Modeling, Experimentation, and Analysis of Data Center Waste Heat Recovery and Utilization

by

Anna Haywood

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved February 2014 by the
Graduate Supervisory Committee:

Patrick Phelan, Chair
Marcus Herrmann
Sandeep Gupta
Steven Trimble
Stefan Myhajlenko

ARIZONA STATE UNIVERSITY

May 2014

ABSTRACT

Increasing computational demands in data centers require facilities to operate at higher ambient temperatures and at higher power densities. Conventionally, data centers are cooled with electrically-driven vapor-compressor equipment. This paper proposes an alternative data center cooling architecture that is heat-driven. The source is heat produced by the computer equipment.

This dissertation details experiments investigating the quantity and quality of heat that can be captured from a liquid-cooled microprocessor on a computer server blade from a data center. The experiments involve four liquid-cooling setups and associated heat-extraction, including a radical approach using mineral oil. The trials examine the feasibility of using the thermal energy from a CPU to drive a cooling process. Uniquely, the investigation establishes an interesting and useful relationship simultaneously among CPU temperatures, power, and utilization levels. In response to the system data, this project explores the heat, temperature and power effects of adding insulation, varying water flow, CPU loading, and varying the cold plate-to-CPU clamping pressure. The idea is to provide an optimal and steady range of temperatures necessary for a chiller to operate. Results indicate an increasing relationship among CPU temperature, power and utilization. Since the dissipated heat can be captured and removed from the system for reuse elsewhere, the need for electricity-consuming computer fans is eliminated. Thermocouple readings of CPU temperatures as high as 93°C and a calculated CPU thermal energy up to 67W_{th} show a sufficiently high temperature and thermal energy to serve as the input temperature and heat medium input to an absorption chiller.

This dissertation performs a detailed analysis of the exergy of a processor and determines the maximum amount of energy utilizable for work. Exergy as a source of realizable work is separated into its two contributing constituents: thermal exergy and informational exergy. The informational exergy is that usable form of work contained within the most fundamental unit of information output by a switching device within a CPU. Exergetic thermal, informational and efficiency values are calculated and plotted for our particular CPU, showing how the datasheet standards compare with experimental values. The dissertation concludes with a discussion of the work's significance.

DEDICATION

To my mother, Gladys, my father, Richard, and my best friend Ron.

ACKNOWLEDGMENTS

The author acknowledges the support of the National Science Foundation through award 0855277. There is also much appreciation extended towards committee members, Dr. Sandeep Gupta, Dr. Marcus Herrmann, Dr. Stefan Myhajlenko, Dr. Steven Trimble, and especially committee chair Dr. Patrick Phelan and researcher Jon Sherbeck whose guidance and expertise have been invaluable in making this work a success.

TABLE OF CONTENTS

Page

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

NOMENCLATURE

COP	Coefficient of performance
CRAC	Computer room air conditioner
CF	Captured fraction of board heat
DCiE	Data Center infrastructure Efficiency
HHF	High heat fraction
PUE	Power usage effectiveness
$\dot{Q}_{CRAC}$	Heat load on the CRAC, W
$\dot{Q}_{C}$	Heat removed by chiller, W
$\dot{Q}_{IT}$	Computer equipment heat, W
$\dot{Q}_{HiT,S}$	High-temperature heat from computer racks in data center to thermal storage, W
$\dot{Q}_{Air,C}$	Low-temperature heat flow of hot air in room removed by heat-driven chiller, W
$\dot{Q}_{EXT,S}$	Heat flow from external source of thermal energy to thermal storage, W
$\dot{Q}_{ext}$	Heat from an external supplemental source of thermal energy, W
$\dot{Q}_{DC,CRAC}$	Heat flow from the data center removed by the CRAC, W
$\dot{Q}_{S,C}$	Heat flow from thermal storage to heat-driven chiller, W
$\dot{Q}_{DC,CT}$	Heat flow from data center to cooling tower when utilizing ambient air, W
$\dot{Q}_{C,CT}$	Heat flow from heat-driven chiller to cooling tower, W

$\dot{Q}_{CT,A}$ Heat flow from cooling tower to ambient, W

$\dot{Q}_{TOT,DC}$ Total heat flowing from data center, W

$\dot{W}_{TOT,DC}$ Total electric power into data center, W

$\dot{W}_{IT}$ Electric power into racks holding computer equipment, W

$\dot{W}_{Loss}$ Electric power loss from PDU to racks, W

$\dot{W}_{Lights}$ Electric power into lights, W

$\dot{W}_{Cool}$ Electric power to operate the data center cooling system, W

$\dot{W}_{CRAC}$ Electric power into CRAC unit, W

$\dot{W}_{AC\_FANS}$ Electric power into condenser fans for CRAC unit, W

$\dot{W}_{CP}$ Electric power into condenser fan for chiller unit, W

CHAPTER 1

INTRODUCTION

1.1. Background

The rise in demand for the important work of data centers has created a noticeable impact on the power grid. The efficiency of data centers has become a topic of concern as the densely packed, energy intensive computer equipment inside a data center are creating power demands that are much higher than those of a standard residence or commercial office building [1]. In fact, data centers can be 40 times more energy intensive than a standard office building and require higher levels of power and cooling [2]. Furthermore, direct power consumption for these datacom facilities is increasing due to growing demand for the services they provide, particularly internet and intranet services [3]. The rise in demand for the important work of data centers has created a noticeable impact on the power grid.

US data centers consume a rising portion of the US electricity supply. According to the August 2, 2007 US Environmental Protection Agency (EPA) report on server and data center energy efficiency to Congress, the US server and data center sector used 61 terawatt hours (TWh) of electricity in 2006 (double the amount consumed in 2000). In 2006, this 61 TW h of electricity represented 1.5% of total US electricity consumption and cost $4.5 billion [3]. At that time, US energy consumption by servers and data centers was predicted to nearly double again by 2011 to 107.4 TW h or $7.4 billion annually [4]. However, an economic downturn beginning in 2008 resulted in fewer server purchases and thus in a reduction in the number of forecasted server installations also

known as the IDC server installation base [5]. This lowered server installation base had the most impact on slowing the growth rate, while the industry's efforts to improve data center efficiency since 2005 had less of an effect [5]. Another noticeable reduction in electricity expenditure came from the benefits of server virtualization which allows users to make better use of each server's capacity and thus limited the number of servers needed [5]. Consequently, the predicted doubling of data center power consumption failed to materialize, though it did grow by 36% nationally and 56% globally, totaling about 2% of US electricity use and 1.3% of World electricity use in 2010 [5].

Nevertheless, growth is still occurring as data centers are growing larger in size, consuming more electricity and dissipating more heat. According to a survey administered by the Association for Computer Operation Managers (AFCOM) and InterUnity Group, data center power requirements are increasing by 8% per year on average, and 20% per year in the largest centers [6]. This trend is leading to more problems with temperature control. As processor manufacturers such as Intel, AMD, IBM, and others continue to deliver on Moore's Law – doubling the number of transistors on a piece of silicon every 18 months – the resulting power density increase within the chips leads to dramatically rising temperatures inside and around those chips [7]. As servers become more power dense, more kilowatts are required to run and to cool them [8].

In 2007, The Uptime Institute's whitepaper entitled Data Center Energy Efficiency and Productivity stated that the 3-year operational and capital expenditures of powering and cooling servers can be 1.5 times the cost of purchasing server hardware [9]. Three years was the chosen period as this is the functional life of most servers [6].

Of the total power consumed by a typical data center, about half is attributed to conventional cooling [3]. Therefore, an effective way to implement energy efficiency in data centers is to address the cooling power necessary for operations. Since approximately 100% of electrical power into a computer device ($\dot{W}_{IT}$) is dissipated as heat ($\dot{Q}_{IT}$), the central idea of this project is to transfer directly that available thermal energy from the highest power components on a server blade (i.e., the CPUs) to drive a heat-activated cooling process, thereby lessening the heat load on and subsequent electrical grid power consumed by the standard vapor-compression refrigeration typical in data centers [10].

A similar effort in CPU heat reuse is the "zero-emission" data center by IBM [11]. Although in both cases, the type of heat being targeted is specifically the concentrated heat directly from each CPU, the difference in our approach is the plan to reuse that heat to drive a cooling process [12].

1.2. Objectives

The primary objective for this dissertation is to examine and then to explore the potential of low-grade data center waste heat to activate an absorption chiller. This primary objective can be categorized as follows:

- Conceptualizing an overall system design
- Modelling the central system under steady state conditions
- Applying and testing a heat capturing scheme
- Establishing an exergetic efficiency of the central system

1.3. Literature Review

[13] has investigated and simulated an alternative cooling methodology in data centers. Their approach centers on directly cooling the chips and auxiliary electronics on server motherboards. They conducted a performance simulation of their proposed hybrid two-phase cooling system with microevaporator elements and simulated five cases with three different working fluids of HFC134a, HFO1234ze and water. They also investigated maximization of exergetic efficiency and specifically conducted an exergy analysis of the energy recovery in the condenser.

An exergy analysis of the cooling cycles, regarding the potential of exergy recovery at the condenser, showed an overall exergetic efficiency of lower than 50%. The researchers then combined their on-chip cooling implementation with waste heat recovery to examine the potential to reduce energy usage in data centers. The results showed that a data center's energy use can be cut by as much as 50% when using a liquid pumping cycle and 41% when using a vapor compression cycle.

[14] adopts a new approach toward optimizing energy flows in data centers. They differ from traditional workflow modeling centers by integrating all three main subsystems of data centers: cooling, power and IT. One of the primary ways they achieve this is through scheduling IT workloads. There are three stages in their process of improving datacenter efficiency: 1) predicting IT and renewable energy demand 2) incorporating those predictions into a system model and 3) testing the model on a scaled-down data center comprised of four server blades. The results from testing this prototype showed that gains in datacenter efficiency are both significant and attainable. In fact, they found that their approach can lead to a reduction in power costs and non-renewable

energy usage of up to 60%. They conclude that their model can be extrapolated to full-scale data centers in geographically diverse areas.

At the server board level, [15] captured run-time power dissipation and thermal characteristics of a 2003 microprocessor immersed in mineral oil. They used infrared cameras with high spatial resolution and high frame rate to create thermal images overlayed on a CPU floorplan block. The mappings showed that the hottest spots occurred in the memory, clock, fetch and execute areas of the block. This information is useful to the scientific community because it breaks down the power dissipation of modern processors into its leakage and dynamic components based on thermal measurements. The measurement setup can help to improve energy and thermal simulation, such as including metal density, to optimize and direct data center cooling.

[16] develops a data center power demand shaping (PDS) technique that enables onsite distributed generation (DG) to provide clean renewable energy to the computing load without compromising workload performance. A DG system refers to a variety of modular, linked renewable energy sources near the point of use. Such electricity generators can be composed of any green energy technology, such as photovoltaics, wind power, fuel cells, or bio-fuel based gas turbines. However, only the outputs of fuel cells and gas turbines are tunable and therefore capable of load following, a key advantage of their PDS design. Load following refers to enabling onsite green energy generation equipment to track the time-varying IT load rather than constraining the IT load to respond to a variable power budget. Thus, they present a fundamentally different and critically important design, which leverages renewable DG to allow near-ideal performance of computing systems. Approaching within 1.2% of the performance of an

ideal power provisioning scheme that can always match the fluctuating IT power demand with no performance overhead, their PDS design applied to DG data centers can realize a savings of over 100 metric tons of carbon emissions annually for a 10MW data center.

[17] responds to the increasing demands of data centers by proposing a tenfold increase in performance per watt of IT operations. They claim that the infrastructure operations now require nearly the same energy as IT operations, and they advocate a three-pronged approach to increasing energy efficiency. First, they propose integrating renewable energy sources wherein the suggested method is to integrate battery systems to store excess energy when it is available and to dispatch the energy when necessary. In addition, using and integrating advanced weather forecasting technologies could enable data centers that use solar and wind energy to more effectively manage and schedule electric load operations.

The second approach the authors advocate is to achieve workload optimization through three steps: 1) use virtualization to reduce the number of servers 2) use an integrated approach for server consolidation 3) match IT workload allocations to the availability of renewable energy sources. Finally, they offer a third method for improving data center energy efficiency. The installation of wireless sensors in a data center to monitor the environment and regulate cooling could translate into as much as a 10% reduction in the data center's overall energy usage.

The studies in this literature review contribute valuable ideas to data center sustainability. However, there are areas or applications they do not address that the work discussed in this dissertation focuses on and implements.

Unlike this dissertation, [13] only simulates direct cooling rather than directly experimenting on the chips themselves. In addition, [13] proposes selling the waste heat to external power plants. This approach does not fully realize the benefits of heat extraction for the data center itself whereas my approach is to recycle the extracted highest quality waste heat within the data center to then cool down the center.

While [14] proposes cooling down data centers through the extraction of waste heat, it does not suggest recycling that waste heat and putting it to work as part of the cooling process. Furthermore, it relies on outside air as a cooling medium. However, this method may not be effective in warm weather climates. My application not only reprocesses the data center waste heat to further cool the data center but also can be effective regardless of ambient conditions outside of the data center.

Although [15] conducts sophisticated measurements of temperature using IR cameras with special filters, they do not measure thermal power directly and do not take into account CPU utilization levels as does my work. In addition to temperature and thermal power monitoring, my experiments simultaneously measure real-time responses of the processor subject to varying levels of CPU loading, an essential element that directly affects temperature and power demand of the processor. Furthermore, taking CPU load into consideration is important because it affects thermal power dissipation and the resulting quality of the heat. Finally, my paper utilizes the most accurate way of gauging a processor's temperature and power simply by using the datasheet values standard for each processor.

[16] is missing on-site absorption chillers amongst its repertoire of renewable energy generators. The small absorption chiller that we employ would be an added

advantage for [16] in two ways: 1) an absorption machine can energize itself with data center waste heat, thus recycling it, and 2) it is auto-tunable to the load so that as there is more power demand and, thus, more heat dissipated and available, there can be more cooling capacity provided.

[17] is theoretical in nature, basing predictions on current trends and projections. In contrast, my work goes a step further by developing a mathematical model and testing one of its components through a small-scale experiment. Also, [17] predicts an equivalent Power Usage Effectiveness (PUE) of 1.1 (90% efficiency) whereas my mathematical model demonstrates that a PUE of less than 1 can be achieved.

1.4. Organization of this Disseration

The rest of this paper is organized as follows. Chapter 2 introduces the initial objective of decreasing grid consumption by easing the cooling load on the typical vapor compression air conditioning unit in a data center. The novel approach of engaging a single-effect lithium bromide-water (Li-Br) absorption refrigeration system to utilize data center waste heat is discussed. Finally, the unique idea of using high quality data center waste heat, specifically the high quality heat dissipated from the CPUs on each server blade, is analyzed for its feasibility to energize the chiller and for its effect on data center performance. Chapter 3 answers questions about how such waste heat flows could be captured by describing the results of experiments on waste heat recovery from CPUs, and reporting the resulting relationship among temperature, power, and CPU utilization levels. Specifically, four liquid-cooling setups and associated heat-extraction experiments

are described with the outcomes of each detailed, including a radical approach using mineral oil.

Chapter 4 introduces the concept of exergy, information exergy, and exergetic efficiency. This chapter explores the exergy of a processor and determines the maximum amount of energy utilizable for work. Exergy as a source of realizable work is separated into its two contributing constituents: thermal exergy and informational exergy. The informational exergy is that usable form of work contained within the most fundamental unit of information output by a switching device within a CPU. Both thermal exergy and informational exergy values are calculated and plotted for our particular CPU as well as the exergetic efficiency for datasheet standards compared with experimental values.

Finally, chapter 5 discusses the significance of the work, offers thoughts for improvements and further exploration and concludes the dissertation.

CHAPTER 2

THERMODYNAMIC FEASIBILITY OF HARVESTING DATA CENTER WASTE HEAT TO DRIVE AN ABSORPTION CHILLER

2.1. Primary Objective and Approach

Computer room air conditioners (CRACs) are the standard for cooling data centers, but these electrically-driven vapor-compression units are responsible for consuming up to 50% of power that a typical data center uses [11]. Since cooling is such a significant portion of data center grid power consumption, the primary objective is to reduce the grid power consumption of the cooling system for a data center (DC).

The primary approach is to employ an absorption refrigeration unit that is energized by thermal energy to provide cooling [12]. Such an approach satisfies the primary objective by reducing DC power consumption in several ways. First, since DC waste heat will be used to drive the chiller, this reduces the amount of cooling power needed to cool that otherwise dissipated waste heat. At the same time, the absorption system provides additional cooling which lessens the load on the CRAC unit.

2.2. Absorption Chiller and Performance Models

As shown in Figure 1, the unit proposed is a 10-ton single-effect lithium bromide-water (Li-Br) absorption refrigeration system to utilize data center waste heat and to reduce the cooling load on the CRAC. The rationale behind choosing this specific model is its availability as a donation from the local utility company. This heat-driven system uses lithium bromide salt as the absorbent and water as the refrigerant.

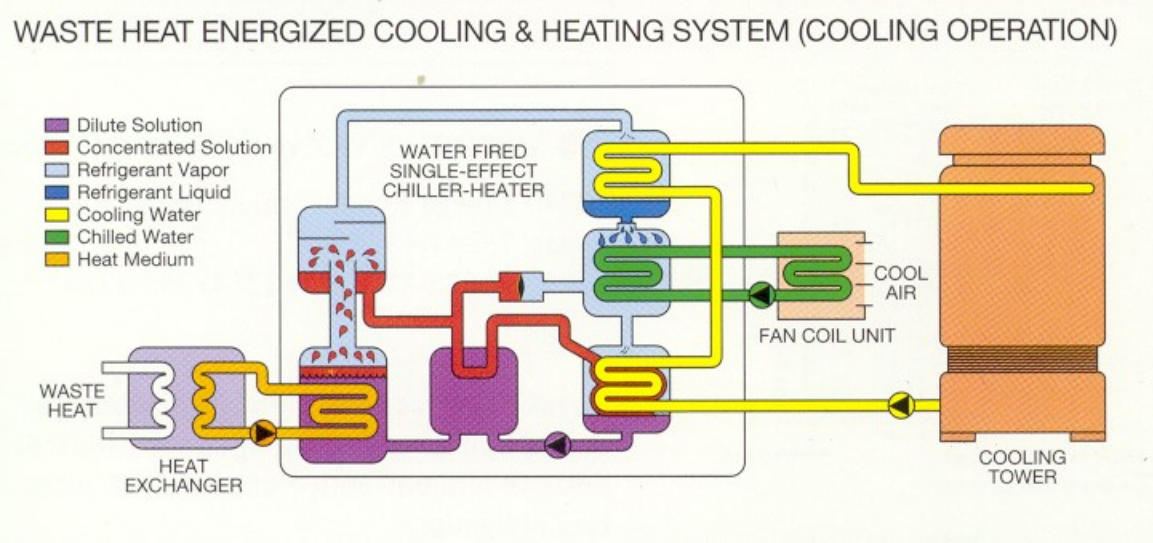

Figure 1. Model WFC-SC10 cooling cycle with permission from Yazaki Energy Systems, Inc. [18].

According to standard specifications (Table 1), at its design point of 88 °C and 50.2 kW_{th} of heat input and a pump power input of 0.21 kW_e the cooling capacity is 35.2 kW_{th} and its coefficient of performance (COP_C) is 0.7. Here, kW_e is taken to mean kilowatt electrical or electric power whereas kW_{th} is kilowatt thermal or heat power. Below the standard specifications, the unit can be energized at lower values of temperature and heat input, favorable to fluctuating output values of CPUs.

Table 1

Specifications in SI Uunits for the Yazaki Water Fired SC10 Chiller under Standard Conditions [18].

Model	WFC-SC10	Standard
Cooling	Capacity (kW_{th})	35.2
	Chilled Water Temp. (°C)	7 Outlet, 12.5 Inlet
Chilled Water	Rated Water Flow (*l*/sec)	1.5
	Evap. Press Drop (kPa)	55.8
	Water Retention Volume (*l*)	17.0
	Heat Rejection (kW_{th})	123.9
Cooling Water	Inlet Temperature (°C)	31 (Standard)
	*Rated Water Flow (*l*/sec)	5.1
	Cond./Abs. Press. Drop (kPa)	84.8
	Water Retention Volume (*l*)	65.9
Heat Medium (HM)	Input (kW_{th})	50.2
	Inlet Temperature (°C) 88 (Standard)	Temperature Range 70 (min) - 95 (max)
	Rated Water Flow (*l*/sec)	2.4
	Generator Press. Drop (kPa)	90.3
	Water Retention Volume (*l*)	20.8
Electrical	Power Supply	208V, 60Hz, 3 ph
	Consumption (W_e)	210

* Minimum cooling water flow

To explore the range of possible efficiencies and cooling capacities within the rated conditions, Figure 2 is devised from Table 1 and a simulator tool provided by Yazaki Energy Systems, Incorporated. Also, the graphs of Figure 2 and Figure 3 help to determine if the CPUs can meet each set of temperatures and heat input requirements to successfully operate the chiller.

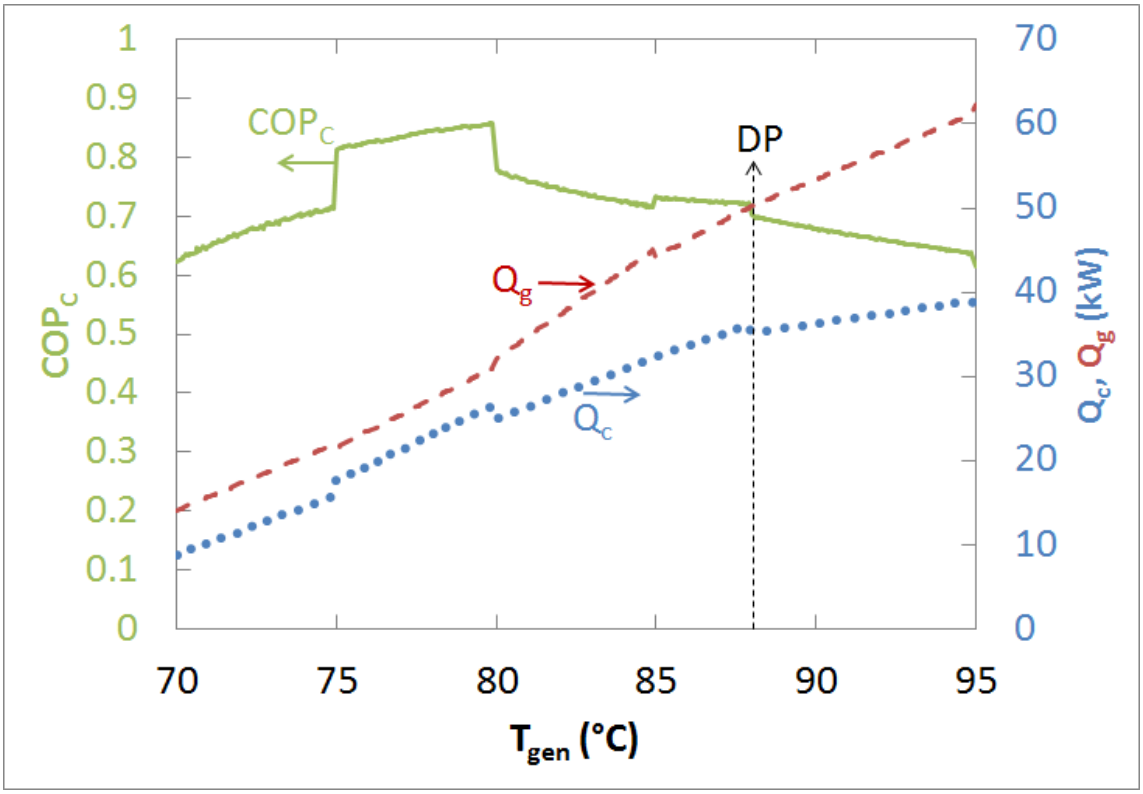

Figure 2. WFC-SC10 performance under standard conditions. DP denotes its design point of $COP_C = 0.7$ and cooling capacity 35.2 kW_{th} at 88 °C, requiring heat input of 50.2 kW_{th}.

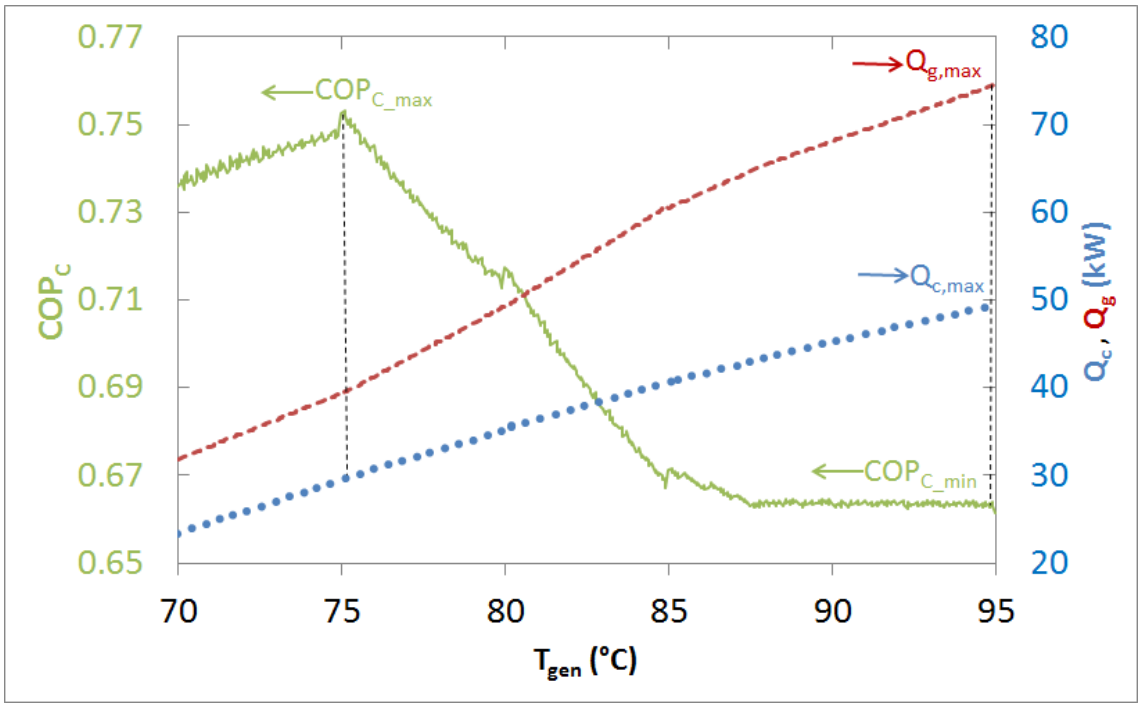

Figure 3. WFC-SC10 performance at maximum conditions, i.e., cooling water temperature lowered to 26.7 °C.

The absorption chiller performance depends on many factors including the cooling water inlet temperature, the heat medium flow rate, and the temperature of the input heat, which determine the resulting cooling capacity. Figure 2 shows the Yazaki WFC-SC10 absorption chiller's performance under standard conditions defined as a heat medium flow of 100% or 2.4 l/s and cooling water temperature of 31 °C. To energize the chiller at this level requires a minimum heat input of 14.1 kW_{th} at 70 °C for a cooling capacity of 8.8 kW_{th} (COP_C = 0.62). Although 8.8 kW_{th} provides low realizable cooling under these conditions, it does demonstrate that an absorption machine can be energized using a minimal amount of waste heat and it will be shown that the CPUs can easily reach these possible requirements.

Notice that cooling capacity significantly increases to 35.2 kW_{th} at its design point of COP_C = 0.7 and 88° C. Heat medium flow remains at 100%. Also notice in Figure 2 that the COP_C reaches a maximum value of 0.86 at lower temperatures. This value corresponds to the least amount of heat input necessary to achieve the highest efficiency. The tradeoff is a lowered cooling capacity of 25.0 kW_{th} at 80 °C. However, this point may be the best compromise to achieving a reasonable cooling capacity for a temperature at which the CPUs can safely operate.

Under maximum conditions, we can achieve a maximum cooling capacity above the WFC-SC10 chiller's rated value. However, as shown in Figure 3, there is a distinct tradeoff between maximum efficiency and maximum cooling capacity. For a cooling water temperature lowered to 26.7 °C, a maximum COPc of 0.753 occurs at 75.1 °C but only delivers a cooling capacity of 29.6 kW_{th}. Compare this cooling capacity with the maximum possible of 49.4 kW_{th} at a minimum COPc of 0.66. In general then, the hotter the heat medium and the lower the cooling water temperatures, the more cooling capacity will be generated. Also, the greater the heat medium flow rate, the more cooling capacity will be generated.

As expected, the COP_C is low at the lowest inlet temperature, but it then peaks in the middle and considerably falls to a minimum value as more input temperature is necessary to generate higher cooling capacities. The ratios of the surface areas of the various tube bundles is such that it is more efficient at releasing refrigerant with the lower temperature up to its design point limit. The COP_C climbs in the middle higher than the design point due to the surface area of the tube bundles. The ratios of surface area of the various tube bundles is such that it is much more efficient at releasing refrigerant with the

lower temperature, but the lower temperature itself cannot release as much total refrigerant so the capacity falls below design capacity. In essence, the system is better at releasing refrigerant with the amount of heat it delivers, but it delivers significantly less total heat. The heat medium generator is made of a particular grade of stainless steel tubing wound into three parallel-flow rows, meaning an inside, middle, and outside tubing coil. At the bottom of the tube bundle, all three of the coils tie into a header. Each coil is wound in a circular pattern and all three are rejoined together at the top of the tube bundle, rejoining in the same configuration. The tube on the outside at the bottom header winds around to be the outside coil in the generator and rejoins the top header at the outside position. The different diameters of each of these bundles means there will be slight differences in how each of these individual coils transfer heat.

At higher temperatures, there is so much input heat that transfer is again limited by the surface area and heat transfer properties of the heat exchange materials and COP_C falls. Then, at the cost of lowered efficiency, the cooling capacity rises to a maximum under each set of environmental conditions. This is advantageous when waste heat is recycled thermal energy; when the heat source is ‘‘free,’’ then capacity is the main consideration.

2.3. CPU Waste Heat Potential

The energy sources to run the absorption chiller will be the waste heat dissipated from the IT equipment inside the data center as well as any external thermal supplementation as necessary [19]. Therefore, the benefits realized will be twofold and interconnected: removing a main source of data center waste heat and, consequently, lessening the load on the CRAC. The data center waste heat to drive the chiller will be

originated from the main heat producing components (processors) on each of the computer server blades. The server blades of choice are the Dell PowerEdge 1855 blades and the processors are two Intel Xeon Nocona CPUs on each blade.

The main challenge is capturing enough high-temperature heat from the processors on each server blade, and then transporting that heat effectively and efficiently to power a Li–Br absorption chiller [20]. Since the energy sources to run the chiller will be the CPUs, the performance graphs in Figure 2 and Figure 3 set the temperature and heat requirements that the CPUs should deliver. From benchtop tests, we discovered that the Xeon CPU can withstand more than the required temperature range and multiple processors can deliver the necessary heat to energize the absorption machine.

2.4. CPU Temperature Data

After installing Dell's Open Manage System Administrator (OMSA) software, we discovered that each processor can operate at temperatures ranging from 10.0 °C to 120.0 °C. The maximum possible temperature is 125.0 °C as shown in Figure 4. According to OMSA, failure would not occur until 125.0 °C. Therefore, the required heat medium temperature range of 70–95 °C for chiller operation is well within the temperature range of the processors' capabilities.

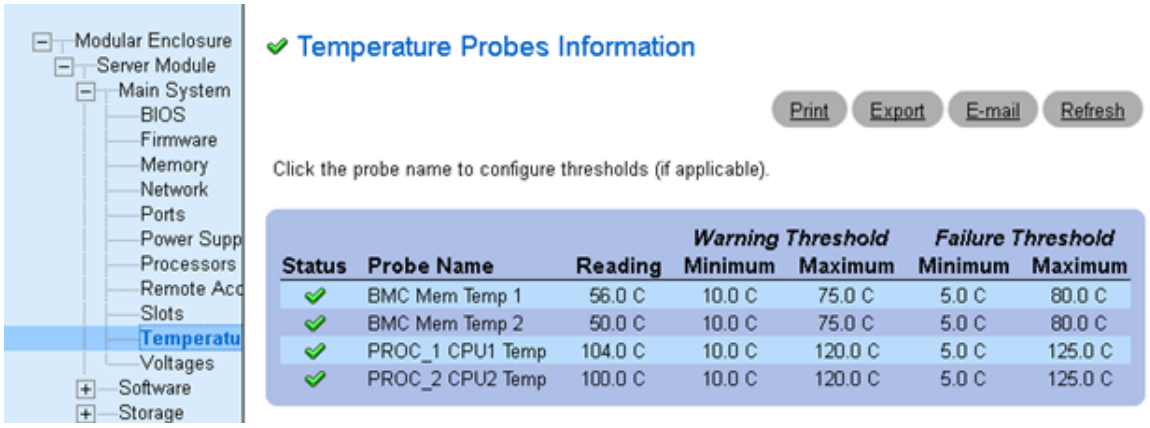

Figure 4. OMSA indicates CPU's capability of reaching up to 125.0 °C.

We tested the CPU temperatures by tasking the CPUs at different levels and modulating the fans. Figure 4 shows the reported CPU temperature at a high tasking level and significantly reduced fan air flow (all but one of the chassis fans removed). Shutting off the fans yielded temperatures above 100 °C and allowed for conservation of power and reduced noise levels. Also, with liquid cooling proposed, fans would not be used and adjusting coolant flow would be the primary modulation method.

To make certain the OMSA readings were reporting accurately, we further tested the CPU by employing a National Instruments FieldPoint data acquisition device and Labview to measure the response from a validation thermocouple affixed directly onto the CPU case. The results over a 5-min test run are shown in Figure 5.

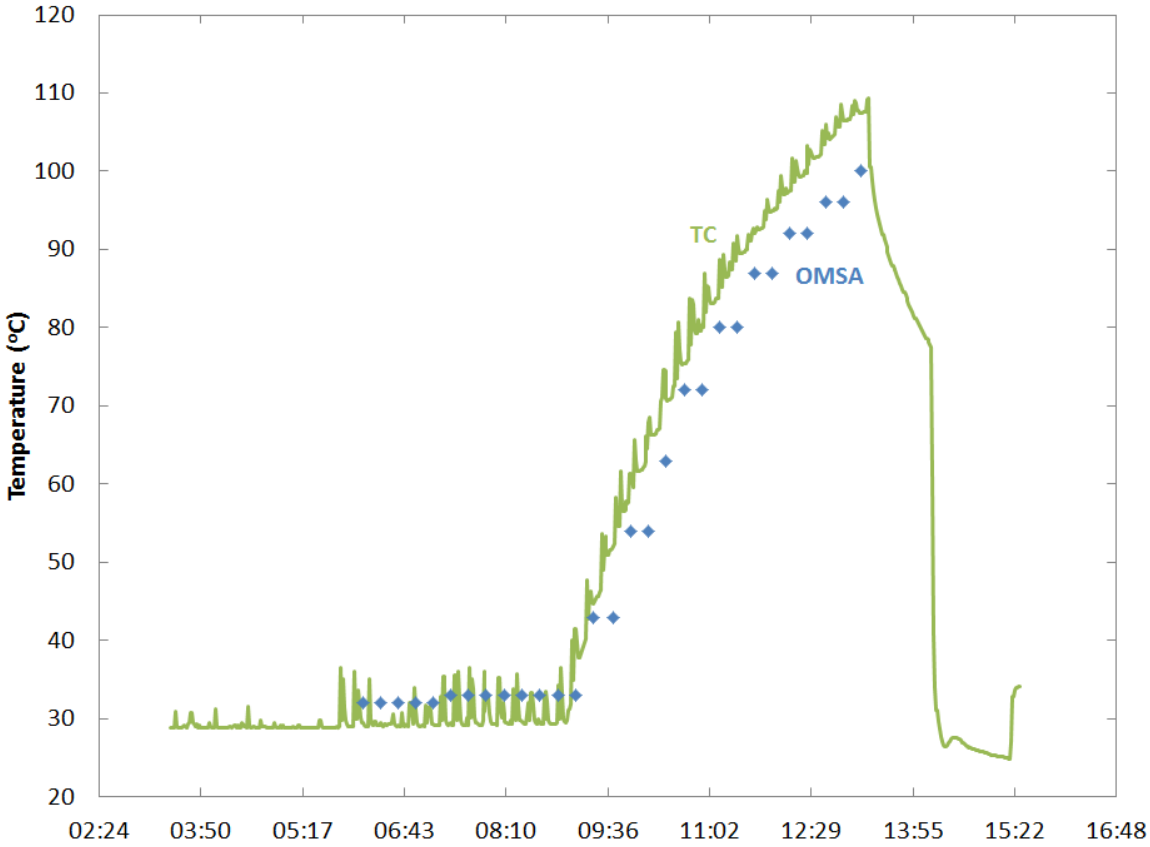

Figure 5. CPUs can operate within required chiller temperature range and beyond.

Figure 5 is a validation test of the OMSA software using a K-type thermocouple (TC) until we reached a target temperature of 110 °C. The Open Manage System Administrator (OMSA) software response lags behind the thermocouple (TC) but shows consistency.

From Table 1, the useable range for chiller operation is between 70 °C and 95 °C. Figure 4and Figure 5 agree that the CPUs are able to operate within the required chiller temperature range. Furthermore, the CPU temperatures can be reasonably stabilized by

modulating the fans as shown in Figure 6. To observe the smaller temperature resolution in Figure 6, we used the TC and Labview setup.

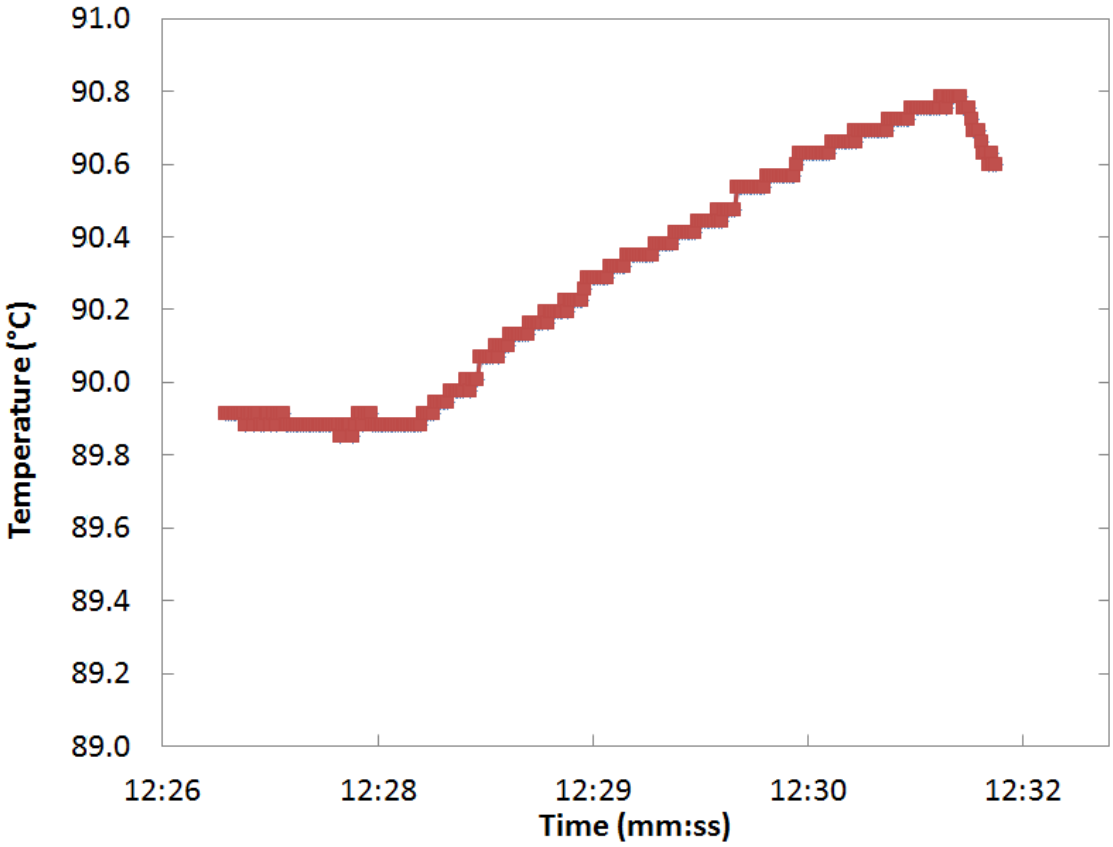

Figure 6. CPU stabilized to an average temperature of 90.2 °C over a 5-min test run.

Taking into consideration a ΔT_{CH} loss between the CPU case and a cold plate interface, we use the general principle that the temperature drop ΔT across a given absolute thermal resistance R_θ is the product of said resistance and a given heat flow Q through it [21]. Applying this product to the application, $\Delta T_{CH} = Q \cdot R_{\theta CH}$, where Q is the

average maximum heat dissipation of the CPU (103 W) and $R_{\theta CH} = 0.1$ K/W (a typical value for a heat-transfer pad) the expected ΔT_{CH} loss is then 10.3 °C [21]. Doubling this value, we arrive at an overall estimated system ΔT loss of about 20 °C.

If we assume a 90 °C level of temperature as shown in Figure 6, a maximum system ΔT loss of 20 °C between the CPU and the heat medium inlet to the chiller suggests this level of CPU temperature would deliver the minimum T_{gen} of 70 °C to energize the Yazaki WFC-SC10 absorption chiller. This ΔT loss of 20 °C is a worst case scenario and the actual temperature loss is expected to be much lower, around half of that value. Thus, a predicted transference ΔT loss of 10 °C between the CPU and the chiller generator input infers a delivery T_{gen} of 80 °C for a 184% increase in cooling capacity (under chiller standard conditions) over the worst case scenario. Although, higher CPU temperatures can be achieved by reducing or eliminating the fans and by tasking the CPU, we plan to run the CPUs within their standard operating range of 70–90 °C in order to ensure that no reliability issues should occur. If necessary, additional external heat supplementation from a sustainable energy source such as solar thermal can contribute to increasing the temperature of the required input heat [22].

2.5. CPU Power

Since nearly 100% of CPU electrical power is dissipated as heat and to determine if we could meet the chiller's heat input requirements, we investigated the thermal design point (TDP) and maximum theoretical power possible. According to Intel specifications, each Intel Xeon Nocona CPU has a TDP of 103 W [23]. While TDP is used as a target point for design solutions, both Intel and AMD agree that TDP is not the maximum

power the CPU may draw or generate as heat but rather the average maximum power that it draws when running typical user applications [24]. There may be periods of time when the CPU dissipates more power than designed, such as during strenuous demands by engineering or scientific applications, such as computer aided design simulations. Under such loads, the CPU temperature will rise closer to the maximum and the CPU power will surpass that of the TDP. TDP is primarily used as a guideline for manufacturers of thermal solutions (heatsinks, fans, etc.) [25]. This guideline ensures the computer will be able to handle essentially all applications without exceeding its thermal envelope, or requiring a cooling system for the maximum theoretical power (which would cost more but in favor of extra headroom for processing power) [24].

Maximum power dissipation is always higher than TDP [26]. Maximum power dissipation by the CPU occurs at the maximum core voltage, maximum temperature and maximum signal loading conditions that the system can handle. Sustainable CPU maximum power dissipation is usually 20–30% higher than the rated TDP value [25] but can be much higher. With the TDP rating of each CPU specified by Intel to be 103 W, this would equate to a maximum possible power dissipation of between 123 W and 144 W per CPU. Referring to Figure 3 and assuming no temperature difference or heat loss, the maximum chiller capacity ‘‘ideally’’ could be reached by using 74.8 kW/0.144 kW = 520 CPUs (rounding up). However, this approach is more conservative as it takes into account a worst case temperature fall of 20 °C and a system heat loss of 15% [11].

According to the Dell Datacenter Capacity Planner v3.04 configuration tool, the system heat is 294.0 W for each PowerEdge 1855 Blade [27]. With two CPUs per blade, we can assume an average maximum power of 206 W. This equates to exactly 70% of the

system board being dominated by the heat coming from the CPUs. This 0.7 value is hence denoted the highest heat fraction or HHF of the system board.

In contrast, the capture fraction (CF) refers to the maximum heat that we estimate can be recovered. With a liquid cooling system in place, we estimate that we can recover 85% of the HHF from each CPU or 59% off the entire system board. This estimate is based on a similar cold plate system currently in use by IBM [11]. Similar to IBM, this heat extraction system will house a heat transfer fluid of water but the extracted heat will be used for cooling rather than for heating. Since the system heat is rated to be 294.0 W for each PowerEdge 1855 blade, the bottom line is that we anticipate being able to capture at minimum $(0.59)\cdot(294\ \text{W}) = 173\text{W}_{th}$ per blade.

Taking the most conservative value in the chiller standard performance graph of Figure 2, the chiller must be energized with a minimum of 14.1 kW_{th} at 70 °C. This conservative stance would call for 14.1/0.173 = 82 server blades and there are more than twice this amount in planned for installation at the ASU data center. For a higher efficiency scenario of $\text{COP}_C = 0.86$ at $T_{gen} = 80$ °C in Figure 2, we would need 169 server blades to deliver 25.0 kW_{th} worth of cooling. Finally, at the design point of 88 °C, the heat medium requirement is 50.2 kW_{th} or 291 server blades. Since there will be access to more than 300 server blades at the ASU BlueCenter data center and since the CPUs can be run as high as 120 °C (Figure 4), all three scenarios can be theoretically accommodated.

2.6. Conceptual System Design for System Analysis

Figure 7 illustrates the heat flow from the data center to various cooling components.

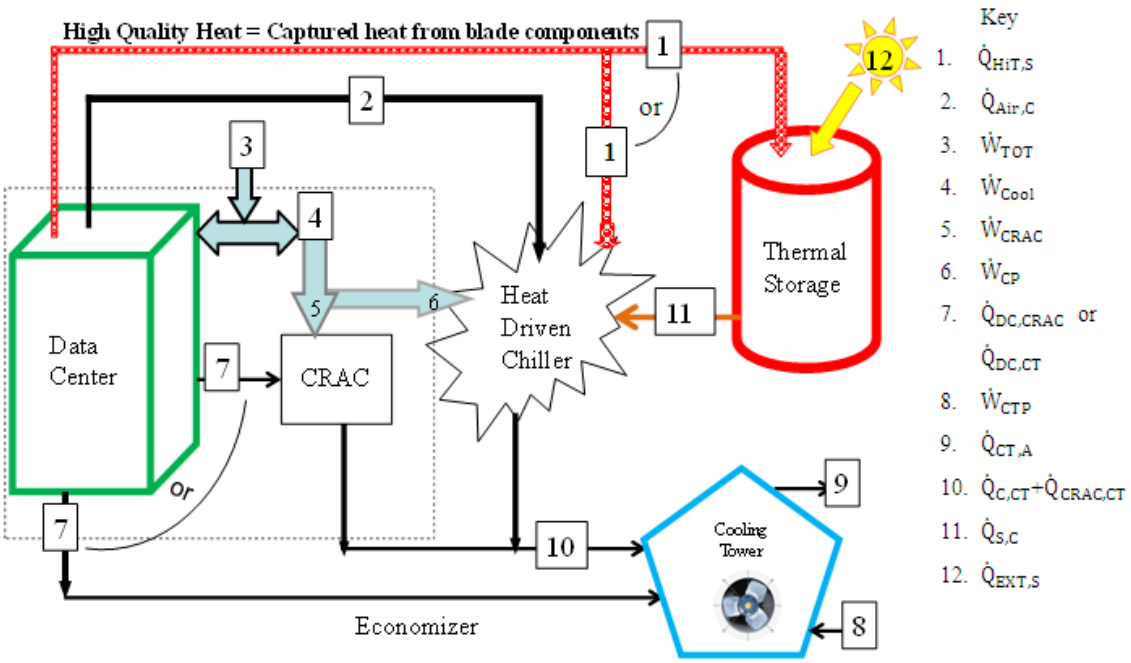

Figure 7. Overall system level diagram showing the work and heat flow paths.

High-temperature, high-quality heat, denoted as $\dot{Q}_{HiT,S}$ in Figure 7, is dissipated by approximately 70% of the components on a server blade inside a standard 42U (200 cm tall) rack. To clarify, each cabinet-style rack can house up to six chassis and each chassis box can hold up to ten server blades [27]. Each rack has a footprint of 0.6 m^2 [27]. To leave plenty of room for the heat extraction equipment as well as switchgear and any PDUs, there can be four chassis per rack equivalent to 80 CPUs per rack. The plan is to have two rows of two racks or 640 CPUs.

As illustrated by lines 1 and 11 in Figure 7, the captured high quality heat $\dot{Q}_{HiT,S}$ from each server blade can be ‘‘piped’’ directly to the Li–Br absorption chiller or, along

with the auxiliary heat source, can be sent to a phase change thermal storage unit for later use [28]. Due to its ability to provide a high energy storage density at a nearly constant temperature, this latent heat storage technique holds the key to the provision of energy when the intermittent sources of energy are not available [29].

A large number of organic and inorganic substances are potential PCMs, based on their melting points and heats of transition. Paraffins, some plastics and acids such as stearic acid are the most practical compounds for thermal storage systems because they are relatively inexpensive, tend to have broad melting ranges, and their properties do not change over time [30]. The main disadvantages of organic PCMs include low thermal conductivity, low durability and high volume change during melting [30]. An ideal PCM may be a combination of organic and inorganic substances [29]. Another idea is to use high thermal conductivity materials within the PCM to increase its apparent thermal conductivity. For example, using finned tubes in which the PCM is placed between the fins can significantly improve heat transfer rates for liquid-based systems [30].

Once enough heat has been supplied to the chiller, it can function to reduce the heat load on the CRAC. As shown by line 2 in Figure 7, a portion of the heated air in the data center, denoted as $\dot{Q}_{\text{Air,C}}$, can be removed by the absorption chiller, which lessens the workload on the CRAC. This ambient heat inside the data center is lower temperature, lower quality heat relative to the higher temperature, higher quality heat of the computing components inside each server blade. This lower temperature air can be cooled by the absorption chiller. Any remaining heat, as illustrated by line 7 within Figure 7, is removed by the CRAC or by the Economizer and rejected to the cooling tower.

One of the most challenging aspects of the proposed system is transferring the thermal energy from the blade components to the generator of the Li–Br chiller. Our approach is for the CPUs to be cold-plated, i.e., in thermal contact with water-cooled heat sinks. The cold plates contain water as the heat-exchange fluid which can then transfer a fraction of high-temperature blade heat to thermal storage to run the chiller. A capture fraction (CF) of 0.85 of this HHF (0.70) is expected to be recovered or about 174 W_{th} per server blade.

The advantages of using water cooling over air cooling include water's higher specific heat capacity, density, and thermal conductivity [31]. For example, the thermal conductivity of liquid water is about 25 times that of air. Even more significant, water is approximately 800 times denser than air, and this greater density allows water to transmit heat over greater distances with much less volumetric flow and reduced temperature difference than that of air. For cooling CPU cores, water's primary advantage is its increased ability to transport heat. Furthermore, since the Yazaki absorption system also runs on water, more efficient water-to-water heat exchangers are used. Also, a smaller flowpath means less area to insulate and less pump power work as compared to using air transport and fans.

2.7. PUE: Power usage effectiveness metric

Besides creatively re-using data center waste heat and reducing the power grid consumption of the CRAC, an additional benefit of applying the system level design of Figure 7 is an improvement in data center efficiency. The most widely accepted benchmarking standard to gauge data center efficiency is power usage effectiveness (PUE). Introduced by The Green Grid in 2006, PUE is commonly used to gauge the total

amount of power consumed by IT equipment relative to total facility power [32]. There is ambiguity as to whether this power is from only non-renewable source or can be from renewable ones as well. In standard practice, reused energy is not factored into the PUE. Nevertheless, as shown in the following carefully worked-out series of equations, the conventional PUE equation can be made to accept waste heat reuse by subtracting this quantity from total power in the numerator.

As shown in Equation (1), PUE is conventionally defined as the ratio of power in to the data center, measured at the utility meter, to the power required to run the IT equipment for computing [32]:

$$\text{PUE} = \frac{\text{Electrical Utility Power}}{\text{IT Equipment Electric Power}} = \frac{\dot{W}_{TOT}}{\dot{W}_{IT}} \quad (1).$$

For example, a PUE of 2.0 means that for every 1 W of electric input to the IT equipment, an additional 1 W is consumed to cool and distribute power to the IT equipment. Thus, the intent of the PUE metric is to characterize the fraction of the total data center power consumption devoted to IT equipment. It compares how much power is devoted to driving the actual computing IT components versus the ancillary support elements such as cooling and PDUs (power distribution units). With PUE, the focus is on maximizing the power devoted to the equipment running applications and minimizing the power consumed by support functions like cooling and power distribution [32]. Thus, the lower the PUE, the more efficient is the data center.

As PUE is defined with respect to the "power at the utility meter," it leaves room to re-use on-site energy generation from waste heat and to use external alternative energy

sources such as solar thermal energy. These previously unaccounted-for energy sources make it possible to lower the PUE to a number below one. That is, the facility power use can actually be a negative number. This is an intriguing prospect as the data center would be generating energy, or put another way, producing exportable power. Therefore, as is suggested by the following steady-state analysis, a very efficient PUE is possible by re-using data center waste heat and a PUE <1 is possible with an external supplemental source of alternative energy.

2.8. Steady-state analysis

The total facility power or total data center electric power consumption ($\dot{W}_{TOT}$) is comprised of three major categories: IT computer rack power ($\dot{W}_{IT}$), power delivery loss from the Power Distribution Units (PDUs) to the racks ($\dot{W}_{Loss}$), and the electric power consumed to operate the data center cooling system ($\dot{W}_{Cool}$). Since we are choosing a "lights out data center," electrical power to the lights ($\dot{W}_{Lights}$) will be off during normal operation and can be neglected compared to the other three terms [33]. Thus, the numerator in equation (1) can be expanded as

$$\dot{W}_{TOT} = \dot{W}_{IT}+\dot{W}_{Loss}+\dot{W}_{Cool} \tag{2}$$

where

$$\dot{W}_{Cool}= \dot{W}_{CRAC}+\dot{W}_{AC\_FANS}+\dot{W}_{CP}. \tag{3}$$

Also, the electric power driving the cooling system ($\dot{W}_{Cool}$) is itself composed of three main terms: the power to run the vapor compression computer room air conditioner

($\dot{W}_{CRAC}$), whose electrical power requirements are determined by the demands of the data center at any given moment; the power to run the CRAC supply and condenser fans ($\dot{W}_{AC\_FANS}$) and any electrical power needed to run the heat-driven chiller ($\dot{W}_{CP}$). The CRAC of choice is an 88 kW_{th} (25-ton) unit which utilizes a supply fan at 5 hp and two condenser fans at 1 hp each for a total $\dot{W}_{AC\_FANS}$ of 7 hp or 5.22 kW_e [34]. The chiller power $\dot{W}_{CP}$ is broken down into the solution pump and condenser fans. Although the electrical power consumption by the Li-Br pump is negligible compared to the power consumed by the vapor compressor of the CRAC, the condenser fans cannot be neglected [35]. Compared to a compression system, the cooling tower will be between 1.5 and 2.5 times larger on an absorption system, which translates into more condenser fan energy at an increase of about 1.25 hp for the fan motor. For example, a 10-ton cooling tower for a vapor compression system may use a ¾ hp fan motor, but a 10-ton WFC-S10 Yazaki LiBr absorber requires a 25-ton cooling tower. As it turns out, a 25-ton cooling tower uses 1-hp fan motor or 0.7456 kW_e. For the absorption system, this increase of electrical consumption equates to about 25% above what the compression system uses.

From equations (1) – (3), the PUE can now be expressed as

$$\mathrm{PUE} = \frac{\dot{W}_{TOT}}{\dot{W}_{IT}} = \frac{\dot{W}_{IT}+\dot{W}_{CRAC}+\dot{W}_{AC\_FANS}+\dot{W}_{CP}+\dot{W}_{LOSS}}{\dot{W}_{IT}} \tag{4}$$

Then, to relate the electric power supplying the CRAC compressor, $\dot{W}_{CRAC}$, to the heat load on the CRAC, $\dot{Q}_{CRAC}$, we can make use of the standard vapor compression COP (coefficient of performance) equation:

$$\dot{W}_{CRAC} = \frac{\dot{Q}_{CRAC}}{COP_{CRAC}} \quad (5).$$

The $\dot{Q}_{TOT}$ total heat load on the CRAC can be reduced by any supplemental chiller cooling capacity, $\dot{Q}_{C,}$

$$\dot{Q}_{CRAC} = (\dot{Q}_{TOT} - \dot{Q}_{C}). \quad (6)$$

Substituting equation (6) into equation (5) and then into equation (4), equation (4) now becomes

$$PUE = \frac{\dot{W}_{IT} + \frac{(\dot{Q}_{TOT} - \dot{Q}_{C})}{COP_{CRAC}} + \dot{W}_{AC\_FANS} + \dot{W}_{CP} + \dot{W}_{Loss}}{\dot{W}_{IT}}, \quad (7)$$

where $\dot{Q}_{TOT}$ is the total heat flow from the data center to be removed by the cooling equipment. A portion of this total heat flow is removed right away by the heat extraction equipment on each server blade, and is denoted as $\dot{Q}_{HiT,S}$:

$$\dot{Q}_{TOT} = \dot{Q}_{IT} - \dot{Q}_{HiT,S} + \dot{Q}_{Loss} \quad (8)$$

However, since the heat output of an electronic device in steady operation is approximately equal to the power input to the device, we can express equation (8) in terms of power as

$$\dot{Q}_{TOT} = \dot{W}_{IT} - \left[\dot{W}_{IT} \cdot HHF \cdot CF\right] + \dot{W}_{Loss} \tag{9}$$

in which $\dot{Q}_{HiT,S} = [\dot{W}_{IT} \cdot HHF \cdot CF]$ and is the portion of rack heat removed from the racks by the heat recovery scheme and sent to the chiller.

As previously stated, *HHF* = 70% of each server blade provides the highest quality heat. An estimated *CF=85%* of *HHF* or 59% can be captured and transferred from the blades to thermal storage to drive the LiBr absorption chiller. With a high enough *CF*, a significant portion of this high-temperature heat can be removed from the original heat load on the CRAC and utilized to drive the chiller.

To continue with the rest of equation (6), the cooling provided by the absorption chiller, $\dot{Q}_C$, is also driven by the heat recovered from the server blades, $\dot{Q}_{HiT,S}$, along with any supplemental external heating, $\dot{Q}_{EXT,S}$:

$$\dot{Q}_C = \mathrm{COP}_C(\dot{Q}_{HiT,S} + \dot{Q}_{EXT,S}), \tag{10}$$

And, since $\dot{Q}_{HiT,S} = [\dot{W}_{IT}\, HHF \cdot CF\,]$,

$$\dot{Q}_C = \mathrm{COP_C}(\dot{W}_{IT} \cdot HHF \cdot CF + \dot{Q}_{EXT,S}) \quad (11)$$

Substituting equations (9) and (11) into equation (7) and simplifying, yields the final general expression for a PUE taking into account parasitic power, recycled thermal energy and alternative energy:

$$PUE = 1 + \frac{\dot{W}_{AC_{FANS}} + \dot{W}_{CP} + \dot{W}_{Loss}}{\dot{W}_{IT}}$$

$$+ \frac{1}{COP_{CRAC}}\left[1 + \frac{\dot{W}_{Loss}}{\dot{W}_{IT}} - HHF \cdot CF - COP_C\left(HHF \cdot CF + \frac{\dot{Q}_{EXT,S}}{\dot{W}_{IT}}\right)\right]. \quad (12)$$

Note that any useful heating power extracted from the data center and supplemented by an external source (e.g., solar) is subtracted from the total electric utility power. Also, note that, depending on the magnitude of the term

$$\frac{COP_C}{COP_{CRAC}}\left(HHF \cdot CF + \frac{\dot{Q}_{ext}}{\dot{W}_{IT}}\right)$$

PUE can become less than one, and the data center can actually export cooling. Here, $\dot{Q}_{ext}$ is essentially the same as $\dot{Q}_{EXT,S}$ but before it reaches any thermal storage. This cooling (removal of heat) is divided by COP_{CRAC} so that it represents an equivalent exported electric power. In other words, external heating can generate excess cooling that

can be "exported," i.e., used to cool adjacent rooms or facilities. The equivalent exported power can be expressed generally as

$$W_{Exported} = \frac{Q_{Excess}}{COP_{CRAC}}, \tag{13}$$

where $$Q_{Excess} = COP_C \left(HHF \cdot CF + \frac{\dot{Q}_{ext}}{\dot{W}_{IT}} \right) \tag{14}$$

2.9. Results and Discussion

To specify working values for each term in the final equation (12), we can apply the values in Table 2 from the 15-m^2 experimental ASU data center running at full capacity. At TDP, with 320 blades rated at 294 W_e per blade, they would deliver a *HHF* of 70% of that value or 205 W_{th}. We can expect to capture 85% of that dissipated heat or 59% from each server blade (174 W_{th}). This captured heat is denoted in Table 2. as $\dot{Q}_{HiT,S}$ and its value for the case of 320 blades is 55.5 kW_{th}. Assuming a 10% power delivery loss of the total power delivered to the data center and its components of $\dot{W}_{IT} + \dot{W}_{CRAC} + \dot{W}_{AC\_FANS} + \dot{W}_{CP}$, the parasitics of $\dot{W}_{AC\_FANS} + \dot{W}_{CP}$ are calculated for a maximum draw of 5.22 kW_e.

Table 2.

Values and Calculations for 15-m^2 ASU Data Center Model Running at Full Load.

Power/Heat	**Datacenter Total**	**Units**
$\dot{W}_{IT}$ (294W * 320 blades)	94.08	kW_e
$\dot{Q}_{HiT,S}$ (capture of 59%)	55.51	kW_{th}
$\dot{W}_{Loss}$ (Power loss of 10%) [25]	9.66	kW_e
$\dot{W}_{AC\_FANS}+\dot{W}_{CP}$	5.22	kW_e
Coefficients of performance		
COP_{CRAC}	4	[36]
COP_C (T_{gen} = 88°C)	0.7	Figure 2
COP_C (T_{gen} = 80°C)	0.86	Figure 2
COP_C (T_{gen} = 70°C)	0.63	Figure 2
PUE results for CF of 85%		
PUE (COP_C = 0.86)	1.16	Figure 8
PUE (COP_C = 0.7)	1.18	Figure 8
PUE (COP_C = 0.63)	1.19	Figure 8
PUE (=1.05Q_{IT})	0.997	Figure 8
PUE (=0.75Q_{IT})	0.996	Figure 8

For the COP_{CRAC}, we selected a standard value from a thermodynamics text for a typical vapor compression unit [36]. Then, for COP_C, we chose three scenarios from Figure 2: at design point, at minimum chiller capacity, and at maximum chiller COP. To isolate how PUE responds to chiller performance under standard conditions, we eliminated any external supplemental heat until the last two rows of Table 2 ($\dot{Q}_{ext}$). Note that even with no external supplementation, the PUE is near the "very efficient" value in Table 3 [37, 38].

Table 3.

Industry Benchmarked Data Center Efficiencies [37, 38].

PUE	**Level of Efficiency**
3.0	33% - Very Inefficient
2.5	40% - Inefficient
2.0	50% - Average Efficiency
1.5	67% - Efficient
1.2	83% - Very Efficient
1.0	100% - IT uses all W_e

As shown in the last two rows of Table 2, to realize a PUE value of less than one at a *CF* of 85%, an amount of supplemental external heat equating to 1.05 times that of the IT heat is necessary at the chiller design point. If the maximum COP_C of 0.86 is selected instead, this 1.05 value can be lowered to 0.75. However, again the tradeoff would be chiller capacity.

With only rack heat driving the chiller, there is a reduction of heat load on the CRAC. Even with no external thermal contribution and driven solely on a portion of data center IT equipment heat, there is a significant reduction of PUE. Figure 8 represents the range of possible PUE values for a high-temperature heat capture fraction, *CF*, ranging from 0% to 100%. *CF* = 0% corresponds to the "business as usual" approach and yields PUE ~ 1.43. As expected, PUE decreases with increasing *CF*, and with increasing COP_C, leading to a PUE as low as 1.16 at the maximum expected capture fraction of 85%. According to the benchmarks shown in Table 3, these PUE values suggest a "very efficient" data center is possible even with zero solar contribution. By employing an external heat source such as solar thermal, we can reach beyond the limitations of Table 3 With an amount of external heat equating to 1.05 times that of the IT heat, the PUE ranges from 1.25 at zero *CF* down to 0.95 at 100% *CF*.

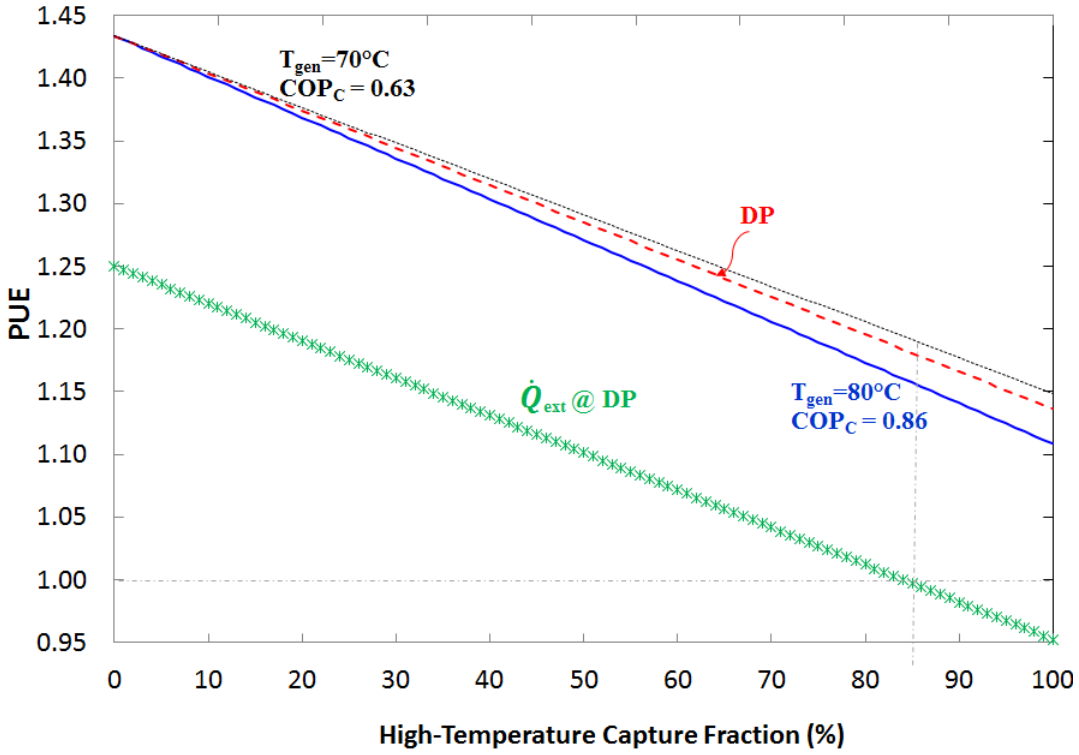

Figure 8. Power usage effectiveness (PUE) versus high-temperature capture fraction (CF). $\dot{Q}_{ext}$ is the external heat supplied to the generator at design point conditions.

Of course, the effectiveness of this approach depends on the *CF* and the factors that affect it. One possible way to achieve an increase in *CF* is to insulate the hotter components from the data center (e.g., using an insulating gel such as Aerogel®). Another possible way to increase the *CF* would be to cold plate more blade components. Also, the size and surface of the cold plate areas may impact the *CF* percentage, as could the thermal conductivity of the materials of construction of each cold plate.

Finally, to capture as much of the dissipated heat as possible, it is important to lower as much as possible the ΔT of the heat capturing scheme. The ΔT_{CH} between the

metal case surrounding the CPU die, T_C, and the cold plate, T_H, is especially variable and the only interface under direct control. A lower ΔT_{CH} can be achieved by reducing the thermal resistance between the device metal case and cold plate (heat sink) interface [21]. One way to reduce these values is to increase thermal contact (or decrease thermal contact resistance) between the two interfaces. Although the top surface of the device case and the bottom surface of the heat sink appear flat, there are many surface irregularities with air gaps that can prevent intimate thermal contact and inhibit heat transport. For a near perfect thermal contact, the two surfaces would need to be match ground and polished to better than mirror quality [39]. Although ideal, this 100% thermal contact would be very difficult to achieve. Instead, a thermal interface material (TIM) can be applied to help reduce both the contact and thermal resistances between the device case and the heat sink.

The heat removal capacity of a TIM is also determined by the TIM's ability to create an intimate contact with the relevant surface. A highly electrically insulating but thermal conductive TIM is essential to good heat capture performance. Fortunately, there are a number of TIM alternatives for electrically insulating the transistor from the heat sink while still allowing heat transfer. Thermally conductive materials used as TIMs come in two general forms: wet dispensed and fabricated [40].

Wet-dispensed TIMs include adhesives, encapsulants, phase-change materials, solder, and thermal compounds. Adhesives such as epoxies provide high adhesive strength; filled acrylic polymers are highly conformable, slightly tacky and are low in contact resistance. Silicone adhesives with fast thermal curing or RTV curing resist humidity, provide good dielectric properties, and are low stress and noncorrosive [40].

Encapsulants are two part silicones, with a flowable liquid that cures to a flexible elastomer at a constant cure rate [40]. Encapsulants can adopt any thickness and offer high mechanical strength compared with other wet dispensed TIMs [41]. Wet-dispensed phase-change materials offer high viscosity and good gap filling after they reach melt temperature. There is, however, some compressive force required to form a very thin bond line [40]. Solder and solder hybrids are low-melt metal alloys that flow at room temperature and provide a very thin bond line [40]. Finally, thermal compounds such as Arctic Silver® 5 thermal grease provide relatively high thermal conductivity and low thermal resistance in an easy to apply paste.

Unlike wet dispensable materials, fabricated TIMs are solids that have been laminated and die-cut to a specific shape for easy application [40]. Fabricated TIM categories include adhesive tapes, pre-cut phase-change materials, thermal insulating pads, and gap filler pads. An advantage of pads is that they can be applied without dispensing and therefore can be easily reworked [41]. For example, Bergquist gap filler pads readily accommodate irregular surfaces with minimal pressure and offer a thermal conductivity as high as 5.0 W $m^{-1}K^{-1}$, which is comparable to Arctic Silver® 5 thermal paste of 8.88 W $m^{-1}K^{-1}$. Lastly, fabricated phase change materials provide high thermal grease performance in a "peel-and-stick" format; compression brings the surface together and causes material flow ([40]).

2.10. Conclusions

Data center electric power consumption is an acknowledged problem that is likely to get worse in the future. The potential exists to utilize the waste heat generated by data centers to drive absorption chillers, which would relieve the cooling loads on the conventional computer room air conditioner (CRAC) and, ultimately, reduce the grid power consumption of the data center. By effectively capturing at least 85% of the heat dissipated from the CPUs on each blade server and efficiently transporting that thermal energy to drive a heat-activated lithium bromide absorption chiller, it is possible for this design to achieve a very efficient PUE (Power Usage Effectiveness) ratio. With external heat supplementation such as that offered by solar thermal, the value of PUE can fall to that of less than one.

Although conventional thinking defies a PUE < 1, the Green Grid has already begun receiving reports of some data centers actually achieving this possibility. For instance, the Finnish public energy company, Helsingin Energia, with their 2 MW eco-efficient datacenter underneath the Uspenski Cathedral in Helsinki, uses geo-thermal cooling with the heat being re-used by local businesses [42]. According to The Times of London, “This centre’s power usage effectiveness - the central measurement of data centre efficiency - will be an unprecedented figure of less than one.” [42].

All this points, ultimately, to the flaws inherent in the conventional PUE metric. A better data center performance metric would take into account the new eco-efficient data center designs and reward waste heat reuse and the use of renewable energy. Fortunately, the Green Grid, the National Renewable Energy Laboratories and Lawrence Berkeley

National Laboratory have responded by developing a new metric, tentatively termed ERE or Energy Reuse Effectiveness [43].

The aim of this new metric is to allow for recognizing and quantifying energy re-use efforts whose benefits are felt outside of the data center's boundaries.

CHAPTER 3

INVESTIGATING A RELATIONSHIP AMONG CPU TEMPERATURE, TASKING AND THERMAL POWER FOR DATA CENTER HEAT RECOVERY

As the average heat density per computer server increases, it has become increasingly difficult to provide the traditional cooling necessary to properly operate a data center. In some cases, the cooling that must be provided is two times what the heat loads truly require, due to inefficiencies within the traditional air cooling system and because the needs of today's data centers have surpassed the capabilities of typical CRAC (computer room air conditioners) units [1, 44]. Because traditional air cooling of datacom centers is reaching its practical limit, liquid cooling is being re-adopted to keep up with processor performance gains [11, 45, 46]. To aid in datacom facility thermal management, several companies have begun to offer alternative cooling solutions by extending liquid media to heat exchangers that are closer to the source of the problem (i.e., the computer equipment that is producing the heat). Companies such as Iceotope™ have liquid cooling jackets that encapsulate each server blade, and the quick connects retain the blades' hot-swappable ability within each data center rack [47]. A few companies submerse either part or the entire computing system in a dielectric liquid that has 1,200 times more heat capacity by volume than air [48]; Hardcore [49] uses a dielectric liquid contained inside the computer casing; Puget uses mineral oil in a custom-built aquarium with the motherboard, processor, power supply and fans submersed [50]; and Green Revolution Cooling (GRC) uses their GreenDEF™ (dielectric white mineral oil) coolant housed in rack baths, into which OEM servers are completely submerged

while the coolant dissipates server heat as it circulates through the system [48]. These approaches have the objectives of localizing cooling paths and offer the possibility of utilizing the captured thermal energy for applications such as space heating, heating swimming pools, hot tubs and alleviating snow and ice on roads or in parking lots. Additionally, there have been measures to capture the highest quality heat dissipated from the hottest components on each server blade, i.e., the CPUs. IBM has done work here using water-cooled heat sinks and piping to capture and transport higher-grade heat from the CPUs to the community [11]. The primary approach presented in this paper is similar to IBM's effort, but is unique in its application of utilizing captured heat from a CPU to drive a cooling process [51]. Another original result is a triad graph characterizing the relationship among the *three* parameters of CPU temperature, thermal power capture, and task loading of the CPU. Similar efforts by Mesa-Martinez et al. [15] and H.F. Hamann et al. [52], feature only *two* parameters at a time such as temperature versus time and power versus time but did not take into account all *three* parameters of temperature, power and loading simultaneously.

It is important to understand this triad relationship in order to efficiently exploit server and/or data center waste heat, such as to drive an absorption cooling system. The steady state analysis outlined in Chapter 2 suggested the thermodynamic feasibility of utilizing data center waste heat, perhaps augmented with solar heating, to serve as the primary energy source for a LiBr absorption cooling system, but left unanswered any questions about how such waste heat flows could be recovered in practice [51]. Here, we address this issue by describing the results of experiments on waste heat recovery from CPUs, and report the resulting relationship among temperature, power, and CPU loading.

3.1. Empirical Objectives

The overall objective of this work is to explore the relationship among CPU temperature, power and utilization during a CPU heat capturing scheme to drive a heat-activated cooling process for BlueCenter, a small-scale data center which is part of the "BlueTool" computer infrastructure project [53]. An initial study, at data-center level, determined how much heat and at what quality (i.e., temperature) was necessary to drive an absorption cooling process [54]. This paper continues on this path by investigating how much heat, and at what quality, can be captured from a CPU.

3.1.1. CPU Model and Frequency

The CPUs recycled were dual Intel® Xeon™ Irwindale processors residing on an IBM eServer xSeries 336 server blade. The Intel® Xeon™ Irwindale CPU has a standard core frequency of 3.6 GHz. This clock frequency is *fixed* by its design, set by the manufacturing process at the factory and represents the maximum frequency at which the CPU can safely and reliably operate. Furthermore, according to the datasheet, individual cores operate only at or below this factory-set frequency [55]. The minimum frequency of the processor is determined during manufacturing by setting bit 18 of the processor's IA32_FLEX_BRVID_SEL register, which is called the Platform Requirement Bit (PRB). For our particular 64-bit Intel® Xeon™ Processor with 2MB L2 Cache Irwindale model, the Platform Requirement Bit (PRB) is set during manufacturing to PRB=1, which means the default setting will be the minimum speed of the processor [55]. Although the datasheet specifies neither the CPU's frequency range nor at what performance level the reduction occurs, a reasonable assumption for this newly added feature in 2005 would be that the CPU throttles down to its minimum frequency during the idle state to save on

unnecessary power consumption. An educated conjecture would be that the minimum frequency at idle would be roughly 1/7 of the standard frequency of 3.6 GHz. This estimation is based on using the idle values of power and voltage in the following formula from the Intel Datasheet [55]:

$$f_{min} = \frac{2 \cdot P_{idle}}{C(VID - 0.06)^2} \quad (15).$$

3.1.2. MIL 217 Reliability Predictor

MIL-HDBK-217 is the original Military Handbook for "Reliability Prediction of Electronic Equipment" and contains failure rate models for the various part types used in electronic systems, such as ICs, transistors, diodes, resistors, capacitors, relays, switches, and connectors. MIL-HDBK-217 is published by the Department of Defense, but has not been updated since 1995. Specifically, the latest version of MIL-HDBK-217 is MIL-HDBK-217F, Notice 2 (217F2) dated February 28, 1995 which predates our Intel Irwindale CPU by a decade and therefore would not apply to any 2005 MOSFET technology. Most commercial electronic product companies are now choosing to use the Telcordia SR-332 handbook for their reliability predictions, but the cost for this documentation is $3200 per educational institution and they do not offer single user licenses.

Therefore, I found another avenue to determine the Mean Time to Failure (MTTF) for our Intel® Xeon Irwindale 3.6 GHz CPU. I contacted Intel directly, and from ticket number 8000880398, Intel Customer Support replied that the Mean Time to Failure for the Intel® Xeon Irwin dale 3.6 GHz CPU is 2.37+E06 hours. In this context,

MTTF would be the number of total hours of service of all CPUs subjected to stress testing divided by the number of CPUs [56]. Therefore, the MTTF time is a reliability predictor only and is not related to exergy efficiency but rather to the redundancy in information theory and indicates how reliable the switching device is [57]. Although not related to exergy, the MTTF value helped to determine the more likely cause of our eventual system failure as the hard drive being exposed to the flood water rather than the CPU itself experiencing failure.

3.2. System Layout, Circulation and Subsystem

To capture waste heat from the ASU Bluecenter data center, water as the liquid coolant is planned to circulate through heat exchangers or cold plates attached to the CPUs and possibly other components on the servers as shown in Figure 9.

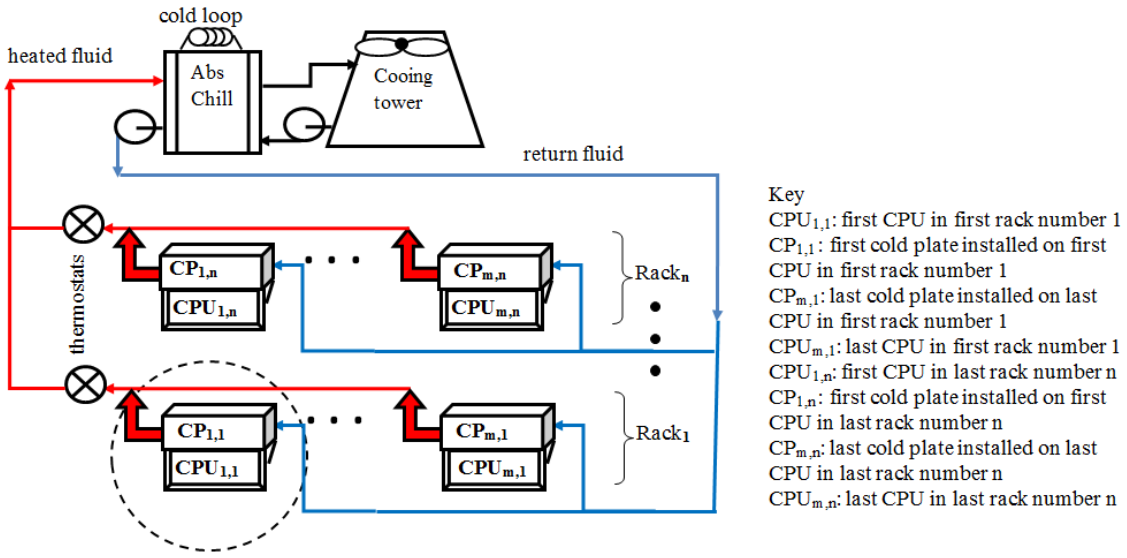

Figure 9. Overall system diagram. Focus for the experimental phase of Chapter 3 is the subsystem within the dashed oval.

If a high enough temperature can be maintained, the heated coolant can be delivered by a central pump to a Yazaki lithium-bromide absorption chiller (Abs Chill in Figure 9), which has a design flow rate of 2.4 $l\ s^{-1}$ (38.0 $gal \cdot m^{-1}$) [18]. This flow must be controlled in order to provide cooling to active components and to maintain a discharge temperature high enough to drive the absorption chiller. Since the tasking of CPUs in a data center can be random and the heat dissipation of each CPU variable, a thermosyphon array was initially designed that would act as a thermal diode. That approach would have allowed a number of servers to be connected to a single thermostatically controlled valve that would adjust the flow rate and maintain the discharge temperature within a narrow range. A wax motor thermostat such as those commonly used in automotive applications was considered. When the Thermal Control Circuit or TCC temperature (maximum safe operating temperature) and 1U (1.75 inches or 4.44 cm) size restrictions of the donated

Intel Irwindale Xeon 3.6 GHz processors and server footprint, respectively, were discovered, this approach had to be abandoned. A more active control scheme was required because the temperature drops in the passive system were unacceptably high. As displayed in Figure 1, the plan is to plumb a number of servers in series with installed liquid-cooling cold plates and use a variable servo-controlled valve to adjust the flow and maintain the discharge temperature. For the purposes of this paper, the focus is the experimental work performed on the subsystem within the dashed oval in Figure 9.

3.3. Measuring Cold Plate Heat Capture Effectiveness

In this first experimental phase**,** a CoolIT Eco© liquid cooling system [56] was connected to one of two Intel Xeon CPUs residing on an IBM xSeries 336 server blade. Each Intel Irwindale Xeon 3.6 GHz processor housed a single core with a specified TDP (thermal design power) of 110 W measured at a T_{CASE} value of 73 °C [23, 55]. According to the Intel® processor's thermal profile B, the T_{CASE} value can be 79 °C at maximum power but is usually limited by TCC activation [55]. Getting around the TCC activation was accomplished by shutting off the fans (eliminating their electrical draw and heat generation) and instead drawing heat quickly from the CPU via a liquid cooling solution.

To gauge the fluid rate, a Sho-Rate 150 Model 1350E flow meter from Brooks® Instrument was employed. This rotameter has a ten-to-one rangeability of float levels and a ±10% accuracy from 10% to 100% of full scale [59]. It was calibrated using hot water from the tap to simulate heated water from the CPU, and the calibration curve is shown in Figure 10.

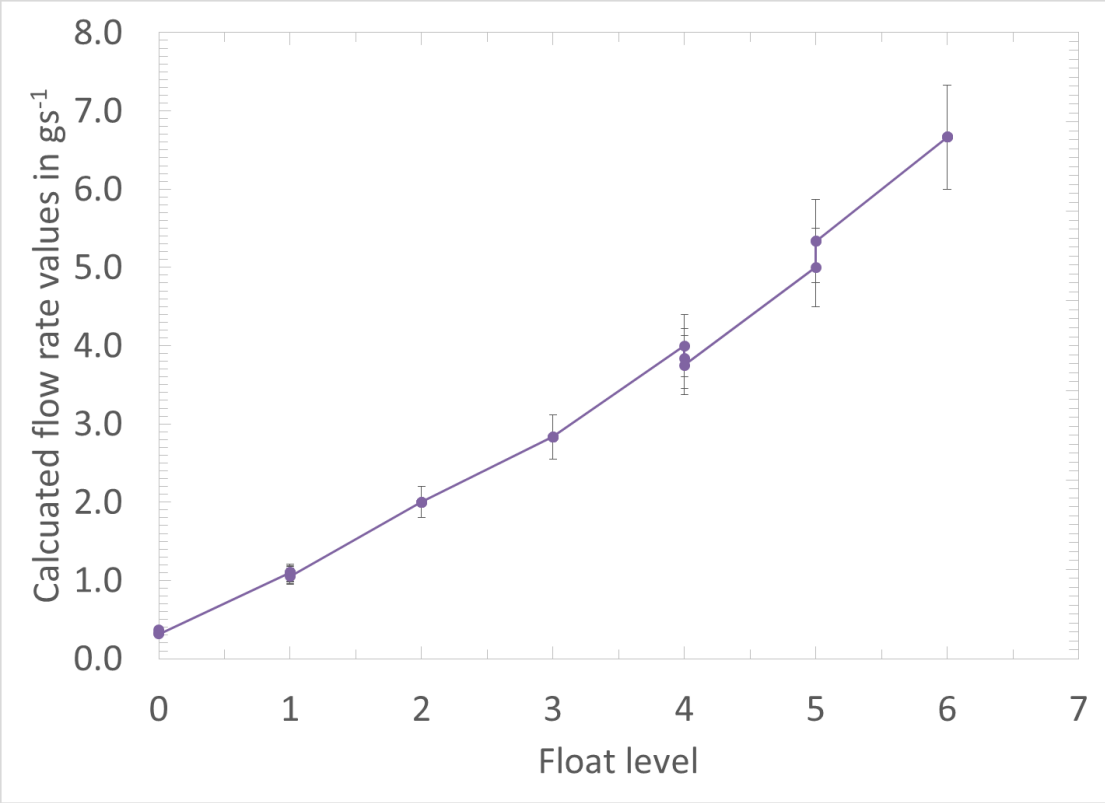

Figure 10. Sho-Rate 150 Model 1350E flow meter calibration curve showing the corresponding water flow rate in $g \cdot s^{-1}$ at each of six distinct float levels. Error bars indicate the ±10% of reading standard accuracy from 10% to 100% of full scale for the 1350 Series [59]. The deviation from linear occurred where instabilities caused the ball float to "dance" at float levels 4 and 5 so that slightly different amounts of fluid were collected at each of the three runs per level.

For each of the 23 runs (at least 3 repeated trials per level), a needle valve was adjusted until the flow rate raised the ball float to the next whole number level and remained steady. A timer ran for 30 seconds while the water was collected. At the end of the 30 seconds, the amount of fluid collected was measured in terms of milliliters (1 *mL* ~

1 g for water with a density ~1 $g \cdot mL^{-1}$) and the value divided by the time. As shown in Figure 10, float level readings ranged from 0 to 6 with the corresponding flow rates in $g \cdot s^{-1}$. The break in linearity curve occurred at levels four and five where the spherical float appeared to be affected by Von Karmen vortices, and it is suspected that the rotameter might have possibly changed flow modes. After installing the calibrated rotameter in the experimental setup of Figure 11, the mass flow rate reading correlated to a 3.3 g s^{-1} water flow rate. This mass flow rate is fairly close to the nearly 4.3 grams per second per CPU calculated for the case when the absorption chiller flow rate and heat input match the design point. To elucidate, at design point, the 10-ton Yazaki chiller flow rate is 2403.8 g s^{-1} with heat input 50.2 kW. At a CF of 85% or 93.5 W_{th} per CPU, this translates into 50.2/.0935 = 537 CPUs and 2403.8/562 = 4.3 g s^{-1} per CPU.

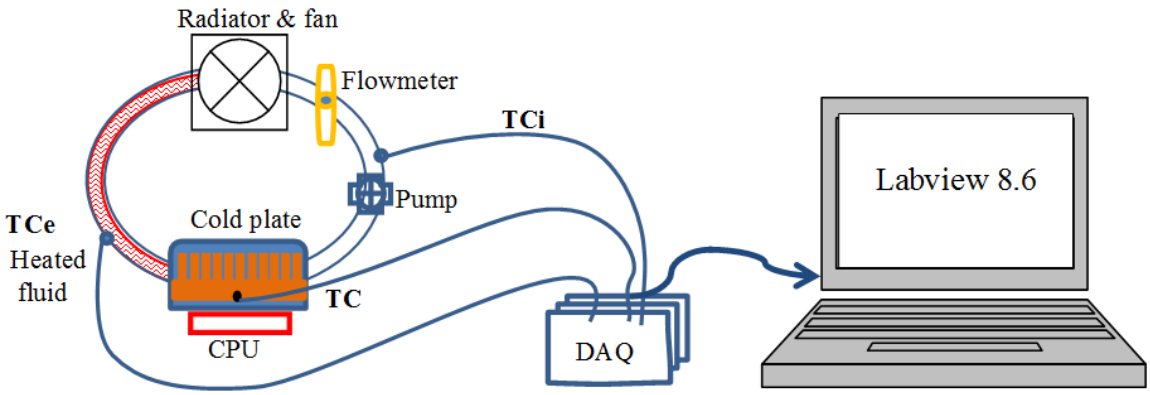

Figure 11. The heart of this experimental thermal extraction setup consists of a flowmeter, pump, radiator and cold plate mounted to one of the server board CPUs. TC, TCi and TCe are thermocouples.

To aid heat conduction across the cold plate-CPU interface, a small spring setup supplied a non-measured pressure to the cold plate affixed to the CPU. To further enhance heat conduction, we applied Biostar® TC-Diamond Thermal Compound (0.56 cm^2 °C W^{-1}) as the TIM (thermal interface material) in between the contact surfaces. Electrically non-conductive, Biostar® Nano Diamond Thermal Compound was applied evenly onto each heat spreader with the included plastic tool. A dot of about 5 cubic millimeters of thermal compound was centered in the middle of each heat spreader and then smoothed in all directions with the plastic tool to achieve a thin, even layer. The method was used to ensure optimum filling of the microscopic interstitial spaces between the contact surfaces of the CPU heat spreader and the bottom of the cold plate [60]. In addition, pressure was applied to further minimize interstices between the surfaces and reduce the *contact resistance* between the cold plate and TIM, TIM and CPU [61].

Although this does not affect the heat *capacity* removal of the cold plate, it does affect the cold plate heat removal effectiveness, i.e., the fraction of CPU heat that can be captured, also known as capture fraction (CF).

As shown in Figure 11, a pump, radiator and cold plate mounted to one of the server CPUs comprises the main components of the thermal extraction part of the system. The flow sequence is as follows: the CPU heats up and transfers thermal energy to the water within the heat-exchanging cold plate. The heated water is pumped to the radiator and cooled by active air flow. The cooled water returns to the heat exchanger and repeats the cycle. Temperature data for the CPU are recorded using one T-type thermocouple epoxied onto the heat spreader of the CPU as indicated by 'TC' in Figure 11. Two additional thermocouples, labeled as 'TCi' and 'TCe' in Figure 11 record the inlet and exit fluid flows, respectively, of the CoolIT Eco© heat extraction liquid cooling unit. To monitor the temperature difference of the fluid between the inlet and outlet ports of the cold plate, we used T-type thermocouples and Labview 8.6 monitoring software.

3.4. Experimental Uncertainty Analysis

To address experimental uncertainties inherent in the measurements, the following general equation was used for determining the uncertainty of a function in more than one variable:

$$w_R = \left(\left(\frac{\partial R}{\partial x_1} w_1 \right)^2 + \left(\frac{\partial R}{\partial x_2} w_2 \right)^2 + \ldots + \left(\frac{\partial R}{\partial x_n} w_n \right)^2 \right)^{1/2} \tag{16}$$

where w_R is the overall uncertainty in the results and w_1, w_2, …,w_n are the individual uncertainties in the independent variables [62]. Also, R is a given function of the independent variables x_1,x_2, x_3,…,x_n [62]. This relation was applied to determine the overall uncertainty of $\dot{Q}(\dot{m}, \Delta T) = \dot{m} C_p \Delta T$ and lead to the following equations,

$$w_{\dot{Q}} = \left(\left(\frac{\partial \dot{Q}}{\partial \dot{m}} \cdot w_{\dot{m}} \right)^2 + \left(\frac{\partial \dot{Q}}{\partial \Delta T} \cdot w_{\Delta T} \right)^2 \right)^{1/2} \tag{17}$$

and

$$w_{\dot{Q}} = \left(\left(C_p \Delta T \cdot w_{\dot{m}} \right)^2 + \left(C_p \dot{m} \cdot w_{\Delta T} \right)^2 \right)^{1/2} \tag{18}$$

or

$$w_{\dot{Q}} = C_p ((\Delta T \cdot w_{\dot{m}})^2 + (\dot{m} \cdot w_{\Delta T})^2)^{1/2} \tag{19}$$

where $w_{\dot{Q}}$ is the absolute uncertainty in the heat transfer product function and $w_{\dot{m}}$, $w_{\Delta T}$, are the uncertainties in the mass flow rates and temperature differences, respectively. The constant C_p is necessary for two reasons. Although it is not a measured quantity and does not influence the fractional uncertainties in the calculation, C_p does affect the absolute uncertainty [63]. Also, the units of $w_{\dot{Q}}$ must be in terms of W_{th} or J s^{-1}, which are dimensionally affected by the units of C_p.

The uncertainty in the mass flow rate, $w_{\dot{m}}$, is a straightforward calculation directly from the rotameter's tolerance specification of ± 10% of measured value. Futhermore, for each experiment, the mass flow rate, $\dot{m}$, is held constant. However, ΔT, the fluid

temperature difference |*TCe* - *TCi*| across the cold plate, is not held constant. Also, since ΔT is a difference, the uncertainty in ΔT or $w_{\Delta T}$ must be computed by summing in quadrature as shown below:

$$w_{\Delta T} = \left(\left(w_{TCe} \right)^2 + \left(w_{TCi} \right)^2 \right)^{1/2} \tag{20}$$

Since the error is *equal* for each individual thermocouple, the absolute uncertainty of the temperature difference across the cold plate is multiplied by a factor of $\sqrt{2}$:

$$w_{\Delta T} = \sqrt{2} w_{TCe} \tag{21}$$

Furthermore, $w_{TC_e} = w_{TC_i}$, and each thermocouple has an inherent uncertainty of ±1°C [64], so equation (21) simply reduces to the following:

$$w_{\Delta T} = \sqrt{2} \tag{22}$$

Finally, with known values and units inserted into equation (19) we arrive at equation (23). Equation (23) then enables us to calculate the absolute uncertainty in the collected thermal power and generate error bars on graphs 12-14, 16, 17, and 20.

$$w_{\dot{Q}} = 4.18 \frac{J}{g \cdot °C} \left(\left(0.1 \cdot \Delta T \cdot \frac{g \cdot °C}{s} \right)^2 + \left(\dot{m} \cdot \sqrt{2} \cdot \frac{g \cdot °C}{s} \right)^2 \right)^{1/2} \tag{23}$$

Figure 12 displays CPU temperature, the *ΔT* of the inlet and exit fluid flows across the CoolIT Eco© heat extraction liquid cooling unit, and the calculated thermal power at idle conditions.

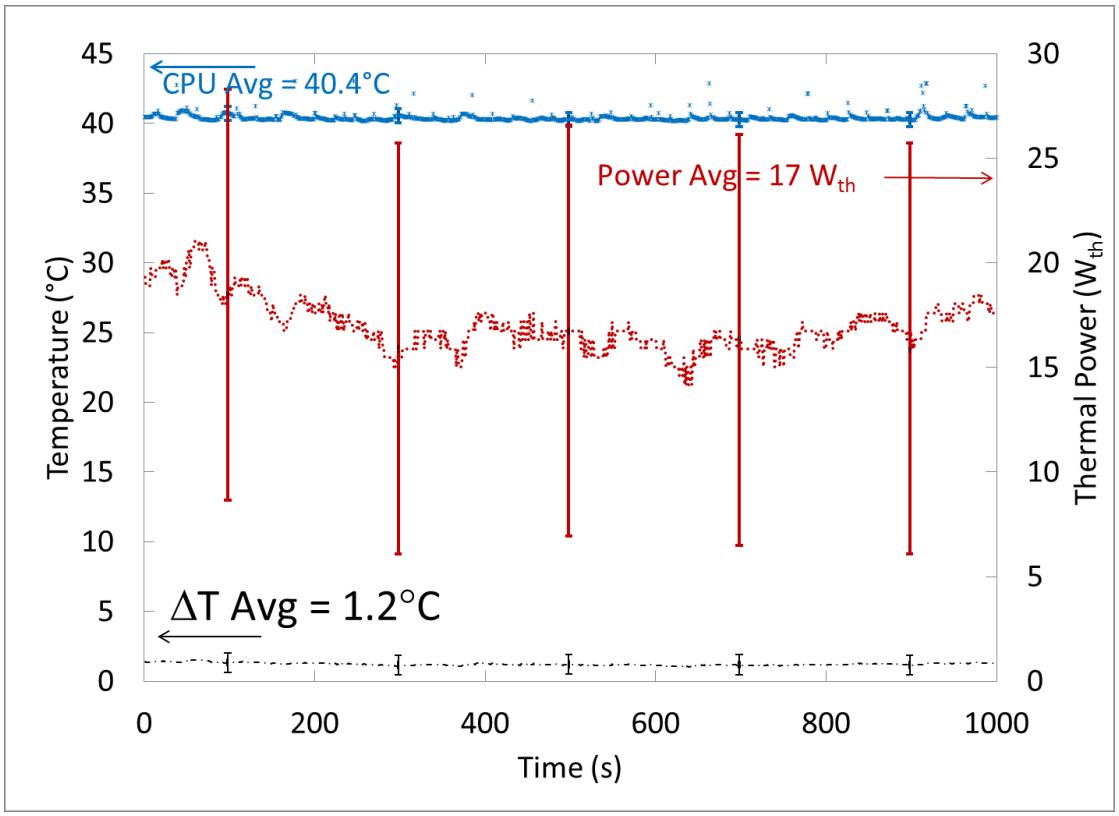

Figure 12. CPU temperature, captured thermal power, and ΔT across the CoolIT © Eco unit at or near idle conditions, with active server fans. Thermal power was calculated from the recorded ΔT and a measured 3.3 g s^{-1} mass flow rate. Error bars indicate the uncertainty of each measured value. Idle conditions correlate with an electric power input of 189 W_e. The error was calculated from the equation and was greatly affected by the C_p factor.

Thermal power was calculated using $\dot{Q} = \dot{m}C_p\Delta T$ where the C_p of water is 4.18 J $g^{-1}K^{-1}$, and the mass flow rate, $\dot{m}$, including uncertainty is (3.3 ±0.33) g s^{-1} as discussed previously. The fluid temperature difference across the CoolIT Eco©, ΔT, is determined from $|TCe\text{-}TCi|$ in Figure 11, with the uncertainty as $\pm\sqrt{2}$ from equation (22). Error bars, which indicate the uncertainty of each measured value, are especially large for the collected heat. This is due in large part to the product of the term $\dot{m} \cdot \sqrt{2}$ and the specific heat of water squared in equation (23). A slower flow rate, tighter thermocouple tolerance, and a lower specific heat such as that of mineral oil ($1.67\frac{J}{g\cdot°C}$) coupled together would lower the error. A reduced temperature difference across the cold plate would have less effect and runs counter to the aim of greater heat transfer and heat capture.

Additionally in Figure 12 the CPU instantaneous temperatures, the average CPU temperature, average ΔT, and average thermal power are denoted. For this first test, the CPU is running at idle and the server fans are active. Idle conditions correlated with an electric power input of 189 W. Figure 12 error bars illustrate specifically a thermocouple uncertainty of ± 1°C, a temperature difference uncertainty of $\pm\sqrt{2}$ °C and a thermal power uncertainty of $\pm w_{\dot{Q}}$, which turns out to be about ±19.6W_{th} from equation (23).

Notice in Figure 12 that the CPU temperatures are relatively low because the processors are idle (not being tasked) and the server fans are providing additional cooling. For later experiments, the processors are tasked and the server fans unplugged to capture more high quality heat.

In preparation for the following tasking phase, it was important to first determine the temperature limitations of the CPU. Therefore, OMSA (Open Manage Systems

Administrator) was installed, which specified that the CPUs could withstand up to 120°C at their warning threshold and 125°C at their maximum failure threshold [65]. This knowledge provided an upper safety limit of at least 120°C.

3.5. Triad Relationship: CoolIT Eco © Results

The CoolIT Eco© heat sink was used in the first experiments. After unplugging the fans, the CPU ran considerably hotter, though still under its threshold limit. The liquid cooling equipment kept the server blade operational throughout the next test while the electric input power surged from a minimum of 180 W at idle to an average of 384 W at maximum tasking load on the CPU. We used the CPU Burn-in v1.01 [66] stability testing program to incrementally increase loading on the CPU to the maximum possible operating temperature that is achievable by using ordinary software. The program continuously monitored for any erroneous calculations during CPU performance under the increasingly stressful tasking conditions from idle to 100% in increments of 25%. The results are shown in Figure 13, again at a flow rate of 3.3 g s^{-1}. Since the ΔT does not have the largest impact on the uncertainty of $\dot{Q}$, $w_{\dot{Q}}$ varied between only ±19.6W_{th} and ±20.4 W_{th} as indicated by the error bars of Figure 13.

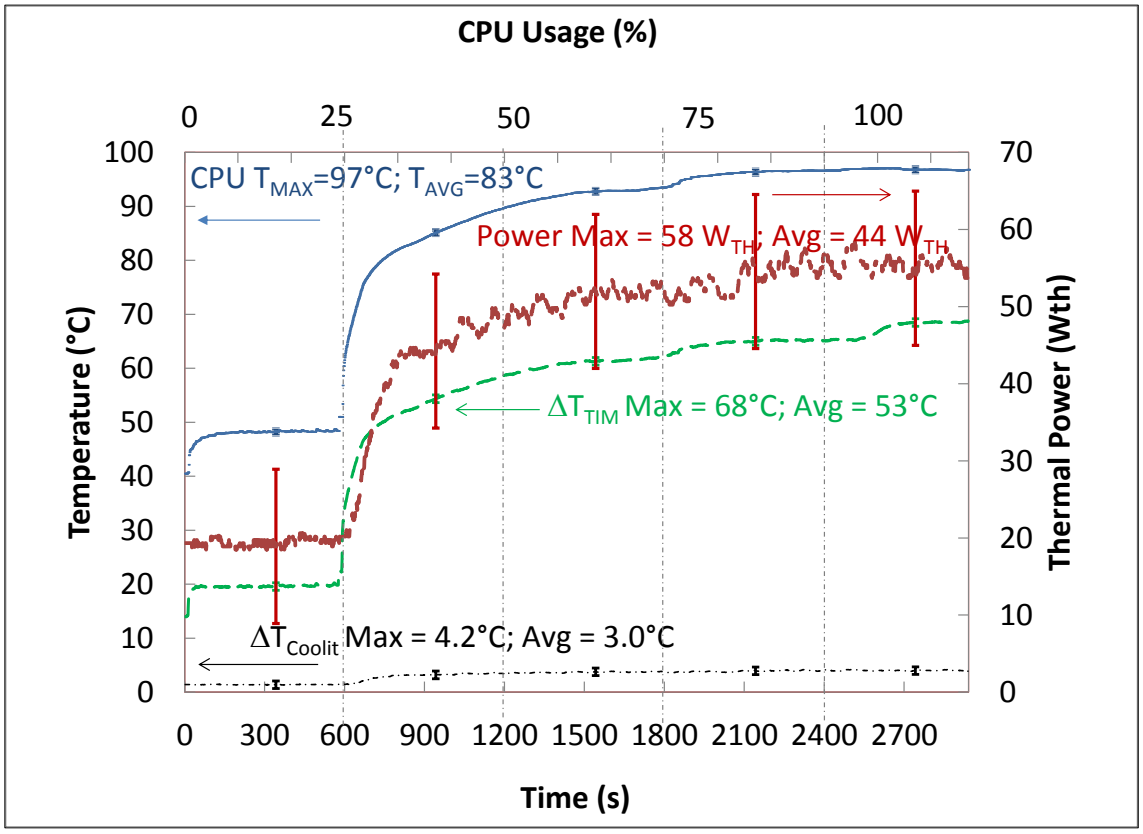

Figure 13. Heat captured, CPU temperature, temperature drop between CPU and CoolIT Eco© cold plate (ΔT_{TIM}), and fluid temperature difference across cold plate (ΔT_{CoolIT}) per CPU at incrementally increasing tasking levels with zero fan cooling. Flow rate is 3.3 gs^{-1} . Electric input power ranged from a minimum of 180 W_e at idle to an average of 384 W_e at 100% full tasking of the CPU. The temperature drop between the copper block and the heat capturing cold plate was unacceptably high across the thermal interface material after 25% utilization. This was primarily due to the inability of the stock cold plate to closely fit the surface of the CPU.

Most importantly, Figure 13 illustrates the triad relationship between temperature and thermal power dissipation of the CPU at increasing tasking levels. I utilized this information to determine how many CPUs would be required to drive the absorption chiller. For instance, the Labview recordings graphed in Figure 13 show that an average of 43 W_{th} can be collected per CPU. To drive the 35.17-kW (10 ton) absorption chiller, all the servers together in the ASU data center would need to dissipate 11.3 kW_{th} minimum of thermal energy. This translates to 11.3 kW / 0.043 kW = 263 CPUs or 132 server blades to provide the required minimum heat to drive the absorption chiller. At peak CPU loading (100% tasking), the number of CPUs required drops to 195. Also, CPU temperatures are at acceptable levels. However, there might be some difficulty with obtaining enough heat when the system is at 0% tasking level since about 665 CPUs would be required (each CPU dissipates only about 17 W_{th} at idle), plus a higher temperature would be critical.

3.6. Triad Relationship: Custom-Built Cold Plate (CCP) Results

Working with Jon Sherbeck, we hypothesized that there may be limitations within the commercial CoolIT© Eco cold plate and decided to explore if we could capture more heat by designing our own cold plate. The CoolIT© Eco gave us a baseline from which we proposed designing and fabricating an improved cold plate with a better fit to the CPU. We inspected the various dimensions and features of the CoolIT© Eco heat exchanger such as frontal area, fin height and base plate thickness and slant. From these inspections, we fabricated an optimized, innovative, and better fitting custom-built cold plate (CCP) by iteratively altering the base design.

From Figure 14, note that the CCP demonstrates higher heat capture values than the CoolIT© Eco as well as lower ΔT_{TIM} values between the CPU and the discharge line of the CCP cold plate. For the TIM, we again used Biostar® TC-Diamond Thermal Compound. The 50% improvement in ΔT_{TIM} for the CCP setup resulted in higher quality heat being transferred more effectively from the CPU to the discharge line of the CCP cold plate. Ideally, a ΔT_{TIM} of zero would allow maximum heat transfer to the chiller, but a ΔT_{TIM} loss of 10°C would still enable the CCP water-cooling setup to deliver, in its discharge line at 50% CPU utilization, the required minimum 70°C temperature to activate the chiller. Therefore, a ΔT_{TIM} under 10°C was a desired deliverable that had to be achieved. In conclusion, while the CCP was a significant improvement over the stock CoolIT© Eco, the ΔT_{TIM} was still too high to be able to deliver the necessary quality heat to the chiller.

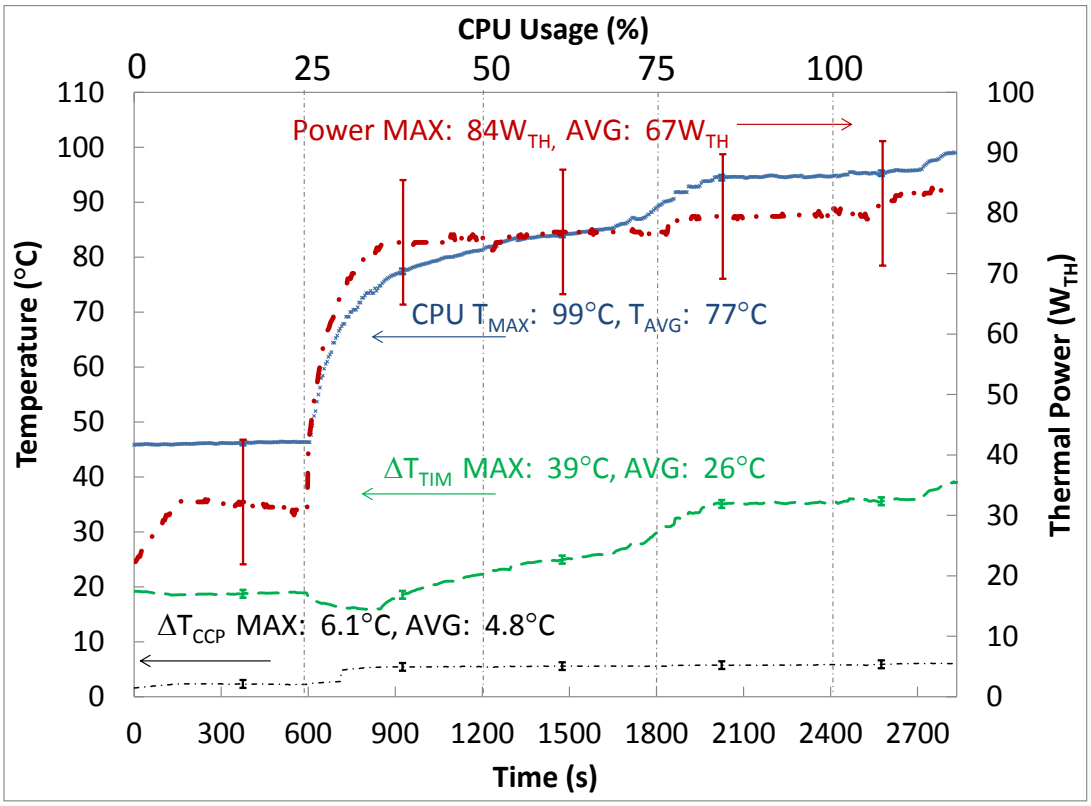

Figure 14. Custom-built Cold Plate (CCP) performance at incrementally increasing tasking levels of the CPU, with zero fan cooling. Flow rate is 3.3 g s^{-1}.The temperature drop between the CPU and the custom built cold plate was still high at 26°C on average across the thermal interface material (ΔT_{TIM}) but about 50% improvement over the initial CoolIT Eco© results. The temperature drop between custom cold plate inlet and discharge is denoted as ΔT_{CCP}.

For the data graphed in Figure 14, the custom-built cold plate or CCP, shows its heat capture performance for varying levels of CPU loading. Also in Figure 14, all three factors of temperature, thermal power dissipation, and tasking levels of the CPU are simultaneously displayed and present a relationship – as tasking levels increase, there is

an increase in power dissipation, CPU temperature, and temperature difference across the cold plate. Note especially the upsurge in power and temperatures after 25% loading on the CPU. If a ΔT_{TIM} drop of under 10°C could be achieved, then we can calculate approximately how many servers it would take to minimally and optimally power the chiller. To determine how many CPUs would be required to drive the absorption chiller for the recordings graphed in Figure 14, an average of 67 W_{th} can be collected per CPU. This translates to 11.3 kW / 0.067 kW = 169 CPUs or 85 server blades to provide the required minimum heat to drive the absorption chiller. At peak CPU loading (100% tasking), the number of CPUs required drops to 135. For the donated 10-ton Yazaki LiBr-water absorption chiller to operate ideally at its design point, the required input generator heat of 50.2 kW_{th} would require 784 CPUs or 625 at peak. There still might be some difficulty with obtaining enough heat when the system is at idle since, according to Figure 14, about 452 CPUs would be needed at minimum at that 0% tasking level. The backup plan would be to use stored thermal energy from a phase change storage medium. Of course, these are the findings of the current system. With improved insulation and applied pressure, the potential is high to collect more of the dissipated heat at an increased temperature. Figure 14 results also suggest that improved heat sink design is necessary to capture waste heat effectively from datacenter server CPUs.

3.7. CPU Simulation: Full Power Results from Copper Block Test Setup using CCP

To test if applied pressure and encapsulating insulation would enable an even higher heat capture, the CCP heat extraction equipment was ported onto the test setup

shown in Figure 15. This is a very similar setup to that displayed in Figure 11 with the exception that the cold plate is now subjected to a known and adjustable clamping pressure. The focus here is on how effectively the CCP thermal extraction system captures heat, so the test block with a known power dissipation value serves to reduce the number of variables.

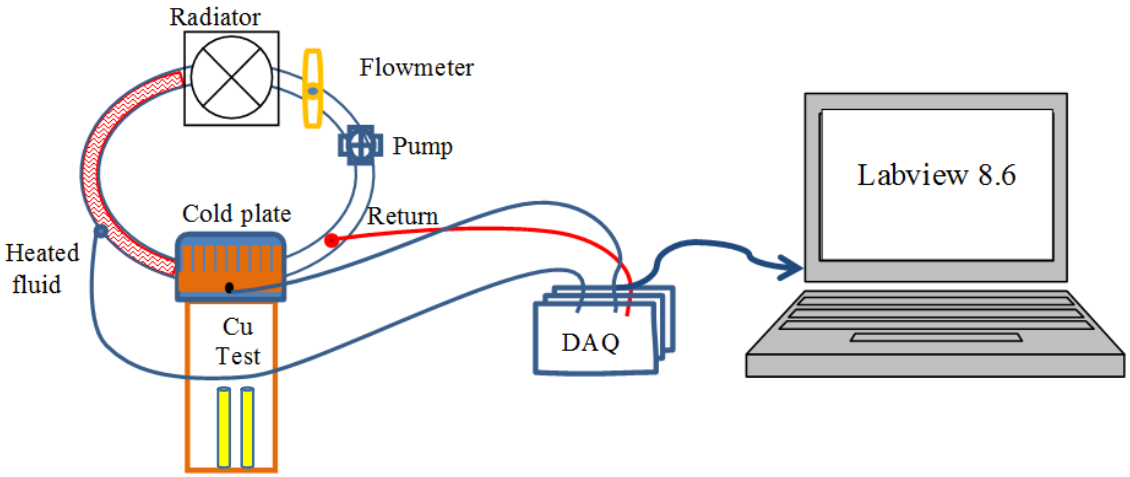

Figure 15. Experimental setup to test the custom-built cold plate (CCP) consisting of a radiator, flowmeter, pump, and pressurized cold plate mounted to a copper test block with an installed cartridge heater.

This simulation was setup to explore the ability of the heat-capturing scheme which consisted of a liquid-cooled heat sink (cold plate), a pressure plate, a cooling radiator (heat exchanger simulating the chiller), insulation, and a pump. A rotameter measured the flow rate. A T-type thermocouple measured the CPU case temperature. Working with Jon Sherbeck, we fabricated and installed a differential thermocouple to

gauge the fluid temperature difference across the cold plate. Another T-type thermocouple was used to monitor the copper block temperature. Labview 8.6 software monitored all of the thermocouples and converted the voltage values into degrees celsius. To maximize heat capture and minimize heat loss, we encapsulated the copper block in 5 mm of aerogel Cryogel 5201 insulation. Additionally, we wrapped foam insulation around the fluid lines and the rotameter as well as the T-shaped copper pipe fitting attached to the pump. We also used a TIM (thermal interface material) consisting of the diamond paste used previously between the top of the copper block and the heat capture surface of the cold plate. A pressure plate was installed onto the cold plate to reduce the contact resistance between the contact surfaces of the copper block, cold plate and sandwiched TIM. A pressure sensor was used to verify the pressure gauge reading. The initial experimental setup is represented in Figure 15.

As shown in Figure 15, a cold plate mounted to a copper test block, a radiator, flowmeter (rotameter), and pump comprised the main components of the thermal extraction part of the system. The rest of the experimental thermal extraction setup consisted of insulation around the bottom and sides of the copper test block and covering the discharge and return lines, differential and standard thermocouples and a data acquisition system. Brass fittings were inserted into the system ports to allow monitoring of the temperature differences between the heated fluid exiting and the return fluid entering the cold plate, as shown in Figure 15. A pneumatic cylinder with a surface area of 5.03 cm^2 and pressure of 689.5 kPa (100 psi) clamped the CCP cold plate to the cold plate, TIM and test block. This translated into 358.5 kPa (52 psi) across the 9.6 cm^2

surface of the CCP. This static pressure is well within Intel's processor loading specifications [55].

For the test run results shown in Figure 16, we ran the cartridge heater inside the copper block at full power or 218 W_{th}. This power emulates both CPUs running at full thermal design power (TDP). To decrease the thermal resistance between the contact surfaces, we applied 689.5 kPa of pressure to the top of the cold plate to reduce any air gaps between the cold plate -TIM- heat block (simulated CPUs) interface. This applied pressure plus the use of a sample of insulating Aerogel around the heat block increased the CF (capture fraction) from a previous value of 67% to the 93% average as shown in Figure 16. The mass flow rate of 4.8 g s^{-1} is directly measured from the rotameter so that the thermal power calculated represented what the custom-made cold plate can capture. On average, the heat captured during this test run was 203 W_{th}. The differential temperature across the cold plate varied by less than ± 1°C indicating that the system had reached steady state.

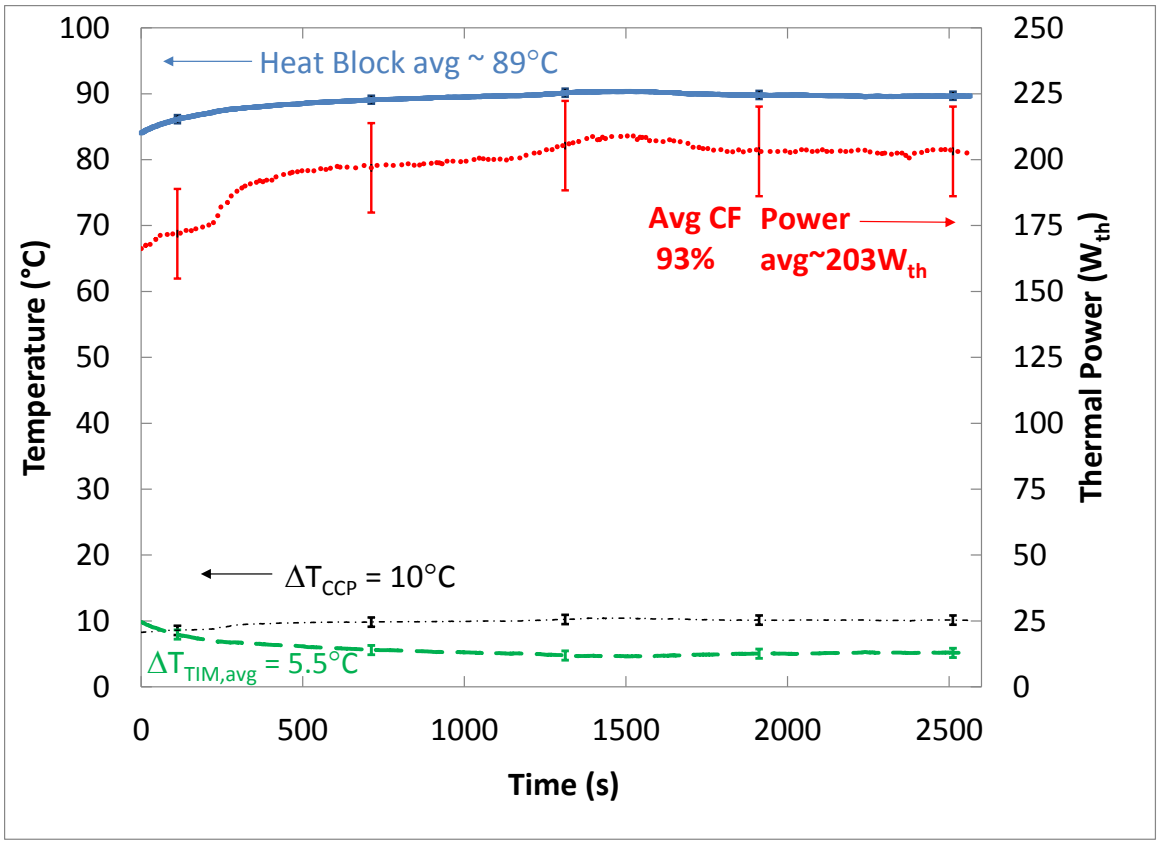

Figure 16. Full-power test results from the system shown in Figure 15. At a mass flow rate of 4.8 g s^{-1} and an applied pressure of 689.5 kPa (100 psi), the average heat captured was 203 W_{th} and the average CF (capture fraction) was 0.93 or 93%. The temperature drop between custom cold plate inlet and discharge is denoted as $\Delta T_{CCP.}$ In contrast, the temperature drop between the copper block and the heat capturing cold plate was noticeably lower at 5.5°C across the thermal interface material (TIM).

This temperature drop was lower than for the CPU liquid cooling experiments and was within the acceptable range of ΔT_{TIM}<10°C, which means that the heat can be transferred with less temperature loss from the source to the cold plate discharge line and eventually to the generator input of the absorption chiller. This improvement in higher

quality heat transfer suggests that applying pressure to the cold plate affixed to the CPU would reduce thermal resistance by minimizing the interstitial gaps and thus providing more intimate contact between the CPU and cold plate. Even better would be to either solder together the cold plate to the CPU or manufacture both together as one seamless unit. Thus, one specific and supported conclusion is that the less gaps there are between the CPU and the cold plate, the less the temperature drop between the CPU to the cold plate discharge line and the greater the quality of heat that can be transferred [60].

3.8. CPU Simulation: Half Power Results From Copper Block Test Setup using CCP

For the test run shown in Figure 17, we used a variac at a 60% setting to decrease the cartridge heater dissipated power to $110W_{th}$. This reduction simulates a single CPU running at its thermal design power (TDP) of $110W_e$ and thus dissipating $110W_{th}$.

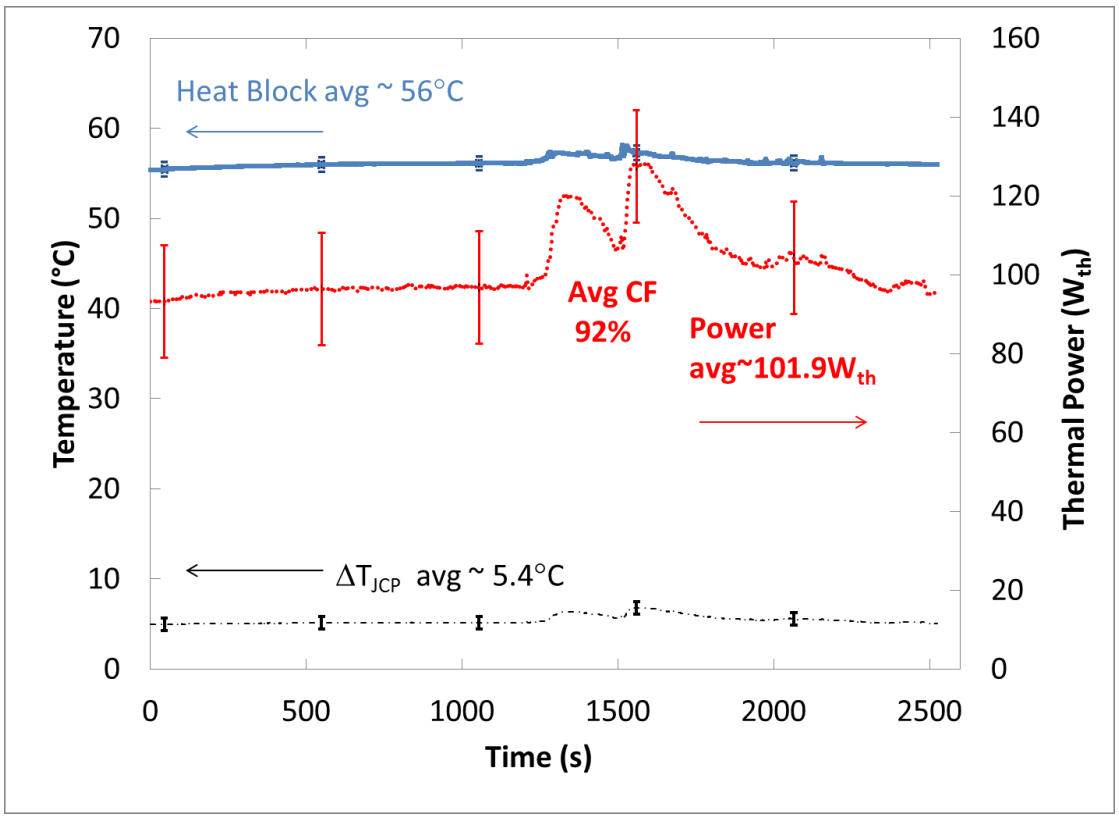

Figure 17. At half power, a mass flow rate of 4.5 g s^{-1} and an applied pressure of 689.5 kPa (100 psi), the average heat captured was 101.9 W_{th} and the average CF (capture fraction) was 0.92 or 92%.

Notice that for both the full power and half power results in Figure 16 and Figure 17, respectively, the flow rate, heat capture and temperature difference across the cold plate have all increased relative to the results from the previous setup in Figure 11. This indicates that applying pressure and insulation to the system is a beneficial addition to enhancing heat capture.

3.9. Discussion

Accessing each CPU's thermal diode temperature and power data via software has been an unexpected difficulty. I tried many different software utilities including Real

Temp, Core Temp, HWiNFO32, HWMonitor, CPU temperature monitor, EVEREST Ultimate Edition and even alternative operating systems such as several versions of Windows and Linux (RedHat, Unbuntu). In many cases, the software simply would not install on this older technology. HWiNFO32 software in particular revealed that the IPMI (Intelligent Platform Management Interface) hardware-level sensor subsystem simply does not support reading CPU temperatures and power reporting also was not available. Nevertheless, in 2011, we attempted using Intel's 2010 CPU Monitoring with DTS/PECI but found that the donated Xeon Irwindale CPUs predated both [67]. PECI (Platform Environment Control Interface) is a standard used for thermal management specific to the Intel Core 2 Duo microprocessors, but since the Xeon Irwindales are each single core and both the CPUs and IBM motherboard predate PECI, PECI does not apply to this older technology. Also, Intel's Xeon Processor -Thermal management did not specify what software can read Thermal Sensor and Thermal Reference Byte information off the SMBus (System Management Bus). We tried Intel's proprietary programs to detect temperatures but, again, with the 2005 IBM motherboard, this did not work.

Only Dell OMSA (OpenManage Server Administrator) was able to provide CPU temperature readings (plus warning and failure thresholds). However, power reporting was still inaccessible and by then we had epoxied thermocouples directly on the CPU casings. Furthermore, Windows Server 2008 and Dell's OMSA added to CPU workload and did not update quickly enough as compared to the thermocouples, which were more responsive and robust. The one advantage of the OMSA software was that each CPU's temperature could be logged automatically with a freeware scripting program for later access. To this end, I used AutoIt scripting program to simulate mouse movements and

keystrokes within the windows environment to automate the logging of CPU temperatures from OMSA into a text file (see Appendix A for the program).

By attaching thermocouples, we were able to take direct temperature readings of the CPU case temperature. Thermocouple readings of CPU temperatures showed a sufficiently high temperature to serve as the input temperature to the absorption chiller [51]. However, although utilizing thermocouples in sync with a data acquisition system (DAQ) provided direct temperature readings of each CPU case temperature, it would be important to make the control system more feasible and streamlined by using unobtrusive system software to monitor CPU temperatures.

Another difficult challenge has been reducing the temperature difference between the exiting coolant flow and the CPU temperature, which is critical to achieving a sufficiently high heat capture fraction. This temperature difference results from three thermal resistances in series as follows: 1) the contact resistance between the CPU and the cold plate, 2) the conduction resistance internal to the cold plate, and 3) the convection resistance of the fluid film. The temperature difference between the CPU and the base of the cold plate is a function of the contact resistance. Contact resistance can be lowered by providing and maintaining very flat and smooth surfaces that can therefore achieve an intimate contact, providing a high pressure holding the two surfaces together, and filling any remaining voids completely with a highly conductive fluid such as the Biostar ® TC-Diamond thermal compound. Because of structural limitations, the contact pressure was limited to the server's original pressure. Providing and maintaining intimate contact between the CPU and cold plate actually entails a compromise between the contact and conduction resistances because thickness in the base plate of the cold plate is

required to maintain a flat surface, while that same thickness adds to the conduction resistance. Working with Jon Sherbeck, we thus optimized our base plate thickness to establish a balance between low contact resistance and structural integrity. From a system perspective, the cold plate flatness can be enhanced by operating the cold plate at the lowest possible pressure by metering the flow upstream of the cold plate, limiting the number of cold plates in series and having them discharge into the circulating pump inlet. The conduction resistance of the cold plate is due to the limited thermal conductivity of the material of construction. Copper is commonly used in fluid heat exchanger applications due to its relatively high thermal conductivity at ~ 400 $Wm^{-1}K^{-1}$, and good corrosion resistance. Reducing the base thickness will reduce the temperature drop between the copper block, and the exiting coolant temperature. The convective heat transfer of the fluid film is the product of the convective heat transfer coefficient, the area, and the ΔT between the surface and the fluid. So, with a given amount of heat dissipated by the CPU, the cold plate heat transfer coefficient, and/or the heat transfer area must be increased, to reduce the ΔT.

This work has successfully accomplished demonstrating and illustrating the triad relationship among CPU utilization, temperature and power. A future endeavor would be to decrease the ΔT between the heat source and the water in the discharge line exiting the Custom-Built Cold Plate (CCP) as well as to decrease the contact resistance between the CCP and heat source (thereby decreasing the thermal resistance across the TIM) while simultaneously maintaining a high heat capture fraction. The applied pressure onto the CCP has already made it possible to decrease the contact resistance while maintaining a high Capture Fraction (CF). Also, another consequence of the test setup shown in Figure

15 is an increase in the ΔT across the CCP. The real challenge will be to concurrently decrease the thermal resistance across the TIM situated between the heat source and the cold plate. To that end, we have begun researching TIMs with a lower thermal resistance than the thermal compounds that we have so far employed. We found that Indium Foil and Heat-Spring® materials can offer a lower thermal resistance compared to thermal grease [68]. At approximately 275 kPa (40psi) and higher, the 0.076 mm (3 mil) thick Heat-Spring® materials exhibit superior performance over thermal grease [66]. This material is used alone, and does not require the use of a grease of any kind, making the process much cleaner. Also, there is no issue of pump-out as compared to using a 0.05 mm (2 mil) layer of thermal compound.

3.10. Mineral Oil Cooling

Eliminating the TIM entirely is another idea via liquid submersion cooling within a thermally, but not electrically, conductive fluid. Dielectric oils such as 3M Fluorinert, transformer, motor, silicone, and even cooking oils are preferable because these liquids have sufficiently low electrical conductivity not to interfere with the normal operation of a computer. Of note, non-purpose oils such as technical grade mineral oil have been successfully used for cooling personal computers [50, 48]. Since computers cooled in this manner do not generally require fans or pumps, and may be cooled exclusively by passive heat exchange between the computer's parts, the cooling fluid and the ambient air, the most efficient possible cooling approach is to immerse the chips and server blade entirely within the dielectric liquid [69, 70]. Directly cooling the CPUs with liquid was thought to bring to near zero the ΔT between the heat source and the water in the

discharge line and the thermal resistance across the TIM no longer would be of concern. Such an approach would still take advantage of better heat transfer properties of liquid cooling compared to air cooling systems and is also an approach that has a high potential for realizing significant energy saving opportunities [70]. Based on this idea, we assembled an immersion system for our server blade as shown in Figure 18. As Figure 18 illustrates, the idea is to submerse a server blade up to the level of the CPUs in a non-conducting liquid media and capture the highest quality dissipated heat via a manifold placed directly over the CPUs (Figure 19). This dielectric fluid can then be circulated and delivered to a heat exchanger, providing the servers with cooling and supplying the Yazaki chiller with consistent heat energy.

The mineral oil setup shown in Figure 18 is principally housed within a 55-gallon metal drum. The tank contained the insulation, the thermocouples, a manifold, a rectangular support made specific to the server dimensions, and then the server blade itself immersed in oil within that support. An important detail was this specially made encasement for the server blade which kept mineral oil expenditure at a minimum and elevated the hard drive above the fluid. Stainless steel is the selected material for the heated oil container because of the relatively high temperatures of interest. Also, steel is strong and can be easily welded, formed, brazed, and soldered. Furthermore, it is low cost, recyclable, and non-corrosive in an oil bath. External to the tank enclosure is the grinder motor that we used as a pump, flexible tubing lines, radiator and rotameter.

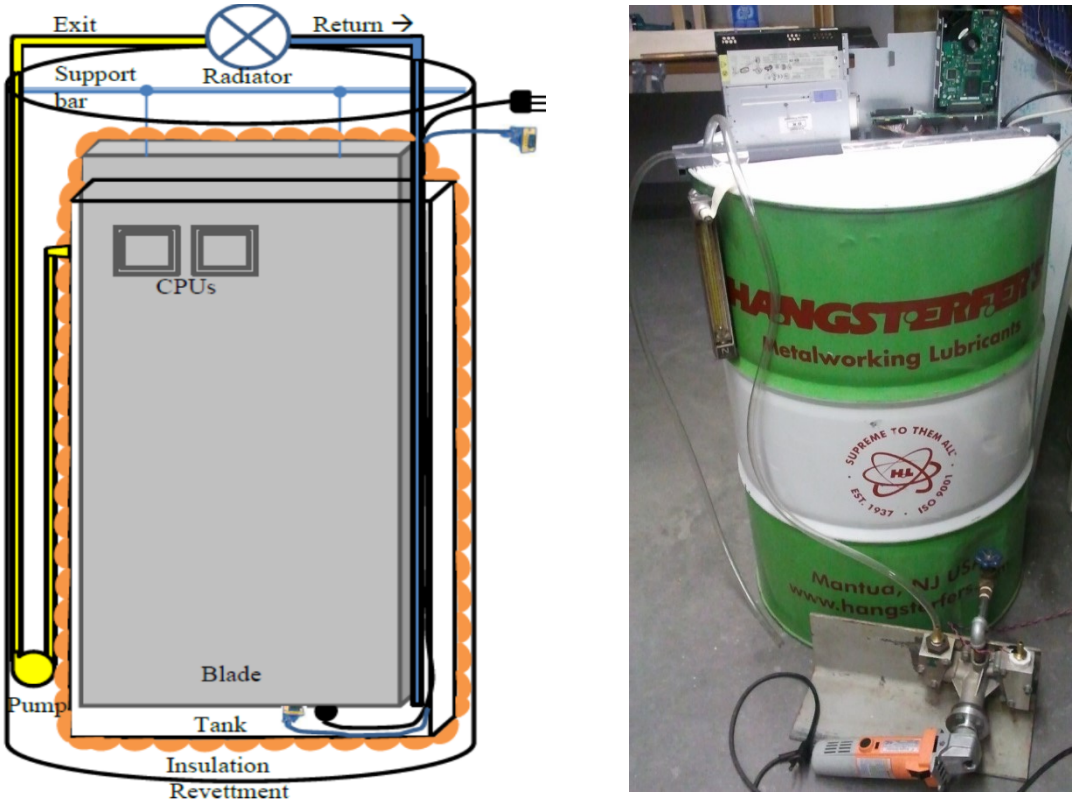

Figure 18. Concept drawing and mineral oil setup to eliminate the TIM altogether and associated ΔT between the CPU and the discharge line. Physical system shown in right picture consisted of tank housing with insulation, immersed server blade inside rectangular encasement, manifold, thermocouples, fittings, tubing lines, orange grinder motor as pump, radiator and rotameter.

To calibrate the rotameter for use with the oil, we set the pump rate, set a timer, flipped the valves on and off on mark to collect a sample, recorded the float position and time interval. We weighed the sample for mass, recycled it into lines and repeated the process for 20 samples. Our average flow rate was 6.75 g s^{-1}.

Figure 19. Manifold fitted over CPUs is used to direct the highest quality thermal energy (i.e., high temperature heat) into the discharge line and ideally to an absorption chiller. The differential thermcocouple plumbed into fitting gauges the discharge temperature. The remaining heat supplied by 1U IBM eServer xSeries 336 server blade is transferred into the clear, odorless, non-toxic, dielectric, technical grade white mineral oil optimized for cooling servers and other data center hardware.

One problem with the mineral oil setup was the grinder motor that we had to use as a pump. Although it proved more powerful than the pump used in the CoolIT© setup, the grinder motor was not made to run for extended periods of time and would lose efficiency and overheat. Also, air would become entrained, filling the tubing lines with bubbles, which had to be continually cleared. This made long runs very difficult to achieve, which was unfortunate because the temperatures of the oil crept more slowly

towards steady state as compared to the trials with the water-cooled systems. Additionally, there was concern of pump work contributing to the heat collected, but we found that the pump work could be rejected by modulating the radiator fan.

In setting up the cabling, we learned that evaporation can pose a problem through a wicking effect up cables that were submerged in the oil [50].Therefore, we used a minimum of cabling, so very little oil was lost this way. The final outcome of our mineral experiment oil is shown in Figure 20. As in Figure 13 and Figure 14, Figure 20 illustrates the triad relationship between temperature and thermal power dissipation of the CPU at increasing tasking levels.

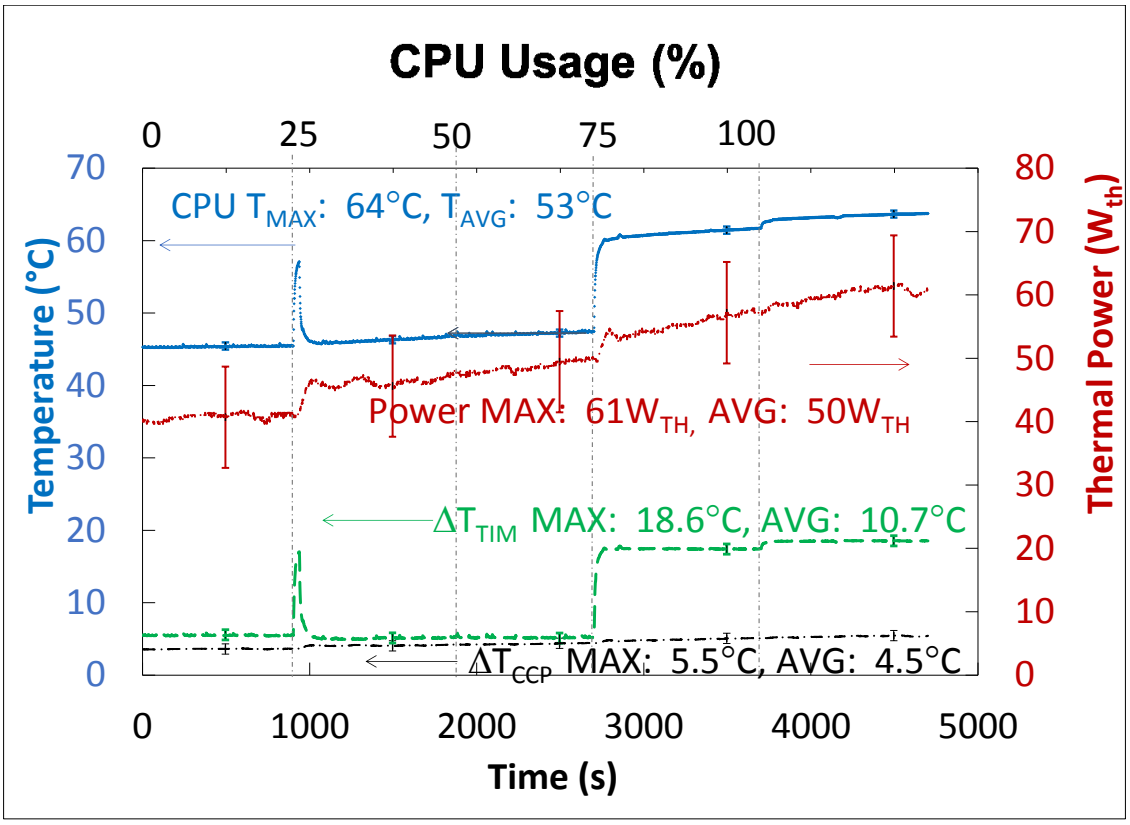

Figure 20. Mineral oil triad graph simultaneously relating CPU utilization, temperature, and thermal power dissipation. Temperatures were more stable and time intervals were much longer than for water-cooled experiments. Thermal power needed more time to reach steady state than for the water-cooled experiments, but there was a safety concern in running hot mineral oil overnight left unattended in the lab so the run time had to be cut short. The thermal power trend especially indicates that the data potentially had more rising to do.

Notice that the time intervals were triple in duration due to the mineral oil's slowness at reaching steady state for thermal power, although the temperatures were stable within each tasking level. The amount of heat rose consistently higher throughout

the experiment as did the delta T between the entrance and exit fluid lines. However, the temperatures and heat values were lower than expected. This was in part due to the periodic equipment failures (grinder motor slowdown and shutdown, kinks in the tubings, air entrainment, CPU throttling, bubbles in the lines) wherein we had to piece the final graph from multiple runs taken over several noncontiguous days. Since there were many incidents of CPU protection throttlings and eventually failure, it is possible that the software program was no longer able to task the CPU to 100% after a certain point. Finally, another reason for the lower values was that although the manifold helped to direct the highest quality heated oil from the CPUs, there was undoubtedly still mixing amongst the rest server blade components dissipating lower quality heat. This had a two-sided effect. While the oil evenly distributes the heat and stabilizes the temperatures and thus the system, the mixing of higher quality heat with lower quality heat compromises the higher quality heat across the collection ports. The ideal would be a steady collection of thermal energy from localized hot zones or a mineral oil bath setup with industry grade equipment as in the case of "the mineral oil PC" from Puget Systems. With this setup, a temperature of 88°C was not only achieved but maintained with absolutely no problems in stability at over a month of usage at those temperatures [50]. This proves that despite our equipment failures, the mineral oil concept is viable and can actually match our chiller's temperature design point (Table 1). Furthermore, the power coming in to this system from the wall was measured to be 820 W or 82% of maximum power [50]. This translates into 410 W for our particular 500 W power supply which is nearly the same or 410 W_{th} dissipated. Compared to the water-cooled experiments, full realization of this thermal energy would drop the number of ASU servers required for minimum chiller

activation to 11.3 kW / 0.41 kW = 29 servers. At the chiller design point of 50.2 kW_{th} of heat medium input, that would be 50.2 kW_{th} / 0.41 kW_{th} = 123 servers, making this experiment the most viable to run our 10-ton Yazaki LiBr absorption cooling unit.

3.11. Conclusion

This work has successfully demonstrated and illustrated a triad relationship among CPU utilization, temperature and power. Similar efforts by others [15, 52] included two parameters such as temperature versus time and power versus time but did not take into account all three parameters of CPU temperature, power and loading simultaneously. It is important to understand this complete triad relationship to more effectively exploit CPU waste heat and optimize it for driving an absorption cooling system.

By attaching thermocouples, we were able to take direct readings of the CPU case temperature and an indirect measurement of the thermal power. CPU temperatures reached as high as 97°C, which surpasses the chiller minimum heat medium inlet temperature requirement of 70°C. The highest CPU thermal power captured was 84W_{th}, which translates into 135 CPUs for the required minimum input heat of 11.3kW_{th} to engage the 10 ton Yazaki lithium-bromide absorption chiller. Since the data center houses at least 300 dual-processor servers, it is possible to generate enough energy to power the chiller.

For a steady power source, both parameters of enough heat and at a high temperature need be present simultaneously and sustainably. Although we were able to repeatedly achieve this high quality heat with the CPU experiments and water-cooled

heat sinks, we were not able to sustain these maximum temperatures for a steady thermal energy supply to the 10 Ton commercial LiBr absorption chiller. The primary reason for this was the limitations inherent in the equipment we used. We employed an IBM eServer xSeries 336 1U standalone server blade with dual Intel Xeon 3.6 GHz processors in our research, which are quite obsolete as compared with current server technology. In addition, some of the measuring equipment, a portion of the heat capturing equipment, and thermal interface material proved substandard at several critical points. For example, we only had thermal paste at our disposal, and, because we lacked automatic equipment, we had to apply the thermal paste manually to the CPU. Thus, we were unable to fill in enough of the thermally resistant interstices between the CPU and coldplate. We also had substantial heat loss due to inadequate insulation and recurrent leaks. Although applying pressure between the coldplate and CPU helped increase the heat capture fraction, the air gaps contributed to a high thermal resistance and insufficient insulation led to a significant difference in ΔT of 26°C across the TIM (ideally, we would like to see this value under 10°C). This resulted in an ineffective method of transporting what was otherwise a sufficient amount of high quality heat energy.

Given the proper resources, the transfer of thermal energy collection amount and quality can be improved by reducing the contact resistance between the coldplate and the CPU. Gwinn [60] outlines two ways to reduce contact resistance: (1) increasing the area of contact spots, accomplished by (a) applying sufficient pressure to raze any peaks of microroughness, or (b) reducing the roughness of the surfaces before the interface is formed by grinding the surfaces to remove non-flatness and then buffing them to diminish micro jaggedness; and (2) using an ideal TIM. For our purposes, an ideal TIM would

consist of an interstitial substance exhibiting the following characteristics: easily deformed by minimal contact pressure to fill in all gaps of both mating surfaces including surface pores, high thermal conductivity, wafer thin yet cohesive enough to remain within the interface (no pump out), and easy to apply and remove [60]. Unfortunately, load constraints on electronic components make it unfeasible to use high contact pressure, and manufacturing highly finished surfaces is not practical due to cost constraints [60].

Despite equipment limitations, we were able to produce and capture CPU-generated high quality heat. This allowed us to meet the goal of demonstrating a unique and useful triad relationship among CPU and system temperatures, thermal power, and CPU tasking levels. The aim was to supply an optimal and steady range of temperatures necessary for a small commercial chiller to operate. While we were able to provide an optimal temperature point, it was not sustainable with the supplied CPU and laboratory equipment. However, the control tests on a well-insulated copper test block with known heat values simulating a CPU showed that a steady capture fraction of high quality heat *can* be achieved. Also, by capturing and removing from the system enough of the dissipated heat, an additional benefit of eliminating the need for electricity-consuming fans was realized.

Among the non-control liquid-cooling experiments, the CoolIT© Eco scheme came the closest to capturing the required quality of heat. Improvements in the TIM and line insulation would certainly improve thermal energy transfer but only to a certain extent. Although it would be ideal to decrease and even eliminate the contact resistance by employing a perfect TIM between the CPU and coldplate, realistically, the mere existence of a TIM material adds a thermal conduction resistance across a measureable

thickness [60]. Therefore, as discussed earlier, the most practical alternative and best approach is to remove the TIM entirely to directly access CPU dissipated heat as with mineral oil immersion. Unfortunately, past idle conditions, the mineral oil experiment that we conducted did not yield the expected high values due to equipment and time concerns, restraints and limitations. However, direct heat access, an absence of contact resistance, and a relatively simple setup suggest that an immersion approach still has the most potential for capturing the heat dissipated by the CPUs at the optimum temperature to serve as input for a commercial chiller. Validating this notion is Puget Systems' 'mineral oil PC' [50]. With their simple design and industry-grade testing and monitoring equipment, Puget Systems completely submerses their Xeon server CPU in technical-grade mineral oil, and achieve an 88°C steady state oil temperature after 12 hours of run time [50]. This lends tangible support that mineral oil submersions have a promising future.

CHAPTER 4

THERMAL AND INFORMATIONAL EXERGY ANALYSIS OF CPU WASTE ENERGY

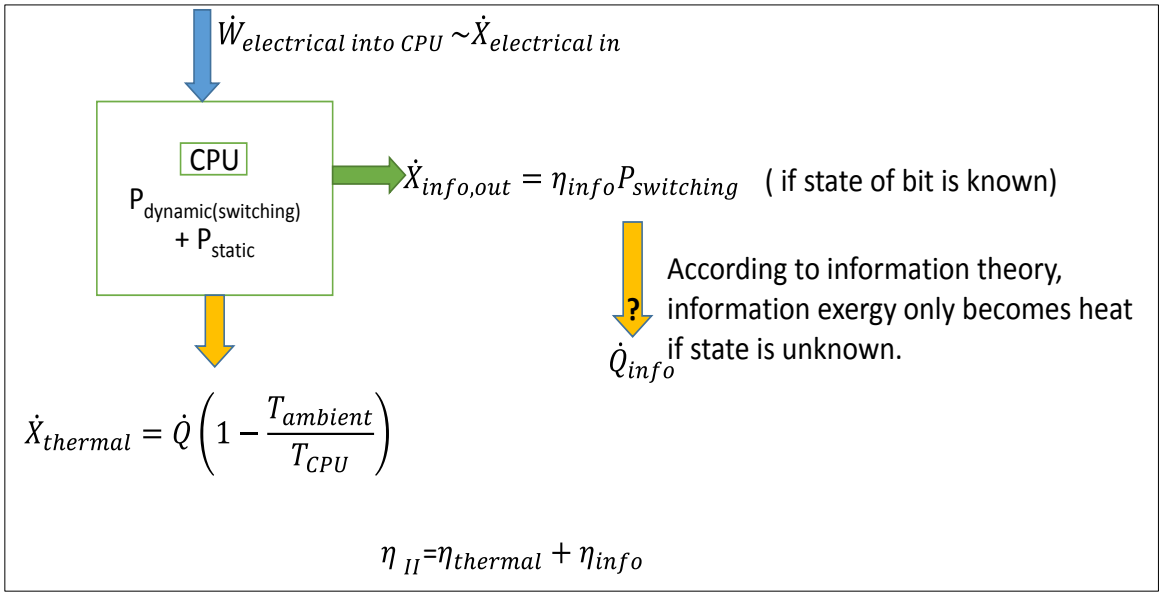

Figure 21. Exergy flow diagram of Intel® Xeon Irwindale 3.6 GHz CPU with 2MB L2 Cache.

This chapter explores the exergy of our specific Intel® Xeon CPU and thereby determines the maximum amount of energy utilizable for work. Not only does this study examine the thermal exergy available from CPU utilization but also the informational exergy offered from processing a binary digit (bit) within the CPU. Unlike the energy balance inherent in the First Law of Thermodynamics, exergy output will not balance the exergy input for real processes. Instead, exergy accounts for the irreversibility of a process and is always destroyed when a process involves a temperature change. [57] According to definition, exergy represents the useful part of energy for a system that

differs in equilibrium from its environment, i.e. the maximum quantity of work that the system can execute until it reaches equilibrium with its environment. [71].

4.1. Thermal Exergy

Useful work can be divided into several categories including mechanical and electrical work and heat delivered for actual use [72]. To analyze the maximal available work for this specific study, a reversible heat engine is utilized that operates between the temperature of the CPU as the hot source and the temperature of the environment as the cold reservoir. The maximal available work or exergy transfer by heat is then

$$\dot{X}_{th} = \dot{Q}_{max}\left(1 - \frac{T_o}{T_{CPU}}\right) \tag{24}$$

where T_o is the cold reservoir temperature of the environment and T_{CPU} the hot reservoir junction temperature of the semiconductor device synonymous with processor core temperature [36]. This exergy analysis is important because it tells us how much productive work can be accomplished by using the maximum available thermal power possible. It is important to realize that exergy represents the upper limit on the amount of work that a device can potentially deliver, not necessarily what it will actually deliver upon installation [73]. This caveat is critical in the following implementation of equation (24) because the limitations in our heat-capturing equipment restricted what we could capture in terms of thermal dissipation from the CPU. For example, the most any one of our heat extraction units could capture was 76% of TDP or 84 W_{th} (Figure 14).

From its Thermal Profile B, our 64-bit Intel® Xeon™ CPU with 2MB L2 cache has a TDP value of 110W at a $T_{CASE\_MAX\_B}$ of 79°C [55]. However, the TDP value from Intel® is not the maximum CPU power output possible nor is $T_{CASE\_MAX\_B}$ the highest temperature possible [74]. In reality, a specified chip's TDP has less to do with the amount of power a chip can potentially use and more to do with the amount of power that must be dissipated by the computer's fan and heatsink while the chip is under sustained load [75]. Actual power usage can be much higher than TDP, so the specification is intended to give guidance to engineers designing cooling solutions [75].

4.1.1. T_{CPU} Directly Scales With CPU Utilization

It is known that CMOS (metal oxide semiconductor) technology microprocessor chips can reach a maximum allowable temperature of 125 °C also known as T_{JMAX} or T_{CPU_MAX} synonymous with the maximum processor core temperature [76]. Although they can be close in value, note that the junction or CPU temperature, T_{CPU}, is not the same physically as the case temperature of the IHS, T_{case}. An incorporated on-die thermal diode measures directly the CPU temperature whereas the case temperature is measured at the geometric top center of the IHS [55]. In operation, T_{CPU} is greater than T_{case} by the amount of heat transferred from the junction to case ($\dot{Q}_{JC}$) multiplied by the junction to case thermal resistance (R_{JC} in KW^{-1}) as shown in equation (25):

$$\Delta T = \dot{Q}_{JC} \cdot R_{JC} \tag{25}$$

and $\Delta T = T_J - T_{case}$ where $T_{CPU} = T_J$ [77].

As per the following analysis, T_{CPU} directly scales with CPU utilization. If CPU utilization is held constant, there can be no variance in T_{CPU}. Beginning with equation (25), T_{CPU} can be determined from

$$T_{CPU} = T_{case} + \dot{Q}_{JC} \cdot R_{JC} \quad (26),$$

where T_{CPU} is directly dependent on the value of T_{case} and the product of $\dot{Q}_{JC} \cdot R_{JC}$ is very small [78].

The diagram in Figure 22 below illustrates the heat flow.

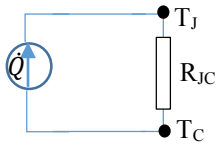

Figure 22. Model of heat flow from processor junction (CPU) to nickel over copper casing (IHS) [78].

The T_{case} values and power values are provided by an Intel datasheet along with the material of the IHS. Then, applying Fourier's Law for heat conduction, the following equation can be derived to find R_{JC}, and is valid as long as the parameters of thickness (z) and thermal conductivity (k) are constant throughout the sample:

$$R_{JC} = \frac{z}{A \cdot k} \quad (27)\ [77].$$

Since the IHS is composed of two different metals (nickel over copper) with uniform but different thermal conductivity and thickness values, I calculated each thermal resistance separately using conductivity tables and mechanical drawings with the thickness for each metal from Intel [55]. Although the thermal conductivity of nickel does vary some depending on temperature, the short temperature range of the CPU renders negligible an effect of ±0.0001 °CW^{-1} on the R_{JC} value of ~0.00808. Finally, putting the calculated R_{JC} into equation (26) showed that T_{CPU} and T_{case} exhibit less than a 1°C difference until about 42% CPU utilization level. By 100%, T_{CPU} rose to 2°C over T_{case}.

Now, to find the maximum possible CPU power possible for our 64-bit Intel® Xeon™ processor, we can utilize the linear relationship presented in the processor's Thermal Profile B as shown in equation (28):

$$T_{case} = 0.32 \cdot P_{n\%} + 43.4 \quad (28),$$

where T_{case} is a linear function of CPU Power, ($P_{n\%}$) [55]. Thermal Profile B was selected due to the fact that the heatsink is volumetrically constrained at a dimension of 1U [55]. With the T_{CPU} set as T_{JMAX} = 125°C, T_{case} in equation (26) can be expressed in terms of CPU power as:

$$T_{CPU} = (0.32\dot{Q}_{JC} + 43.4) + \dot{Q}_{JC} \cdot R_{JC} \quad (29).$$

Finally, to determine our maximum CPU power at maximum CPU core temperature, we can use

$$\dot{Q}_{JC} = \frac{(T_{CPU} - 43.4)}{(0.32 + R_{JC})} \quad (30).$$

and replace numerical values for T_{JMAX} and R_{JC}~0.00808. At a maximum allowable processor temperature of 125°C verified in Figure 4, the CPU could potentially supply 248 W_{th}. This maximum possible power value allows for significantly higher exergetic content than at the standard 110W_{th} TDP and corresponding 79°C case temperature [55].

With power values determined at the extremes of no load and maximal 100% loading, it would be useful to determine power values for each CPU utilization level in between 0% and 100%. This importance is highlighted by the fact that the microprocessor is the only component of the server system that has a marked effect on system level power draw based on its utilization [79]. The CPU architecture has been optimized to enable large parts of the silicon to shut down when in idle states and as the server power scales linearly with CPU utilization, it makes sense that the processor power scales linearly as well [79]. According to the green grid® and Liu et al., an estimate of power draw at any specific CPU utilization ($P_{n\%}$) can be calculated using the following formula:

$$P_{n\%} = (P_{max} - P_{idle})\frac{n}{100} + P_{idle} \quad (31),$$

where P_{idle} corresponds to 22W_e drawn at no load and P_{max} to 248W_e in the CPU's fully utilized state [79, 14]. This formula shows that CPU power scales linearly with CPU

utilization. If CPU utilization ($n\%$) is held constant, then CPU power, ($P_{n\%}$), must also be a constant (since P_{max} and P_{idle} are always constants). Replacing $P_{n\%}$ in equation (28) with a constant value then yields a constant T_{case} and hence a constant T_{CPU}. Therefore, holding CPU utilization as constant would produce an η_{II} versus T_{CPU} graph with only one point. The reason is that T_{CPU} depends on CPU utilization.

Equation (31) has proved very useful and accurate in modeling power consumption since other components' activities are either static or correlate well with main processor activity [14]. Additionally, making use of equations (28) and (29) enables T_{case} and T_{CPU} to be determined as well for each integer utilization level from 0% to 100% as shown in Figure 23.

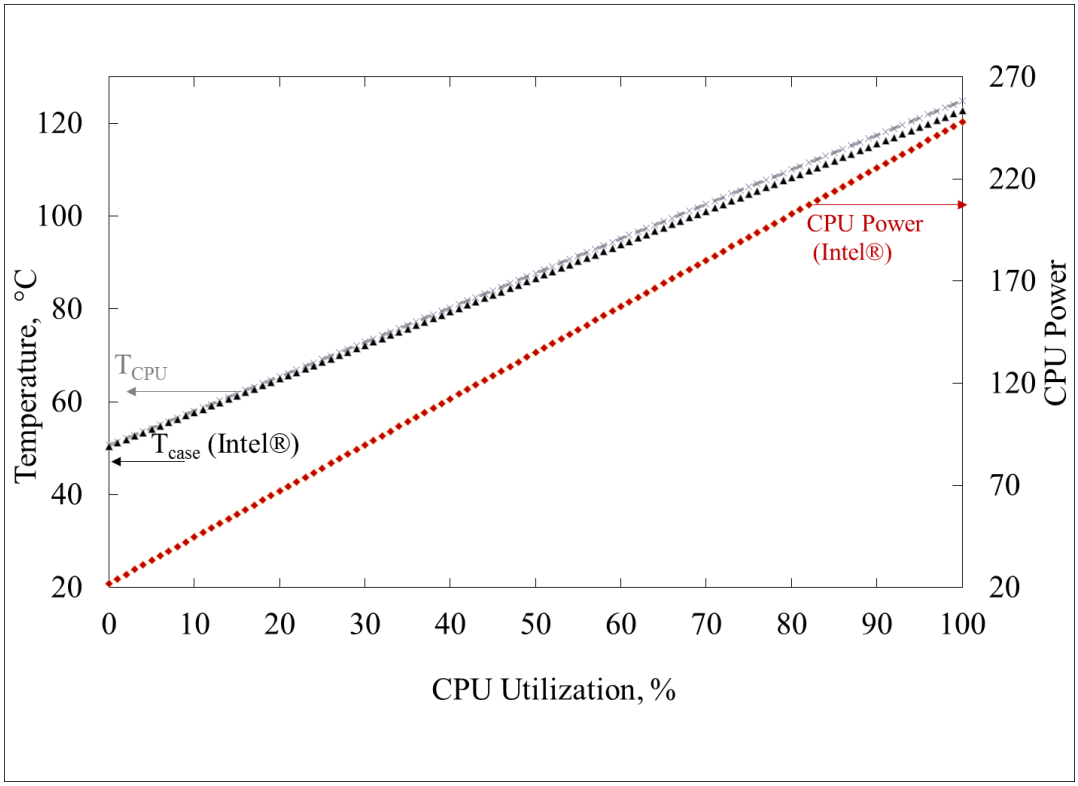

Figure 23. Graph plotting case and CPU die temperatures and processor power values, employing Intel's datasheet values and the Thermal Profile B linear equation [55]. The power estimation formula from the green grid® and Liu et al. formed the basis for the interpolated data. Maximum CPU power possible at the T_{CPU} temperature limit of 125°C is 248 W [79, 14].

4.1.2. Exergy Transfer by Heat, $\dot{X}_{th}$

To analyze the maximal available work for this specific study, a reversible heat engine is utilized that operates between the temperature of the CPU as the hot source and

the temperature of the environment as the cold reservoir. The maximal available work or exergy transfer by heat is then [36]

$$\dot{X}_{th} = \dot{Q}_{max}\left(1 - \frac{T_o}{T_{CPU}}\right) \tag{32}$$

where T_0 is the cold reservoir of the environment and T_{CPU} the hot reservoir junction temperature of the semiconductor device synonymous with processor core temperature. This exergy analysis is important because it indicates how much productive work can be accomplished by using the maximum available thermal power possible.

With the power values for all utilization levels calculated and the linearly related T_{case} and T_{CPU} values as well, it is a straightforward process to replace these values into equation (24) for transfer of heat by exergy, $\dot{X}_{\text{th}}$. This same process can be applied to the Intel® datasheet values, to both water-cooling experiments, and to the mineral oil experiment with the resulting graph displayed in Figure 24.

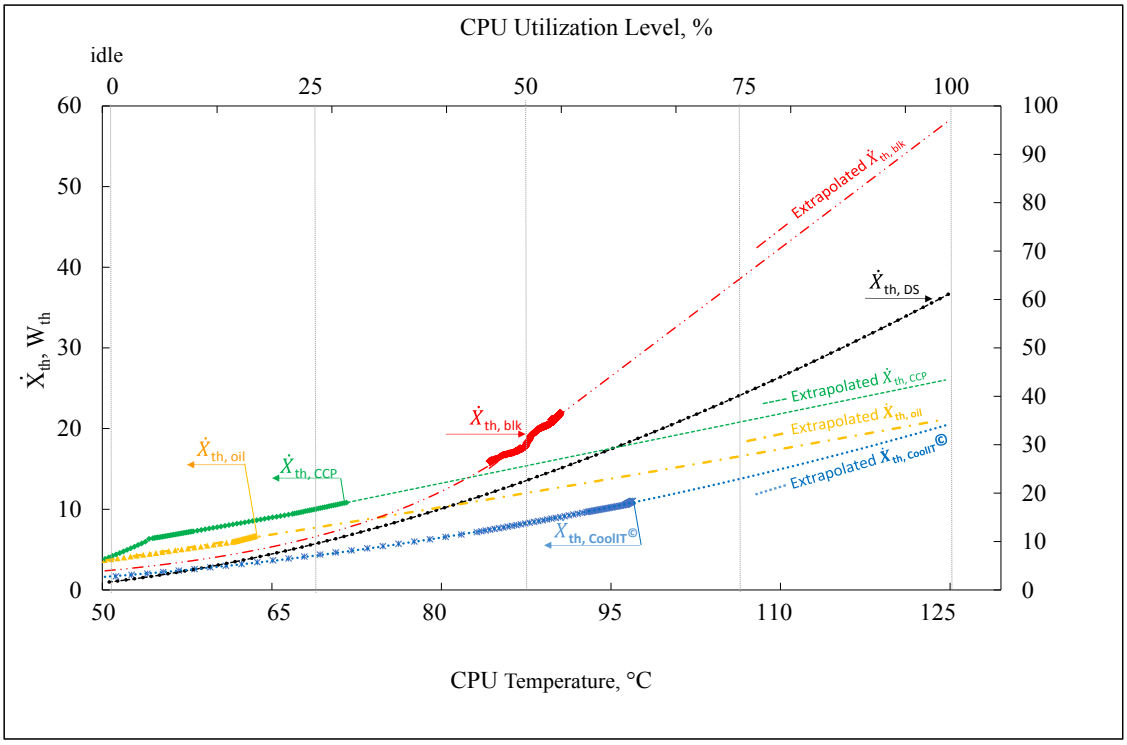

Figure 24. Exergy graph showing thermal exergy power datasheet values $(\dot{X}_{th,DS})$, the heater block experiment $(\dot{X}_{th,blk})$, the two water-cooled experiments $(\dot{X}_{th,CoolIT©})$ and $(\dot{X}_{th,CCP})$, and the mineral oil experiment $(\dot{X}_{th,oil})$versus heat source temperature. Since the temperatures varied for each experiment, the datasheet CPU temperature range (50-125°C) served as the reference for comparison among all the data sets. To make this comparison one-to-one, the experimental data sets were extrapolated beyond their measured values to correlate to this baseline temperature range. The $\dot{X}_{th,blk}$ yielded the most exergy since its capture fraction was greatest (Figure 16) and it had the highest quality heat.

For thermal exergy, the maximum power possible is assumed hence the full CPU power consumption possible extrapolated from Intel datasheet values using Thermal

Profile B. Following 25% CPU utilization, the Intel® data with an assumed lossless heat capturing system have notably higher exergy transfer values, but there is one instance during idle where the heat collected was higher as in the mineral oil. Still, the usable range for the chiller is really after $T_{CPU} \geq 70°C$. The thermal exergy transfer for the liquid-cooling experiments would most likely be higher if we had better heat capturing equipment and perhaps more run time. For example, with their industry-grade equipment and also a Xeon server CPU, Puget Systems' "mineral oil PC" reached 88°C steady state oil temperature after 12 hours of run time [50]. Furthermore, the system was stable and did not crash when running stress tests for 48 hours at these temperatures with a 820W draw (½*820W * 70% per CPU) draw from the wall [50]. A server blade typically would be even higher, so we would only need at most 50.2 kW/0.574 kW = 88 servers to supply the chiller at its design point with this high quality heat.

4.2. An Introduction to Informational Exergy

Within science, information is just as fundamental as energy and matter [80]. Information theory provides a definition for an important connection between exergy and information. Information theory is a branch of applied mathematics, electrical engineering, and computer science involving the quantification of information and was developed by Claude E. Shannon to find the fundamental limits on data processing operations [81].

John von Neumann first asserted that the "thermodynamic *minimum* of energy per elementary act of information" was $kT\ ln2$, where k is Boltzmann's constant, 1.38054×10^{-23} JK^{-1}, and T is expressed in absolute temperature [82] . In 2000, Meindl and

Davis proved this result for an ideal SETE MOSFET (single electron thermal energy metal oxide semiconductor field effect transistor) at the limit of its capacity for binary signal discrimination (i.e., assuming the minimum possible gate charge of a single electron, q) [83]. For an ideal MOSFET, the minimum allowable supply voltage at which it can still operate has been theoretically determined to be

$$V_{min} = \frac{2ln2 \cdot kT}{q}, \tag{33}$$

where q=1.602$\times$ 10^{-21} J and T is absolute temperature in K [83]. Thus, this minimum supply voltage delivers 36 mV to the MOSFET, half of which is dissipated as heat and half becomes available as stored energy on the driven load capacitor for switching states [84]. Then, the available switching energy on the load capacitance of a single MOSFET is

$$E_{sw} = \frac{1}{2}C_{eff}V^2_{DD} \tag{34}$$

or, since $C_{eff} = \frac{q}{V_{DD}}$, in terms of charge,

$$E_{sw} = ½\ q\ V_{DD} \tag{35}$$

where C_{eff} is the effective load capacitance, q=1.602× 10^{-21} J, and V_{DD} is the supply voltage that enables the MOSFET to operate [84].

4.2.1. X_{info} compared with P_{sw} and minimum power

Replacing V_{DD} in (35) with V_{min} in (33) results in

$$E_{sw,min} = kT\ ln2 \tag{36}$$

which is the *minimum* signal energy transfer during a binary switching transition of a MOSFET [83]. Thus, this theoretical minimum energy available to process a fundamental unit of information, i.e., to do useful work, is the minimum informational exergy or $X_{info,min}$ = 2.9 × 10^{-21} J. This analysis is further supported by information theory which firmly establishes the connection between exergy, energy and information as

$$X_{info} = E_{sw} I \quad (37),$$

where X_{info} is the informational exergy, E_{sw} is the energy available to switch states, and I the information in binary digits (bits) [80]. Thus, the minimum informational exergy contained in one bit is

$$X_{info,min} = E_{min} \quad (38),$$

where $E_{min} = (kTln2)$ and k is the Boltzmann constant 1.38×10^{-23} m^2 kg s^{-2} K^{-1} [80, 83].

Therefore, at room temperature, the amount of exergy contained in 1 bit of information is about 2.9×10^{-21} [J]. This order of magnitude is typical as information usually has a very small exergy value. Exergetic power of information recovered can be calculated by multiplying by frequency and this is defined as $\dot{X}_{info\_out}$.

In contrast, $\dot{X}_{info\_in}$ is the exergetic power equal to energy per the time for one switch or switching power, which is directly related to the available energy to process bits as

$$P_{sw} = E_{sw} f \quad (39).$$

In terms of information exergy, this relationship would indicate that P_{sw} is equivalent to $\dot{X}_{info\_in}$, the actual exergetic power that is supplied to generate 1 bit of

information. However, that single bit of information ideally could be produced from only the minimum power of $P_{sw\_min} = E_{min}f$, and this minimum power is what is equivalent to $\dot{X}_{info\_out}$. Informational exergetic efficiency can now be defined as η_{info}, the ratio of $\dot{X}_{info\_out}$ recovered to supplied $\dot{X}_{info\_in}$,

$$\eta_{info} = \frac{\dot{X}_{info\_out}}{\dot{X}_{info\_in}} \quad (40).$$

For an ideal system, $\dot{X}_{info\_in} = \dot{X}_{info\_out}$ and only the minimum power need be supplied to generate 1 bit. However, because of electrical resistance in practice, much more than the minimum power is required to process a bit or $\dot{X}_{info\_in} >> \dot{X}_{info\_out}$, resulting in low information exergetic efficiency in the range of 10^{-11}.

4.2.2. Determining Informational Exergy for a Real Transistor

The workhorse of information creation is the transistor. When transistors change states, there is a switching power required to charge and discharge the load capacitances [84]. About half of the energy delivered from the power supply is available for these transistors to switch states, thereby creating the fundamental units of information commonly known as binary digits (bits). Physically implemented with a CMOS two-state device such as a MOSFET, a bit can have one of two values usually represented by a 0 at the low state and 1 at the high state.

Although MOSFETS have been demonstrated operating with power supplies under 100 mV, these do not actually minimize energy in a real CMOS process [84]. At this extremely low switching energy, they run so slowly that the leakage energy

dominates [84]. Therefore, the true minimum energy point is at a higher voltage that balances switching and leakage energy [84]. For example, a single real MOSFET transistor in our 90 nm 1.25 V process draws a minimum of 2×10^{-16} J from the supply when switching for an available information exergy of half that or 1×10^{-16} J. This energy value compares favorably to [84] on page 201 and was calculated from equation (34), in which $V_{DD} = 1.25$ V and the effective capacitances are determined from $C_{eff} = \alpha C$ and the values provided in Table 4.

Table 4.

Parameter values from two sources to calculate $P_{switching}$ for the 64-bit Intel® Xeon® Processor 3.60 GHz, 2M Cache, 800 MHz FSB code named Irwindale [84, 23]

Lithography	**Transistor type**	**Transistor count (169 million tot.)**	**Avg width (λ = feature size/2)**	**Activity factor, α α_{max}=0.5**
90nm	1. Logic transistors 2. Memory transistors	1. 10% or 16.9 million 2. 90% or 152.1 million	1. 12 λ 2. 4 λ	1. α=0.1 (typical) 2. α=0.05 (typical)
Die size: 135 mm² (1.35 cm²)				
$P_{sw} = \frac{1}{2}\alpha CV^2_{DD}f = \dot{X}_{info}$	**C (gate+diffusion+wire)**		**V^2_{DD}**	***clock freq***
	1 fF / μm (gate)	0.8 fF / μm (diffusion) 0.2 fF / μm (wire)	1.25 V	3.6 GHz

While $X_{info,min} = 2.9 \times 10^{-21}$ J is the theoretical *minimum* exergy out contained in a bit from a SETE MOSFET, we want to determine the *maximum* exergy in from a single real MOSFET transistor on our particular CPU. Active power is the power utilized for the

chip to carry out useful work, and it is often dominated by dynamic or switching power, P_{sw} [84].

In terms of exergy, only half the energy supplied is available for the useful work of switching states and thus processing bits of information [84]. Therefore, the exergetic power, $\dot{X}_{info\_in}$, of a MOSFET transistor on a CPU happens when the circuit is actively switching the load and creating bits of information so that

$$\dot{X}_{info\_in} = \frac{1}{2} C_{eff} V_{DD}^2 f \quad (41),$$

or

$$\dot{X}_{info\_in} = \frac{1}{2} \alpha C V_{DD}^2 f \quad (42),$$

where α is the activity factor, C is the true capacitance of the node (transistor), V_{DD} is the supply voltage, and f is the clock frequency [76]. If the entire CPU is completely turned off, the activity factor and switching power go to zero [84].

4.2.3. Activity Factor and True Capacitance

The activity factor, α, is the probability that a transistor transitions from 0 to 1, because that is the only time that it uses power [84]. The clock frequency has an activity factor α = 1 because it rises and falls with every cycle, thus transitioning twice per cycle [85]. The activity factor is substantially lower for logic circuits, and most data have a maximum activity factor of 0.5 since it can transition at most once each cycle, either only on the rising portion or falling portion of the timing clock waveform [84]. Truly random data has an activity factor of 0.25 as it changes state only once every other cycle, a fourth

of the clock waveform [84]. “Static CMOS logic has been empirically determined to have activity factors closer to 0.1 because some gates maintain one output state more often than another and because real data inputs to some portions of a system often remain constant from one cycle to the next” [84]. By including these activity factors to expand equation (42) as shown below,

$$\dot{X}_{info\_out} = \frac{1}{2}\left(\alpha_{logic}\, C_{logic} + \alpha_{mem}\, C_{memory}\right) V^2_{DD} f \qquad (43),$$

true gate and memory load capacitances can be attained for a more accurate and precise calculation of the switching power for the transistors inside a particular CPU [84]. The correct capacitance for each type of gate is determined by multiplying together the number of capacitors (a percentage of the number of transistors), the average transistor width, half the feature size (λ /2), and the sum of the individual transistor capacitances (gate + diffusion +wire capacitance). With the parameter values of Table 4, and from an example that logic gates comprise about 10% (and memory gates the remainder) of the total number of transistors on a 100 nm process die,

$$C_{logic} = (10\% \text{ of } 169 \text{ million})(12\lambda)(0.045\ \mu m/\ \lambda)(2fF/\mu m) = 18.25\ nF \qquad (44)$$

and

$$C_{memory} = (90\% \text{ of } 169 \text{ million})(4\lambda)(0.045\ \mu m/\ \lambda)(2fF/\mu m) = 54.75\ nF \qquad (45)$$

for our particular 90 nm feature sized CPU [23, 84, 86] . Finally, inputting equations (44) and (45) into equation (43) along with our values from Table 4. yields the informational exergetic power as 12.8 W_{info} at the typical activity factor levels listed in Table 4. [23, 84]. This would translate into 0.3 μW_{info} and 0.05 μW_{info} per logic and memory gates, respectively. However, if the both transistor activity levels were at the maximum activity factor of α=0.5 for data, this would increase the exergetic power by an order of magnitude to 1.5 μW_{info} and 0.5 μW_{info} per logic and memory gates, respectively. That translates into a maximum total exergetic power in of 103 W_{info}, which is about 41% of the 248 W ceiling presented in Figure 23.

4.2.4. Maximum Single Bit Operations Per Second

Interestingly, equation (37) from information theory compares favorably to this maximum 103 W_{info} exergetic power by inputting the maximum number of bits that our CPU can process per clock cycle or SBOPs (single bit operations per second) at peak perfomance. The peak performance is the maximal theoretical performance a computer can achieve, calculated as the machine's frequency, in cycles per second, times the number of operations per cycle it can perform [87]. First, using the formula for computing the theoretical maximum performance of a CPU in GFLOPS (10^9 floating point operations),

$$\text{GFLOPS} = \#\text{cores} * \frac{GHz}{core} * \frac{FLOPs}{Hz} \tag{46}$$

where cores refer to the number of physical processors installed, GHz is the clock frequency, and FLOPs is the floating point operations per second [88]. In general, a core

can execute a certain number of FLOPs every time its internal clock ticks, which are called cycles and measured in Hertz (Hz) [88]. Most modern microprocessors can do four FLOPs per clock cycle, that is, 4 FLOPs per Hz [88]. For our particular server blade system, we have two cores, 3.6 GHz per core, and 4 FLOPs per Hz. Substituting these values into (46) then yields 28.8 GFLOPS or 14.4 GFLOPS per CPU core.

Using LINPACK, a software library for performing numerical linear algebra on digital computers, we were able to validate this 14.4 GLOPS theoretical number as a reasonable estimate [89]. A processor's computing performance measured by the LINPACK benchmark consists of the number of 64-bit floating-point operations, generally additions and multiplications, a computer can perform per second, also known as FLOPS [87]. Running Windows version of LINPACK on our machine, our experimental value of 12.4 GFLOPS comes within 14% of the calculated theoretical value thereby confirming 14.4 GFLOPS as a reasonable number. To obtain SBOPS, we can estimate one FLOP to be equivalent to 1000 SBOPS (single bit operations), so we simply multiply GFLOPS by 1000 to attain 14.4 TSBOPS or 14.4 trillion bits that can be processed per second at peak processing performance [90, 91, 92]. Finally, inputting these 14.4 trillion bits into equation (37) with $E_{sw} = 2.9 \times 10^{-21}$ J per bit computes to 41.76 nJ of available information exergy or $\dot{X}_{\text{info}} = X_{\text{info}} * f = 150$ W of exergetic power.

Taking the most conservative value from equation (43) of 103 W_{info}, we can scale this peak exergetic power in terms of percentage of transistors in operation corresponding to the CPU utilization range of 0% to 100% and the maximum available exergetic power and exergetic efficiency are shown in Figure 25. This means that, although the activity

factor is peak at α=0.5 (one signal per cycle), only a certain percentage of the 169 million transistors, corresponding to CPU utilization, are actively processing a bit.

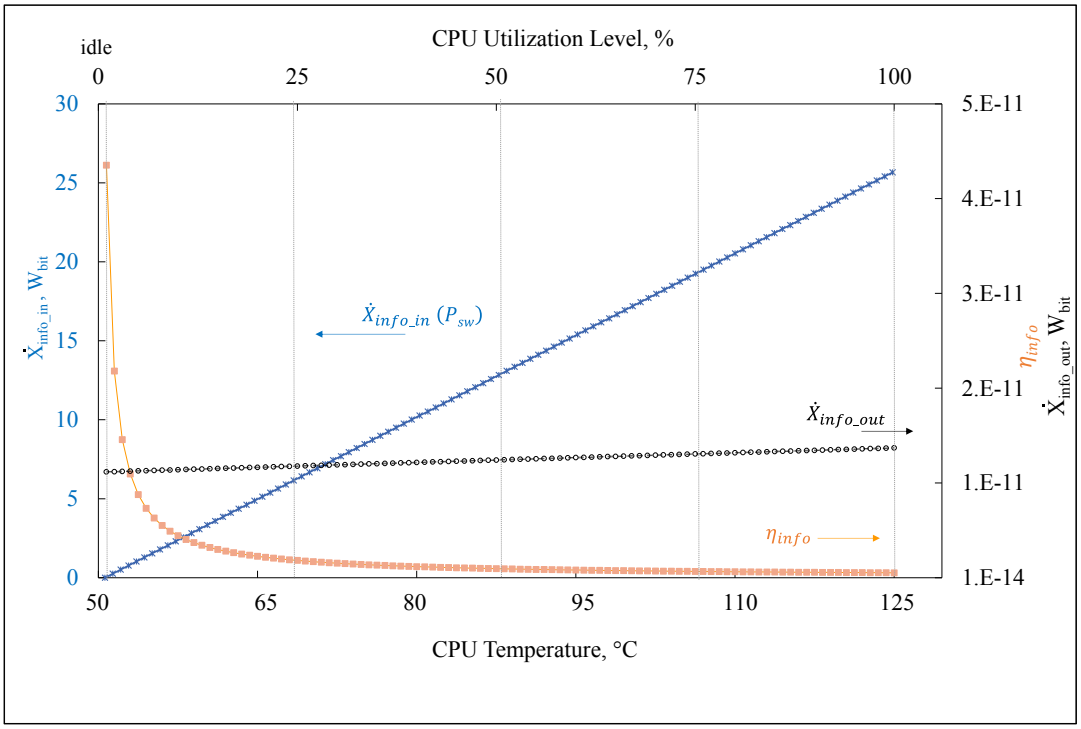

Figure 25. Exergetic power in to process bits (creating information) and exergetic power out inherent in those bits plotted against CPU temperature. Since the same processor was used in the $\dot{X}_{info}$ calculations, the plotted values are consistent for the datasheet and the liquid cooling experiments. Informational exergetic efficiency is η_{info} as the ratio of recovered $\dot{X}_{info\_out}$ to supplied $\dot{X}_{info\_in}$.

Information must be stored and transported safely. To reach this we must use redundancy (over-explicitness) in codes and in copying, which implies excess dissipation of energy to make the process irreversible (safe) enough [79]. Notice that η_{info} is a very

small value and this is typical due to the electrical resistance in real circuits. Also, notice the lower the temperature, the more efficient is the information transfer.

4.3. Exergetic Efficiency

To determine the overall exergetic efficiency, we again employ the Second Law of Thermodynamics with the results shown in Figure 26. The overall exergetic efficiency can then be defined as the sum of the individual exergetic efficiencies of η_{info} and η_{th}. This results in the following equation for Second Law exergetic efficiency, η_{II}, as:

$$\eta_{II} = \frac{\dot{X}_{info\_out}}{\dot{X}_{info\_in}} + \frac{\dot{X}_{th\_out}}{\dot{X}_{e\_in}} \quad (47)$$

or

$$\eta_{II} = \eta_{info} + \eta_{th} \quad (48),$$

where $W_e = \dot{X}_e$ [80].

4.3.1. η_{II} for an Ideal System

An ideal system would suggest that $\dot{X}_{info\_in} = \dot{X}_{info\_out}$ and $\eta_{info} = 1$ so that only the minimum power draw would occur to process each bit. Furthermore, since there would be no electrical resistance and thus no heat dissipation in an ideal system (a hypothetical superconducting CPU for instance), $\eta_{th} = 0$. Therefore, for the ideal case, the Second Law exergetic efficiency would be dependent only on η_{info} and at its maximum of $\eta_{II} = 1$. However, in practice for real systems at temperatures above

absolute zero, $\dot{X}_{info\_in}$ is never equal to $\dot{X}_{info\_out}$, since there would always be some form of electrical resistance and therefore heat dissipation. Also, since η_{info} is typically very small for real systems so that $\eta_{II} \sim \eta_{th}$ as shown in Figure 26.

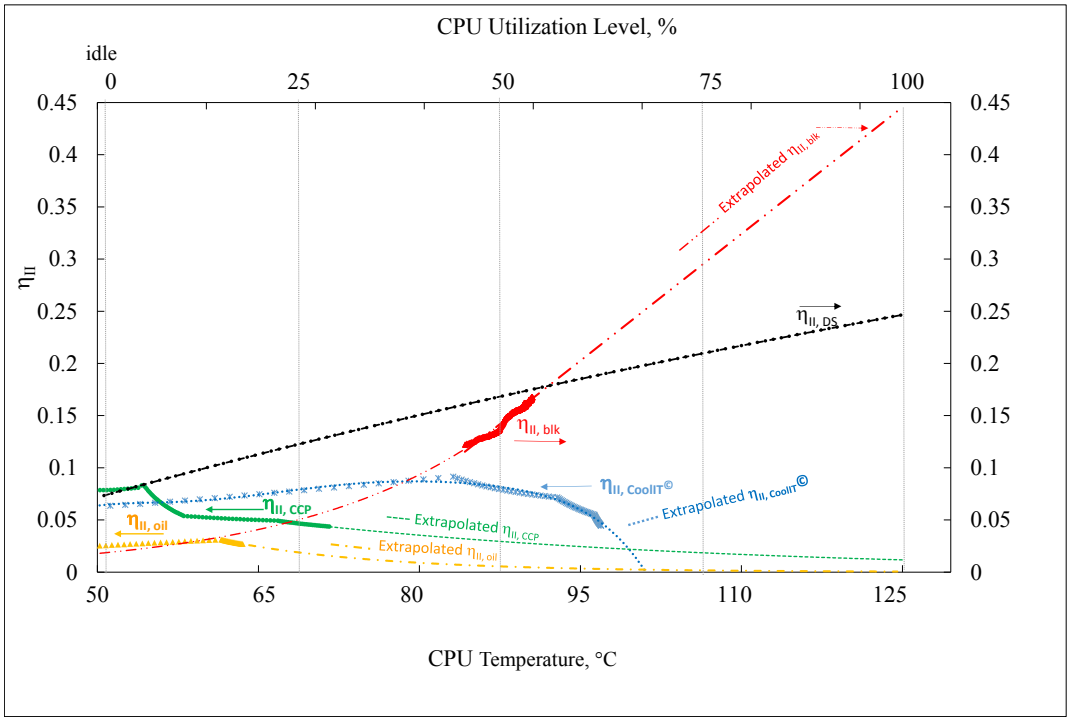

Figure 26. Second law exergetic efficiency plotted against CPU temperature. Since the values of η_{info} are many orders of magnitude smaller as shown in Figure 24, $\eta_{II} \sim \eta_{th}$. Of the three the experiments conducted on the CPU, note that the exergetic efficiency deteriorates for the experimental values. This is due to the increasing ΔT_{TIM} drops and the inability to capture that increased heat as the temperature of the CPU increases. The datasheet and heater block exergetic efficiency values continue to rise as expected.

4.3.2. Using the η_{II} to Optimize a System

As shown in Figure 24 and Figure 26, a reduced ΔT_{TIM} and higher capture fraction would be factors that would optimize a system and therefore increase η_{II}. Furthermore, to approach the ideal system of $\eta_{II} = 1$ we could utilize its underlying components of η_{info} and η_{th} to design and create a processor with low to no electrical resistance. Specifically, we could strive to match $\dot{X}_{info\_in}$ and $\dot{X}_{info\_out}$ as closely as possible which would simultaneously reduce heat dissipation. Other factors that the η_{II} results can point to for maximizing improvements in computer processor design would be electrical line lengths, feature size, and power draw. Essentially, it is all about lowering the Ohms, so that short connections, large areas and even superconductors are of interest. Of course, if no heat dissipation occurred, no chillers or even CRACs would be needed, and we would have solved the initial problem of 50% of data center electricity being spent on the cooling system, except for the expense of maintaining any necessary cryogenic temperatures.

The overall exergetic efficiency shows that improvement in the thermodynamic performance of the heat capturing system is vital, since less than 50% of the available exergy is netted. This is due mostly to the inability of the setup to capture all the available transferred heat. Exergy analysis can identify the need for efficiency improvements and point out the thermodynamics losses attributable to insufficient equipment. More far-reaching advantages of such analysis are the potential to evaluate green technology aspects such as environmental impact or sustainable development. Exergy efficiency can help manage growth of electricty consumption and is a cost-effective way to transition to a resilient and more sustainable economy with a lower carbon footprint [72].

CHAPTER 5

CONCLUSIONS

With the economy's gradual recovery (more server purchases), the arrival of cloud computing and the explosion of mobile devices continually accessing the internet, the response has been a rapid growth in the number and size of data centers worldwide. This growth is a problem as datacenters' rising energy demands are increasingly impacting electricity expenditure both in the U.S. at 2% and globally at 1.3% [5]. Approximately half of that electricity disbursement goes towards cooling [11]. This dissertation addresses both issues of growth in numbers of computer servers and subsequent electricity expenditure by using them to improve the sustainability of a data center. It proposes a solution by proposing, analyzing and testing a method of recycling waste generated by data center servers to provide a renewable energy source for use in cooling the data center. To the best of the author's knowledge, the approaches and methods toward improving data center sustainability set forth in this dissertation, such as integrating energy reuse into the PUE equation, graphing CPU temperature, power and loading simultaneously, the mineral oil experimental design and implementation, and the informational exergy analysis are distinct from other approaches and methods.

For one, the first stage in this dissertation's development takes the novel approach of utilizing CPU waste heat to drive an absorption cooling process. The donated heat-driven cooling system is a 10-ton (35.2 kW_{th}) LiBr-H_2O Yazaki absorption SC10 water fired single-effect chiller. The usable temperature range for this Yazaki absoption chiller is between 70 °C and 95 °C with its design point at 88 °C. Further examining chiller

performance for different scenarios under which it can successfully function illustrates a notable variance in COP_C values and cooling capacities. There is a general rule that the hotter the heat medium and the lower the cooling water temperatures, the more cooling capacity will be generated. That rule is certainly applicable here since at 95 °C, and with cooling water temperature lowered to 26.7 °C, the cooling capacity delivered is 49.4 kW_{th}, a substantial 40% over the design point cooling capacity of 35.2 kW_{th}. One tradeoff is efficiency: at higher temperatures, heat transfer in the steel coils is limited by the surface area and heat transfer properties of the heat exchange materials and COP_C falls. Then, at the cost of lowered efficiency, the cooling capacity rises to its maximum. Fortunately, since the thermal source under consideration is an abundance of free waste heat, cooling capacity can override efficiency as the main objective.

To elicit the maximum cooling capacity of 49.4 kW_{th}, the remaining tradeoff would be more server blades required to generate the 75 kW_{th} input heat at 95 °C necessary to power the chiller. At 432 server blades required (174 W_{th} per blade), that is more than the 300 or so available within the BlueCenter at ASU. However, this issue too may be resolved since the linear formula for the processor's Thermal Profile B within the Intel® datasheet suggests a 218 W_{th} power dissipation per CPU. This translates into 436 W_{th} per server blade for a 95 °C input temperature to the chiller, which includes the conservative estimate of a -20 °C ΔT delivery temperature loss. In addition, the informational exergetic power contributes 12.8 W_{info} at the typical activity factor level. With 59% of each blade's heat assumed capturable based on a similar cold plate system used by IBM® [11], that equates to a sum total of 264 W_{th} per server blade or a 56.7%

reduction in the number of server blades required or 188 blades to operate the chiller for its maximum possible cooling delivery of 49.4 kW_{th}.

A pivotal contribution from this dissertation is the critical analysis and careful reworking of the PUE equation to take into account the involvement of the absorption cooling process plus any supplemental solar thermal input. Since PUE is defined with respect to the power read at the utility meter, it is perceived that a renewable energy source could affect this input facility power value by potentially reducing it. Theoretically, the meter could even show a negative number through exportation of electricity from a renewable source back onto the grid such as with a solar system. In our case, this exportation would be in the form of cooling. Nonetheless, the point is that these previously unaccounted-for energy sources make it possible to lower the PUE number by subtracting the absorption cooling contribution and waste heat reuse from the facility power in the numerator of the PUE equation. Even more, with solar thermal supplementation, the PUE metric can be a value lower than one. A fractional PUE value results in a data center infrastructure efficiency (DCiE) that is greater than 100% [38].

The experimental stages of this project examined the source of high quality waste heat within a data center, the CPUs on each server blade. Initial testing took place on a borrowed Dell® server blade with Xeon Nocona processors so the assessments had to be conducted cautiously to maintain the computers in working order. This initial testing phase focused on temperature determination and Dell's® OMSA software showed the CPU temperature range to be from 10.0 °C up to 125.0 °C. Shutting off the fans and fully tasking the CPU confirmed this top end. Unfortunately, the chiller functions within a much more narrow range, and since the CPUs within a datacenter can often be idle (~30-

40 °C), it would be advantageous to have a cooling architecture that can be energized at the lower end of the CPU temperature spectrum.

Subsequent experimental phases delved into how much CPU waste heat can be captured. We could not conduct this set of more rigorous experiments on the same borrowed Dell® server blades as in the initial phase of temperature evaluation. Instead, we were supplied with several donated IBM® server blades containing Intel® Xeon Irwindale CPUs. Fortunately, the temperature ranges turned out to be the same. However, the TDP number for the Irwindales is 6.7% higher than for the Noconas and the thermal profiles are different.

Utilizing the linear formula within the Intel® datasheet for the Irwindale processors, we calculated what the power levels should be and used that as a comparison point against our experimental data. Unfortunately, the outdated 90 nm technology CPUs and obsolete server blade motherboard predated both DTS and PECI, so there was no accessible means to measure power directly. Instead, the measured flow rate of the coolant and the measured temperatures across the cold plate were input into the heat transfer equation so that thermal power was indirectly calculated from our data. Having to rely on substandard equipment and dealing with issues such as high contact resistance and a lack of proper insulation, it is remarkable that we were able to collect useful data. With industry-grade monitoring and capturing equipment, the author believes that the data would show much higher values of temperature and heat, possibly even exceeding those in the Intel® datasheet. This confidence is based on the fact that, even with equipment limitations, two of the experiments already started off with idle power values higher than those specified in the Intel® datasheet [55].

The liquid cooling experimental data graphed in figures 13, 14 and 20 illustrate the fluctuating nature of the temperature and heat collection for the operating range of our CPU. This fluctuation indicates that, while not impossible, it can be difficult to consistently achieve high quality heat to drive a 10-ton absorption chiller needing 50.2 kW_{th} at 88°C to optimally run. However, the chiller *can* operate with a minimal heat input of 11.3 kWth at 70°C. This lower requirement offers more available opportunities for the limited heat extraction scheme used in the water-cooled CPU experiments due to the large ΔT_{TIM} drop observed between the CPU and cold plate. As indicated in Figure 13, there is a significant loss in heat quality shown as the ΔT_{TIM} between the CPU and CoolIT Eco© cold plate, which averages 53°C with a high of 68°C. This was primarily due to the inability of the CoolIT Eco© cold plate to closely fit the surface of the CPU which led to an excessive amount of contact resistance and thus thermal resistance for the subsequent heat capture.

To improve upon this inherent limitation within the commercial CoolIT© Eco cold plate, we designed and fabricated our own cold plate. The CoolIT© Eco gave us a baseline from which we proposed designing and fabricating an improved cold plate with a better fit to the CPU. We inspected the various dimensions and features of the CoolIT© Eco heat exchanger such as frontal area, fin height and base plate thickness and slant. From these inspections, we fabricated an optimized, innovative, and *better fitting* custom-built cold plate (CCP) by iteratively altering the base design.

From Figure 14, note that the CCP demonstrates higher heat capture values than the CoolIT© Eco as well as lower ΔT_{TIM} values between the CPU and the discharge line of the CCP cold plate. For the TIM, we again used Biostar® TC-Diamond Thermal

Compound. The 50% improvement in ΔT_{TIM} for the CCP setup over the CoolIT© scheme resulted in higher quality heat being transferred more effectively from the CPU to the discharge line of the CCP cold plate. Ideally, a ΔT_{TIM} of zero would allow maximum heat transfer to the chiller, but a ΔT_{TIM} loss of 10°C would still enable the CCP water-cooling setup to deliver, in its discharge line at 50% CPU utilization, the required minimum 70°C temperature to activate the chiller. Therefore, a ΔT_{TIM} under 10°C was a desired deliverable that had to be achieved. While the ΔT_{TIM} was still too high to consistently deliver the necessary quality heat to the chiller CCP, our redesigned cold plate was a significant improvement over the stock CoolIT© Eco, suggesting that further improvements to the system could further reduce the ΔT_{TIM} under 10°C so that consistent delivery of high quality heat would be feasible.

The additional improvements that we chose to implement were applying pressure to the cold plate and encapsulating the heat source within an insulated dewer. Since the CPU could not be encapsulated, we utilized a copper heater block as a heat source nearly equivalent to the CPU's maximum power dissipation. To test if applying pressure and encapsulating insulation would achieve the desired goal of ΔT_{TIM} <10°C, the CCP heat extraction equipment was ported onto this copper heater block with a known power dissipation value of 218 W_{th}. The focus here was on observing the effect of applied pressure and encapsulating insulation on ΔT_{TIM} and the resulting amount of high quality heat captured as shown in Figure 16.

For the experiments on the copper block in Figure 16, the ΔT_{TIM} loss between the heat source and cold plate was significantly lower than that of the CPU experiments. The

copper block ΔT_{TIM} temperature drop of 5.5°C was within the acceptable range of $\Delta T<10°C$, which means that the heat can be transferred with less temperature loss from the source to the cold plate discharge line and ultimately to the generator input of the absorption chiller. This ΔT_{TIM} temperature drop represented a 78% improvement over the CCP-CPU water-cooled experiment in Figure 14 and an 89% improvement over the CoolIT©-CPU experiment in Figure 13.

Furthermore, in the discharge line, the heater block could provide 203 W_{th} of heat at 83°C, which is a 63% increase in heat quality compared to the previous CCP system with no applied pressure or encapsulating insulation. To operate the chiller, 248 heat sources would suffice to drive it near its optimum design point of 50.2 kW_{th} and 88°C, which was not even possible with the previous setup. Thus, with applied pressure and insulation, the copper block experiment demonstrated the feasibility of achieving consistent and reliable delivery of high quality heat to the cold plate discharge line that would meet the requirements to near-optimally run our absorption chiller.

In addition, this improvement in higher quality heat transfer suggests that applying pressure to the cold plate affixed to the CPU would reduce thermal resistance by minimizing the interstitial gaps and thus providing more intimate contact between the CPU and cold plate. Even better would be to either solder the cold plate directly to the CPU or manufacture both together as one seamless unit. Thus, we can conclude that reduced interstices between the CPU and the cold plate result in a lesser drop in temperature between the CPU and the cold plate discharge line, as well as a higher quality of heat that can be transferred [60].

Since reducing the ΔT_{TIM} was so critical in obtaining the necessary quality of heat, eliminating the TIM entirely was another idea we tested. We did this by using liquid submersion cooling within a thermally, but not electrically, conductive fluid. We predicted that directly cooling the CPUs with mineral oil would bring to near zero the ΔT between the heat source and the oil in the discharge line, and the thermal resistance across the TIM no longer would be of concern. Based on this notion, we assembled a mineral oil immersion system for our server blade to submerse a server blade up to the level of the CPUs in a non-conducting liquid media. To capture the highest quality dissipated from the CPU, we placed a manifold directly over the CPUs to direct the CPU heat to the cold plate discharge line. The idea was that this dielectric fluid can then be circulated and delivered to a heat exchanger, providing the servers with cooling and supplying the Yazaki chiller with consistent heat energy. As with the previous CPU experiments, the processor was tasked in 25% increments up to its maximum utilization, and the results are shown in Figure 20.

With the numbers that we were able to achieve, the mineral oil experiment reduced the ΔT_{TIM} drop between heat source and discharge line by 61% over the CCP experiment and 81% over the CoolIT© experiment. Due to this reduction in ΔT_{TIM}, the heat quality temperatures in the oil discharge line were 12.3°C higher on average than for the CoolIT© experiment and only 8.7°C lower than for the CCP experiment. Interestingly, the heater block experiment still proved far superior in all respects exhibiting the lowest temperature drop at 5.5°C and the highest quality heat ($203W_{th}$ at 89°C) and capture fraction of 93%. This is most likely because the heater block employed

focused pressure and thorough insulation whereas the entire board immersed in mineral oil exhibited a spreading effect on the heat thus lowering the overall temperature.

While contact resistance was not a factor for the mineral oil experiment, the spreading out of the high quality heat despite the manifold we used lowered the quality of heat that we could extract from the system. To transfer as much focused CPU temperature and heat as possible, we had installed a manifold but it needed redesigning to be able to better focus on the CPU temperatures. We had planned several improvements to the system such as a stronger pump, a baffle and a better manifold. Unfortunately, the equipment had to be restarted several times during the course of the experiment (the two spikes in Figure 20) and soon failed due to external flooding circumstances so the planned improvements could not be implemented.

While we were unable to conduct the mineral oil experiment in the manner we had hoped, our data values were supported by research from the Royal Institute of Technology (KTH) in Stockholm using the Green Revolution Cooling (GRC) CarnotJet System. The GRC system immerses server blades in a bath of non-conductive (dielectric) oil with a water loop externally run to an evaporative cooling tower to reject the heat to the atmosphere. It is also possible to reuse the heat from the water as shown in the tests in Stockholm that reliably produced an end product of water at 50°C [93].

Via email correspondence on April 6, 2012, GRC confirmed that the ΔT was 22°C and could be lowered to 18°C with heat sinks designed for oil. This indicated a mineral oil temperature that could be 72°C. In parallel, KTH was in the process of achieving a final water temperature as high as 70°C which could be used to consistently stimulate an absorption chiller like ours at its low end. Another cooling application that

can take advantage of the relatively low but consistent temperature and heat output of mineral oil submersed CPUs is power for an adsorption machine such as the silica gel water pair 7.5kW Sortech unit that requires a lower range of 55°C to operate. Finally, such temperatures could simply produce hot tap water, a useful commodity all year around.

Of all the cooling schemes presented in this work, mineral oil is the only one with the potential to eliminate the need for the chiller and CRAC, and thus could be the ultimate solution to eliminate the 50% cooling expenditure which initially motivated this project. No CRACs or chillers are needed as the heat can be exported via the discharge line. Given that oil has a greater heat-absorption capacity than air, the oil can be run at a higher inlet temperature while still effectively cooling the computers. This increases server reliability. In the true spirit of sustainability, submersion data centers, as they are called, require significantly less infrastructure and build time than conventional data centers. Overall, there is a potential cooling power savings of 95% over conventional data centers with 10-25% reduction in server power, increased server reliability and a dramatic reduction in infrastructure costs [70, 71].

In conclusion, the data from the heater block experiment offer the most promise to drive an absorption chiller because there was more heat captured at a higher temperature. Moreover, the heater block experiment indicates that a CPU run at a medium temp of 89°C (28% below its maximum of 125°C) would allow enough high quality heat to be transferred to the cold plate discharge line for sufficient thermal energy to power a 10-ton chiller near its optimum level. The way to deliver this high quality thermal energy consistently is with server scheduling in a data center. That is, with server virtualization

now being common practice, data center scheduling can seamlessly switch between CPUs to consistently capture the ones outputting high quality heat at a target value.

5.1 Significance of the Work

My work has gained recognition among those conducting research on data center sustainability, as evidenced by several citations found in the scientific literature. In fact, the previously published portions of the work presented in this dissertation are now considered a contribution to the scientific community and are a part of the recognized literature on the subject of sustainable data center cooling technologies [94].

Gupta et al [95, 96] recognize this work as a step towards implementing a mini sustainable data center they refer to as a BlueCenter. They also designate the proposed LiBr absorption cooling architecture discussed in this dissertation as a segment of their BlueTool main project and call it BlueCooling.

Brady et al. also refer to my proposal of an absorption chiller for reuse of heat within a data center [97]. Furthermore, they cite my suggestion for an alternative metric to the PUE in support of their critical assessment of the limitations inherent in using PUE as a standard metric in comparing performances of data centers [97]. They point out that despite its limitations and ambiguousness the PUE continues to be incorrectly used and misused as a universal measurement of data center efficiency [97]. Nonetheless, there is a paucity of literature which critically assess the PUE. They imply that my work is one of the few that does [97].

Woodruff et al. discuss the reuse of waste heat within the context of a new strategy to surmount the challenges of data center power consumption and waste heat

production [98]. They label this strategy "environmentally opportunistic computing" (EOC) [98]. My notion of using DC waste heat to help power a cooling system is one of several concepts they cite as having begun to attract attention as others seek ways to reduce the high energy costs of running data centers, a great part of which comes from cooling.

In their case study of a Finnish data center's energy development performance, Lü et al. [99]stress that "new perspectives and methods aimed at both tackling the problems and exploiting the potential for energy efficiencies are needed." Among the new approaches they reference is my work on an alternative cooling architecture. They specifically cite it as a method that can be directly applied to their case study as a means to achieving a needed improvement in energy efficiency in the data center they studied.

5.2 The Way Forward

This dissertation has demonstrated that it is conceivable to recycle CPU waste heat to sustainably cool a data center. It is encouraging to see that this work has already served as a starting point for further research into data center waste heat reuse, alternative cooling technologies, and critical assessments of current data center metrics (PUE) [94, 95, 96, 97, 98, 99]. It is hoped that this work will continue to inspire further research into understanding and refining data center energy efficiency, achieving successful waste heat reuse and promoting data center sustainability on larger scales. For one, larger scale experimentation with industry-grade equipment would help to improve the numbers and efficiently transport optimum quality waste heat to power a heat-activated cooling system. Also, development of a more comprehensive and unambiguous performance

metric would better assess DC efficiency in its use of both renewable and non-renewable energy resources [97].

Despite the challenges, the development of further research into improving the sustainability of data centers through more effective and efficient cooling systems is a source of optimism. If the work presented within this dissertation proves to be a launching point, rather than an end point, to continuing investigations into the need and design for alternative data center cooling systems, then it will have been well worth the effort.

REFERENCES

[1] M. K. Patterson, "The effect of data center temperature on energy efficiency.," in *ITHERM 2008 11th intersociety conference on thermal and thermomechanical phenomena in electronic systems*, Orlando, 2008.

[2] J. D. Mitchell-Jackson, "Energy needs in an internet economy: a closer look at data centers.," University of California, Berkeley, 2001.

[3] J. G. Koomey, "Worldwide electricity used in data centers.," *Environ Res Lett.,* no. 3, pp. 1-8, 2008.

[4] US Environmental Protection Agency, "Report to congress on server and data center energy efficiency public law 109431.," 2007.

[5] J. G. P. Koomey, "Growth in data center electricity use 2005 to 2010," Analytics Press, Oakland, CA, 2011.

[6] K. G. Brill, "Moore's law economic meltdown.," *Forbes,* 2008.

[7] K. G. Brill, "The invisible crisis in the data center: The economic meltdown of moore's law.," The Uptime Institute White Paper, 2007.

[8] T. Vanderbilt, "Data center overload.," *The New York Times Magazine,* 08 June 2009.

[9] K. G. Brill, "The invisible crisis in the data center: How IT performance is driving the economic meltdown of Moore's law.," in *The Uptime Institute Symposium 2007*, Orlando, 2007.

[10] G. Rabah, "Investigation of the potential of application of single effect and multiple effect absorption cooling systems.," *Energy Conversion and Management,* vol. 51, pp. 1629-1636, 2010.

[11] T. S. B. R. E. &. M. B. Brunschwiler, "Toward zero-emission data centers through direct reuse of thermal energy.," *IBM J Res Dev,,* vol. 53, no. 11, pp. 1-3, 2009.

[12] Y. M. L. F. A. &. P. P. E. Gupta, "Comparative analysis of thermally activated, environmentally friendly cooling systems.," *Energy Conversion and Management,* vol. 49, pp. 1091-1097, 2008.

[13] J. B. Marcinichen, J. A. Olivier and J. R. Thome, "On-chip two-phase cooling of datacenters: Cooling system and energy recovery evaluation," *Applied Thermal Engineering,* vol. 41, p. 36–51, August 2012.

[14] Z. Liu, Y. Chen, C. Bash, A. Wierman, D. Gmach, Z. Wang, M. Marwah and C. Hyser, "Renewable and cooling aware workload management for sustainable data centers," in *SIGMETRICS '12 Proceedings of the 12th ACM SIGMETRICS/PERFORMANCE joint international conference on Measurement and Modeling of Computer Systems*, New York, 2010.

[15] F. Mesa-Martinez, J. Nayfach-Battilana and J. Renau, "Power Model Validation through Thermal Measurements," in *ISCA '07: Proc. Of The 34th Annual ISCA*, San Diego, 2007.

[16] C. Li, R. Zhou and T. Li, "Enabling distributed generation powered sustainable high-performance data center," in *High Performance Computer Architecture (HPCA2013), 2013 IEEE 19th International Symposium on*, Shenzhen, 2013.

[17] L. . J. Klein, S. Bermudez, H.-D. Wehle, S. Barabasi and H. F. Hamann, "Sustainable data centers powered by renewable energy," in *Semiconductor Thermal Measurement and Management Symposium (SEMI-THERM), 2012 28th Annual IEEE*, San Jose, CA, 2012.

[18] Yazaki Energy Systems, Incorporated, "Yazaki Aeroace: Water Fired Chiller/Chiller-Heater.," 2009. [Online]. Available: www.yazakienergy.com/docs/SB-WFCS-1009.pdf.

[19] K. A. Joudi and Q. J. Abdul-Ghafour, "Development of design charts for solar cooling systems. Part I: computer simulation for a solar cooling system and development of solar cooling design charts.," *Energy Conversion and Management,* Vols. 44-2, pp. 313-339, January 2003.

[20] A. H. Ali, P. Noeres and C. Pollerberg, "Performance assessment of an integrated free cooling and solar powered single-effect lithium bromide-water absorption chiller," *Solar Energy,* vol. 82, no. 11, pp. 1021-1030, 2008.

[21] "Thermal resistance," 13 May 2013. [Online]. Available: http://en.wikipedia.org/wiki/Thermal_resistance. [Accessed 2011].

[22] M. Mazloumi, M. Naghashzadegan and K. Javaherdeh, "Simulation of solar lithium bromide–water absorption cooling system with parabolic trough collector," *Energy Conversion and Management,* vol. 49, no. 10, pp. 2820-2832, 2008.

[23] Intel® Support, "64-bit intel® Xeon® processor 3.60 Ghz, 2MB cache, 800 MHz FSB," [Online]. Available: http://ark.intel.com/products/28019/64-bit-Intel-Xeon-Processor-3_60-GHz-2M-Cache-800-MHz-FSB.

[24] "CPU power dissipation," 13 November 2013. [Online]. Available: http://en.wikipedia.org/wiki/CPU_power_dissipation. [Accessed 2011].

[25] G. Shvets, "Thermal design power (TDP)," 15 October 2013a. [Online]. Available: http://www.cpu-world.com/Glossary/T/Thermal_Design_Power_(TDP).html. [Accessed 2011].

[26] G. Shvets, "Minimum/Maximum power dissipation," 15 October 2013b. [Online]. Available: http://www.cpu-world.com/Glossary/M/Minimum_Maximum_power_dissipation.html.

[27] Dell™ Inc., "Dell™ Poweredge™ Rack Systems," 2003. [Online]. Available: http://www.verinet.dk/dell_rack_system.pdf.

[28] A. M. Tayeb, "A simulation model for a phase-change energy storage system: Experimental and verification," *Energy Conversion and Management,* vol. 34, no. 4, p. 243–250, April 1993.

[29] A. M. Tayeb, "Organic-inorganic mixtures for solar energy storage systems," *Energy Conversion and Management,* vol. 36, no. 10, pp. 969-974, October 1995.

[30] M. M. Farid, A. M. Khudhaira, S. A. K. Razack and S. Al-Hallaj, "A review on phase change energy storage: materials and applications," *Energy and Conversion Management,* vol. 45, no. 9-10, p. 1597–1615, June 2004.

[31] "Water cooling," 27 September 2013. [Online]. Available: http://en.wikipedia.org/wiki/Water_cooling. [Accessed 2011].

[32] ASHRAE, Best practices for datacom facility energy efficiency. (2 ed.), 2nd ed., Atlanta: ASHRAE, 2009.

[33] ASHRAE, High Density Data Centers: Case Studies and Best Practices, Atlanta: ASHRAE, 2008.

[34] L. Lu, W. Cai, Y. . C. Soh, L. Xie and S. Li , "Hvac system optimization—condenser water loop," *Energy Conversion and Management,* vol. 45, no. 4, p. 613–630, 2004.

[35] F. W. Yu and K. T. Chan, "Modelling of the coefficient of performance of an air-cooled screw chiller with variable speed condenser fans.," *Building and Environment,* vol. 41, no. 4, pp. 407-417, 2006.

[36] Y. A. Çengel and M. A. Boles, Thermodynamics, An Engineering Approach, 6th ed., Boston: McGraw Hill, 2008.

[37] P. M. Curtis, Maintaining Mission Critical Systems in a 24/7 Environment, 2nd ed., vol. Volume 61 of IEEE Press Series on Power Engineering, Hoboken, New Jersey: John Wiley & Sons, 2011, p. 512.

[38] the green grid®, "Green Grid Data Center Power Efficiency Metrics: PUE and DCiE," 2008.

[39] R. Elliot, "ESP - heatsink design and transistor mounting," 10 March 2013. [Online]. Available: http://sound.westhost.com/heatsinks.htm. [Accessed 2011].

[40] Fabrico™, "Thermal management of leds in automotive applications," 2013. [Online]. Available: http://www.fabrico.com/Collateral/Documents/English-US/FAB_Application_Sheet_Transportation_1_Thermal_LEDs.pdf. [Accessed 2011].

[41] J. Harris, "Empfasis - tech tips: Thermal interfaces & materials," August 2006. [Online]. Available: http://www.empf.org/empfasis/2006/aug06/tech_tips-0806.html.

[42] R. Pagnamenta, "Computer power provides heat for helsinki," *The Times,* 11 February 2010.

[43] B. Tschudi, O. Vangeet, J. Cooley and D. Azevedo, *ERE: A metric for measuring the benefit of reuse energy from a data center,* M. Patterson, Ed., The Green Grid, 2010.

[44] ASHRAE, Datacom Equipment Power Trends and Cooing Applications, Atlanta: ASHRAE, 2005.

[45] ASHRAE, Liquid Cooling Guidelines for Datacom Equipment Centers, Altanta: ASHRAE, 2006.

[46] D. Beaty and R. Schmidt, "Back to the future: Liquid cooling," *ASHRAE Journal,* vol. 46, no. 12, pp. 42-46, December 2004.

[47] Iceotope Corp., "Iceotope offer full time free cooling for ICT anywhere," December 2011. [Online]. Available: http://www.iceotope.com/.

[48] Green Revolution Cooling, "Introduction to fluid submersion," December 2012. [Online]. Available: http://www.grcooing.com/technology/Introduction-to-fluid-submersion/.

[49] G. M. Ung, "Exclusive: Oil immersion cooling goes mainstream with hardcore," pp. 1-3, Dec. 2008. [Online]. Available: (11/2013)," pp. 1-3, 25 December 2008.

[50] Puget Custom Computers, "Mineral Oil Submerged Computer; Our Most Popular Custom PC," December 2011. [Online]. Available: http://www.pugetsystems.com/submerged.php.

[51] A. M. Haywood, J. Sherbeck, P. Phelan, G. Varsamopoulos and S. S. Gupta, "Thermodynamic feasibility of harvesting data center waste heat to drive an absorption chiller," *Energy Conversion and Management,* vol. 58, pp. 26-34, June 2012.

[52] H. F. Hamann, J. Lacey, A. Weger and J. Wakil, "Spatially-resolved imaging of microprocessor power (SIMP): hotspots in microprocessors," in *Proceedings of 10th Intersociety Conf. on ITHERM*, 2006.

[53] S. S. Gupta, G. Varsamopoulos, A. M. Haywood, P. Phelan and T. Mukherjee, "Handbook of Energy-Aware and Green Computing," in *Bluetool: Using A Computing Systems Research Infrastructure Tool To Design And Test Green And Sustainable Data Centers*, vol. 2, S. Rank and I. Ahmad, Eds., Chapman and Hall/CRC Press, 2011, pp. 745-760.

[54] A. M. Haywood, J. Sherbeck, P. Phelan and G. Varsamopoulos, "A Sustainable Data Center with Heat-Activated Cooling," in *12th IEEE Intersociety Conference on ITHERM*, Las Vegas, 2010.

[55] Intel®, "64-bit Intel® Xeon® Processor with 2MB L2 Cache Datasheet," September 2005. [Online]. Available: http://www.intel.com/content/dam/doc/datasheet/64-bit-xeon-2mb-l2-cache-datasheet.pdf.

[56] R. B. Abernathy, The New Weibull Handbook, 4th ed., Robert B Abernethy, 2000.

[57] "Exergy," 23 December 2013. [Online]. Available: http://en.wikipedia.org/wiki/Exergy.

[58] CoolIT Systems Inc., "CoolIT Systems - CoolIT Systems," January 2012. [Online]. Available: http://www.coolitsystems.com/.

[59] Brooks® Instrument, "Sho-Rate™ "50" Model 1350E and Sho-Rate™ "150" Model 1355E," March 2008. [Online]. Available: http://www.brooksinstrument.com/downloads/Product%20Documentation/Variable%20Area%20Meters%20Low%20Flow%20Glass%20Plastic%20Tube%20Purgemeters/Data%20Sheets/DS-VA-1350E-eng.pdf.

[60] J. P. Gwinn and R. L. Webb, "Performance and testing of thermal interface materials," *Microelectronics Journal,* vol. 34, no. 3, p. 215–222, March 2003.

[61] F. P. Incropera and D. P. DeWitt, Fundamentals of Heat and Mass Transfer, 4th ed., New York: Wiley, 1996.

[62] J. Holman, "Uncertainty Analysis," in *Experimental Methods For Engineers*, 7th ed., New York, McGraw-Hill, 2001, pp. 51-55.

[63] Washington University in St. Louis, "Uncertainties & Error Analysis Tutorial," 22 December 2013. [Online]. Available: http://physics.wustl.edu/introphys/Phys117_118/Lab_Manual/Tutorials/ErrorAnalysisTutorial.pdf.

[64] American Limits of Error ASTM E230-ANSI MC 96.1, "Thermocouple tolerances," [Online]. Available: http://www.omega.co.uk/temperature/pdf/Thermocouple_Tolerance.pdf.

[65] Dell©, "Dell© OMSA 6.3," May 2010. [Online]. Available: Available with permission: ftp://ftp.us.dell.com.

[66] M. Mienik, "CPU Burn-in v1.01," July 2011. [Online]. Available: http://www.cpuburnin.com.

[67] M. Berktold and T. Tian, "Monitor CPU Temperature with DTS and PECI," September 2010. [Online]. Available: http://www.intel.com/content/www/us/en/intelligent-systems/cpu-monitoring-dts-peci-paper.html.

[68] Indium Corporation©, "Compressible TIM (SMA-TIM) Heat-Spring® - Thermal Interface Materials," 2010. [Online]. Available: http://www.indium.com/TIM/solutions/heatspringtim1.php.

[69] "Computer cooling," 26 December 2013. [Online]. Available: http://en.wikipedia.org/wiki/Computer_cooling#Liquid_submersion_cooling.

[70] L. Chu, J. Weale and P. Rumsey, "Direct Liquid Cooling," in *Environmental Energy Technologies Division 2006 CommHandbook*, 2006.

[71] G. Wall, "Exergy - A useful concept within resource accounting," University of Goteborg, Goteborg, 1977.

[72] K. Tweed, "greentechefficiency," 7 January 2014. [Online]. Available: http://www.greentechmedia.com/articles/read/forget-energy-efficiency-think-exergy.

[73] E. Rathakrishnan, Fundamentals of engineering thermodynamics, 2nd ed., New Delhi: Prentice-Hall of India Private Limited, 2005.

[74] "Thermal design power," 21 December 2013. [Online]. Available: http://en.wikipedia.org/wiki/Thermal_design_power.

[75] A. Cunningham, "The technical details behind Intel's 7 watt Ivy Bridge CPUs," 14 January 2013. [Online]. Available: http://arstechnica.com/gadgets/2013/01/the-technical-details-behind-intels-7-watt-ivy-bridge-cpus/.

[76] "CMOS," 9 January 2014. [Online]. Available: http://en.wikipedia.org/wiki/CMOS.

[77] "Junction temperature," 30 July 2013. [Online]. Available: http://en.wikipedia.org/wiki/Junction_temperature.

[78] "Thermal resistance," 28 December 2013. [Online]. Available: http://en.wikipedia.org/wiki/Thermal_resistance. [Accessed 2011].

[79] the green grid®, "White_Paper_7_-_Five_Ways_to_Save_Power.pdf," 2008. [Online]. Available: http://www.thegreengrid.org/~/media/WhitePapers/White_Paper_7_-_Five_Ways_to_Save_Power.pdf.

[80] G. Wall, "Exergetics," in *EOLSS Our Fragile World*, M. K. Tolba, Ed., Oxford, Eolss Publishers, 2001.

[81] "Information Theory," 6 January 2014. [Online]. Available: http://en.wikipedia.org/wiki/Information_theory. [Accessed 19 January 2014].

[82] J. v. Neumann, Theory of Self-Reproducing Automata, Urbana, IL: Univ. Illinois Press, 1966.

[83] J. Meindl and J. Davis, "The fundamental limit on binary switching energy for terascale integration (TSI)," *JSSC,* vol. 35, no. 10, pp. 1515-1516, October 2000.

[84] N. H. Weste and D. M. Harris, CMOS VLSI Design: A Circuits and Systems Perspective, 4th ed., Boston: Addison-Wesley, 2011.

[85] K. Moiseev, A. Kolodny and S. Wimer, "Timing-aware power-optimal ordering of signals," *ACM Transactions on Design Automation of Electronic Systems,* vol. 13, no. 4, 4 September 2008.

[86] D. M. Harris, *Lecture 18: Design for Low Power - CMOS VLSI Design,* Claremont, CA, 2004.

[87] "Linpack Benchmarks," 14 December 2013. [Online]. Available: http://en.wikipedia.org/wiki/LINPACK_benchmarks. [Accessed 20 January 2014].

[88] M. P. Fernandez, "Nodes, Sockets, Cores and FLOPS, Oh, My," 2011. [Online]. Available: http://en.community.dell.com/techcenter/high-performance-computing/w/wiki/2329.aspx. [Accessed 2012].

[89] "LINPACK," 16 November 2013. [Online]. Available: http://en.wikipedia.org/wiki/LINPACK. [Accessed 2012].

[90] V. Henson, "The Kernel Hacker's Bookshelf: Ultimate Physical Limits of Computation," 18 June 2008. [Online]. Available: http://lwn.net/Articles/286233/. [Accessed 2012].

[91] jzbiciak, "1000 single bit operations per FLOP," 23 June 2008. [Online]. Available: http://lwn.net/Articles/287045/. [Accessed 2013].

[92] S. Lloyd, "Ultimate physical limits to computation," *Nature,* vol. 406, pp. 1047-1054, August 2000.

[93] E. Pfanner, "Liquid-Cooled Supercomputers, to Trim the Power Bill," *The New York Times,* p. B1, 11 February 2014.

[94] K. Ebrahimi, J. F. Gerard and A. S. Fleische, "A review of data center cooling technology, operating conditions and the corresponding low-grade waste heat

recovery opportunities," *Renewable and Sustainable Energy Reviews,* vol. 31, p. 622–638, March 2014.

[95] G. Varsamopoulos and S. K. S. Gupta, "Energy Proportionality and the Future: Metrics and Directions," in *2010 39th International Conference on Parallel Processing Workshops*, San Diego, 2010.

[96] S. K. Gupta, R. R. Gilbert and A. Baner, "GDCSim - An Integrated Tool Chain for Analyzing Green Data Center Physical Design and Resource Management Techniques," in *IEEE International Green Computing Conference (IGCC)*, Orlando, 2011.

[97] G. A. Brady, N. Kapur, J. L. Summers and H. . M. Thompson, "A case study and critical assessment in calculating power," *Energy Conversion and Management,* vol. 76, p. 155–161, December 2013.

[98] J. Z. Woodruff, P. Brennerb, A. P. Buccellatoc and D. . B. Go, "Environmentally opportunistic computing: A distributed waste heat reutilization approach to energy-efficient buildings and data centers," *Energy and Buildings,* vol. 69, pp. 41-50, 2014.

[99] X. Lü , T. Lu and M. Viljanen, "Field analysis for sustainable data centres," *International Journal of Sustainable Building Technology and Urban Development,* vol. 3, no. 2, pp. 120-128, 9 August 2012.

APPENDIX A

AUTOIT PROGRAM SCRIPT FOR AUTOMATIC LOGGING OF CPU TEMPERATURES

AutoIt is an open-source scripting program by AutoIt Consulting, Ltd (http://www.autoitscript.com/) that was used to simulate keystrokes and mouse movements to automate the logging of CPU temperatures from OMSA into a text file. Specifically, OMSA would be periodically refreshed (~5sec), then the updated CPU temperature of interest would be copied from the screen output in the Firefox browser and pasted into a Notepad text file.

```
;Script Start -
   WinActivate("[CLASS:MozillaWindowClass]", "")
$i = 0
While $i <= 3600 ; 30 minute run.


   ; Click Refresh button
   MouseClick("left", 745, 248)
   Sleep(2000)

   ; Left click drag from 637,359 to 655,359
   MouseClickDrag("left", 375,392, 404,392)


   Send("^c")
   ;WinActivate("[CLASS:XLMAIN]", "")
   WinActivate("[CLASS:Notepad]", "")
   Send("^v")
   Send("{ENTER}")
   Send("^s")  ;Save each entry
```