTime Institute, White paper, 2005.

[5] G. Koutitas, “Challenges for energy efficiency in local and regional data centers,” 2010.

[6] R. Brown, E. Masanet, B. Nordman, B. Tschudi, A. Shehabi, J. Stanley, J. Koomey, D. Sartor, and P. Chan, “Report to congress on server and data center energy efficiency, public law, 109-431,” United States Environmental Protection Agency Energy Star Program, 2007.

[7] J. Koomey, “Growth in data center electricity use 2005 to 2010,” *N. Y. Times*, vol. 49, no. 3, 2011.

[8] “Results Of Annual North American Data Center Demand Survey,” Digital Realty Trust, Inc., Survey, Apr. 2013.

[9] “N. America Campos Survey Results,” Digital Realty Trust, Inc., Jan. 2013.

[10] B. Lajevardi, K. R. Haapala, and J. F. Junker, “Real-time monitoring and evaluation of energy efficiency and thermal management of data centers,” *J. Manuf. Syst.*, 2014.

[11] B. K. Paul, R. B. Peterson, and W. Wattanutchariya, “The effect of shape variation in microlamination on the performance of high-aspect-ratio, metal microchannel arrays,” in *Microreaction Technology: Industrial Prospects*, Springer, 2000, pp. 53–61.

[12] B. K. Paul, “Micro Energy and Chemical Systems (MECS) and Multiscale Fabrication,” in *Micromanufacturing and nanotechnology*, Springer, 2006, pp. 299–355.

[13] C. Bash, C. D. Patel, and R. K. Sharma, “Dynamic thermal management of air cooled data centers,” presented at the Thermal and Thermomechanical Phenomena in Electronics Systems, 2006, pp. 445–452.

[14] I. Sauciuc, R. Prasher, J.-Y. Chang, H. Erturk, G. Chrysler, C.-P. Chiu, and R. Mahajan, “Thermal performance and key challenges for future CPU cooling technologies,” in *ASME 2005 Pacific Rim Technical Conference and Exhibition on Integration and Packaging of MEMS, NEMS, and Electronic Systems collocated with the ASME 2005 Heat Transfer Summer Conference*, 2005, pp. 353–364.

[15] M. K. Patterson, “Energy efficiency metrics,” in *Energy Efficient Thermal Management of Data Centers*, Springer, 2012, pp. 237–271.

[16] M. K. Patterson and D. Fenwick, “The State of Data Center Cooling: A review of current air and liquid cooling solutions,” *Intel White Pap. Digit. Enterp. Group*, 2008.

[17] “Results Of Annual North American Data Center Demand Survey,” Digital Realty Trust, Inc., Survey, Mar. 2012.

[18] “The green grids opportunity: decreasing data center and other IT energy usage patterns,” The Green Grids, Technical Report, 2007.

[19] L. Stahl and C. Belady, “Designing an alternative to conventional room cooling,” presented at the IEE conference publication, 2001, pp. 109–115.

Chapter 2

Real-time Monitoring and Evaluation of Energy Efficiency and Thermal Management of Data Centers

Babak Lajevardi, Karl R. Haapala, and Joseph F. Junker
School of Mechanical, Industrial, and Manufacturing Engineering
Oregon State University

This chapter appeared in:
Journal of Manufacturing Systems, Accepted 18 June 2014, doi:10.1016/j.jmsy.2014.06.008

Chapter 2: Real-time Monitoring and Evaluation of Energy Efficiency and Thermal Management of Data Centers

2.1 Abstract

The rapid growth of IT industry and the miniaturization of semiconductors have resulted in substantial increases in energy consumption, power density of IT equipment and, subsequently, heat dissipated from data center racks. Metrics have been proposed to overcome these energy efficiency and thermal management challenges. Measuring the performance of the data centers using a combination of wisely chosen metrics can increase the opportunity for considerable energy reduction. A variety of metrics developed and applied for such evaluation are reviewed herein. The energy and cooling efficiency of a small data center is then evaluated by applying several metrics. To perform the analysis, real-time monitoring of 25 parameters over a period of six weeks was performed through design and implementation of a wireless monitoring network. The results are analyzed and current energy efficiency and thermal management issues are discussed with respect to the relative effectiveness of the various metrics.

2.2 Introduction

Real-time, automated energy monitoring and control of manufacturing processes offers the potential to reduce energy use and improve environmental performance [1]. The ever-increasing demand of IT support, however, along with the developments in semiconductor miniaturization have led to higher density processors and a sharp increase in the energy consumption and the heat dissipated per unit volume of racks in data centers [2]. A data center is a facility housing networked computers and servers as well as associated infrastructure, for the purposes of storage and management of large amount of data and information [3].

The United States Environmental Protection Agency (EPA) reported that the annual energy consumption of U.S. data centers is approximately 61 billion kWh, which is about 1.5% of the total U.S. electricity consumption [4]. It is also reported that data center electricity consumption in 2006 was nearly double that in 2000 [4]. Energy in data centers is consumed by two main categories of equipment: IT equipment and infrastructure that supports the IT facilities and provides reliable cooling and the thermal environment needed for IT equipment to operate.

Since the power consumed by IT equipment is converted into heat dissipated through the racks [5], reliable thermal management is imperative to provide an adequate environment for IT devices. Due to rapid growth of the IT equipment miniaturization and significant increase in rack level power densities, many thermal management challenges have been rising over the past few decades [6][7][8]. Poor thermal management lowers energy efficiency and can lead to higher risks of server failures and lower IT equipment longevity. The data center cooling system often accounts for a significant portion of total energy consumption, and cooling cost drives the total operational cost of a typical data center [9]. It is estimated that about five times the server cost is spent on cooling and supporting infrastructure when a $1500 server is operated in an adequate thermal environment [10]. A successfully implemented thermal management system can significantly reduce operational cost by increasing the energy efficiency [6]. Thus, many efforts have been made to increase the efficiency and effectiveness of cooling system and thermal management of data centers.

The power usage effectiveness (PUE) of data centers, one of the most practiced metrics, was studied by Lawrence Berkeley National Laboratory [11]. Benchmarking 22 data centers revealed a PUE drop of 16% in 2005 when compared with that in 2003 (1.95 to 1.63). Despite all the efforts that have been taken, energy consumption of cooling systems is still a major concern, and there is plenty of room for efficiency improvements. In traditional data centers, there is additional pressure to cool the ever-increasing rack power densities using novel strategies due to inefficiencies of conventional thermal management. The first step in implementing new strategies is by evaluating data center performance. Thus, a challenge is to effectively monitor the energy consumption and environmental conditions.

Microelectromechanical system (MEMS) technology [12] has empowered wireless sensor networks (WSNs) by introducing more reliable, smaller, and inexpensive sensors that allow wider utilization of ad hoc wireless systems in industrial applications [13]. WSNs facilitate better insight into the industrial systems by real-time monitoring and provide the opportunity to improve the efficiency and productivity by evaluation and control of industrial operations. The monitoring applications of WSNs include environmental monitoring (indoor/outdoor), power monitoring, process automation, and structural monitoring, among other types of monitoring [14].

In this research, a data center in the city hall of Gresham, Oregon was evaluated. The data center consists of a row of seven racks, which was monitored by installing a wireless sensor network (WSN) to collect energy use, temperature, and humidity data for a period of six weeks. Energy efficiency and thermal management of the data center was evaluated using a combination of energy and thermal metrics. This study forms a pilot project for installation and performance monitoring of a new data center cooling technology to be installed for the data center [15]. The current split system AC units will be replaced with a rooftop mounted indirect evaporative cooling unit. Energy loads will be evaluated with the existing system (baseline) and new system over extended periods to account for seasonal variation. Ultimately, the goal is to evaluate and compare data center energy efficiency and thermal management performance in each case.

The following sections review a variety of metrics to be applied to data collected for the data center. The equipment configuration and setup is then described and results are presented. Finally, conclusions are drawn from the work and opportunities for future work are discussed.

2.3 Energy Efficiency and Thermal Management Metrics for Data Centers

Higher energy efficiency leads to lower operational costs in data centers. In order to improve and optimize the energy consumption and thermal management in data centers, appropriate metrics are imperative to evaluate their efficiency and performance. Measuring the performance of a data center based upon a standard metric provides the opportunity to track improvements and changes, to estimate the impact of changes, and to draw comparisons to other technologies and data center configurations.

A variety of metrics have been proposed [8] to quantify data center efficiency and performance. In this study, metrics were selected which would enable better insight into energy efficiency and thermal management issues from among the most widely used metrics. The metrics reviewed below help in understanding the operational health and the load on different types of equipment. The objective is to measure baseline performance of the data center, so performance-related impacts of future changes, such as installation of a new cooling system, can be evaluated. Metrics have been previously introduced to evaluate the performance of servers inside the racks, which are not the focus of this paper.

2.3.1 Power Usage Effectiveness (PUE) and Data Center Infrastructure Energy (DCiE) Metrics

Two primary metrics, power usage effectiveness (PUE) and data center infrastructure energy (DCiE), were introduced by the Green Grid industry consortium over the past decade [16] to measure data center energy efficiency. PUE is defined as the total energy delivered to the data center over the total energy drawn by the IT equipment. IT equipment energy is defined as "the energy consumed by equipment that is used to manage, process, store, or route data within the compute space" [16]. PUE is the most widely-used metric and can be calculated using Eq. 2.1.

$$PUE=\frac{P_{inf.}+P_{IT}}{P_{IT}} \tag{2.1}$$

In this equation, $P_{inf.}$ is the power input into the supporting infrastructure, mainly the cooling system, and P_{IT} is the power consumed by the IT equipment in the racks.

Ideally, PUE would hold a value of 1.0, meaning all the power into the data center is consumed by the IT equipment. However, in reality, due to the heat dissipated, energy consuming cooling strategies are imperative to reject heat from the racks. Additional power used for rack cooling purposes increases the value of PUE as suggested by Eq. 1. Higher values of PUE imply inefficiency in cooling systems and thermal management of data center. An average data center PUE of 2.0 is reported by the U.S. Department of Energy, while several efficient data centers have reported a PUEs of about 1.1 [17].

DCiE represents a reciprocal of PUE and, thus, can be calculated using Eq. 2.2.

$$\mathrm{DCiE}=\mathrm{PUE}^{-1}=\frac{P_{IT}}{P_{inf.}+P_{IT}} \tag{2.2}$$

As seen in the above equations, PUE and DCiE measure the portion of the total power into the data center that is consumed by the IT equipment and infrastructure.

2.3.2 Rack Cooling Index (RCI)

The rack cooling index (RCI) proposed by Herrlin [18] measures the degree to which the IT equipment inside the racks are maintained in the rack intake air temperature range recommended by American Society of Heating Refrigerating and Air-Conditioning Engineers (ASHRAE) [19]. Thus, the RCI metric evaluates how effectively an adequate environment is provided for the racks and is expressed by the range defined by Eqs. 2.3 and 2.4, below.

$$\mathrm{RCI}_{HI}=\left[1-\frac{\sum\left(T_{intake}-T_{max\text{-}rec}\right)_{T_{intake}>T_{max\text{-}rec}}}{(T_{max\text{-}all}-T_{max\text{-}rec})n}\right]\times 100\% \tag{2.3}$$

$$\mathrm{RCI}_{LO}=\left[1-\frac{\sum\left(T_{min\text{-}rec}-T_{intake}\right)_{T_{intake}<T_{min\text{-}rec}}}{(T_{min\text{-}rec}-T_{min\text{-}all})n}\right]\times 100\% \tag{2.4}$$

RCI_{HI} is the rack cooling index value at the high end of the recommended temperature spectrum, RCI_{LO} is the value at the low end of the recommended temperature spectrum, T_{intake} is the rack intake air temperature, n is the total number of intakes, $T_{max-rec}$ is the maximum recommended temperature, $T_{max-all}$ is the maximum allowable temperature, $T_{min-rec}$ is the minimum recommended temperature, and $T_{min-all}$ is the minimum allowable temperature.

According to ASHRAE, the recommended and allowable ranges for rack intake temperature is 18-25°C (64-77°F) and 15-32°C (59-90°F), respectively [19]. An RCI of 100% reflects intake temperatures within the recommended range. Lower percentages of RCI_{HI} imply that heat rejection from the racks is not effective and there is a possibility of hot spots within the racks. Similarly, lower percentages of RCI_{LO} indicate that the racks are overcooled, which suggests low cooling power efficiency due to poor thermal management.

2.3.3 Return Temperature Index (RTI)

In order to evaluate the air management effectiveness of data centers, the return temperature index (RTI) was proposed [20], which can be calculated using Eq. 2.5.

$$RTI=\left[\frac{\left(T_{Return}-T_{Supply}\right)}{\Delta T_{Rack}}\right]\times 100\% \tag{2.5}$$

T_{Return} is the temperature of air leaving the data center, T_{Supply} is the supplied air temperature, and ΔT_{Rack} is the temperature difference between the rack intake and exit air. RTI assesses the extent to which the air bypasses the rack equipment, as well as the air recirculation in the racks. Air bypass and recirculation impact data center thermal management and energy performance. Bypassed air does not contribute to cooling the IT equipment and depresses the temperature of the air leaving the room. Likewise, air recirculated through the racks produces hot spots in the IT equipment, which in turn increases the temperature of the air returned to the cooling system.

Therefore, higher deviations from an ideal RTI (100%) imply a poor air management system in the data center. Recirculation dominates when an RTI of above 100% is obtained, indicating an elevated return air temperature. Similarly, an RTI of below 100% due to return air depression suggests air bypass as the primary reason for poor air management performance.

2.3.4 Supply and Return Heat Indices (SHI, RHI)

In order to improve air management and prevent mixing of cold and warm air streams, it is imperative to separate the cold and hot aisles of the data center using containment strategies. It can be noted that the cold aisle is located on the air intake side of a row of racks, while the hot aisle is located on the air exit side. Effective containment strategies maximize the temperature differences between the data center supply air and air returned to the cooling system, which minimizes the overall cooling load. Lower cooling loads lead to higher efficiency data centers. Sharma et al. [2] proposed the supply and return heat indices to measure the level of separation of supplied and returned air streams. The supply heat index (SHI) is the ratio of sensible heat gained in the cold aisle to the heat gained at the rack exit (Eq. 2.6).

$$SHI=\frac{T_{intake}-T_{Supply}}{T_{exit}-T_{Supply}} \tag{2.6}$$

T_{intake} is the rack intake air temperature, T_{exit} is the temperature of rack exit air, and T_{Supply} is the temperature of air supplied to the data center. Lower values of SHI suggest less mixing of warm and cold streams in the cold aisle due to effective containment strategies.

The return heat index (RHI) is defined as the ratio of heat extracted by the cooling system to the sensible heat gained at the rack exit (Eq. 2.7).

$$RHI=\frac{T_{return}-T_{Supply}}{T_{exit}-T_{Supply}} \tag{2.7}$$

RHI measures the degree to which supply air is mixed with the return air stream. A high RHI value implies that insignificant mixing of rack exit and cold aisle air streams takes place before the air is returned to the cooling unit. RHI and SHI hold values between 0 and 1. Higher RHI and lower SHI values imply effective separation of the cold and hot aisles. The next section introduces the data collection system implemented to evaluate the above metrics for an actual data center over a period of time.

2.4 Experimental Setup

A wireless monitoring network was developed and installed at the data center located in the Gresham City Hall (Figure 2.1). The network includes data collection nodes within the data center, as well as those in close proximity to the data center cooling system located on the roof. All evaluation nodes on the roof and in the data center were connected in a single wireless network.

The monitoring equipment on the roof, including a dry bulb temperature sensor, a relative humidity (RH) sensor, and current transducers, provides the outdoor air status as well as the cooling system load. The nodes that log data from temperature/RH sensors in the data center enable measuring the quality of rack intake and exit air, as well as the air supplied to the room and returned to the cooling units at the HVAC air duct vents. The IT equipment load is measured by monitoring the power draw of each of the 14 rack power cords (two cords per rack).

Each of the data loggers was scheduled to record the data with an interval of one minute. In order to increase the robustness of the network, a router was added into the monitoring system to improve the communication path from the equipment on the roof to the receiver. A receiver connected to a laptop in data center collects data transferred by data loggers and saves it to the database.

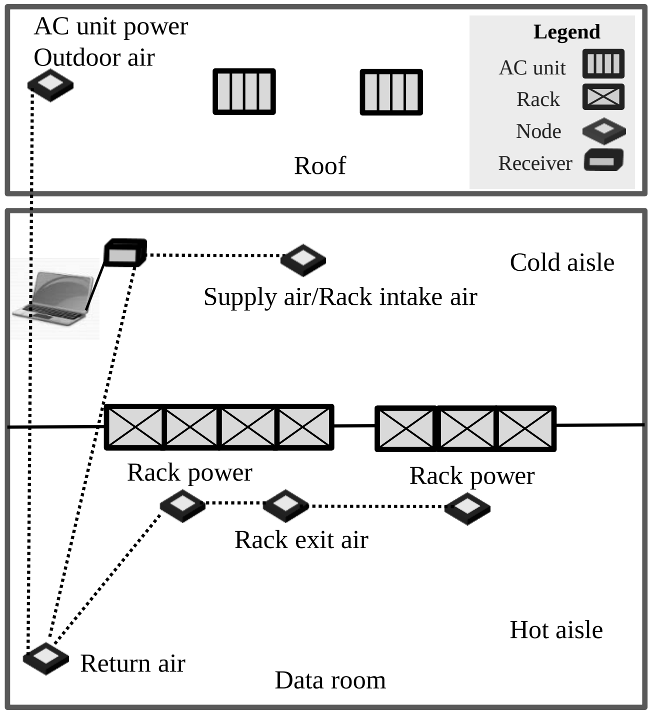

Figure 2.1. Schematic top view of the wireless sensor network (WSN) installed at the Gresham City Hall (Gresham, OR).

Recorded data from nodes in the wireless network are automatically saved to a single file. The network is programmed to save a copy of the updated file to a local drive, send a copy to the Oregon State University Energy Efficiency Center via email, and save a copy to an Oregon State University FTP address every 24 hours. An alarm is set to alert the research team when a logger reading is out of range or a node is identified as missing from the network. Also, a “heartbeat” alarm notifies researchers every 12 hours that the network is active and the receiver is collecting and recording data from the nodes.

With the system thus implemented, it is possible to collect detailed data for the IT equipment and cooling system for extended time periods, which can then be used to evaluate the efficiency of the data center using the metrics defined above. These results are provided in the next section for a period extending from late summer to early fall, which allows for some seasonal variation effects to be elucidated.

2.5 Results and Discussion

The energy efficiency and thermal management metrics discussed above are used and the data center baseline performance is evaluated and summarized based upon the data collected over a monitoring period of six weeks. A total of 1,512,000 data points for all of the 25 parameters (over 60,000 data points for each parameter) were collected. Figure 2.2 demonstrates the measured PUE over the warmest and coolest weeks of the monitoring period. As can be seen in the figure, the average PUE tends to slightly decrease with a reduction of outdoor temperature (1.34 to 1.33).

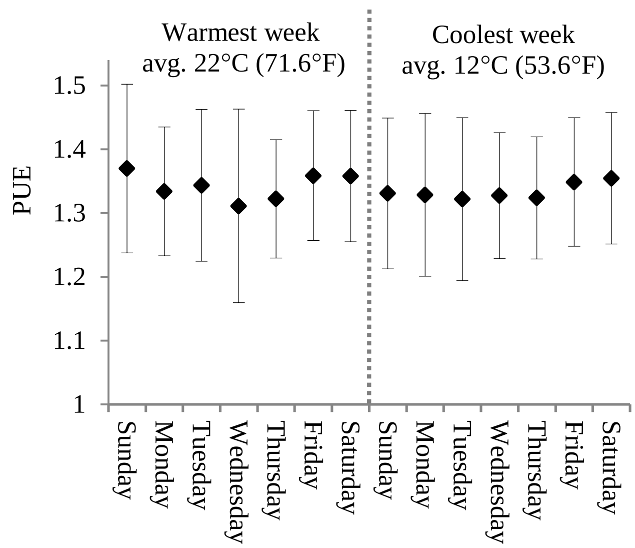

Figure 2.2. Average and standard deviation of PUE over the warmest and coolest weeks of the monitoring period.

An overall average PUE of 1.33 (DCiE of 75%) was measured, which suggests that about 0.35W of power is consumed to condition the data center air and remove the heat from the racks for every watt of electricity delivered to the IT equipment.

To examine the impact of the outdoor temperature and cooling power on PUE, the trend lines of the relationships are plotted over the warmest day of monitoring period in Figure 2.3. As the plot illustrates, a temperature rise of 10°C corresponds to a cooling load increase of about 300W, which in turn elevates the PUE value by about 0.02.

Due to significant variations in the cooling power (relative standard deviation of 38%) and slight variations of the IT load (relative standard deviation of 1%) a greater effect of outdoor temperature on cooling load and PUE was expected. The PUE analysis implies that the cooling system is operating sub-optimally and there is a potential for improvement of energy efficiency. In order to investigate the cooling efficiency of the data center the thermal metrics are evaluated. Figure 2.4 provides a picture of how the temperature of the air entering the racks fulfills the ASHRAE guidelines, using the RCI method.

The calculated RCI_{HI} of 100% indicates the absence of overheating. In other words, the rack intake air temperature falls within the ASHRAE recommended range over the entire monitoring period. Calculated RCI_{LO} values below 100% indicate that the racks are cooled below the low temperature recommended by ASHRAE during the monitoring period.

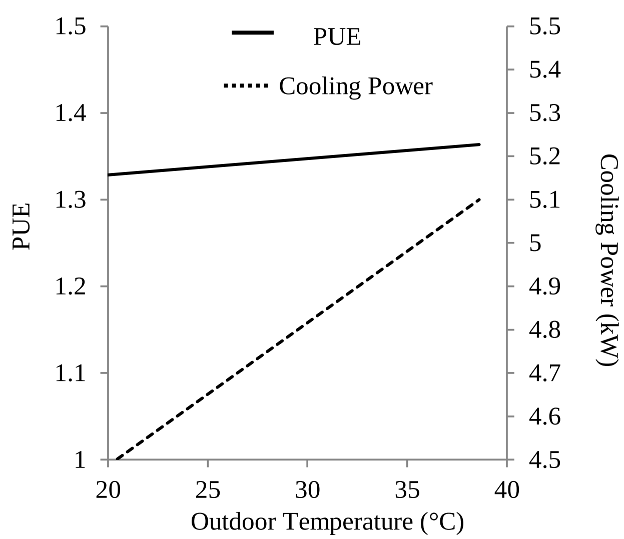

Figure 2.3. The impact of outdoor temperature and cooling power on PUE.

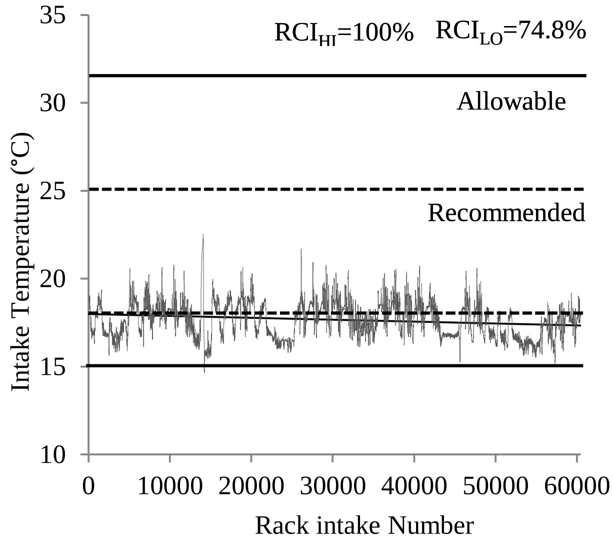

Figure 2.4. IT rack intake temperature compared to ASHRAE environmental guidelines [19].

A similar analysis can be conducted by evaluating collected intake air temperature and relative humidity data using a psychrometric chart (Figure 2.5). It can be seen from the chart that the data mostly fall below the recommended temperature range suggested by ASHRAE (none are above), which confirms the results of the rack cooling indices.

The psychrometric chart suggests that a shift in rack intake air to the right within the recommended envelope (higher temperature and lower relative humidity) will enable a considerable saving in energy due to a reduced cooling system load. Overcooling the racks elevates the cooling load and lowers the power efficiency, which in turn increases the PUE value.

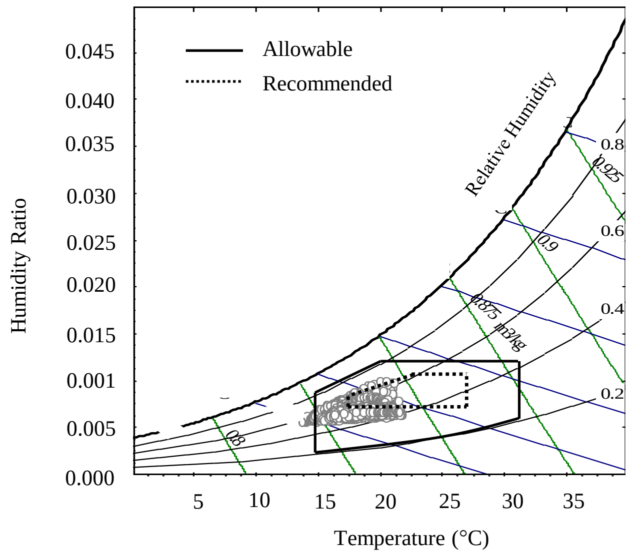

Figure 2.5. Psychrometric chart of IT rack intake temperatures with ASHRAE guideline envelopes shown.

Return temperature index (RTI) values are plotted in Figure 2.6 for the warmest and coolest weeks. RTI provides a picture of the effectiveness of air management strategies in the data center, such as hot aisle/cold aisle isolation and air bypass prevention. Deviations from ideal RTI (100%) suggest presence of air recirculation and bypass due to ineffective air management strategies. Average RTIs of about 75% and 58% were observed at the warmest and coolest weeks, respectively. An RTI below the ideal value of 100% suggests that a portion of the air supplied to the racks bypasses the IT equipment and enters the hot aisle, resulting in a temperature reduction of the air exiting the data center. During the cooler week, the RTI deviation from 100% is greater, implying that air management is less efficient due to larger bypassed air.

In order to evaluate the containment strategy employed, supply and return heat indices are measured. Hot/cold aisle containment curtains are placed in line with the racks, while the racks are fitted with doors that facilitate air flow over the IT equipment. The racks are relatively open, internally.

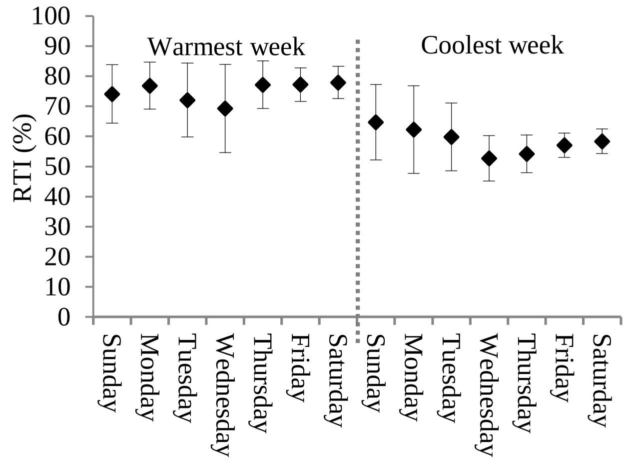

Figure 2.6. Average and standard deviation of RTI over the warmest and coolest weeks of the monitoring period.

As Figure 2.7 illustrates, averages of about 16% and 74% were calculated for SHI and RHI respectively, which indicate the level of separation of the cold and hot aisles. The deviations of the measured SHI and RHI from ideal values of 0 and 1, respectively, illustrate that air flow effectiveness can be improved with other containment strategies.

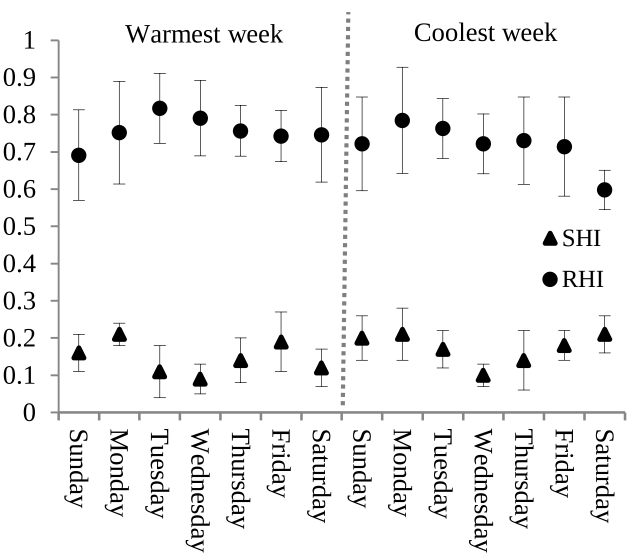

Figure 2.7. Daily average and standard deviation of SHI and RHI over the warmest and coolest weeks of the monitoring period.

2.6 Conclusions and Future Work

Monitoring of energy use and environmental conditions is imperative in order to assess data center energy efficiency and thermal management effectiveness. Related metrics that have been developed by the IT community enable measurement of performance and evaluation of the impact of IT equipment and infrastructure changes on data center performance. Selecting and continuously evaluating the appropriate metrics will enable better insights into data center efficiency improvement strategies.

A wireless sensor network (WSN) was established at the selected data center in Gresham, Oregon to facilitate remote monitoring of real-time data for equipment power use and the inside and outside environmental conditions. Through the application of various metrics that have been developed by the IT community, this monitoring network provided the opportunity to identify the performance issues that could be addressed in redesign to improve in energy efficiency and thermal management of the data center. As a result, the measurement and analysis of energy efficiency and thermal management metrics for the data center suggests the potential for performance improvement through a change in cooling system and hot/cold aisle containment strategies.

It was observed that the IT racks were overcooled for over 25% of the monitoring period. Air delivered to the racks was often below the ASHRAE recommended guideline envelope for temperature and humidity ratio, and often outside both limits for relative humidity. The low return temperature index (RTI) indicated that a considerable portion of the air delivered to the racks bypassed the IT equipment in the racks. The value of supply and return heat indices (16% and 74%, respectively) showed inadequate separation of hot/cold aisles. Thus, hot/cold aisle containment and air flow management can be improved to enhance the overall efficiency of the center.

Acknowledgements

The authors wish to express gratitude to Oregon BEST, the Portland Development Commission, and the School of Mechanical, Industrial, and Manufacturing Engineering and the Energy Efficiency Center at Oregon State University for support of this research. The authors wish to thank Julie Larsen and Janelle Serniotti from the City of Gresham and Hamed Seyedmahmoudi from Oregon State University for their assistance in setting up the energy monitoring system, as well as Michael Dexter from the OSU Energy Efficiency Center for assisting in psychrometric chart preparation.

References

[1] A. Vijayaraghavan and D. Dornfeld, "Automated energy monitoring of machine tools," CIRP Annals-Manufacturing Technology, vol. 59, no. 1, pp. 21–24, 2010.

[2] R. K. Sharma, C. E. Bash, and C. D. Patel, "Dimensionless parameters for evaluation of thermal design and performance of large-scale data centers," presented at the 8th ASME/AIAA Joint Thermophysics and Heat Transfer Conference, 2002, pp. 1–11.

[3] S. F. Anderson, "Improving Data Center Efficiency," Energy Engineering, vol. 107, no. 5, pp. 42–63, 2010.

[4] R. Brown, E. Masanet, B. Nordman, B. Tschudi, A. Shehabi, J. Stanley, J. Koomey, D. Sartor, and P. Chan, "Report to congress on server and data center energy efficiency, public law, 109-431," United States Environmental Protection Agency Energy Star Program, 2007.

[5] Y. Joshi and P. Kumar, Energy efficient thermal management of data centers. Springer, 2012.

[6] B. Kenneth, "Heat Density Trends in Data Processing, Computer Systems, and Telecommunications Equipment," UpTime Institute, White paper, 2005.

[7] C. Bash, C. D. Patel, and R. K. Sharma, "Dynamic thermal management of air cooled data centers," presented at the Thermal and Thermomechanical Phenomena in Electronics Systems, 2006, pp. 445–452.

[8] G. Koutitas and P. Demestichas, “Challenges for Energy Efficiency in Local and Regional Data Centers,” Journal of Green Engineering, vol. 1, p. 32, 2010.

[9] “The green grids opportunity: decreasing data center and other IT energy usage patterns,” The Green Grids, Technical Report, 2007.

[10] L. Stahl and C. Belady, “Designing an alternative to conventional room cooling,” presented at the IEE conference publication, 2001, pp. 109–115.

[11] S. Greenberg, E. Mills, B. Tschudi, P. Rumsey, and B. Myatt, “Best practices for data centers: lessons learned from benchmarking 22 data centers,” Proceedings of the ACEEE Summer Study on Energy Efficiency in Buildings in Asilomar, CA. ACEEE, August, vol. 3, pp. 76–87, 2006.

[12] J. W. Gardner and V. K. Varadan, Microsensors, MEMS and smart devices. John Wiley & Sons, Inc., 2001.

[13] C.-Y. Chong and S. P. Kumar, “Sensor networks: evolution, opportunities, and challenges,” Proceedings of the IEEE, vol. 91, no. 8, pp. 1247–1256, 2003.

[14] J. Yick, B. Mukherjee, and D. Ghosal, “Wireless sensor network survey,” Computer networks, vol. 52, no. 12, pp. 2292–2330, 2008.

[15] S. Law, “Gresham startup says it can slash energy costs at power-thirsty server farms,” Portland Tribune, 2013.

[16] V. Avelar, Azevedo, and A. French, “PUETM: A comprehensive examination of the metric,” The Green Grid, White paper 49, 2012.

[17] J. Bruschi, P. Rumsey, R. Anliker, L. Chu, and S. Gregson, “Best practices guide for energy-efficient data center design,” U.S. Department of Energy, Energy Efficiency & Renewable Energy, Federal Energy Management Program, [Washington, DC], 2010.

[18] M. K. Herrlin, “Rack cooling effectiveness in data centers and telecom central offices: The rack cooling index (RCI),” Transactions-American Society of Heating Refrigerating and Air conditioning Engineers, vol. 111, p. 725, 2005.

[19] “2011 Thermal guidelines for data processing environments–expanded data center classes and usage guidance,” ASHRAE TC 9.9, White paper, 2011.

[20] M. K. Herrlin, “Improved data center energy efficiency and thermal performance by advanced airflow analysis,” presented at the Digital Power Forum, 2007, pp. 10–12.

Chapter 3

A metric for evaluating energy efficiency and thermal management effectiveness in data centers

Babak Lajevardi, Karl R. Haapala, and Joseph Junker
School of Mechanical, Industrial, and Manufacturing Engineering
Oregon State University

To be submitted to Journal of Manufacturing Systems

Chapter 3: A Metric for Evaluating Energy Efficiency and Thermal Management Effectiveness in Data Centers

3.1 Abstract

Despite the development and application of various metrics for evaluating the energy efficiency and thermal management effectiveness of data centers, a holistic metric to evaluate the impact of simultaneous data center level and rack level changes, i.e., for infrastructure and information technology (IT) equipment, respectively, has not emerged. The most commonly applied metric, power use effectiveness (PUE), is limited due to both infrastructure and IT loads appearing in the denominator of the expression. Thus, if energy improvements are made to IT equipment without simultaneous changes to infrastructural energy use, PUE will increase, indicating a reduction in performance. To address this deficiency, a metric and measurement approach is developed to evaluate energy efficiency and thermal management effectiveness in data centers. This approach can assist decision makers in evaluating data center performance based on a single metric, rather than necessitating measurement, analysis, and evaluation of a disparate set of metrics. Data collected from a wireless monitoring network developed and installed at a data center located in Gresham, OR is analysed to validate the new metric and compare its performance to common existing metrics. Metric effectiveness and level of response were investigated to examine the capability of the new metric in capturing the impact of changes in the data center.

3.2 Introduction

Advancements in information technology (IT) equipment miniaturization have led to higher rack power densities, which have resulted in data center cooling challenges over the past few decades [1]-[3]. As a result, this has led to significant increases in heat dissipated per unit volume of IT racks within data centers. The electrical energy consumed by IT equipment is converted almost entirely to heat, which is dissipated into the racks and surrounding data center space. Reliable thermal management strategies are needed to effectively remove the heat from the data center and provide an adequate environment for servers [4]. A significant portion of data center energy is consumed to remove the heat generated by IT equipment (e.g., servers), often using conventional air conditioning technology. This has led to a sharp increase in the energy consumption of data centers [5], and the energy consumed for thermal management drives the operational cost of a typical data center [6], [7]. It is estimated that about five times the server cost is spent on cooling and supporting infrastructure when a server is operated in an adequate thermal environment [10]. In addition to higher energy costs, ineffective thermal management strategies increase the risk of server failures and can lead to poor server performance and increased cooling energy use. A significant increase in energy efficiency is expected when successful and effective thermal management strategies are implemented, which in turn reduces the total operational cost of data centers [1], [11].

The United States Environmental Protection Agency (U.S. EPA) reported that the energy consumption of data centers in 2006 nearly doubled that in 2000. It is also reported that about 1.5% (61 billion kWh) of the total annual electricity consumption in the United States is consumed by data centers [8]. An average data center rack power density of 8.5 kW per rack was reported in 2012 with a growth rate of about 8% over the previous year (7.9 kW) [9]. Due to the significant increase in rack power density in recent years, demands for effective data center cooling strategies have substantially increased. This demand has created an opportunity for development of technologies and systems for energy efficiency and cost savings.

Despite all the efforts that have been undertaken over recent decades, energy consumption of data center cooling systems remains a major concern, and opportunities remain for efficiency improvements. In conventional data centers, there is additional pressure to cool the ever-increasing rack power densities using novel strategies due to inefficiencies of conventional thermal management systems, which often rely on direct expansion (DX) air conditioning (AC).

In a DX-AC system, the air is directly cooled while passing over a refrigeration-cooling coil. The refrigerant is pumped through coils to transfer heat. The refrigerant absorbs heat from the air inside the data center by changing state from liquid to gas (in an expansion valve). It then rejects the heat to the outside air by changing state from gas to liquid (in a condenser). A fan blows over the coils carrying refrigerant, which provides the cooled air back to the data center [4]. In order to compensate for the lack of sufficient humidity control, the indoor air temperature is often lowered and the run time of the system increased, which in turn increases the energy use and lowers the energy efficiency [12].

In order to overcome the cooling challenges due to increasing heat dissipated by high density servers, it is imperative to incorporate innovative cooling solutions that consume less energy [12]. One potential technology is evaporative cooling. An evaporative cooling system cools air through evaporation of water. Water has a high enthalpy of vaporization (the amount of heat gained or released through evaporation at constant pressure), which allows the rapid removal of heat. Evaporative cooling is an isenthalpic process, a process in which the enthalpy does not change [12]. A schematic of the indirect evaporative cooling principle is shown in Figure 3.1. The evaporatively cooled surface provides an area for heat exchange to cool the primary stream. A secondary air stream flows across the wetted outer surface and removes heated vapor.

For an evaporative cooling system, the heat exchanger surface cooling the air is not as cold as that of a DX cooling system. The dew point of the air is not depressed as far and the air is not dried to the same degree as in a DX cooling system. As a result less re-humidification should be required.

In order to ensure reliable and continuous operation of data centers at a low cost, it is imperative to improve energy efficiency. The first step in implementing efficiency improvement strategies is to effectively evaluate the energy consumption and data center environment by measuring its performance. Energy efficiency measurement of data centers is associated with significant challenges due to limitations of existing efficiency measurement techiques and evaluation metrics.

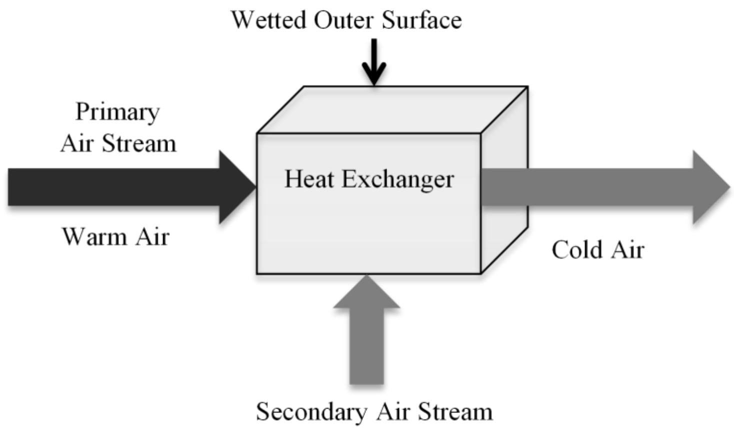

Figure 3.1. Schematic of indirect evaporative cooling

A variety of metrics have been proposed to quantify data center efficiency and performance [13]. In this study, metrics were selected and reviewed from among the most widely used metrics to gain insight into energy efficiency and thermal management issues. The metrics reviewed support understanding the cooling system load and its performance, as well as its operational health. Despite the number of metrics that have been developed to measure data center energy efficiency, a holistic metric to evaluate the impact of changes at both the rack level and center level is absent from literature. The metrics currently used fail to effectively track changes and improvements and are, as such, limited in their ability to assist evaluation of the impact of changes on overall data center performance.

The objective of this study is to measure and evaluate the performance of a data center and investigate the capability of a holistic efficiency metric as compared to existing metrics. The holistic data center efficiency metric developed herein provides insight into the impact of changes inside the racks on energy efficiency and on thermal management effectiveness. The metric enables tracking potential improvement and allows for estimating the impact of changes on the efficiency of data centers. In this research, a data center in the city hall of Gresham, Oregon was monitored and the performance was evaluated using a variety of widely used energy and thermal metrics. The data center was cooled using a new indirect evaporative cooling system.

In the following sections, the most commonly used data center energy efficiency and thermal management effectiveness metrics are reviewed and a performance metric, data center efficiency (ε_{DC}), is developed and demonstrated. The metric will link the thermal and energy efficiency by estimating the portion of the energy utilized to discharge the heat dissipated inside the racks. The monitoring system and evaluation method are then described and results are presented based on the implementation with the Gresham data center. Finally, conclusions and opportunities for future research are discussed.

3.3 Current Energy Efficiency and Thermal Management Metrics

Power use effectiveness (PUE) was introduced by The Green Grid industry consortium [14] as a primary metric for data centers, and has been widely adopted within the IT industry. PUE is defined as the ratio of total power draw of the data center to the total power consumed by the IT equipment within the racks (Eq. 3.1). PUE is the most commonly used data center energy efficiency metric.

$$PUE=\frac{P_{inf.}+P_{IT}}{P_{IT}} \tag{3.1}$$

In this equation, $P_{inf.}$ is the power draw of supporting infrastructure, mainly the cooling system, and P_{IT} is the power consumed by the IT equipment in the racks. The U.S. Department of Energy (U.S. DOE) reported an average U.S. data center PUE of over 2.0 while PUE values of 1.1 and lower have been reported by several highly efficient data centers [15]. Data center infrastructure energy (DCiE) is a metric that represents a reciprocal of PUE, and is sometimes used to measure energy efficiency [15].

The rack cooling index (RCI) proposed by Herrlin [16] measures the degree to which an adequate environment is provided in the racks for IT equipment as proposed by the ASHRAE [17]. The RCI evaluates how effectively the IT equipment inside the racks is maintained within the recommended rack intake temperature range, and is expressed by Eqs. 3.2 and 3.3.

$$RCI_{HI} = \left[1 - \frac{\sum \left(T_{intake} - T_{max\text{-}rec} \right)_{T_{intake} > T_{max\text{-}rec}}}{(T_{max\text{-}all} - T_{max\text{-}rec})n} \right] \times 100\% \tag{3.2}$$

$$RCI_{LO} = \left[1 - \frac{\sum \left(T_{min\text{-}rec} - T_{intake} \right)_{T_{intake} < T_{max\text{-}rec}}}{(T_{min\text{-}rec} - T_{min\text{-}all})n} \right] \times 100\% \tag{3.3}$$

RCI_{HI} is the rack cooling index value at the high end of the recommended temperature range, RCI_{LO} is the value at the low end of the recommended temperature range, T_{intake} is the rack intake air temperature, n is the total number of intakes, $T_{max\text{-}rec}$ is the maximum recommended temperature, $T_{max\text{-}all}$ is the maximum allowable temperature, $T_{min\text{-}rec}$ is the minimum recommended temperature, and $T_{min\text{-}all}$ is the minimum allowable temperature. According to ASHRAE, the recommended and allowable ranges for rack intake temperature are 18-25°C (64-77°F) and 15-32°C (59-90°F), respectively [17].

The return temperature index (RTI) was proposed to measure air management effectiveness [18], [19]. RTI evaluates the degree to which cooling air bypasses the rack equipment, and captures the effect of air recirculation within the racks. Bypassed air does not contribute to rack cooling and lowers the temperature of the air returning to the air cooling system. Likewise, hot spots can be produced due to air recirculation, which in turn reduce efficiency and performance of the data center. RTI can be calculated using Eq. 3.4.

$$RTI=\left[\frac{\left(T_{Return}-T_{Supply}\right)}{\Delta T_{Rack}}\right]\times 100\% \tag{3.4}$$

T_{Return} is the temperature of air leaving the data center, T_{Supply} is the temperature of air entering the data center, and ΔT_{Rack} is the temperature difference between the rack intake and exit air.

In order to improve data center thermal management effectiveness, isolation of cold and hot aisles is imperative. Arranging IT equipment in separated cold and hot aisles using containment strategies prevents mixing of the air streams. The level of separation of cold and hot air streams can be measured by the supply heat index (SHI) and return heat index (RHI) proposed by Sharma et al. [5]. SHI is defined as the ratio of sensible heat gained in the cold aisle to the heat gained at the rack (Eq. 3.5). RHI is defined as the ratio of heat extracted by the cooling system to the heat gained at the rack exit (Eq. 3.6).

$$SHI=\frac{T_{intake}-T_{Supply}}{T_{exit}-T_{Supply}} \tag{3.5}$$

$$RHI=\frac{T_{return}-T_{Supply}}{T_{exit}-T_{Supply}} \tag{3.6}$$

T_{intake} is the rack intake air temperature, T_{exit} is the rack exit air temperature, and T_{Supply} is the temperature of air supplied to the data center. Lower values of SHI and RHI suggest more effective separation of cold and hot aisles and less mixing of air streams.

Despite the metrics that have been developed to measure the performance of data centers, a holistic metric to evaluate the impact of changes in both rack level and center level is absent from literature.

Table 3.1. Overview of existing data center efficiency metrics

Metric	**Area of focus**	**Advantages**	**Disadvantages**
Power Use Effectiveness (PUE) [14]	Infrastructure energy efficiency (includes IT equipment)	• Simple and easy to measure (energy of IT equipment and supporting infrastructure) • Evaluates the impacts of data center infrastructure changes on energy use	• Fails to evaluate the thermal management effectiveness • Not able to track changes to IT equipment well • Does not provide insight into the air flow inefficiencies
Rack Cooling Index (RCI) [16]	Thermal management effectiveness (rack intake air quality)	• Simple and easy to measure • Determines the degree to which an adequate environment is provided to the IT equipment	• Fails to evaluate the cooling load and overall data center efficiency • Not able to evaluate the air flow inefficiencies and mixing of air streams
Return Temperature Index (RTI) [19]	Thermal management effectiveness (rack bypass and recirculation)	• Simple and easy to measure • Assesses the air bypass of IT equipment • Evaluates the air recirculation in the racks	• Fails to evaluate the cooling load and overall data center efficiency • Not able to evaluate changes to infrastructure and IT equipment
Supply and Return Heat Indices (SHI, RHI) [5]	Thermal management effectiveness (rack bypass and recirculation)	• Simple and easy to measure • Assess the mixing of air streams • Evaluate the cold/hot aisle containment effectiveness	• Fail to evaluate the cooling load and overall data center efficiency • Not able to evaluate changes to infrastructure and IT equipment
Data Center Efficiency Metric (this work)	Infrastructure and IT equipment energy efficiency and thermal management effectiveness	• Assesses air flow inefficiencies • Evaluates changes to infrastructure and IT equipment	• 8 parameters need to be measured, dry bulb temperature and relative humidity (4 nodes)

As seen in Eq. 3.1, improvements in IT equipment that result in less power dissipated within the racks inversely affect the PUE. In other words, a higher PUE value (lower data center energy efficiency) will be obtained when less power is consumed by servers and other IT equipment when energy use of supporting infrastructure remains the same. PUE as the most widely used metric, for example, does not provide insight to thermal management issues and has deficiencies in evaluating the impact of changes within a data center. As discussed by Patterson [20], PUE does not reflect improvements in IT energy efficiency, and its use is limited to energy efficiency assessment of the supporting infrastructure of data centers.

3.4 Efficiency Metric Model Development

Measuring the performance of a data center using a holistic metric allows tracking improvements and changes, estimating the impact of the changes, and comparisons to other technologies and average industry performance.

Mixing of warm and cold air streams is a major source of inefficiency in data centers. In order to improve air flow management and prevent mixing of cold and warm air streams, it is imperative to separate the cold and hot aisles using containment strategies [5]. As seen in Figure 3.2, the rack intake air is pulled from the cold aisle and the rack exit air is discharged to the hot aisle.

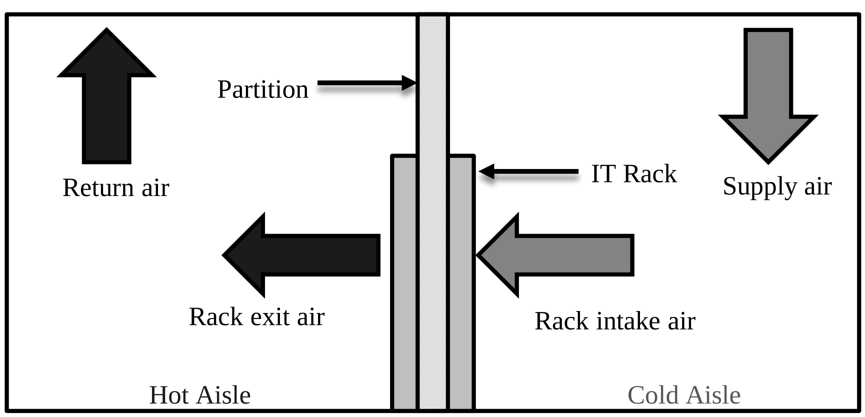

Figure 3.2. The hot aisle/cold aisle (HACA) strategy for data center cooling.

The warmed air is then discharged from the center through outlet vents and cooled, most often using a DX-AC system [12]. The cooled air re-enters the space as supply air for rack cooling. Effective containment solutions maximize the temperature difference between the data center supply air and air returned to the cooling system [21], which in turn minimizes the overall cooling load. Although improvements of 25-30% in cooling efficiency have been reported by using containment solutions, containment does not necessarily enable the highest possible efficiency [22]. An effective data center cooling solution should also provide an adequate environment for IT equipment by directing the correct volume of air at the right temperature into the racks.

As seen in Figure 3.3, the energy balance (first law of thermodynamics) in the center suggests that overall heat gain in a data center, E_c, would be equal to the energy dissipated by the racks, E_d, as shown in Eqs. 3.6, 3.7 and 3.8.

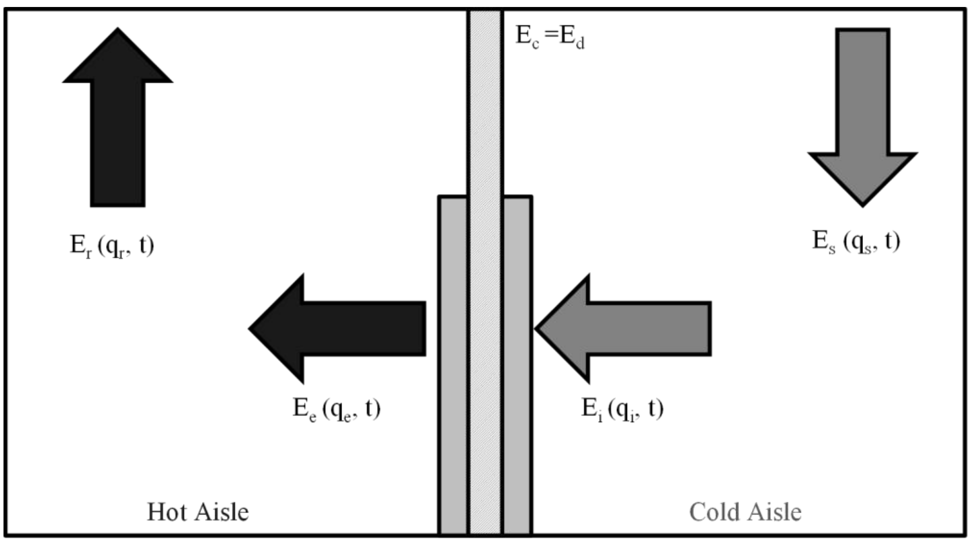

Figure 3.3. The energy balance in data center

$$E_c = E_d \tag{3.7}$$

$$P \times t\,[\mathrm{kJ}] = q_s \times \rho \times \left[\left(h_a + \omega \times h_v \right)_r - \left(h_a + \omega \times h_v \right)_s \right] \times t \tag{3.8}$$

$$q_s\left[{}^{m^3}\!/\!{}_{s}\right] = \frac{P}{\rho\times\left[\left(h_a+\omega\times h_v\right)_r - \left(h_a+\omega\times h_v\right)_s\right]} \tag{3.9}$$

where P is the heat power dissipated into the racks, t is the dissipation time period considered, q_s is volumetric air flow rate, ρ is the density of dry air, h_a is the enthalpy of dry air, ω is the humidity ratio, h_v is the enthalpy of moisture, $(h_a+\omega\times h_v)_r$ is the enthalpy of the return air (H_r), and $(h_a+\omega\times h_v)_s$ is the enthalpy of the supplied air (H_s).

Similarly, the energy balance inside the racks can be expressed as shown in Eqs. 3.10 and 3.11.

$$E_d = E_e - E_i \tag{3.10}$$

$$P\times t\,[kJ] = q_{ing}\times\rho\times\left[\left(h_a+\omega\times h_v\right)_e - \left(h_a+\omega\times h_v\right)_i\right]\times t \tag{3.11}$$

where, q_{ing} is the volumetric air ingested by the IT equipment inside the racks, ρ is the density of dry air, $(h_a+\omega\times h_v)_i$ is the enthalpy of the rack intake air (H_i), and $(h_a+\omega\times h_v)_e$ is the enthalpy of the air exiting the racks (H_e). Thus, the volumetric flow rate of the ingested air inside the racks can be calculated using Eq. 3.11.

$$q_{ing}\left[{}^{m^3}\!/\!{}_{s}\right] = \frac{P}{\rho\times(H_e - H_i)} \tag{3.12}$$

Thus, q_{ing} suggests the ideal volumetric flow rate needed to be supplied to the data center in order to remove the total heat generated within the racks.

The energy balance in cold and hot aisle can be expressed as shown in Eqs. 3.13 and 3.14.

$$E_d = (E_i - E_s) + (E_e - E_i) \tag{3.13}$$

$$P \times t\,[kJ] = \left[q_s \times \rho \times (H_i - H_s) + q_i \times \rho \times (H_e - H_i)\right] \times t \tag{3.14}$$

The volumetric flow rate of rack intake air, which shows the portion of supplied air provided to the racks to cool down the servers, can be calculated using Eq. 3.15.

$$q_i\,\left[{}^{m^3}\!/\!_{s}\right] = \frac{P \times (H_r - H_i)}{(H_e - H_i) \times (H_r - H_s)} \tag{3.15}$$

Thus, efficiency of air supplied to the aisles (ε_A) can be calculated as the ratio of the flow rate of cold air provided to the racks (q_i) to the flow rate of air supplied to the room (q_s) as shown in Eq. 3.16.

$$\varepsilon_A = \frac{q_i}{q_s} = \frac{H_r - H_i}{H_e - H_i} \tag{3.16}$$

Similarly, the efficiency of air supplied to the racks (ε_R) can be calculated as the ratio of the flow rate of air ingested by the IT equipment (q_{ing}) to the flow rate of rack intake air (q_i) as shown in Eq. 3.17.

$$\varepsilon_R = \frac{q_{ing}}{q_i} = \frac{H_r - H_s}{H_r - H_i} \tag{3.17}$$

The ratio of air ingested by racks to the supplied air indicates the amount of air supplied that is not used for cooling the IT equipment.

Thus, overall efficiency of the data center (ε_{DC}) can be calculated as a product of the efficiencies of air supplied to the aisles (ε_A) and racks (ε_R), as shown in Eq. 3.18.

$$\varepsilon_{DC} = \varepsilon_A \times \varepsilon_R \quad (3.18)$$

The air flow delivered to the racks and the enthalpy changes within the racks will be estimated using the real-time measured parameters (dry bulb temperature, relative humidity, rack power, and cooling system power). In addition, the state (temperature and relative humidity) of the air returned to the cooling system and supplied to the data room can be used to determine the enthalpy changes within the cooling system.

The existing metrics that have been proposed to measure energy efficiency of data centers fail to evaluate the impact of changes at the rack level and data center level simultaneously, thus are limited in their ability to assist evaluation of the impact of changes on overall data center performance. The data center efficiency metric developed can provide insight into improvements within both the IT equipment and the supporting cooling system, individually or together. This allows for evaluation of the impact of changes on the overall data center performance, which is not possible using the common metrics above. The metric also captures the inefficiencies within the racks due to ineffective IT equipment layout, which may lead to a possible recirculation.

3.5 Experimental Setup

In order to evaluate energy efficiency and thermal management effectiveness for the data center, a wireless monitoring network was developed and installed using data collection nodes within the data center and on the roof of the building, where cooling equipment is located. The data center consists of a row of seven IT racks. The nodes within the data center collected temperature and relative humidity data for air supplied to the data center, at the rack intakes and exits, and returned to the cooling system. The IT equipment load was measured by monitoring the power consumption of all 14 rack power cords (two cords per rack). The sensors on the roof monitored the status of outdoor air dry bulb temperature and relative humidity, power consumed by the racks as well as evaporative cooling system. The robustness of the system was increased by adding a router into the wireless monitoring system. The data transferred by the data loggers was collected using a receiver connected to a laptop in the data center and saved for subsequent analysis.

3.6 Results and Analysis

The data collected during June-August 2014 from the wireless monitoring network were used to quantify the data center efficiency metric developed above and to compare it to several existing metrics.

As seen in Figure 3.4 the sensitivity of the new metric (ε_{DC}) from the average new metric value was 10% with a standard deviation of 1% for a 1kW change in cooling power. The RTI, SHI, and RHI average chnages fall in a range between 1% and 6%. Average changes of other metrics is less than 1% which indicates lower sensitivity to conditions that lead to increased cooling demand.

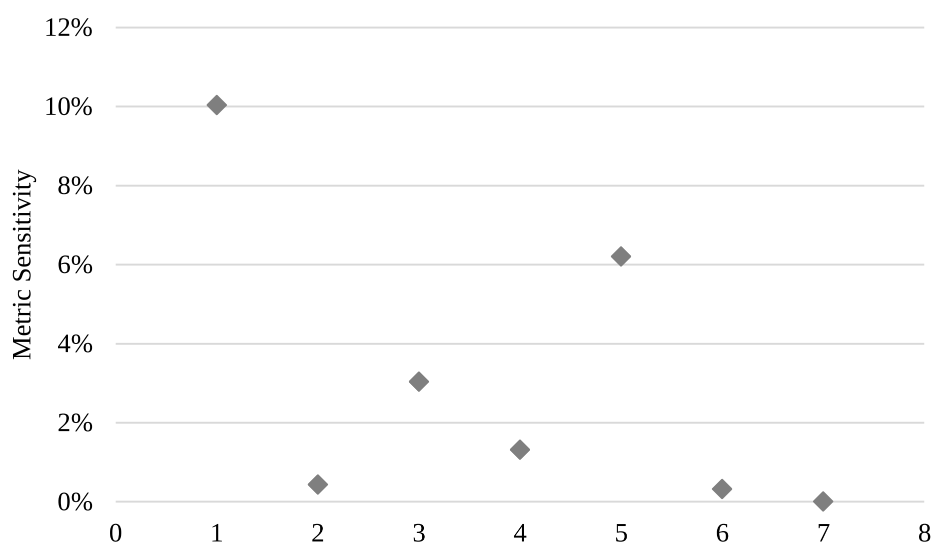

Figure 3.4. Metric sensitivity for a 1kW change in cooling power.

To investigate the effects of changes in cooling power on data center efficiency, the relation between power consumed by cooling system and values of data center metrics was examined, as shown in Figure 3.5. The results indicate that the new data center efficiency metric increases from about 55% to about 85% when the cooling power decreases from 4kW to 1kW (an average predicted efficiency improvement of 10% for every 1kW reduction in cooling power). None of the other metrics exhibited this consistent behavior. Direct correlation with cooling power and continuous data center efficiency improvement is also suggested by the trend of the developed metric (ε_{DC}). The change in PUE was fairly consistent, and demonstrated an improvement of about 10% for a cooling power reduction of 3kW (an average improvement of about 3% for every 1kW reduction in cooling power).

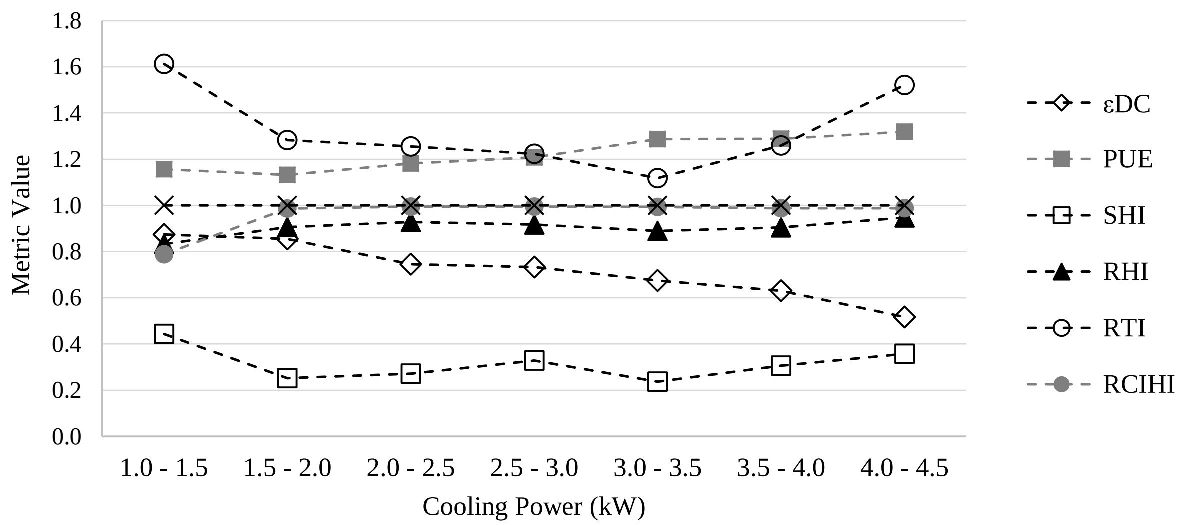

Figure 3.5. Effects of cooling power on data center efficiency.

The response levels of the selected metrics and their trends against changes in cooling power reveal that the data center efficiency metric (εDC) allows better insight into the inefficiencies, which, in turn, enables evaluation of changes within the data center.

In addition to conditions within the data center, outdoor weather conditions impact the performance of data center cooling. In order to investigate the effect of weather conditions, the data center efficiency metric values corresponding to different conditions [4] within different regions of the ASHRAE envelope are indicated in the psychrometric chart shown in Figure 3.6. Region A of the ASHRAE envelope returns the highest values of efficiency when the enthalpy of outdoor air falls within the ASHRAE envelope due to lower enthalpy changes needed to provide adequate environment, which, in turn, minimizes the load on cooling system.

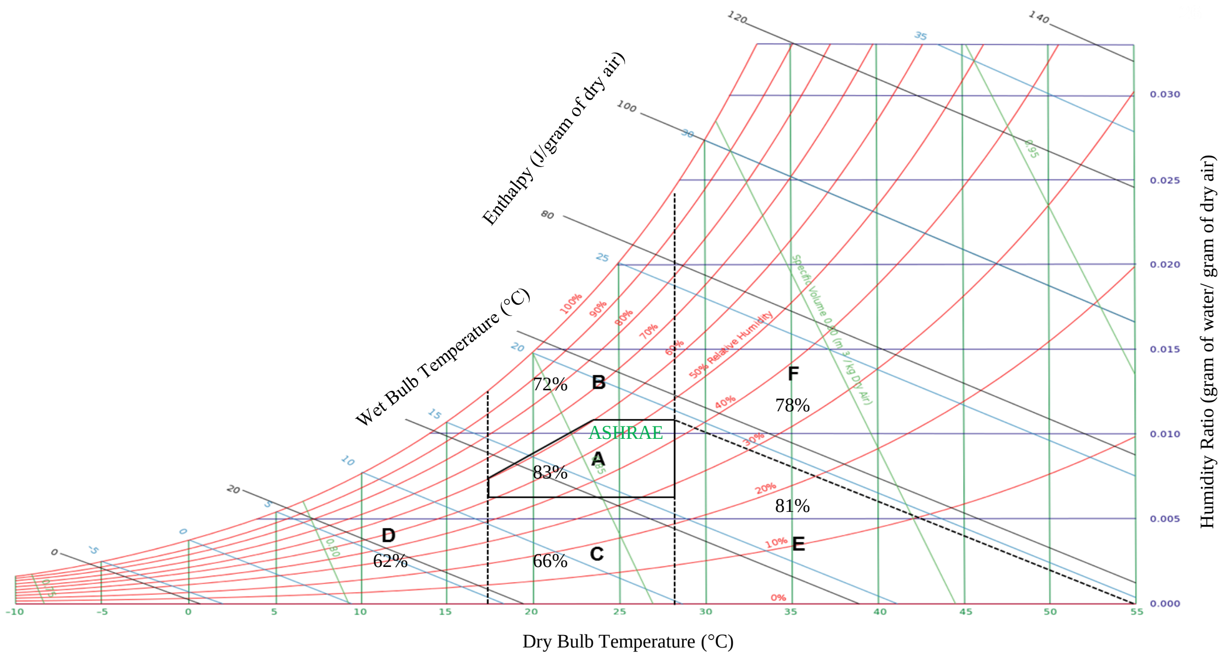

Figure 3.6. Data center efficiency under varying weather conditions.

Operating at higher wet bulb temperatures (right hand side of the ASHRAE envelope) increases the window of sensible cooling. Higher sensible cooling window allows conditioning the outdoor air at constant enthalpy through humidification which, in turn, enables lower cooling loads. The developed efficiency metric values are in good agreement with the cooling technology incorporated and suggest higher energy savings can be achieved when higher sensible cooling is enabled by enthalpy of outdoor air regions E and F).

3.7 Conclusions

In order to evaluate data center energy efficiency and thermal management effectiveness, it is essential to measure the performance by means of effective metrics which reflect the impact of changes in operational environment and infrastructure power consumption. Various metrics have been developed to assess the efficiency and evaluate the impact of changes to data center infrastructure and IT equipment. Some of these are widely applied in industry to track performance and assist strategic decision making. Due to the limitations and inadequacies of the currently used metrics, this research focused on developing an efficiency metric which allows monitoring and measurement of data center performance more precisely, and which is more responsive to the changes and improvements.

To assess their performance, various widely used metrics were quantified using data measured with a wireless monitoring network established in a small data center located in Gresham, Oregon. These results were compared to a data center efficiency metric developed as a part of this research. The volumetric air flow delivered to and exiting from the data center and the IT equipment racks, as well as the related enthalpy changes were calculated using the real-time measured temperature and relative humidity parameters. The level of metric variations in response to cooling power changes were investigated and compared. In addition, the state (temperature and relative humidity) of the air returned to the cooling system and supplied to the data room was used to calcuate the enthalpy changes within the cooling system.

The results demonstrated that the developed metric is more responsive than commonly used metrics to changes in cooling power and environmental conditions. The developed metric suggested average efficiency improvement of 10% for every 1kW reduction in cooling system power. Continuous data center efficiency improvement was measured when the power consumed by cooling system decreased from 4kW to 1kW. None of the prior metrics exhibited a comparative level of sensitivity and accuracy to the newly developed metric.

Acknowledgements

The authors wish to express gratitude to the City of Gresham, Oregon BEST, and the Portland Development Commission for support of this research.

References

[1] B. Kenneth, “Heat density trends in data processing, computer systems, and telecommunications equipment,” UpTime Institute, White paper, 2005.

[2] C. Bash, C. D. Patel, and R. K. Sharma, “Dynamic thermal management of air cooled data centers,” presented at the Thermal and Thermomechanical Phenomena in Electronics Systems, 2006, pp. 445–452.

[3] G. Koutitas and P. Demestichas, “Challenges for Energy Efficiency in Local and Regional Data Centers,” Journal of Green Engineering, vol. 1, p. 32, 2010.

[4] Y. Joshi and P. Kumar, Energy efficient thermal management of data centers. Springer, 2012.

[5] R. K. Sharma, C. E. Bash, and C. D. Patel, “Dimensionless parameters for evaluation of thermal design and performance of large-scale data centers,” presented at the 8th ASME/AIAA Joint Thermophysics and Heat Transfer Conference, 2002, pp. 1–11.

[6] “The green grids opportunity: decreasing datacenter and other IT energy usage patterns,” The Green Grids, Technical Report, 2007.

[7] L. Stahl and C. Belady, “Designing an alternative to conventional room cooling,” 2001.

[8] R. Brown, “United States Environmental Protection Agency Energy Star Program, Report to congress on server and data center energy efficiency, public law, 109-431,” 2008.

[9] “Results of Annual North American Data Center Demand Survey,” Digital Realty Trust, Inc., Survey, Apr. 2013.

[10] L. Stahl and C. Belady, “Designing an alternative to conventional room cooling,” presented at the IEE conference publication, 2001, pp. 109–115.

[11] W. Tschudi, E. Mills, S. Greenberg, and P. Rumsey, “Measuring and Managing-Data Center Energy Use,” Ernest Orlando Lawrence Berkeley NationalLaboratory, Berkeley, CA (US), 2005.

[12] Y. Joshi and P. Kumar, “Emerging Data Center Thermal Management and Energy Efficiency Technologies,” in Energy Efficient Thermal Management of Data Centers, Springer, 2012, pp. 569–611.

[13] G. Koutitas, “Challenges for energy efficiency in local and regional data centers,” 2010.

[14] V. Avelar, Azevedo, and A. French, “PUETM: A comprehensive examination of the metric,” The Green Grid, White paper 49, 2012.

[15] J. Bruschi, P. Rumsey, R. Anliker, L. Chu, and S. Gregson, “Best practices guide for energy-efficient data center design,” U.S. Department of Energy, Energy Efficiency & Renewable Energy, Federal Energy Management Program, [Washington, DC], 2010.

[16] M. K. Herrlin, “Rack cooling effectiveness in data centers and telecom central offices: The rack cooling index (RCI),” Transactions-American Society of Heating Refrigerating and Air conditioning Engineers, vol. 111, p. 725, 2005.

[17] T. ASHRAE, “9.9 (2011) Thermal guidelines for data processing environments–expanded data center classes and usage guidance,” Whitepaper prepared by ASHRAE technical committee (TC), vol. 9, 2011.

[18] M. K. Herrlin, “Improved data center energy efficiency and thermal performance by advanced airflow analysis,” presented at the Digital Power Forum, 2007, pp. 10–12.

[19] M. K. Herrlin, “Airflow and cooling performance of data centers: two performance metrics,” ASHRAE transactions, vol. 114, no. Part 2, 2008.

[20] M. K. Patterson, “Energy efficiency metrics,” in Energy Efficient Thermal Management of Data Centers, Springer, 2012, pp. 237–271.

[21] S. Gondipalli, B. Sammakia, S. Bhopte, R. Schmidt, M. K. Iyengar, and B. Murray, “Optimization of cold aisle isolation designs for a data center with roofs and doors using slits,” in ASME 2009 InterPACK Conference collocated with the ASME 2009 Summer Heat Transfer Conference and the ASME 2009 3rd International Conference on Energy Sustainability, 2009, pp. 841–850.

[22] P. Kumar and Y. Joshi, “Fundamentals of Data Center Airflow Management,” in Energy Efficient Thermal Management of Data Centers, Springer, 2012, pp. 39–136.

Chapter 4
Manufacturing Energy Analysis of a Microchannel Heat Exchanger for High-Density Servers

Babak Lajevardi, Ian C. Garretson, Brian K. Paul, and Karl R. Haapala
School of Mechanical, Industrial, and Manufacturing Engineering
Oregon State University

A form of this chapter appeared in:
Proceedings of NAMRI/SME, Paper 118, Vol. 43, 2015.
To be submitted to Journal of Manufacturing Systems

Chapter 4: Manufacturing Energy Analysis of a Microchannel Heat Exchanger for High-Density Servers

4.1 Abstract

The power consumed by servers is transformed into heat dissipated within IT equipment racks. With the emergence of high-density servers, the increased levels of heat dissipation have challenged conventional air cooling as a means for maintaining IT equipment at an adequate temperature. Liquid cooling has been used to supplement conventional air cooling approaches, providing significantly lower cost-of-ownership for high-density servers. However, the environmental implications of liquid cooling are not well understood. In this paper, cradle-to-gate analyses are performed on the manufacturing of a microchannel heat exchanger to consider the environmental impact of producing microchannel heat exchangers. Microchannel heat exchangers offer a smaller form factor over conventional heat exchangers providing a potential means to lower the environmental impact of direct cooling schemes. The conceptual design and manufacturing of a microchannel heat exchanger is discussed for liquid cooling the warm exiting air from server racks. Material and manufacturing methods for producing the design are compared using environmental analyses. Diffusion bonding analyses are supported by prior work. A laser welding process model is developed to help choose laser welding parameters. It is found that both aluminum and laser welding significantly lower the life cycle energy impact compared with stainless steel and diffusion bonding. Improvements attributed to aluminum are due to lower heat capacities and densities during smelting and joining. Improvements caused by laser welding are due to significant reductions in PCM requirements due to design for weldability considerations.

4.2. Introduction

Due to miniaturization of IT equipment, incorporation of high density servers, and a rapid growth in rack level power densities, data center thermal management challenges have been rising over the past few decades [1]–[3]. For instance, an average rack power density of 8.5 kW was reported in 2012 with a growth rate of about 8% over the previous year (7.9 kW) [4]. As a result of significant increases in rack power density, demands for effective data center cooling strategies have also significantly increased in recent years. This demand has created an opportunity for developing innovative technologies and systems for energy efficiency improvement and cost savings. Since the power consumed by servers is converted into heat dissipated through the racks [5], reliable cooling is required to provide an adequate environment for IT devices. Poor thermal management lowers energy efficiency and can lead to higher risks of server failures and lower IT equipment longevity [6]. The cooling system accounts for a significant fraction of the total data center energy use, and drives the total operational cost of a typical data center [7]. It is estimated that five times the server cost is needed for cooling and supporting infrastructure to provide an adequate thermal environment [8].

Due to system inefficiencies, the cooling energy load is more than that needed to cool IT equipment [9], [10]. In data centers, the risk of local hot spots is reduced by operating the cooling system at full capacity to prevent overheating of IT equipment [11]. Insufficient cooling capacity can be compensated by either increasing the air volume flow rate or decreasing the setpoint temperature. Increasing the supplied air flow results in higher loads on cooling system fans, while reducing temperature increases energy consumption of the chiller in the air conditioning system. Air temperature and humidity cannot be individually adjusted and must be tuned with respect to the air flow rate to avoid equipment damage [12].

An alternative to the converntional cooling system would be to supplement the system through direct liquid cooling of the rack exit air using a heat exchanger. Liquid cooling can lower the exit air temperature prior to being discharged into the hot aisle from the IT rack. This method has been demonstrated by circulating chilled water through a rear-door heat exchanger (RDHX) [13]–[15]. It was demonstrated that liquid cooling could reduce the cost of ownership by about 40% compared to the conventional air-cooled solutions in cooling of racks with power densities over 20kW. However, the environmental implications of this approach were not explored.

In other cooling applications, microchannel heat exchangers (MCHXs) have demonstrated three to five times higher heat fluxes when compared to conventional heat exchangers [16]. MCHXs have shorter diffusional distances in microchannels provide enhanced heat and mass transfer, which combine with the high surface area-to-volume ratios in microchannel arrays to radically reduce the size and weight of components. Other advantages of MPT include better process control, reaction selectivity, and safety [17]. The most common fabrication approach for MPT components is microchannel lamination, or microlamination, in which thin layers (laminae) of metal or polymer are patterned with microchannel features, registered (aligned), and bonded to produce monolithic components [18]. This type of MPT array can be used as a heat sink in thermal management for electronics cooling, such as discussed here for cooling of high-density servers. The most widely used microlamination architecture involves the photochemical machining and diffusion bonding of stainless steel laminae. It is expected that MCHXs could provide reductions in the life cycle energy requirements of a direct cooling approach using conventional heat exchanger technology.

Paul et al. [19] investigated the required bonding conditions and the limits of the channel header width to fin thickness aspect ratio for the diffusion bonding of stainless steel laminae. Experiments demonstrated that, for a given set of bonding conditions, laminae material type, dimensions, and other parameters, a maximum fin width-to-thickness aspect ratio in the channel header was required to permit hermeticity (sealing). Hermeticity is especially critical in data center operations to avoid damage to equipment, data center outages, and loss of critical and irreplaceable data. Prior work has established that poor yields in diffusion bonding due to failure in hermeticity often drive the costs of MCHXs, along with the long cycle times necessary to ensure proper sealing [20]. Laser welding has been demonstrated as an alternative bonding technology within microlamination architectures [21]. In manufacturing, the potential advantages of laser welding over diffusion bonding include faster cycle times, layer-to-layer evaluation of hermeticity (leading to higher yields), and smaller heat affected zones.

The present research applies a weld depth model for investigating the application of laser keyhole welding to the microchannel lamination of stainless steel 316 and 3003 aluminum alloy for a hypothetical MCHX design for cooling of a high density (blade) server. The Cumulative Energy Demand (CED) and Global Warming Potential (GWP) for material production and manufacturing of the MCHX device designs are evaluated using process-based life cycle assessment (LCA) methods. Photochemical machining is selected as the patterning process, while the joining methods considered are diffusion bonding and laser welding. The analysis is conducted using a commercial LCA software (SimaPro 8) and takes advantage of available life cycle inventory databases for most processes. A process model for laser welding was developed, as described below, and used to assist in quantifying process energy use for different welding parameters required by the MCHX designs. The following sections describe the laser welding model development, the LCA procedure, and the analysis results.

The primary objective of this paper is to compare the environmental impacts of manufacturing microchannel heat exchangers using laser welding as an alternative to diffusion bonding. The cumulative energy demand (CED) and global warming potential (GWP) are determined from a life cycle perspective to evaluate the relative environmental impacts of each alternative on the basis of energy use and carbon footprint.

4.3. Model Development

In order to determine the laser parameters to use for a specific welding application, it is necessary to understand the process capabilities and limitations. One of the limitations to consider is the energy loss due to reflection and heat conduction into the material. Energy losses due to reflection are negligible (~2%); once the keyhole is formed it acts as black body. Thus, immediate formation of the keyhole only necessitates consideration of losses due to heat conduction, as described below. In the laser welding of metal foils, immediate formation of the keyhole (Figure 4.1) by the laser beam is essential to avoid unnecessary heat transfer into the workpiece due to low absorption. The figure depicts a lap joint formed by two laminae (a top shim and bottom shim). The laser beam is a moving energy source that creates a molten pool of metal within both shims that coalesces and solidifies to form a joint. Analytical models for a moving energy source include solutions for moving point and moving line sources [22], [23].

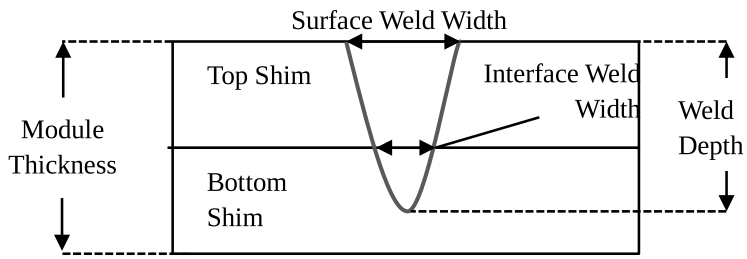

Figure 4.1. Schematic diagram of a transverse cross-section of a keyhole weld.

When welding in keyhole mode, it is assumed that a cylindrical melt pool is formed around the keyhole (Figure 4.2). The keyhole is essentially a black body with very high absorption characteristics and offers insignificant surface heat losses. Therefore, a two-dimensional line source is used for depicting keyhole welds. Based on the energy balance of a line heat source, which assumes high laser energy absorptivities as encountered in laser keyhole welding [24], a melting ratio, ε_m, was defined by Swift-Hook and Gick [25] as the quotient of the normalized scanning velocity, Y, over the normalized power, X (Eqs. 4.1-4.3):

$$\varepsilon_m = \frac{Y}{X} \tag{4.1}$$

$$X = \frac{P_L}{D_w \cdot k \cdot (T_m - T_0)} \tag{4.2}$$

$$Y = \frac{V_s \cdot W_{ws}}{\alpha} \tag{4.3}$$

where P_L is the laser power, D_w is the weld depth, k is the thermal conductivity, T_m is the melting temperature, T_0 is the ambient temperature, V_s is the laser scanning velocity, W_{ws} is the weld width at the surface of the weldment, and α is the thermal diffusivity of the material.

The melting ratio defines the proportion of energy input used to melt a unit volume of weldment in the fusion zone compared with energy lost by convection and conduction. Swift-Hook and Gick [25] proposed that the melting ratio for laser welding at high scanning velocities approaches a limit of 0.483 suggesting that the theoretical fraction of input energy needed to melt the weldment in the fusion zone is 48.3%.

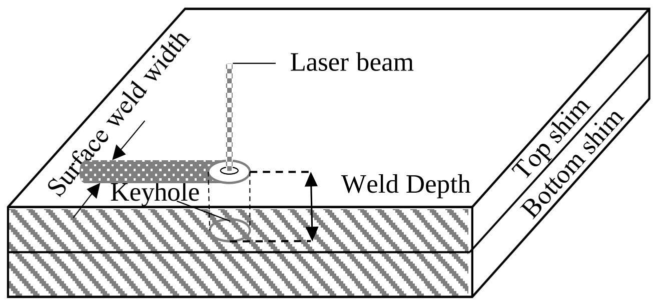

Figure 4.2. Schematic of a moving line energy source forming a keyhole in laser welding of two metal shims (lap weld configuration).

Using Equations 4.1, 4.2, and 4.3 above, the weld depth can be expressed as:

$$D_w = \left(\frac{P_L}{V_S}\right) \cdot \left(\frac{\alpha}{k \cdot (T_m - T_0)}\right) \cdot \left(\frac{\varepsilon_m}{W_{ws}}\right) \tag{4.4}$$

For a given weld depth and a base material, Equation 4.4 can be used to calculate the laser operating parameters (power to scanning velocity ratio [J/mm]), which in turn determines the required process energy use for the total device length.

4.4 Application

As mentioned above, manufacturing of a microchannel heat exchanger (MCHX) for liquid cooling of high density servers is being explored herein. The primary purpose of the study is to compare the relative performance of laser welding with a conventional MPT bonding technology, diffusion bonding, from a life cycle perspective using Cumulative Energy Demand (CED) and Global Warming Potential (GWP) as key indicators. Since yield varies between the two processes, the effect on these indicators of the materials (316 stainless steel and 3003 aluminum alloy) and a typical patterning process, photochemical machining (PCM), is also considered.

The laser welding of thin sheet metals requires low heat input to avoid large thermal gradients and material distortion. Laser welding with a smaller spot size can reduce the energy input into the workpiece by increasing the local intensity of the beam, producing higher power densities [26]. Energy input into the material is the ratio of power to scanning velocity. Increasing laser power enables higher scanning velocities, which, in turn, conserves energy transmission (energy intensity). The fiber laser combines superior beam quality, a small spot size, and a large effective working distance leading to good process repeatability at high power densities and large work envelopes [27].

To conduct the energy analysis herein, it is assumed that a 1000W continuous wave (CW) Yb (ytterbium) fiber laser is used. The laser beam moves via a galvanometer scanner, which offers very high scanning velocities. Thin metal foils of stainless steel 316 and aluminum 3003 with a thickness of 500 µm were selected as the base material of the MCHX considered for process evaluation and subsequent energy studies.

To create the MCHX design, a high-density server (30.2 mm height, 288 mm width, and 480 mm length) was selected to be liquid cooled. A MCHX module was conceived consisting of two alternating 30.2 mm by 30.2 mm laminae (Figure 4.3). The laminae are photochemically machined and then either diffusion bonded or laser welded. As seen in Figure 4.3a, the top lamina consists of embossed rings about two through-holes which will server as a header for manifolding the liquid coolant. The area between the two bosses provides a heat transfer area for air flowing from the server to be cooled. The bottom lamina is a flat shim with a welding boss around the perimeter used to help create a hermetic seal with the underside of the adjoining top shim (Figure 4.3b). Joining of these two laminae produces a cavity through which the liquid coolant will flow. Repeating of this unit cell will lead to a heat exchanger of suitable size.

It should be noted that the device is designed for laser welding and would need to incorporate internal bosses for diffusion bonding to avoid distortion of the laminae under bonding pressures.

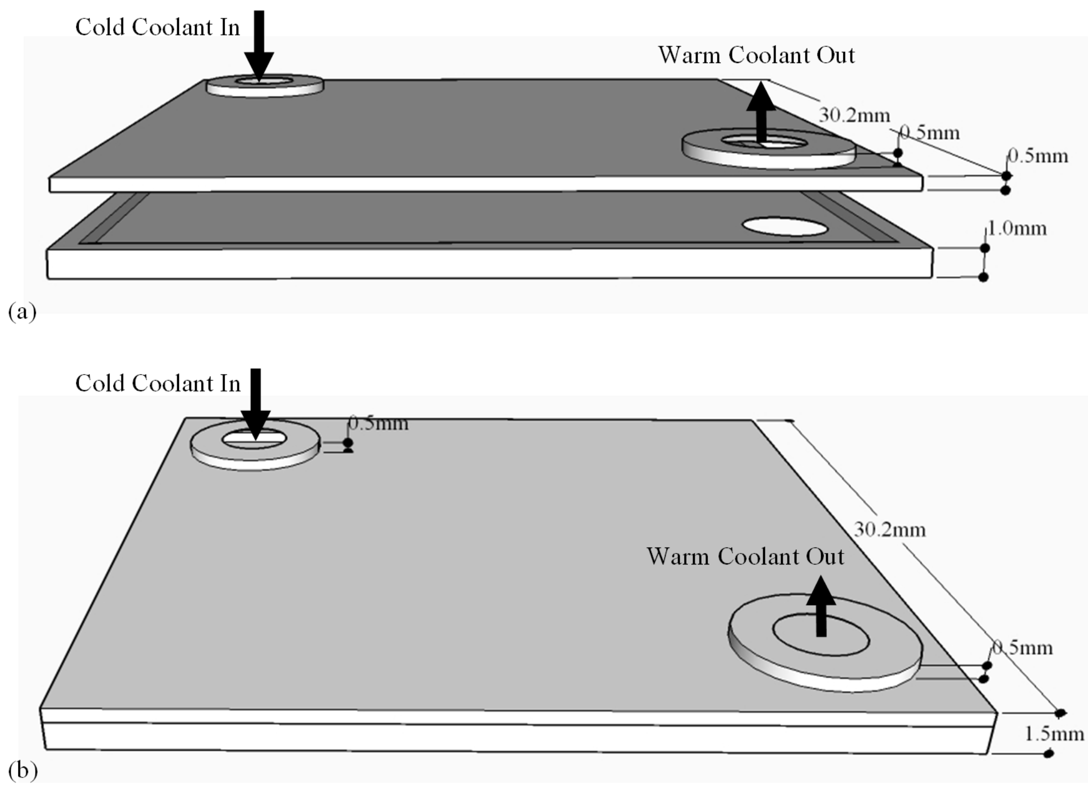

Figure 4.3. Lamina designs for microchannel heat exchanger (a) before bonding and (b) after bonding.

Figure 4.4 demonstrates a high density data server with the rear-installed MCHX device. Several key process requirements exist for the application of laser keyhole welding to microchannel lamination architectures. First, weld joints must be hermetically sealed. Second, weld joints must be strong enough to resist operating pressures within the device. Third, the welding process must not thermally warp or otherwise distort the delicate microchannel fins leading to flow maldistribution across the microchannel array.

In prior work, it was demonstrated that using a CW Yb fiber laser for joining metal foils enables the above-mentioned requirements when a weld depth of about 90% of material thickness at high scanning velocity (700 mm/s) is provided [28]. Equations 1 to 4 are used to estimate the energy required for laser welding of the MCHX. Given the energy requirements for laser welding, an environmental analysis could proceed for the manufacturing scenarios introduced above. The next section describes the energy analysis methods applied.

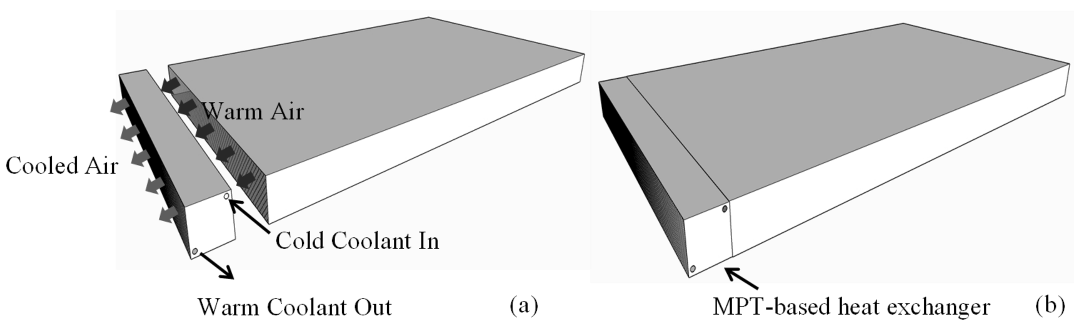

Figure 4.4. (a) High-density data server and rear microchannel heat exchanger (MCHX) assembly and (b) MCHX installed on a high density data center server.

4.5 Energy Analysis Method

To evaluate the impacts of energy use in manufacturing the MPT-based MCHX device, process-based life cycle assessment (LCA) is used. The impact assessment methods chosen were cumulative energy demand (CED) and global warming potential (GWP), which comprise a single impact indicator (energy) and multi-impact indicator (greenhouse gas emissions). Patterning is analyzed when using photochemical machining to pattern the shims in combination with diffusion bonding or laser welding as alternate joining processes.

The use of two different base materials, stainless steel (SS 316) and aluminum (Al 3003), is also evaluated. While functional variation exists, the use phase was not evaluated and device designs are assumed to be functionally equivalent. To assess overall manufacturing system impacts, each unit process is first considered individually to understand material and energy requirements (process inputs and outputs). Total impact is then considered for the entire process flow, by aggregating the unit process-level results.

Figure 4.5 shows the manufacturing process flow for producing the MCHX device under the alternate process flows, and indicates the study system boundaries. The manufacturing process flow is modeled to examine the three fundamental operations: 1) mining, metals production, and forming the shims (metals upstream), 2) pattering the shims (PCM), and 3) joining the shims. The figure shows that devices of the same material type are assumed to use the same PCM process, while the joining process varies; the devices are either laser welded (Figure 4.5a) or diffusion bonded (Figure 4.5b).

Upstream processing is impacted by material choice, as well as downstream processes (due to variation in joining process yields). Because the objective of the study is to investigate the energy-related impacts of alternative bonding processes, more detailed models of these processes are applied. Energy-related impacts of other processes were assumed to be negligible in the creation of a MCHX device; the omitted processes included inspection, finishing, and packaging of the device.

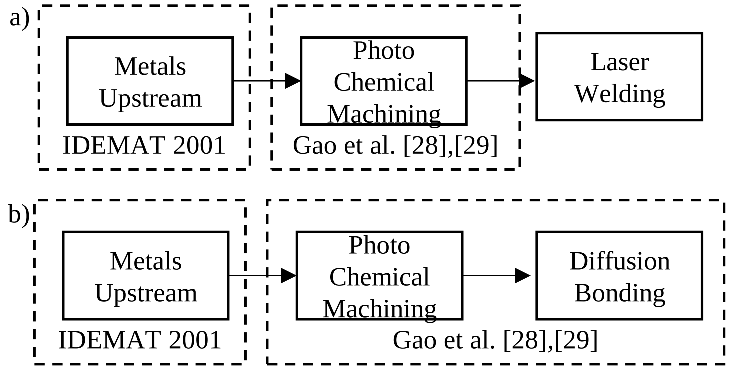

Figure 4.5. The manufacturing process flow for the MCHX devices. a) Laser welding flow and b) Diffusion bonding flow.

For creating a cradle-to-gate process flow model of the MCHX devices, a variety of sources within the LCA software (SimaPro 8) used for environmental impact assessment were used to capture the data required. There was not one database that contained all the necessary data to comprehensively model the processes.

The eco-invent, IDEMAT 2001, Franklin USA 98, and ELCD databases were used to support life cycle inventory (LCI) construction. While the eco-invent database was used for a majority of the processes, IDEMAT 2001 was used to assess the upstream impacts of the metals evaluated (stainless steel 316 and aluminum 3003), Franklin USA 98 was used to assess electricity impacts for the United States, and ELCD was used for the assessment of chemicals required for photochemical machining (PCM).

The IDEMAT 2001 database contains accurate life cycle inventories (LCI) for the synthesis of a variety of ferrous and non-ferrous alloys. Synthesis includes mining, refining, smelting, and forming required to make stock material. To machine the shims to their correct geometries, photochemical machining (PCM) was the selected process, and is used in both process models. The two joining processes (i.e., laser welding and diffusion bonding) are the basis of comparison between the two process flows. Both PCM and diffusion bonding were modeled using the life cycle inventory models developed by Gao et al. [29], [30]. The laser welding model is provided above in Section 4.3.

From a life cycle perspective, PCM requires significant levels of energy, largely due to the production of chemicals required by the process, i.e., sodium hydroxide, hydrochloric acid, polymethyl methacrylate, sodium carbonate, ferric chloride, sodium chlorate, and deionized water. The PCM process energy for each device was 21.25 MJ, which is 1.5% to 2.2% of the cumulative energy use, when considering materials production energy.

The model used for PCM machining was adapted from Gao et al. [29], [30], with quantities of materials assumed to be scaled to the mass of the device (where the stainless steel device was 526.4 g and the aluminum device was 177.66 g). Design details for the assemblies can be found in Table 4.1. An LCI of the chemicals used in PCM and the diffusion bonding process parameters can be found in Gao et al. [29], [30].

Table 4.1. MCHX device data

Material	Units	Inventory
SS 316 input	grams	2,071.01
SS 316 waste	grams	1,544.61
Al 3003 input	grams	699.84
Al 3003 waste	grams	522.18
Diffusion bonding yield	percent	71.25
Laser welding yield	percent	100.00
Density of aluminum	g/cm^3	2.70
Density of stainless steel	g/cm^3	7.99

As mentioned above, two methods were used to assess the environmental impacts of energy use for manufacturing the two MCHX device designs: cumulative energy demand (CED) [31] and global warming potential (GWP). CED (in MJ eq.) was selected to complement the laser welding energy calculations to evaluate energy use. CED is used to quantify the equivalent of primary energy consumption for an entire process chain [32]. GWP (in kg CO_2 eq.) was performed using the IPCC 2007 GWP 100 method [31].

The GWP measure indicates the relative radiative forcing of a gas emitted to the atmosphere normalized to that of carbon dioxide over a period of time [33]; here the study integrated the values over a period of 100 years. After compiling the inventory of material and energy inputs and outputs for the processes considered, environmental impact assessment was completed using LCA software (SimaPro 8). The results of the impact analysis are presented below.

4.6 Results and Discussion

The results for Global Warming Potential (GWP, kg CO_2 eq.) are reported in Table 4.2 and Figure 4.6, and in Table 4.3 and Figure 4.7 for Cumulative Energy Demand (CED, MJ eq.). It can be seen that GWP and CED are linearly correlated, since energy is primarily from fossil-based electrical energy, which results in emissions of carbon dioxide and other greenhouse gases.

Table 4.2. Global Warming Potential (kg CO_2 eq.) for the selected design and manufacturing scenarios.

	Total	Upstream	PCM	Joining
Stainless Steel 316 (LW)	71.24	12.89	58.35	0.003
Aluminum 3003 (LW)	66.60	9.23	57.37	0.002
Stainless Steel 316 (DB)	108.46	18.09	81.89	8.47
Aluminum 3003 (DB)	95.24	12.95	80.52	1.76

A reduction in diffusion bonding (DB) energy consumption of about 79% (140.77 MJ to 29.30 MJ) can be achieved when aluminum is used as the base material in place of stainless steel. Relatively lower diffusion bonding cycle time and temperature for aluminum enables a significant reduction in energy consumption when compared to stainless steel, which should be expected as the diffusion bonding process is highly sensitive to material properties, e.g., thermal conductivity [20].

Table 4.3. Cumulative Energy Demand (MJ eq.) for the selected design and manufacturing scenarios.

	Total	Upstream	PCM	Joining
Stainless Steel 316 (LW)	1182.37	209.91	972.42	0.050
Aluminum 3003 (LW)	1102.24	147.70	954.51	0.032
Stainless Steel 316 (DB)	1800.16	294.60	1364.79	140.77
Aluminum 3003 (DB)	1576.26	207.30	1339.66	29.30

Due to significantly higher cycle times of diffusion bonding compared to laser welding, the diffusion bonding process has remarkably higher energy requirements. The laser welding energy requirements and carbon dioxide equivalent emissions are miniscule compared to those of the patterning process. This is largely driven by reduced process time due to the laser speed, which travels at 700 mm/s resulting in a total on-time of 14.5 seconds for welding the entire device. Note that this assessment does not include setup time between welding shims; the machine will be in a lower energy, standby mode during setup.

As seen in Figures 4.6. and 4.7. the GWP and EDW of patterning process (PCM) is significanltly higher when diffucsion bonding used as joining method. Poor yields in diffusion bonding [20] lowers the effcieicny of the PCM process by increasing the material as well as the energy required for the patterning process.

Laser welding is a layer-to-layer joining method (each shim is welded sequentially), which forces the laser to be turned off in-between each weld in order to place the next shim on top of the stack. While the energy consumption of the laser welding process falls to 36% (0.050 MJ to 0.032 MJ) when welding aluminum rather than stainless steel, the welding process energy is negligible compared to the total process energy utilized in each scenario evaluated.

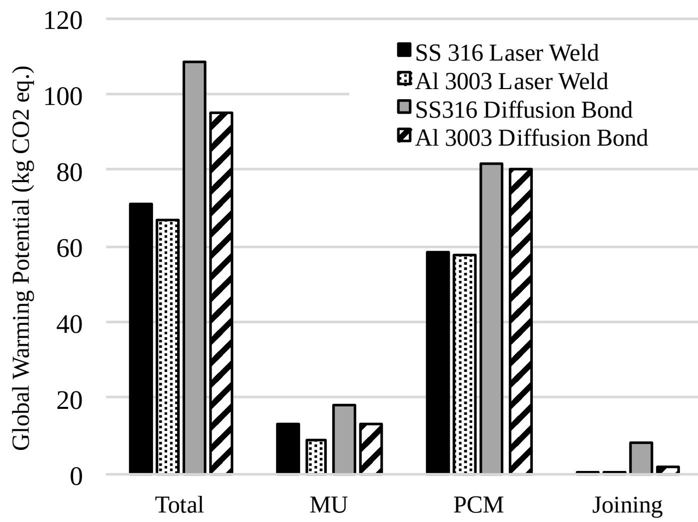

Figure 4.6. Global Warming Potential (kg CO_2 eq.) for the selected design and manufacturing scenarios.

In summary, the CED and GWP of the upstream and manufacturing processes considered for the MCHX device are significantly lower when selecting aluminum as the material and laser welding as the bonding method. Substituting laser welding for diffusion bonding allows a reduction of 35% and 30% in CED for stainless steel 316 and aluminum 3003, respectively. Faster cycles times of laser welding reduces the total process energy when compared to diffusion bonding. For a given weld depth, the cycle time of the materials with lower melting temperature increases, which in turn the required process energy decreases.

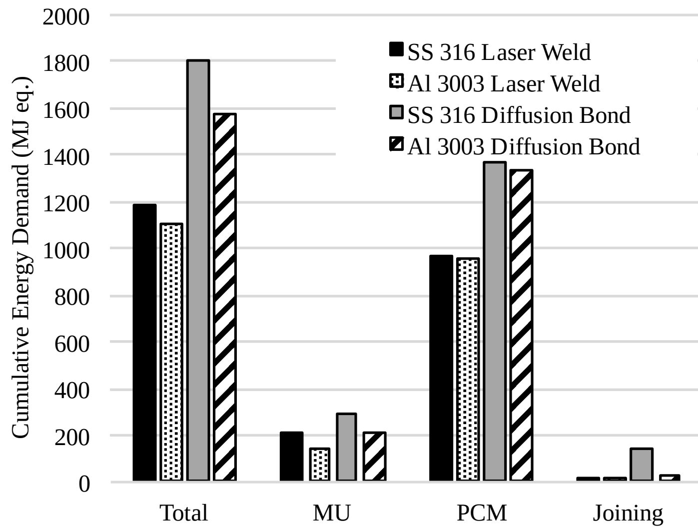

Figure 4.7. Cumulative Energy Demand (MJ eq.) for the selected design and manufacturing scenarios.

4.7 Conclusions (summarize primary learning)

A microchannel heat exchanger (MCHX) was proposed for the liquid cooling of high-density servers. A cradle-to-gate analysis was performed for two contrasting methods of producing the devices. A laser welding process model was developed for choosing welding parameters compared with typical parameters used for diffusion bonding. To produce the device, laser welding was found to offer a lower environmental impact than diffusion bonding largely due to the reduced processing time, reduced energy use, and improved process yield compared to diffusion bonding. The analysis revealed the the manufacturing energy of the device is very sensitive to the bonding step mainly due to the lower yield associated with diffusion bonding.

In order to understand the capabilities and limitations of the laser welding process, and select the applicable laser operating parameters, a weld depth model is developed and used for parameter selection. Manufacturing process environmental analysis was conducted for making the device and revealed a considerable reduction in Cumulative Energy Demand and Global Warming Potential when laser welding is used in place of diffusion bonding. In addition, it was determined that aluminum alloys would have a better environmental performance than stainless steel.

Future work should focus on evaluating the environmental impacts of the full life cycle of the device in comparison with conventional cooling systems. Given the unknown level of performance and operational lifetime of the device, the selected material and device geometry could have a significant impact on unit process and overall life cycle environmental impacts. Analysis should compare use phase efficiency of various devices to consider the energy (heat) transfer of the designs. In addition, other metrics should enter into the analysis to consider the sustainability performance more comprehensively (e.g., additional environmental impacts, economic competitiveness, and relative social impacts).

Acknowledgements

The authors gratefully acknowledge the support of this research by the U.S. Army (W15P7T08CV201).

References

[1] B. Kenneth, “Heat density trends in data processing, computer systems, and telecommunications equipment,” UpTime Institute, White paper, 2005.

[2] C. Bash, C. D. Patel, and R. K. Sharma, “Dynamic thermal management of air cooled data centers,” presented at the Thermal and Thermomechanical Phenomena in Electronics Systems, 2006, pp. 445–452.

[3] G. Koutitas, “Challenges for energy efficiency in local and regional data centers,” 2010.

[4] “Results of Annual North American Data Center Demand Survey,” Digital Realty Trust, Inc., Survey, Apr. 2013.

[5] Y. Joshi and P. Kumar, Energy efficient thermal management of data centers. Springer, 2012.

[6] B. Lajevardi, K. R. Haapala, and J. F. Junker, “Real-time monitoring and evaluation of energy efficiency and thermal management of data centers,” J. Manuf. Syst., 2014.

[7] “The green grids opportunity: decreasing data center and other IT energy usage patterns,” The Green Grids, Technical Report, 2007.

[8] L. Stahl and C. Belady, “Designing an alternative to conventional room cooling,” presented at the IEE conference publication, 2001, pp. 109–115.

[9] J. Mitchell-Jackson, J. G. Koomey, B. Nordman, and M. Blazek, “Data center power requirements: measurements from Silicon Valley,” Energy, vol. 28, no. 8, pp. 837–850, 2003.

[10] J. D. Mitchell-Jackson, “Energy needs in an internet economy: A closer look at data centers,” University of California, 2001.

[11] A. J. Shah, C. E. Bash, C. D. Patel, and V. P. Carey, “Exergy analysis of data center thermal management systems,” J. Heat Transf., vol. 130, no. 2, p. 021401, 2008.

[12] P. Kumar and Y. Joshi, “Fundamentals of Data Center Airflow Management,” in Energy Efficient Thermal Management of Data Centers, Springer, 2012, pp. 39–136.

[13] R. Schmidt, R. C. Chu, M. Ellsworth, M. Iyengar, D. Porter, V. Kamath, and B. Lehman, “Maintaining datacom rack inlet air temperatures with water cooled heat exchanger,” in ASME 2005 Pacific Rim Technical Conference and Exhibition on Integration and Packaging of MEMS, NEMS, and Electronic Systems collocated with the ASME 2005 Heat Transfer Summer Conference, 2005, pp. 663–673.

[14] A. Heydari and V. Gektin, “Reliability challenges in design and operation of high heat powered processors,” in ASME 2005 Pacific Rim Technical Conference and Exhibition on Integration and Packaging of MEMS, NEMS, and Electronic Systems collocated with the ASME 2005 Heat Transfer Summer Conference, 2005, pp. 1665–1668.

[15] T. Tsukamoto, J. Takayoshi, R. R. Schmidt, and M. K. Iyengar, “Refrigeration heat exchanger systems for server rack cooling in data centers,” in ASME 2009 InterPACK Conference collocated with the ASME 2009 Summer Heat Transfer Conference and the ASME 2009 3rd International Conference on Energy Sustainability, 2009, pp. 915–921.

[16] B. K. Paul, R. B. Peterson, and W. Wattanutchariya, “The effect of shape variation in microlamination on the performance of high-aspect-ratio, metal microchannel arrays,” in Microreaction Technology: Industrial Prospects, Springer, 2000, pp. 53–61.

[17] B. K. Paul, “Micro energy and chemical systems (MECS) and multiscale fabrication,” in Micromanufacturing and Nanotechnology, Springer, 2006, pp. 299–355.

[18] B. K. Paul and R. B. Peterson, “Microlamination for microtechnology-based energy, chemical, and biological systems,” in ASME International Mechanical Engineering Congress and Exposition, Nashville, Tennessee, 1999, vol. 39, pp. 45–52.

[19] B. K. Paul, P. Kwon, and R. Subramanian, “Understanding limits on fin aspect ratios in counterflow microchannel arrays produced by diffusion bonding,” J. Manuf. Sci. Eng., vol. 128, no. 4, pp. 977–983, 2006.

[20] B. Lajevardi, S. D. Leith, D. A. King, and B. K. Paul, “Arrayed microchannel manufacturing costs for an auxiliary power unit heat exchanger,” in Proceedings of the 2011 Industrial Engineering Research Conference, Reno, Nevada, 2011, pp. 21–25.

[21] T. Klotzbücher, T. Braune, A. G. Arlt, and G. Kolb, “Nd: YAG-Laser Welding with Dynamic Beam Forming,” Laser Tech. J., vol. 7, no. 5, pp. 28–31, 2010.

[22] D. Rosenthal, The theory of moving sources of heat and its application to metal treatments. ASME, 1946.

[23] D. Rosenthal, “Mathematical theory of heat distribution during welding and cutting,” Weld. J., vol. 20, no. 5, p. 220s–234s, 1941.

[24] P. G. Klemens, “Heat balance and flow conditions for electron beam and laser welding,” J. Appl. Phys., vol. 47, no. 5, pp. 2165–2174, 1976.

[25] D. T. Swift-Hook and A. E. F. Gick, “Penetration welding with lasers,” Weld. J., vol. 52, no. 11, p. 492s–499s, 1973.

[26] S. V. Pawar, “Fabrication of precipitation-hardened aluminum microchannel cooling plates for adsorption-based hydrogen storage systems,” Oregon State University, Corvallis, OR, 2013.

[27] L. Quintino, A. Costa, R. Miranda, D. Yapp, V. Kumar, and C. J. Kong, “Welding with high power fiber lasers–a preliminary study,” Mater. Des., vol. 28, no. 4, pp. 1231–1237, 2007.

[28] B. Lajevardi, “Laser keyhole welding for microlaminating a high-temperature microchannel array,” Oregon State University, Corvallis, OR, 2012.

[29] Q. Gao, J. Lizarazo-Adarme, B. K. Paul, and K. R. Haapala, “An Economic and Environmental Assessment Model for Microchannel Device Manufacturing: Part 2 – Application,” J. Clean. Prod., p. in review, ND.

[30] Q. Gao, J. Lizarazo-Adarme, B. K. Paul, and K. R. Haapala, “An Economic and Environmental Assessment Model for Microchannel Device Manufacturing: Part 1 – Methodology,” J. Clean. Prod., p. in review, ND.
[31] “Ecoinvent Data v.3.0,” Ecoinvent Centre, Swiss Centre for Life Cycle Inventories, Dübendorf, 2014.
[32] M. Patel, “Cumulative energy demand (CED) and cumulative CO2 emissions for products of the organic chemical industry,” Energy, vol. 28, no. 7, pp. 721–740, Jun. 2003.
[33] IPCC, 2006 IPCC Guidelines for National Greenhouse Gas Inventories. Japan: IGES, 2006.

Chapter 5: Conclusions

The rapid growth of the information technology (IT) industry and the miniaturization of semiconductors have resulted in substantial increases in energy consumption and power density of IT equipment, and heat generated within data center racks. The cold air supplied to the racks passes through the servers and discharges the dissipated heat to the hot aisle. The warmed air is discharged through outlet vents and cooled using an air conditioning system. Effective data center thermal management allows higher data center energy efficiency which in turn lowers total cost of ownership (TCO). Selecting and continuously evaluating the appropriate metrics will enable better insights into data center efficiency improvement strategies. In order to improve data center energy efficiency, efforts have been focused on novel center-level and rack-level cooling technologies to remove the heat generated by high-density servers. The operational energy performance of a data center evaporative cooling system and the manufacturing energy requirements for a server-scale microchannel heat exchanger (MCHX) are investigated in this research.

5.1. Summary

The evaluation of data center efficiency and measuring the impact of incorporating novel center-level and rack-level cooling technologies to remove the heat generated by high-density servers are investigated in this research. Energy monitoring and analysis was conducted to evaluate an evaporative cooling system installed at a data center located in Gresham, OR. To support this analysis a holistic energy efficiency and thermal management effectiveness metric was developed. A comprehensive metric was not previously reported in literature. An effective holistic metric allows better insight into the inefficiencies and enables higher level decision makers to take more effective investment strategies. Thus, a holistic metric and measurement approach was developed to evaluate the impact of changes for data center infrastructure and information technology (IT) equipment. A wireless sensor network (WSN) was established at the selected data center in Gresham, Oregon to facilitate remote monitoring of real-time data for equipment power use and the inside and outside environmental conditions.

In addition, energy analysis for the production of a MCHX device to liquid cool the warm exiting air from server racks is conducted. A weld depth model is developed and used to understand the capabilities and limitations of the laser welding process. Manufacturing process environmental analysis was conducted for making the device. Process-based life cycle assessment (LCA) used to evaluate the impacts of energy use in manufacturing the MPT-based MCHX device. Results were compared for photochemical machining and two joining methods: diffusion bonding and laser welding.

5.2. Conclusions

Evaluation of data center efficiency revealed that air delivered to the racks was often below the ASHRAE recommended guideline envelope and considerable portion of the air delivered to the racks bypassed the IT equipment in the racks due to inadequate separation of hot/cold aisles. Thus, hot/cold aisle containment and air flow management can be improved to enhance the overall efficiency of the center. Measuring the developed metric and comaring to the existing metrics it was found that the developed metric is more responsive to changes in cooling power and environmental conditions than commonly used metrics. The developed metric suggested average efficiency improvement of 10% for every kW reduction in cooling power and continuous data center efficiency improvement was measured when the power consumed by cooling system decreased from 4kW to 1kW. Average improvement of about 3% for every kW reduction in cooling power was suggested measuring the PUE. The responses of the other metrics did not not show a continuous correlation with the cooling power.

Further, the evaporative cooling technology was shown to be more efficient and effective than conventional cooling technology.

Liquid cooling has been demonstrated as an effective strategy to provide a reliable environment for servers and to reduce the load on conventional cooling systems. While microchannel process technology (MPT)-based devices offer a space-efficient approach to liquid cooling of high-density servers, MPT device manufacturing has been shown to energy intensive. Laser welding is shown to reduce process time, reduce energy use, and improve process yield, and thus reduce environmental impacts compared to diffusion bonding.

5.3. Contributions

Due to the limitations of existing metrics to evaluate the performance and the impact of changes in data centers, a new metric developed which can more effectively evaluate the energy efficiency and thermal management of data centers. The evaporative cooling technology installed at the selected data center was evaluated using the existing and developed metrics.

Direct cooling of rack exit air enables this balance by reducing the load on the cooling fans and chiller. Using a heat exchanger, liquid cooling can lower the exit air temperature prior to being discharged into the hot aisle from the IT rack. Application of microchannel heat exchangers in liquid cooling of high density servers was investigated in this research and manufacturing energy analysis for the production of a MCHX device to liquid cool the warm exiting air from server racks was conducted. A process model for laser welding is developed, which enhances energy analysis.

5.4. Opportunities for Future Work

Future work could explore the potential for developing a real-time optimization method to improve overall data center performance through operational adjustments in response to real-time performance measurement data. Monitoring over an extended time period could be leveraged in predicting power consumption of IT equipment which in turn allows quick response to environmental conditions surrounding IT equipment. Development and deployments of statistical predictive models could increase the effectiveness of real-time optimization methods by faster responses to environmental changes. Predictive modeling could also reveal the need for maintenance, changes to infrastructure or operating conditions, or new IT equipment technology.

Future work should also focus on development and deployment of MCHX applications in innovative data center cooling solutions. Analysis should explore the potential for performance optimization of the existing technologies. Development of optimization methods could improve the efficiency of MCHXs cooling applications, which, in turn, increases the energy efficiency of data centers. Future work could also focus on evaluating the environmental impacts of the full life cycle of the microchannel heat exchangers (MCHXs) for liquid cooling of high density servers in comparison with conventional cooling systems. Given the unknown level of performance and operational lifetime of the device, the selected material and device geometry could have a significant impact on unit process and overall life cycle environmental impacts. Analysis should compare use phase efficiency of various devices to consider the energy (heat) transfer of the designs. In addition, other metrics should enter into the analysis to consider the sustainability performance more comprehensively (e.g., additional environmental impacts, economic competitiveness, and relative social impacts).

Appendix A
Data Center Cooling System Evaluation

Babak Lajevardi, Karl R. Haapala, and Joseph Junker
School of Mechanical, Industrial, and Manufacturing Engineering
Oregon State University

This appendix was provided as a white paper for IT Aire of Gresham, OR, July 2014

Appendix A: Data Center Cooling System Evaluation

A1.1 Abstract

Worldwide, many organizations are pursuing higher energy efficiency by reducing power consumption of their processes, systems, and supporting infrastructure. Likewise, higher energy efficiency is recognized to enable lower operational costs of data centers, which are common and proliferating throughout industry, academic institutions, and governmental organizations. More efficient data centers not only lower total cost of ownership (TCO), but also enable organizations to better manage increasing computing and network demands.

Data center power use and cooling challenges are two major issues facing organizations and, in particular, enterprises that operate conventional data centers. The rapid growth of the information technology (IT) industry and the miniaturization of semiconductors have resulted in substantial increases in energy consumption and power density of IT equipment, and, subsequently, heat generated by data center equipment contained within data center racks. Heat generated inside the racks must be removed from the data center to avoid damaging IT equipment (e.g., servers). In this white paper, the impact of replacing a conventional data center cooling system with a new cooling system design on data center energy efficiency is examined.

The energy efficiency of these two technologies has been evaluated for a data center in the city hall of Gresham, Oregon through the installation of the new cooling system, while maintaining the original, conventional system as operational. The data center consists of a row of seven racks of IT equipment within a small, air conditioned room. The data center was monitored by installing a wireless sensor network (WSN) to collect energy use, temperature, and humidity data. To perform the energy efficiency analysis, real-time monitoring of 25 parameters over an extended period of time was performed.

The conventional, split direct expansion (DX) air conditioning (AC) system was supplemented with a rooftop mounted indirect evaporative cooling unit (IT Aire DCS, Gresham, OR). Energy efficiency and thermal management have been assessed for each system by applying various data center performance metrics. The metrics were evaluated for the original system (baseline) and new system (IT Aire DCS) over extended periods. The results were analyzed, and it was concluded that the IT Aire system reduced energy use by 75%, on average, and improved power usage effectiveness (PUE) from 1.25 to 1.05 (1.00 is ideal). In addition, the new system showed that the overcooling evident with the original system was eliminated, which saves energy and can extend the life of the IT equipment. Additional details of the systems, monitoring, and analysis are discussed in this white paper.

A1.2 Introduction

A data center is a facility housing networked computers and servers as well as associated infrastructure, for the purposes of storage and management of large amounts of data and information [1]. The ever-increasing demand of information technology (IT) support along with developments in semiconductor miniaturization have led to higher density processors and a sharp increase in the energy consumption and the heat dissipated per unit volume of racks in data centers [2]. Energy in data centers is consumed by two main categories of equipment: the IT equipment and the infrastructure that supports the IT facilities, including the systems that provide reliable cooling and an adequate environment for the IT equipment. Since the power consumed by the IT equipment is converted into heat dissipated inside the racks [3], a significant increase in rack level heat density and resulting thermal management challenges have been rising over the past few decades [4], [5].

The United States Environmental Protection Agency (EPA) reported that the annual energy consumption of U.S. data centers in 2005 was approximately 61 billion kWh, or about 1.5% of the total U.S. energy consumption [6]. It was also reported that data center energy consumption in 2005 was nearly double that in 2000 [6]. According to Koomey [7], data center energy consumption significantly increased worldwide and in the U.S. from 2000 to 2010. As seen in Figure A1.1, worldwide and U.S. data centers consumed about 240 and 80 billion kilowatt-hours (kWh) of energy, respectively, in 2010, and this consumption was on a steady increase for the past decade. This trend is expected to continue with increasing data storage and processing demands.

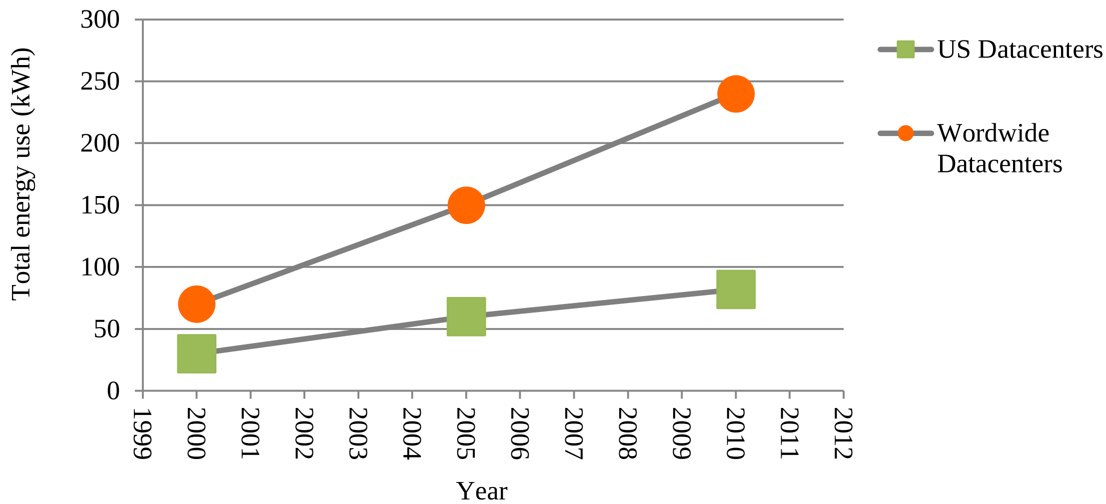

Figure A1.1. Worldwide and U.S. data center energy use

One of the most common metrics, power usage effectiveness (PUE), was also studied by the EPA [6]. Benchmarking more than 100 North American data centers revealed an average PUE of 1.91 in 2007 (a PUE of 2.00 implies that for every 2 kWh of energy supplied to the data center, 1 kWh is used by the IT equipment). Digital Realty Trust, Inc., a provider of data center solutions, surveyed more than 300 North American corporations (annual revenue of higher than $1 billion or at least 5000 employees) in 2013 [8]. An average PUE value of 2.9 was reported, while only 20% of the data centers reported a PUE of less than 2.0 and 9% stated a PUE of 4.0 or greater [9]. Despite all of the efforts during past decade to improve the energy efficiency of data centers, the reported PUEs suggest that the energy efficiency of North American data centers is decreasing. The continuous growth in rack power density and the high degree of inefficiencies in the existing data centers imply that additional investment in innovative data center cooling solutions is of great importance.

A1.3 Background

An average data center rack power density of 8.5 kW was reported in 2012 with a growth rate of about 8% over the previous year (7.9 kW) [8]. Due to the significant increase in rack power density in recent years, demands for effective data center cooling strategies have substantially increased. This demand has created an opportunity for development of technologies and systems for energy efficiency and cost savings. Thus, many efforts have been made to increase the efficiency and effectiveness of thermal management in data centers. To promote industrial adoption, the performance of these systems must be evaluated and compared to existing and new technologies.

The cooling system accounts for a significant portion of total data center energy consumption, and cooling cost drives the total operational cost of a typical data center [10]. It is estimated that about five times the server cost is spent on cooling and supporting infrastructure when a server is operated in an adequate thermal environment [11]. Since the power consumed by IT equipment is converted into heat dissipated through the racks [3], reliable cooling is imperative to provide an adequate environment for IT devices. Poor thermal management lowers energy efficiency and can lead to higher risks of server failures and lower IT equipment longevity.

Despite all the efforts that have been undertaken over recent decades, energy consumption of data center cooling systems has remained a major concern, and opportunities remain for efficiency improvements. In conventional data centers, there is additional pressure to cool the ever-increasing rack power densities using novel strategies due to inefficiencies of conventional thermal management systems, which often rely on direct expansion (DX) air conditioning (AC). The first step in implementing efficiency improvement strategies is to evaluate data center performance. Thus, a challenge is to effectively monitor data center energy consumption and environmental conditions in the data center and outdoors (heat is ultimately transferred to the external environment).

In this study, a data center in the City Hall of Gresham, Oregon was monitored and evaluated by installing a wireless sensor network (WSN) to collect energy use, temperature, and humidity data. The data center consists of a row of seven IT equipment racks in a room cooled with an existing split DX-AC system. The existing system was then supplemented with a rooftop mounted indirect evaporative cooling unit, an IT Aire DCS (Data-center Cooling Solution), developed and manufactured by IT Aire Inc. of Gresham, OR. Energy efficiency and thermal management of the data center was evaluated using a combination of effective energy and thermal metrics. Performance metrics were evaluated for both the original (baseline) and new (IT Aire) systems over an extended period. Ultimately, the goal was to compare data center energy efficiency and thermal management performance in each case.

The following sections review the two data center cooling technologies and metrics used to evaluate the performance of the data center. The equipment configuration and setup is then described and results are presented and discussed.

A1.4 Data Center Cooling Technology

Due to rapid growth of IT equipment miniaturization and a significant increase in rack level power densities, data center thermal management challenges have been rising over the past few decades. An effective data center cooling solution provides an adequate environment for the IT equipment by directing the correct volume of air at the right temperature into the racks. The cold air supplied to the data center passes through the racks and removes the heat generated by IT equipment to the hot aisle. The warmed air is discharged through outlet vents and cooled, most often using a direct expansion air conditioning system.

A1.4.1 Direct Expansion Air Conditioning (DX-AC) Systems

In a DX-AC system the air is directly cooled while passing over the refrigeration-cooling coil. The refrigerant is pumped through coils to transfer heat. The refrigerant absorbs heat by changing state from liquid to gas and reject it to the outside air by state change from gas to liquid. A fan blows over the coils carrying refrigerant, which cools down the passing air and provide the cooled air to the data room [3].

Indoor humidity control significantly affects the energy use of DX-AC systems. In DX-AC systems the indoor moisture load is provided mainly from outdoor air, which significantly varies with changes in weather conditions. In order to compensate for the lack of sufficient humidity control, the indoor air temperature is often lowered and the run time of the system increased, which in turn increases the energy use and lower the energy efficiency [12].

Thus, in order to overcome the cooling challenges due to increasing heat dissipated by high density servers, it is imperative to incorporate innovative cooling solutions that consume less energy [12]. One potential technology is evaporative cooling.

A1.4.2 Indirect Evaporative Cooling Systems

An evaporative cooling system cools air through evaporation of water. Water has a large enthalpy (the amount of heat gained or released through a process at constant pressure) of vaporization, which allows the rapid removal of heat. Evaporative cooling is an isenthalpic process (a process in which the enthalpy does not change) and takes place at constant wet bulb temperature and constant enthalpy [12]. A schematic of the indirect evaporative cooling principle is shown in Figure. A secondary air stream flows across the wetted outer surface and removes heated vapor. The evaporatively cooled surface provides an area for heat exchange to cool the primary stream.

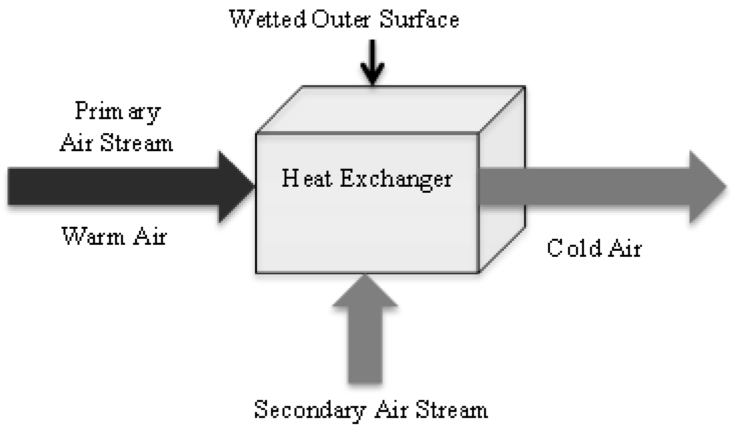

Figure A1.2. Schematic of indirect evaporative cooling

For an evaporative cooling system the heat exchanger surface cooling the air is not as cold as that of a DX cooling system. The dew point of the air is not depressed as far and the air is not dried to the same degree as in a DX cooling system. As a result less re-humidification should be required.

The indirect evaporative cooling system developed by IT Aire enables precise cooling for data centers [13]. To maintain the IT equipment within an environment recommended by ASHRAE, control systems are incorporated that allow real time air flow and humidity adjustments. A polymer heat exchanger in the primary cooling section of the IT Aire cooling system removes heat from the data center returned air. The heat transfer is empowered by passing an outdoor air stream over the wetted medium (outer surface) of the heat exchanger.

A1.5 Methodology

Monitoring of energy use and environmental conditions is imperative in order to assess data center energy efficiency and thermal management effectiveness. Measuring the performance of a data center based upon standard metrics provides the opportunity to track improvements and changes, to estimate the impact of changes, and to draw comparisons to other technologies and data center configurations.

A variety of metrics have been proposed to quantify data center efficiency and performance [5]. In this study, metrics were selected from among the most widely used metrics to enable better insight into energy efficiency and thermal management issues. The objective is to measure the performance of the data center before and after installation of the IT Aire cooling system. The metrics reviewed below support understanding the load on and performance of the cooling system, as well as its operational health.

A1.5.1 Power Usage Effectiveness (PUE) and Data Center Infrastructure Energy (DCiE)

Two primary metrics, power usage effectiveness (PUE) and data center infrastructure energy (DCiE), were introduced by the Green Grid industry consortium over the past decade [14] to measure data center energy efficiency. PUE is defined as the total energy delivered to the data center divided by the total energy drawn by the IT equipment. IT equipment energy is defined as "the energy consumed by equipment that is used to manage, process, store, or route data within the compute space" [14]. PUE is the most widely-used metric and can be calculated using Eq. A1.1.

$$\mathrm{PUE}=\frac{P_{inf.}+P_{IT}}{P_{IT}}=1+\frac{P_{inf.}}{P_{IT}} \tag{A1.1}$$

In this equation, $P_{inf.}$ is the power input into the supporting infrastructure, mainly the cooling system, and P_{IT} is the power consumed by the IT equipment in the racks. Ideally, PUE would hold a value of 1.0, meaning all the power into the data center is consumed by the IT equipment. In reality, however, due to the heat generated, energy consuming cooling strategies are imperative to reject heat from the racks. Additional power used for rack cooling purposes increases the value of PUE as suggested by Eq. A1.1.

DCiE represents a reciprocal of PUE and, thus, can be calculated using Eq. A1.2.

$$\mathrm{DCiE}=\mathrm{PUE}^{-1}=\frac{P_{IT}}{P_{inf.}+P_{IT}} \tag{A1.2}$$

As seen in the above equations, PUE and DCiE measure the portion of the total power input to the data center that is consumed by the IT equipment.

A1.5.2 Rack Cooling Index (RCI)

The rack cooling index (RCI) proposed by Herrlin [15] measures the degree to which the IT equipment inside the racks are served with air in the rack intake air temperature range recommended by ASHRAE (formerly the American Society of Heating Refrigerating and Air-Conditioning Engineers) [16]. Thus, the RCI metric evaluates how effectively an adequate environment is provided for the racks and is expressed by the range defined by Eqs. A1.3 and A1.4, below.

$$RCI_{HI}=\left[1-\frac{\sum\left(T_{intake}-T_{max\text{-}rec}\right)_{T_{intake}>T_{max\text{-}rec}}}{(T_{max\text{-}all}-T_{max\text{-}rec})n}\right]\times 100\% \tag{A1.3}$$

$$RCI_{LO}=\left[1-\frac{\sum\left(T_{min\text{-}rec}-T_{intake}\right)_{T_{intake}<T_{max\text{-}rec}}}{(T_{min\text{-}rec}-T_{min\text{-}all})n}\right]\times 100\% \tag{A1.4}$$

RCI_{HI} (Eq. A1.3) is the index value at the high end of the recommended temperature range, RCI_{LO} (Eq. A1.4) is the value at the low end of the recommended temperature range, T_{intake} is the rack intake air temperature, n is the total number of intakes, $T_{max\text{-}rec}$ is the maximum recommended temperature, $T_{max\text{-}all}$ is the maximum allowable temperature, $T_{min\text{-}rec}$ is the minimum recommended temperature, and $T_{min\text{-}all}$ is the minimum allowable temperature.

According to ASHRAE, rack intake temperatures of 64-77°F (18-25°C) and 59-90°F (15-32°C), are defined as the recommended and allowable ranges, respectively [16]. An RCI of 100% reflects intake temperatures within the recommended range. Lower percentages of RCI_{HI} imply that heat rejection from the racks is not effective and there is a possibility of hot spots within the racks. Similarly, lower percentages of RCI_{LO} indicate that the racks are overcooled, which suggests low cooling power efficiency due to poor thermal management. Table A1.1 shows the corresponding values of RCI at rack intake temperatures within and outside the ASHRAE recommended envelope.

Table A1.1. Rack cooling index (RCI) values at different intake temperature ranges.

Rack Intake Temperature (T_{intake})	**RCI**	**Description**
$64°F \leq T_{intake} \leq 77°F$	RCI_{HI}=100%, RCI_{LO}=100%	IT equipment is maintained within ASHRAE recommended envelope
$T_{intake} \leq 64°F$	RCI_{HI}=100%, <u>RCI_{LO}≤100%</u>	IT equipment is overcooled
$T_{intake} \geq 77°F$	<u>RCI_{HI}≤100%</u>, RCI_{LO}=100%	Ineffective heat removal

A1.6 Experimental Setup

The first step in evaluating data center performance is to effectively monitor energy consumption and environmental conditions. Wireless sensor networks (WSNs) facilitate gaining insight into industrial systems through real-time monitoring and provide the opportunity to improve efficiency through evaluation and control of industrial operations. A wireless sensor network (WSN) was established at the data center under study to facilitate remote monitoring of real-time data for equipment power use and the inside and outside environmental conditions (Fig. A1.3). The network included data collection nodes within the data center, as well as in close proximity to the cooling system units located on the roof. All evaluation nodes on the roof and in the data center were connected in a single wireless network.

The monitoring equipment on the roof, including a dry bulb temperature sensor, a relative humidity (RH) sensor, and current transducers, provided the outdoor air status as well as the cooling system power load. The nodes that logged data from temperature and RH sensors in the data center enabled measuring the quality of rack intake and exit air, as well as the air supplied to the room and returned to the cooling units at the HVAC air duct vents. The IT equipment load was measured by monitoring the power draw of each of the 14 rack power cords (two cords per rack).

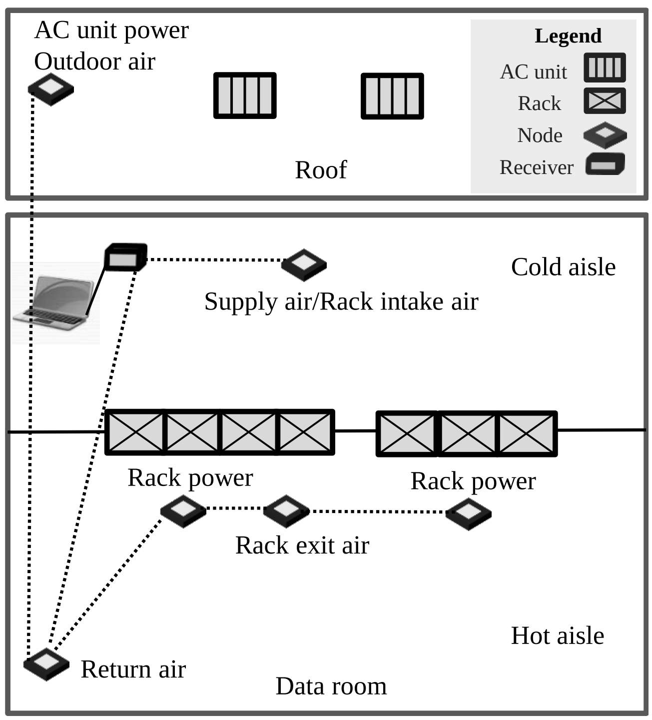

Figure A1.3. Schematic top view of the wireless sensor network (WSN) installed in the data center under study (City Hall, Gresham, OR).

Due to the changes in the layout and equipment within the data center over the course of the study, portable data loggers were selected, which could be deployed anywhere in the center. Data loggers were capable of recording data in internal memory, so during power outages data could be recorded. The internal memory could store up to 40,000 readings. Each of the data loggers was scheduled to record the data with an interval of one minute. Recorded data from nodes in the wireless network was automatically saved to a single file.

The network was programmed to save a copy of the updated file to a local drive, send a copy to the Oregon State University Energy Efficiency Center via email, and save a copy to an Oregon State University FTP address every 24 hours. An alarm was set to alert the research team if a logger reading was out of range or if a node was identified as missing from the network. Also, a “heartbeat” alarm notified researchers every 12 hours that the network was active and the receiver was collecting and recording data from the nodes.

In order to increase the robustness of the network, a router was added into the monitoring system to improve the communication path from the equipment on the roof to the receiver. The receiver was connected to a laptop in data center; it collected data transferred by data loggers and saved it to a database. The implemented wireless network used low-power 2.4 GHz radio signals to transmit data across the network and to the receiver. The signals lost strength due to obstructions in the communication path. When a data logger lost its signal, the nodes which routed through it would automatically search for a new path to transfer data to the receiver.

A1.7 Results and Discussion

The energy efficiency and thermal management metrics discussed above are used and the data center performance is evaluated and summarized based upon the data collected over a one-week monitoring period each for the original DX-AC system and new IT Aire DCS. A total of 504,160 data points for each of the 25 parameters were collected. As seen in Figure A1.4, an overall average PUE of 1.25 (DCiE of 80%) was measured for original system, which suggests that about 0.25 kW of power use is consumed to condition the data center air and remove the heat from the racks for every kilowatt of electricity delivered to the IT equipment. A reduction of 0.2 (1.25 to 1.05) in average PUE value was observed when cooling with the new IT Aire system.

The IT Aire cooling system enables a DCiE of about 95%, which suggests that only 5% of total power consumption is used to cool the IT equipment racks.

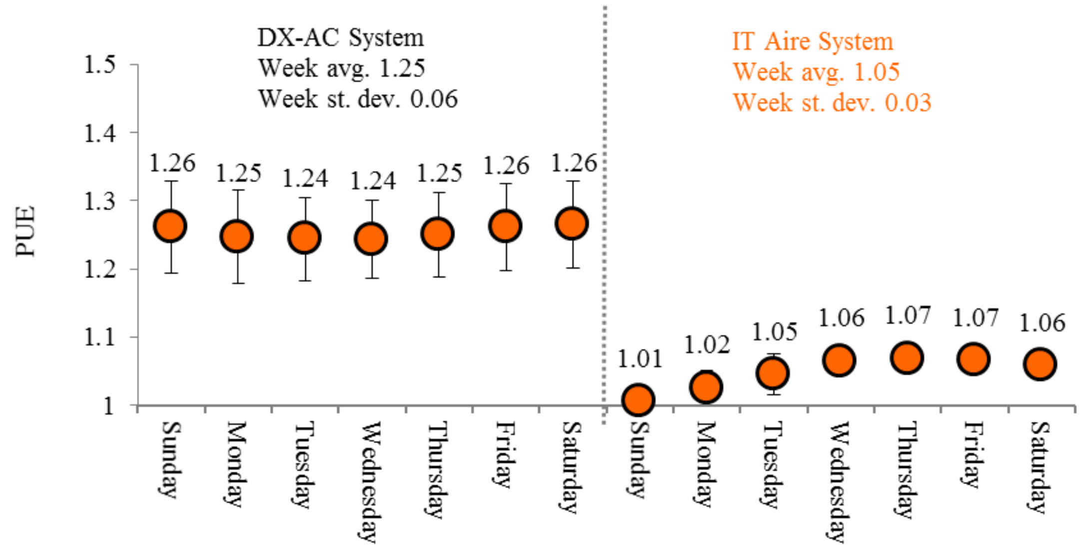

Figure A1.4. Measured PUE over two weeks of the monitoring period (one week before and one week after installation of the IT Aire indirect evaporative cooling system).

In order to evaluate the degree to which the IT equipment inside the racks is maintained in the rack intake air temperature range recommended by ASHRAE, the rack cooling index (RCI) was calculated. As seen in Figure A1.5, the RCI_{HI} of 100% for both systems indicates the absence of overheating.

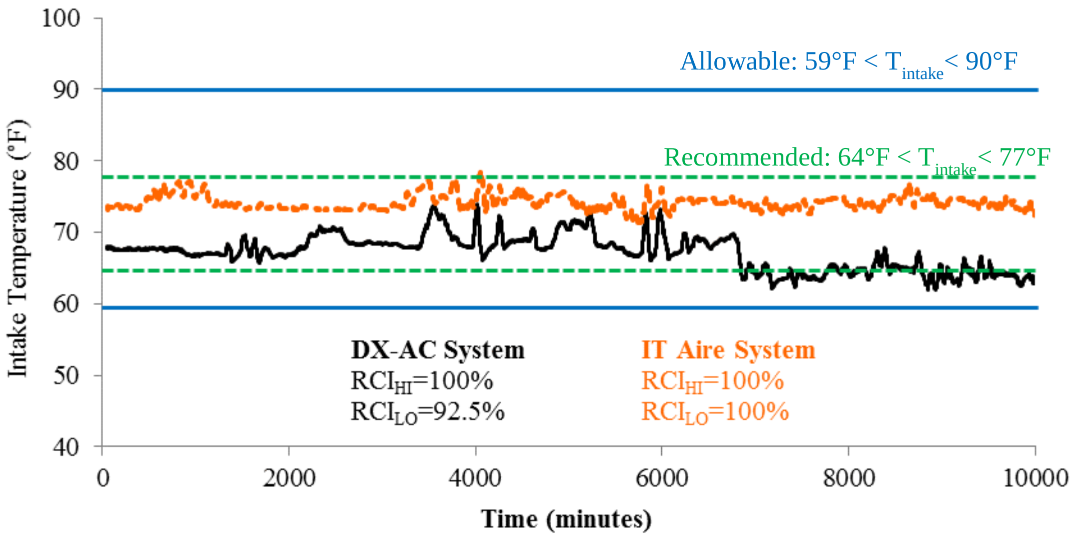

Figure A1.5. Rack intake temperatures compared to ASHRAE environmental guidelines.

The RCI_{LO} value of 92.5% for the existing (DX-AC) system indicates that the racks are cooled below the low temperature recommended by ASHRAE during 7.5% of the monitoring period. The calculated RCI_{HI} and RCI_{LO} values for the IT Aire system indicate that the rack intake air temperatures fell within the ASHRAE recommended range over the entire monitoring period. Both systems operate within the allowable range, however, indicating that data center environmental conditions should be conducive to equipment operational health. The IT Aire cooling system utilizes evaporative cooling technology and controls that adapt the cooling system capacity to the load. Thus, the IT Aire cooling system reduces the wasted energy, facilitating a significant reduction in cooling energy consumption.

This reduction in energy use of the IT Aire cooling system compared to the energy consumption of the original DX-AC system is shown in Figure A1.6. An average energy savings of about 75% was observed when the cooling system was switched to the IT Aire evaporative indirect cooling system compared to the existing direct expansion air conditioning system.

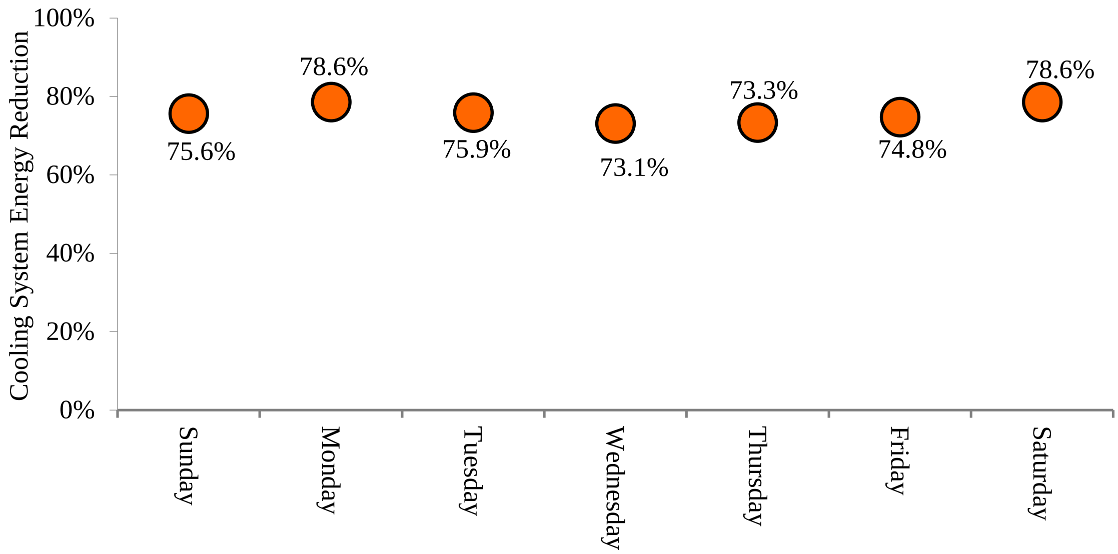

Figure A1.6. Energy savings after installation of the IT Aire indirect evaporative cooling system.

A1.8 Bin Analysis

Average PUE values for the DX-AC system and IT Aire cooling system as a function of outdoor dry bulb and wet bulb temperatures are illustrated in Figure A1.7 and Figure A1.8, respectively. Due to the different technologies used by two cooling systems, the PUE values are plotted as a function of dry bulb for DX-AC and wet bulb for IT Aire system to reflect the efficiency variations by changes in weather conditions. Bins of 5°F were selected to better highlight the variations in efficiency performance within each system. As seen in the figures, the average PUE in both cases increased as the outdoor temperature increased, which indicates cooling needs increased, as expected. PUE values between 1.3 and 1.45 were observed when running the DX-AC system at dry bulb temperatures above 50°F. PUE values between 1.12 and 1.18 were achieved when running the IT Aire system at wet bulb temperatures above 50°F.

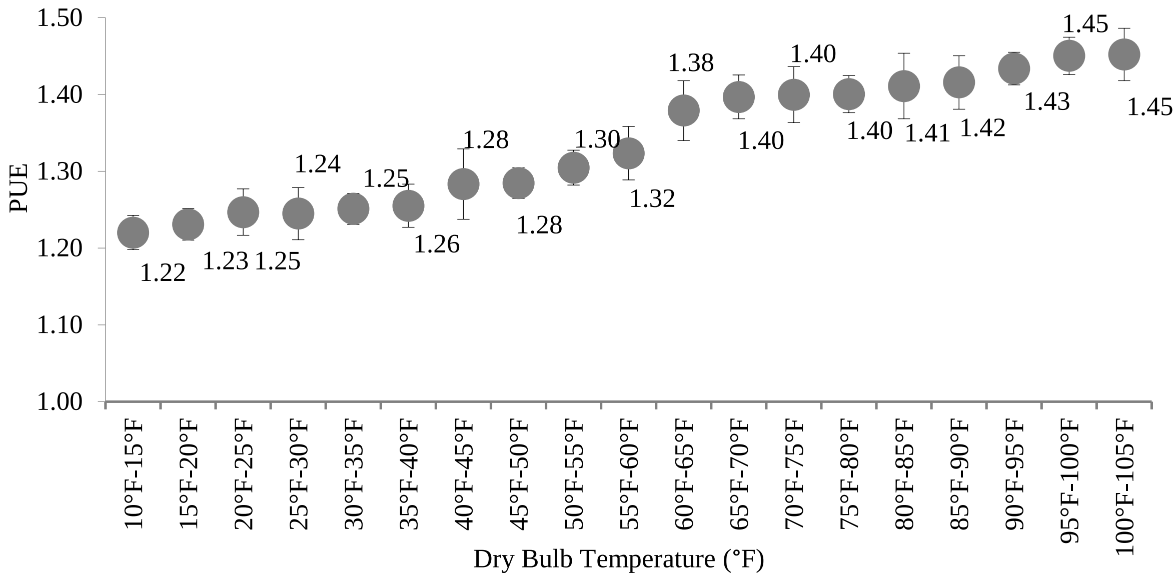

Figure A1.7. PUE as a function of dry bulb temperature for the DX-AC system.

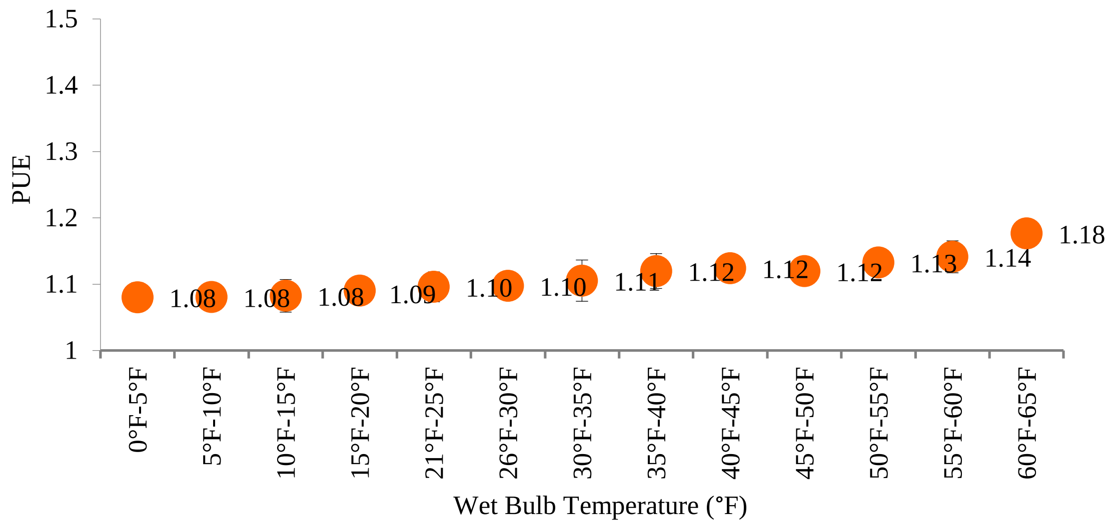

Figure A1.8. PUE as a function of wet bulb temperature for the IT Aire indirect evaporative system.

Using the data collected by the outdoor air status data logger, the dry bulb and wet bulb temperature distributions over a one-year period are plotted in Figure A1.9 and Figure A1.10, respectively. As seen in Figure A1.9, for about 70% of the monitoring period, the dry bulb temperature falls at temperatures above 50°F. The bin analysis suggests PUE values between 1.3 and 1.45 when running the DX-AC system at temperatures above 50°F as shown in Figure A1.7. The bin distribution in Figure A1.10 suggests that about 47% of the monitoring period, the wet bulb temperatures fall above 50°F while the corresponding PUE values varies between 1.12 and 1.18 for the IT Aire system as illustrated in

Figure A1..

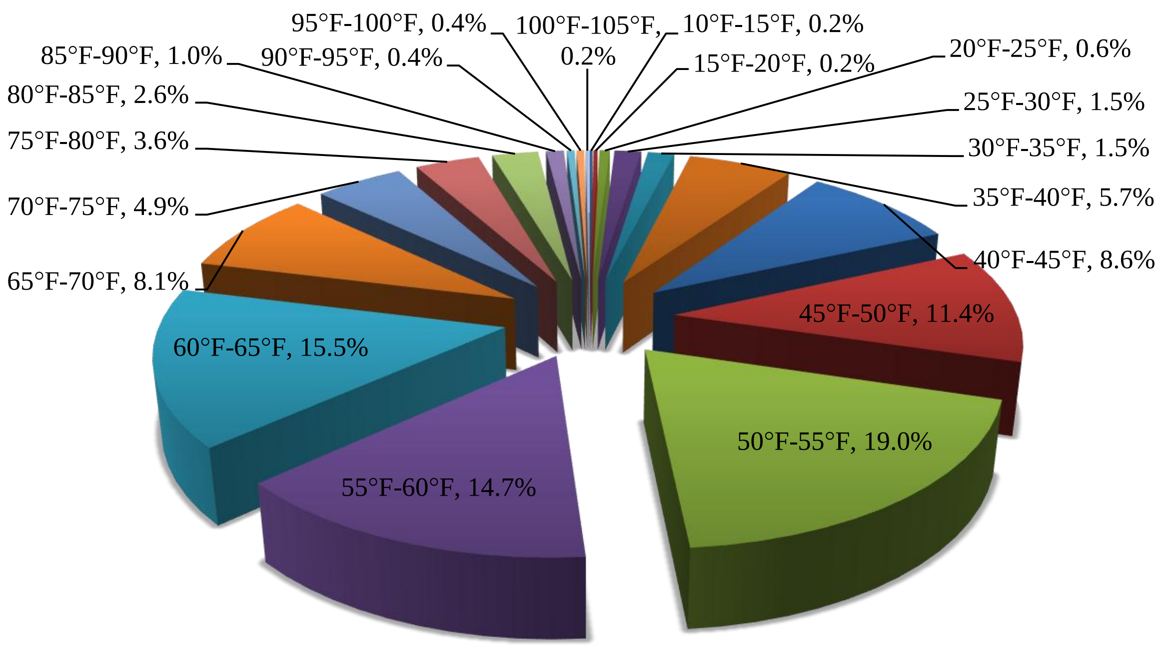

Figure A1.9. Dry bulb temperature bin distribution over a monitoring period of one year.

In order to compare the annual efficiency performance of the two cooling systems, the average PUEs calculated for bins of outdoor dry bulb and wet bulb temperatures were used and total annual energy use by each system were estimated.

The calculated annual energy consumption of DX-AC system and IT Aire cooling system are plotted in Figure A1.11. An energy savings of about 72% is expected when switching to from direct expansion air conditioning to indirect evaporative cooling.

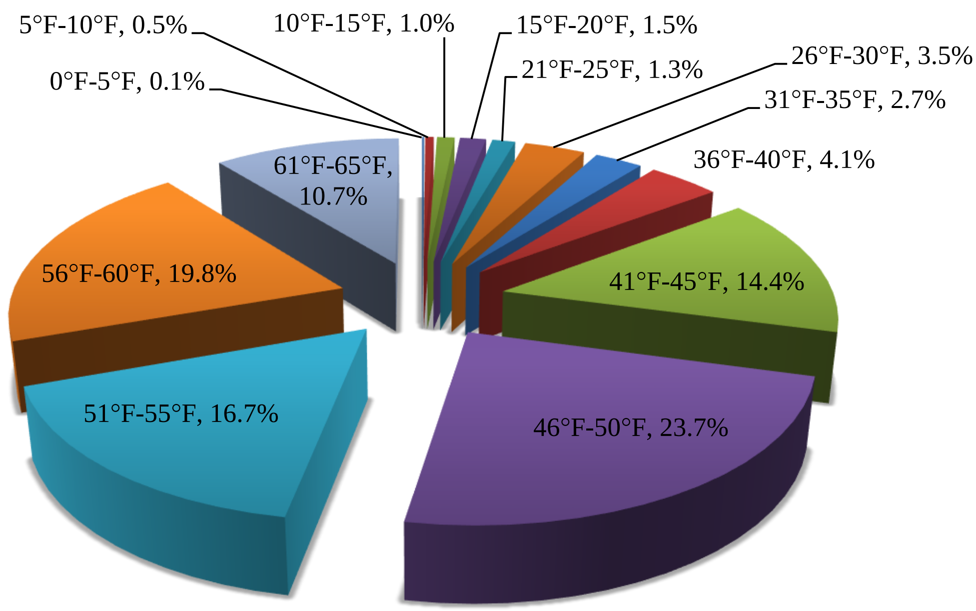

Figure A1.10. Wet bulb temperature bin distribution over a monitoring period of one year.

A similar analysis could be performed with historical dry bulb and wet bulb data for any location to estimate potential energy savings with a switch to an IT Aire indirect evaporative cooling system from a standard refrigerated cooling system.

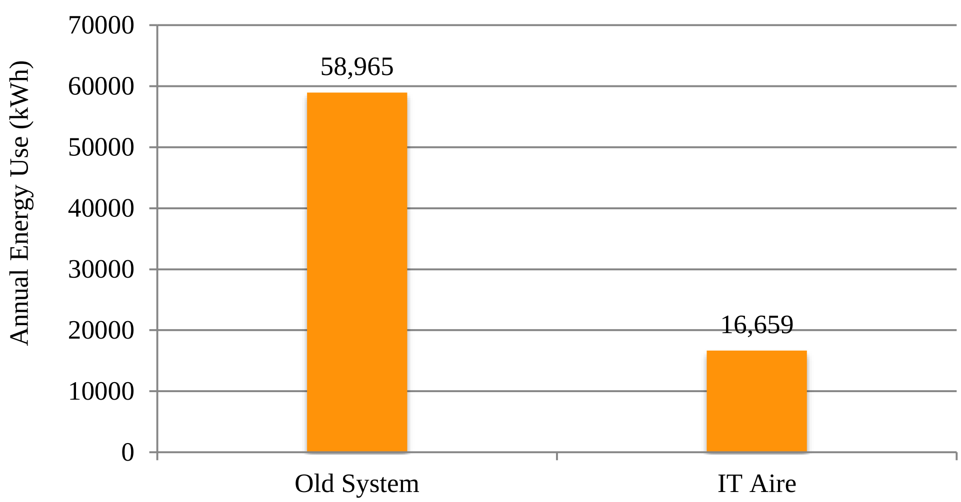

Figure A1.11. Estimated annual energy consumption of the two cooling systems.

A more in-depth discussion of the methods and results, as well as other data center energy efficiency metrics are reported in the following publications:

1. Lajevardi, B., K.R. Haapala, and J.F. Junker, 2014, "An Energy Efficiency Metric for Data Center Assessment," *Proceedings of the IIE/ISERC*, Paper I657, May 31-June 3, Montreal, Quebec, Canada.
2. Lajevardi, B., K.R. Haapala, and J.F. Junker, 2014, "Real-time Monitoring and Evaluation of Energy Efficiency and Thermal Management of Data Centers," *Proceedings of the NAMRI/SME*, Vol. 42, Paper NAMRC42-4465, June 9-12, Detroit, MI.
3. Lajevardi, B., K.R. Haapala, and J.F. Junker, 2014, "Real-time Monitoring and Evaluation of Energy Efficiency and Thermal Management of Data Centers," *SME Journal of Manufacturing Systems*, in press.

Acknowledgements

The authors gratefully acknowledge the support of this project from Oregon BEST, the Portland Development Commission (PDC), and Energy Trust of Oregon (ETO) through the Oregon BEST Commercialization Program. The project would not have been possible without their funding support. The authors also wish to express their sincere appreciation for the hours of assistance they received from everyone involved at the City of Gresham and IT Aire, as well as their availability for questions and willingness to participate in this lengthy experiment, which was nearly two years in the making.

References

[1] S. F. Anderson, "Improving Data Center Efficiency," Energy Engineering, vol. 107, no. 5, pp. 42–63, 2010.

[2] R. K. Sharma, C. E. Bash, and C. D. Patel, "Dimensionless parameters for evaluation of thermal design and performance of large-scale data centers," presented at the 8th ASME/AIAA Joint Thermophysics and Heat Transfer Conference, 2002, pp. 1–11.

[3] Y. Joshi and P. Kumar, Energy efficient thermal management of data centers. Springer, 2012.

[4] B. Kenneth, "Heat Density Trends in Data Processing, Computer Systems, and Telecommunications Equipment," UpTime Institute, White paper, 2005.

[5] G. Koutitas, “Challenges for energy efficiency in local and regional data centers,” 2010.

[6] R. Brown, E. Masanet, B. Nordman, B. Tschudi, A. Shehabi, J. Stanley, J. Koomey, D. Sartor, and P. Chan, “Report to congress on server and data center energy efficiency, public law, 109-431,” United States Environmental Protection Agency Energy Star Program, 2007.

[7] J. Koomey, “Growth in data center electricity use 2005 to 2010,” The New York Times, vol. 49, no. 3, 2011.

[8] “Results Of Annual North American Data Center Demand Survey,” Digital Realty Trust, Inc., Survey, Apr. 2013.

[9] “N. America Campos Survey Results,” Digital Realty Trust, Inc., Jan. 2013.

[10] “The green grids opportunity: decreasing data center and other IT energy usage patterns,” The Green Grids, Technical Report, 2007.

[11] L. Stahl and C. Belady, “Designing an alternative to conventional room cooling,” presented at the IEE conference publication, 2001, pp. 109–115.

[12] Y. Joshi and P. Kumar, “Emerging Data Center Thermal Management and Energy Efficiency Technologies,” in Energy Efficient Thermal Management of Data Centers, Springer, 2012, pp. 569–611.

[13] “IT Aire System.” [Online]. Available: http://itaire.com/it-aire-system. [Accessed: 16-Jul-2014].

[14] V. Avelar, Azevedo, and A. French, “PUETM: A comprehensive examination of the metric,” The Green Grid, White paper 49, 2012.

[15] M. K. Herrlin, “Rack cooling effectiveness in data centers and telecom central offices: The rack cooling index (RCI),” Transaction-American Society Of Heating Refrigerating Aand Air Conditioning Engineers, vol. 111, no. 2, p. 725, 2005.

[16] ASHRAE, TC. “9.9 (2011) Thermal guidelines for data processing environments–expanded data center classes and usage guidance,” Whitepaper prepared by ASHRAE technical committee (TC), vol. 9, 2011.