Final Technical Report

Recovery Act: Economizer Based Data Center Liquid Cooling with Advanced Metal Interfaces

Prepared For

US Department of Energy
DOE Award Number DE-EE0002894
Project Period: 1/31/2010 – 8/31/2012

Principal Investigator:

Timothy Chainer
IBM TJ Watson Research Center
1101 Kitchawan Road
Yorktown Heights, NY 10598
914-945-2641
tchainer@us.ibm.com

Recipient Organization:

IBM TJ Watson Research Center
1101 Kitchawan Road
Yorktown Heights, NY 10598

11/30/2012

Acknowledgment

This report is based upon work supported by the U. S. Department of Energy under Award Number DE-EE0002894. We would also like to thank DOE Project Officer Debo Aichbhaumik, DOE Project Monitors Darin Toronjo and Chap Sapp and DOE HQ Contacts Gideon Varga and Bob Gemmer for their support throughout the project.

Disclaimer

Any findings, opinions, and conclusions or recommendations expressed in this report are those of the author(s) and do not necessarily reflect the views of the Department of Energy.

The results are based on measurements in an experimental environment. The actual results that any user will experience will vary depending upon considerations such as the actual environment where the equipment is operated and other external factors than can affect cooling requirements. Therefore, no assurance can be given that an individual user will achieve estimated energy savings stated here. IBM may not offer the products, services or features discussed in this document, and the information may be subject to change without notice.

Document Availability

Reports are available free via the U.S. Department of Energy (DOE) Information Bridge Website: <http://www.osti.gov/bridge>

Reports are available to DOE employees, DOE contractors, Energy Technology Data Exchange (ETDE) representatives, and Informational Nuclear Information System (INIS) representatives from the following source:

Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831
Tel: (865) 576-8401
FAX: (865) 576-5728
E-mail: reports@osti.gov
Website: <http://www.osti.gov/contract.html>

© Copyright IBM Corp. 2012
ALL RIGHTS RESERVED

Table of Contents

Acknowledgement.....	i
Disclaimer.....	i
Document Availability.....	i
List of Acronyms.....	iv
List of Figures.....	v
List of Tables.....	ix
1. Executive Summary.....	1
2. Introduction.....	3
3. Background.....	4
4. Results and Discussion.....	8
Chapter 1: System Design.....	8
1.1 System Overview.....	8
1.2 Temperature Excess between Outdoor Air and Server Node Coolant.....	12
1.3 Flow Modeling and Measurement.....	13
1.4 System Monitoring and Control.....	17
Chapter 2: Node Liquid Cooling Design and Advanced Metal Interfaces.....	19
2.1 Overview of Liquid Cooled Server.....	20
2.2 Liquid Cooled Server Design.....	22
2.3 Comparison with Typical Air Cooled Server.....	28
2.4 Advanced Metal Interfaces for Module Liquid Cooling.....	35
2.5 Advanced “Compliant” Heat Sinks.....	39
Chapter 3: System Modeling and Simulations.....	40
3.1 Energy Based System Model.....	40
3.1.1 Design Impact Study.....	44
3.2 Dynamic System Model.....	46
3.2.1 Simulation Model Application Example: Exploratory Control.....	48
Chapter 4: System Characterization, Operation and Performance Projection.....	50
4.1 System Characterization Studies.....	50
4.2 Day Long Operational Runs.....	61
4.3 System Servo Control.....	69
4.4 Long Term Continuous Run with Dynamic Servo Control.....	71
4.5 Energy Based System Model Validation.....	75
4.6 Model Simulation for Simple Control Methods.....	79
4.7 Performance Prediction for Typical Year and Geographical Locations.....	82
5. Benefit Assessment.....	87
6. Commercialization.....	88
7. Accomplishments.....	89
8. Conclusions.....	90
9. Recommendations.....	91
10. References.....	92

List of Acronyms

ASHRAE	- American Society of Heating, Refrigerating and Air-Conditioning Engineers
BCW	- Building Chilled Water
BMC	- Baseboard Management Controller
CFD	- Computational Fluid Dynamics
COP	- Coefficient of Performance
CPU	- Central Processing Unit
CRAC	- Computer Room Air-Conditioning
CRAH	- Computer Room Air Handlers
CTE	- Coefficient of Thermal Expansion
DAQ	- Data Acquisition system
DELC	- Dual Enclosure Liquid Cooling
DIMM	- Dual Inline Memory Module
DIO	- Digital Input/Output
DOE	- Department of Energy
DTS	- Digital Thermal Sensor
ECR	- End Cold Rail
EPA	- Environmental Protection Agency
FCR	- Front Cold Rail
GPM	- Gallons per minute
HE or HX	- Heat Exchanger
IBM	- International Business Machines
IOH	- Input/Output Hub
IPMI	- Intelligent Platform Management Interface
IT	- Information Technology
LBNL	- Lawrence Berkeley National Laboratory
LGA	- Land Grid Array
LMTI	- Liquid Metal Thermal Interface
LPM	- Liters per minute
MCR	- Middle Cold Rail
MWU	- Modular Water Unit (also referred as Buffer Unit)
NREL	- National Renewable Energy Laboratory
OEM	- Original Equipment Manufacturer
OHE	- Outdoor Heat Exchanger (also referred as Dry Cooler)
PCB	- Printed Circuit Board
PFC	- Pin-fin Compliant
PG	- Propylene Glycol
PI	- Proportional-Integral
PLC	- Programmable Logic Controller
PUE	- Power Usage Effectiveness
R&D	- Research and Development
RPM	- Revolutions per minute
TIM	- Thermal Interface Material
VFD	- Variable Frequency Drive

List of Figures

Figure 1-1. Highly energy efficient chiller-less, liquid cooled system for data center cooling.....	2
Figure 1-2. Efficient cooling solution for the volume server (a) Plan view of the liquid cooled server (b) Server liquid cooling structures for processor and memory cooling.....	2
Figure 2-1. Typical data center (a) total and (b) cooling energy breakdown [11].....	4
Figure 3-1. Traditional chiller plant based data center cooling loop.....	5
Figure 3-2. Schematic of typical data center facility cooling infrastructure.....	6
Figure 3-3. Schematic of the Dual Enclosure Liquid Cooling (DELC) system.....	7
Figure 4.1-1. Schematic representation of the energy efficient dual enclosure data center liquid cooling test facility. IT servers are warm water cooled with the heat ultimately rejected to ambient air via a liquid-to-air heat exchanger.....	9
Figure 4.1-2. Photographs of data center cooling hardware (a) External dry cooler unit, (b) Internal (lab) piping layout and instrumentation.....	10
Figure 4.1-3. Rack liquid cooling design, (a) Photograph of the front of the rack, (b) Plan view schematic of rack internals [18], (c) Plumbing of server node liquid cooling to rack manifolds.....	10
Figure 4.1-4. Hybrid air-water cooled 1U server designed for intake of 45°C water and 50°C air.....	11
Figure 4.1-5. Approach temperature differences for the three heat exchanger devices in the loop (a) Dry cooler external air-to-liquid heat exchanger coil, (b) Buffer unit liquid to liquid heat exchanger, (c) Rack Side Car air to liquid heat exchanger.....	13
Figure 4.1-6. Flow model of external system cooling loop.....	14
Figure 4.1-7. Flow model of rack level water cooling piping.....	15
Figure 4.1-8. Predicted nodal flow distribution from bottom to top of rack for 3/4 inch diameter supply and return manifolds for various rack flow rates.....	15
Figure 4.1-9. Node flow-pressure drop simulator for experimentally determining flow through an individual node.....	15
Figure 4.1-10. Comparison of node simulator hose assembly's flow-pressure drop characteristics with flow-pressure drop characteristics of actual node containing parallel channel cold plates.....	16
Figure 4.1-11. Flow –pressure drop correlation for node simulator hose test assembly.....	16
Figure 4.1-12. Node flow measurement test set-up.....	16
Figure 4.1-13. Measured flow through simulated node as a function of node position measured from bottom of rack.....	16
Figure 4.1-14. Photo of PLC control unit used to monitor and control the data center test facility. The PLC is connected to a remote computer with monitoring and control software... ..	17
Figure 4.1-15. Server side data collection process and protection measures.....	18
Figure 4.2-1. Photographs of (a) Standard air –cooled version of IBM x 3550 M3 server. (b) Hybrid air/liquid-cooled version with cold-plates for CPUs and cold rails for DIMM Cooling ...	19
Figure 4.2-2. Liquid cooled components inside the server; (a) cold plate for the	

microprocessor module, (b) DIMM with conduction spreader, (c) front cold rail, (d) middle cold rail and, (e) end cold for memory liquid cooling.....	20
Figure 4.2-3. Node cooling sub-assembly for partially liquid cooled server.....	21
Figure 4.2-4. Illustration of DIMM-spreader assembly into liquid cooled server node.....	22
Figure 4.2-5. Candidate node liquid cooling loop designs (a) Tubed cold plates for processor module cooling and cold rails for DIMM cooling (b) parallel channels based cold plates for processor module cooling and cold rails for DIMM cooling (c) Tubed cold plates for processor module cooling only and (d) Parallel channels based cold plates for processor module cooling only.....	23
Figure 4.2-6(a) Comparison of DIMMs temperature delta with respect to the coolant inlet temperature for air cooled DIMMs with that for liquid cooled DIMMs. (b) Server node liquid cooling assembly with cold plates for the micro-processors cooling and cold rails for DIMM liquid cooling. 6 DIMMs in the front bank (slot numbers 2,3,5,6,8 and 9) and 6 DIMMs in the rear bank (slot numbers 11,12,14,15,17 and 18).....	24
Figure 4.2-7. Hydrodynamic performance of the node liquid cooling loop designs.....	25
Figure 4.2-8. (a) Variation of thermal resistance as a function of flow rate for the tubed cold plate and parallel channels based cold plate. (b) Temperature contours at a chip-cold plate package cross-sectional plane passing through the center of the package for tubed cold plate (c) for parallel channels based cold plate with a liquid coolant flow rate of 0.238 gpm and inlet temperature of 45 °C.....	26
Figure 4.2-9. Numerical prediction of the flow distribution between the front and middle cold rails and of the pressure drop from the cooling loop inlet to the inlet of the first cold plate for a design flow rate of 0.9 liter per minute.....	27
Figure 4.2-10. Numerical prediction of variation of flow distribution among the front cold rail (FCR) and middle cold rail (MCR) as a function of the total flow rate through the server liquid cooling loop.....	27
Figure 4.2-11. Experimental measurement of variation of flow distribution among the front cold rail (FCR) and middle cold rail (MCR) as a function of the total flow rate through the server liquid cooling loop.....	28
Figure 4.2-12. DIMM spreader sub-assembly made up of two copper spreader plates that are mechanically attached to a DIMM card using spring clips with a thermal interface material between the spreader and the DIMM card.....	28
Figure 4.2-13. Power and device temperature data for partially water cooled node (a) Server power, (b) Fan power, (c) CPU lid temperature, (d) DIMM temperature.....	30
Figure 4.2-14. Comparison of estimated junction temperature for a liquid cooled server with a typical air cooled server (a) when the CPUs are exercised at 90% and (b) when the memory modules are exercised.....	31
Figure 4.2-15. Comparison of DIMM temperatures for a liquid cooled server with a typical air cooled server (a) when the CPUs are exercised at 90% and (b) when the memory modules are exercised.....	32
Figure 4.2-16. Comparison of server power and fan power consumption for a liquid cooled server with a typical air cooled server (a) when the CPUs are exercised at 90% and (b) when the memory modules are exercised.....	33

Figure 4.2-17. Comparison of IOH temperatures for a hybrid air/liquid cooled server with a typical air cooled server (a) when the CPUs are exercised at 90% and (b) when the memory modules are exercised.....	34
Figure 4.2-18. Difference between the thermal resistance path for the conventional approach and the direct attach approach.....	36
Figure 4.2-19. Chip and laminate with the residual materials removed and ready for LMTI... ..	36
Figure 4.2-20 (a) Server with single de-lidded module installed with a commercially available heat sink. (b) Server with single de-lidded module installed with a commercially available heat sink (close-up).....	37
Figure 4.2-21. DTS data for varying coolant temperatures for a variety of exercise conditions on a single processor module.....	37
Figure 4.2-22. Comparison of one CPU core temperature average for standard thermal grease and direct attach LMTI implementations.....	38
Figure 4.2-23. Core temperature variability within the 42 servers of the complete demonstration system showing the lower operating temperature of LMTI equipped nodes.....	39
Figure 4.2-24. Pin Fin Compliant heat sink cooled module structure compared with standard rigid heat sink liquid cooled module structure.....	40
Figure 4.3-1. Schematic of the system model.....	41
Figure 4.3-2. Flow diagram of system model operation.....	42
Figure 4.3-3. Design choices: (a) Single Loop, (b) Dual Loop, (c) High-cost high performance cold-plate1 and (d) Low-cost low performance cold-plate2.....	45
Figure 4.3-4. Variation of thermal resistance as a function of flow rate for different Cold-plates and different coolants.....	45
Figure 4.3-5. Cooling power usage comparison at different outdoor ambient temperatures for different design choices.....	45
Figure 4.3-6. Limiting components temperature variation as function of outdoor ambient air temperature.....	46
Figure 4.3-7. Overall data center Cooling System & Control Simulation Model.....	47
Figure 4.3-8. Rack Sub-system Model with Side Car, CPU/Cold Plates and Memory/Cold Rails.....	48
Figure 4.3-9. Data center Cooling System Control Components.....	49
Figure 4.3-10. Comparison of rack inflow and rack outflow fluid temperature set-point control.....	49
Figure 4.4-1. Schematic representation of the liquid cooled IBM X-3550 Server.....	51
Figure 4.4-2 (a) CPU coldplate to rack inlet water thermal resistance and (b) DIMM spreader to rack inlet water thermal resistance. Both show a power decay trend with internal loop water flow rate and an increase in the thermal resistance as the inlet water temperature reduces.....	52
Figure 4.4-3. Approach thermal resistance surface for the buffer unit showing the thermal resistance as a function of internal and external liquid flow rates. The thermal resistance improves with increasing external (cold) loop flow rate and reducing internal (hot) loop flow rate.....	53

Figure 4.4-4. (a) Heat exchanger conductance and (b) approach thermal resistance as functions of internal and external loop flow-rates and heat exchanger configuration. 1x C: single heat exchanger in counter flow, 2x C-P: double heat exchanger in counter-parallel flow and 2x C: double heat exchanger in full counter flow configuration.....	54
Figure 4.4-5. Approach thermal resistance as functions of internal and external loop flow-rates and external loop coolant. Addition of 20% and then 50% propylene glycol to the external loop results in an increase in the thermal resistance, with the effect growing stronger at higher internal loop flow rates.....	55
Figure 4.4-6. Approach thermal resistance surface for the dry-cooler unit showing the thermal resistance as a function of external liquid flow rate and dry-cooler fan speed. The thermal resistance improves strongly with increasing fan speed (and air flow) as well as with reducing external (hot) loop flow rate and follows a power trend with both parameters..	56
Figure 4.4-7. Relation between fan speed in RPM and the calculated air flow rate. The trend is almost linear (shown as a dotted line) though there is considerable spread due to the small temperature differences at higher fan speeds resulting in greater uncertainty in the determined air flow rate.....	57
Figure 4.4-8 (a) Approach thermal resistance as functions of external loop flow-rate, fan speed and external loop coolant. Addition of 20% and then 50% propylene glycol (PG) to the external loop results in an increase in the thermal resistance (b) Dry Cooler conductance versus flow rate.....	57
Figure 4.4-9. (a) Internal loop pressure drop and (b) internal loop pump power as a function of the flow rate. Addition of the second heat exchanger results in an increase in the pressure drop and thus energy required to provide a given flow rate.....	58
Figure 4.4-10. (a) External loop pressure drop and (b) external loop pump power as a function of the flow rate. Addition of the second heat exchanger as well as addition of propylene glycol results in an increase in the pressure drop and thus energy required to drive a the external coolant at a given flow rate.....	59
Figure 4.4-11. Power consumed by the dry-cooler fans as a function of fan speed follows a cubic trend. The power consumed by the fans is comparable with the pumps below 750RPM beyond which they progressively consume a large percentage of the total cooling power.....	59
Figure 4.4-12. (a) Thermal resistance and (b) power consumption at a low 4 GPM internal loop while external loop flow rate and dry cooler fan speed are varied.....	60
Figure 4.4-13. (a) Thermal resistance and (b) power consumption at a high 8 GPM internal loop while external loop flow rate and dry cooler fan speed are varied.....	61
Figure 4.4-14. Temperature trace for a summer (a) and two fall days (b), (c) the first being rainy and the second overcast and dry. Flow rates set to 7.2 GPM both on the internal and external loop with fans set to vary from 170 to 500 RPM as the pre-buffer temperature rises from 30 °C to 35 °C. Cooling power was very similar on all three days ~430 W and is only 3.3% of the total IT power.....	62
Figure 4.4-15. Temperature trace for the low cooling power test with lowered flow rates of 4 GPM on the internal and external flow loops. Fans set at a constant 170 RPM. Cooling consumes 210 W, 1.6% of the IT power.....	63
Figure 4.4-16. Temperature trace during a four day period over the course of which a freak	

fall snow storm occurred. The trace shows the dramatic impact of adding propylene glycol to the external loop coolant to avoid freeze damage.....	65
Figure 4.4-17. Server component data showing (a) hottest core DTS numbers for CPU 1 and CPU 2, (b) DIMMs temperature for each of the 12 DIMMs and (c) system fans rpm for one of the server from a sample 22 hours run.....	66
Figure 4.4-18. Variation of temperature from the outdoor air to the server components.....	67
Figure 4.4-19. Frequency distribution of CPU 1 and CPU 2 DTS numbers and maximum DIMM temperatures at t=12.4 hrs(cooler) and t=20.7 hrs(warmer) from the 22 hours test run...	68
Figure 4.4-20. Three zone control algorithm for cooling energy minimization for the DELC system (a) control flowchart (b) graphical representation of the three distinct control zones.....	69
Figure 4.4-21. Transient response of rack inlet water temperature to changes in power and set-point.....	71
Figure 4.4-22 (a) IT Power and (b) ambient air temperature over the course of the 62 dayrun ...	72
Figure 4.4-23. Key data center temperatures during the 62 day run with dynamic servo control implemented.....	72
Figure 4.4-24. Equipment speeds as driven by the servo control in response to measured IT power and ambient air temperature.....	73
Figure 4.4-25. Equipment power consumption over the 62 day run showing the peaks during the hotter periods and flat minimums during the cooler portions of the run window.....	73
Figure 4.4-26. Impact of variable workload (top) on the pre-rack water temperature (middle) and cooling equipment power use (bottom).....	74
Figure 4.4-27. Cooling power as a fraction of IT power over the 62 day run period. Hotter days and days with lower workload result in comparatively higher cooling power fraction.....	75
Figure 4.4-28. Pre-MWU temperature prediction and comparison with experimental data.....	76
Figure 4.4-29. Pre-Rack temperature prediction and comparison with experimental data.....	76
Figure 4.4-30. System model prediction of CPU and DIMM temperatures and comparison with experimental data.....	77
Figure 4.4-31. CPU temperature prediction and comparison with experimental data for test case 1 with the bars showing the temperature variability across all servers.....	77
Figure 4.4-32. DIMM temperature prediction and comparison with experimental data for test case 1 with the bars showing temperature variability across all servers.....	77
Figure 4.4-33. Facility side temperature prediction and comparison with experimental data for a day long summer run.....	78
Figure 4.4-34. Typical server CPU and DIMM temperature prediction and comparison with data for a day long run.....	78
Figure 4.4-35. Cooling Power prediction and comparison with experimental data for a day long summer run.....	79
Figure 4.4-36. Outdoor Heat Exchanger fans speed prediction and comparison with experimental data for a day long run.....	79
Figure 4.4-37. Control methods - Fan and pump speed as a function of post cooler water temperature, (a) Dry cooler fan control with fixed seeds for external and internal pumps, (b)	

Control of dry cooler fans and external pump with fixed speed for internal pump, (c)	
Control of dry cooler fans, external pump, and internal pump.....	80
Figure 4.4-38. Cooling power and coolant temperatures for different control methods (a)	
Cooling power usage by dry cooler fans, external pump, and internal pumps (b) Rack inlet	
water temperatures, (c) Rack inlet air temperatures.....	81
Figure 4.4-39. A simple graphical user interface for the system model. System model	
prediction of the key server components temperature, of the power consumption and annual	
average power consumption and annual energy and cost savings for a typical year in	
Poughkeepsie, NY. The typical outdoor air temperature profile was obtained from NREL	
database [21].....	82
Figure 4.4-40. System model prediction for a typical year in Raleigh, NC. The typical	
outdoor air temperature profile was obtained from NREL database [21].....	83
Figure 4.4-41. Psychometric chart illustrating the ASHRAE recommended guideline classes	
for air cooled IT equipment.....	84
Figure 4.4-42. Outdoor dry bulb temperatures for nine US cities having different climates.....	85
Figure 4.4-43. Temperature and power variation for warm and cool climates, (a) Coolant	
temperatures for Dallas (warm), (b) Cooling device (fan and pump) speeds for Dallas	
(warm), (c) Coolant temperatures for San Francisco (cool), (d) Cooling device (fan and	
pump) speeds for San Francisco (cool).....	86
Figure 6-1. IBM System X iDataPlex Direct Water Cooled dx360 M4 Server.....	88
Figure 9-1. Overview of renewable energy based net-zero data center.....	91

List of Tables

Table 1. Thermal test chamber data for air cooled servers.....	29
Table 2. Thermal test chamber data for partially water cooled servers.....	29
Table 3. One to one comparison of the water cooled node with its air cooled counter-part.....	35
Table 4. Node level thermal data comparing results for lidded modules with thermal grease with	
results for the same modules with direct attach LMTI in the DELC server rack environment.....	38
Table 5. Summary of test conditions during the three comparative standard tests as well as a	
fourth low cooling power test.....	64
Table 6. Steady state test cases for system model validation.....	76
Table 7. Average coolant temperatures for nine US cities for model simulation using Case 3	
control for August 15, NREL [21] typical year temperature data.....	85
Table 8. Average fan and pump speeds and cooling powers from model simulation using Case 3	
control Nine US cities, August 15, Use of NREL typical year temperature data [21].....	86
Table 9. Energy/cost savings for a 1 MW IT load, for a typical August 15 day, Traditional	
versus DELC data center.....	87

1. Executive Summary

Data centers are typically large warehouse-like facilities that house thousands of computer servers to process the growing volume of electronic transactions throughout the world. Data center power usage, which in 2010 amounted to 2% of the United States energy consumption [1], has been growing rapidly as new Internet applications are developed, requiring computing resources such as high-density cloud computing server farms.

All of the electricity consumed by data centers is ultimately turned into heat. This heat is removed using fans, pumps and air conditioning equipment, which comprise roughly a quarter of the total data center power consumption. In addition to energy usage, data centers in the US consume roughly 0.625 to 1 gallon of water for every kilowatt-hour of electricity [2, 3]. For example, a 10 megawatt data center can use up to 150,000 – 240,000 gallons of water a day [2, 3] and roughly 2.5 megawatts of power to cool the equipment in the data center [4].

Project Goal

Recognizing the growing demand for electrical and water resources to support data centers, IBM -- as part of a US Department of Energy cost shared grant -- has undertaken a project to develop a highly energy efficient chiller-less, liquid cooled system for data center cooling that would reduce the cooling energy usage with the additional goal of eliminating daily consumption of water.

Results Summary

IBM engineers and scientists at the IBM T.J. Watson Research Center, and in IBM Systems and Technology Group in Poughkeepsie and Raleigh, have reported that trial runs with the experimental direct liquid cooling system located in Poughkeepsie New York, developed under a DOE Advanced Manufacturing Office (AMO) American Recovery and Reinvestment Act (ARRA) award have demonstrated up to approximately a 90% reduction in the energy required for cooling when compared to traditional chiller-based data centers [5-9]. A longer term study over a nine week period performed from May to July 2012 was consistent with the trial runs achieving up to 90% reduction in average cooling power [10]. In addition, the experimental direct liquid cooling system did not consume water to achieve this reduction in cooling energy. When compared to a traditional 10 MW data center, which typically uses 25% of its total data center energy consumption for cooling [4], the experimental measurements show that this technology could potentially enable a cost savings of roughly between \$800,000-\$2,200,000/year (assuming electricity costs of 4¢ to 11¢ per kilowatt-hour) through the reduction in electrical energy usage.

Technical Approach

The technical approach taken by the IBM team was to develop a closed loop liquid cooling system as shown in Figure 1-1 below, which does not use energy intensive vapor-compression refrigeration based cooling. All the cooling required is achieved by using the outside air environment. This cooling approach provides a chiller-less cooling system with all year round “economizer” operation to achieve up to 90% reduction in cooling energy compared to a chiller based system [2-6].

The system as shown in Figure 1-1(left-side) includes a sealed equipment rack, which was constructed to both contain and extract the heat generated from the IT equipment housed inside. A liquid loop was constructed to transport the heat from the Sealed Rack to an Outdoor Heat

Exchanger, which is placed outside the data center. The system operates using only the outside air environment for cooling and offers two other key advantages. First, the system transports heat from the data center to the outside environment without exposing the equipment to the outside air that contains humidity and other contaminants which degrade IT equipment reliability. Secondly, the Outdoor Heat Exchanger rejects the heat to the outside air environment through a radiator and fan system operating without consuming external water.

Figure 1-1. Highly energy efficient chiller-less, liquid cooled system for data center cooling.

Figure 1-2. Efficient cooling solution for the volume server (a) Plan view of the liquid cooled server (b) Server liquid cooling structures for processor and memory cooling.

A new method of efficiently extracting the heat from the servers was required to make this system work effectively. The IBM team engineered a new efficient cooling solution for volume servers by bringing liquid coolant directly into the servers, replacing the traditional air cooled systems. As shown in Figure 1-2(a), liquid coolant was routed into the IT equipment by installation of a liquid cooling assembly. As shown in Figure 1-2(b) the structures route the liquid coolant to cold plates that are attached to the processors and to cold rails attached to the cold plates of the memory in each individual server in the rack. This approach removed roughly up to 70% of the heat by more efficient direct conduction of the heat from the most energy intensive processor and memory components to the liquid coolant. The remaining heat was removed by air flow and an air-to-liquid heat exchanger inside the sealed rack. Prior to entering the server liquid cooling assemblies, recirculating liquid cooled by passing through the Outdoor Heat Exchanger enters the rack heat exchanger extracting heat from the recirculating air within the rack. It was critical to improve the thermal efficiency of the design to successfully achieve the goals. In addition, since the rack is sealed and the air flow through each server is reduced, an additional benefit is quieter operation than a comparable air cooled machine.

Commercialization Activities

The successful IBM experimental demonstration of this cooling technology is already having an impact on IBM products. In June, the Leibniz Superconductor Center in Germany announced the world's fastest commercially available hot-water-cooled supercomputer built with IBM System x iDataPlex Direct Water Cooled dx360 M4 Servers, including some of the technologies developed with this award. The technologies are also drawing additional external interest for direct liquid cooled volume servers.

2. Introduction

In response to the growing use of electrical and water resources used by data centers, IBM as part of a US Department of Energy cost shared grant, initiated a project to develop a highly energy efficient data center cooling system. The new cooling system would use a chiller-less, liquid cooled system for data center cooling which would reduce the cooling energy required to cool a data centers from a current 25-30% [4,11-14] as shown in Figure 2-1(a) below to only 5% or less of the total data center energy. The new technologies proposed by IBM would replace state of the art Refrigeration Chiller Plants and CRACs (Computer Room Air Conditioning) with a liquid cooling system using the outdoor air ambient environment, thus eliminating the use of chillers which are the largest cooling energy usage component as shown in Figure 2-1(b).

In 2010, there were roughly 33 million computer servers installed worldwide. In the US alone, about 12 million computer servers were installed of which 97% were volume servers and 3% were Mid-range and High-end servers [1]. Thus, to maximize the energy impact of reducing cooling energy usage this project focused on the largest segment of the server market which is the Volume server segment. IBM System x3550 M3 Volume server was chosen to demonstrate the new technology.

The energy usage of data centers reported by the EPA [15] was 61 billion kWhrs in 2006 which had doubled since 2000 and was projected to double again by 2011. The EPA [15] reported that Volume servers, the fastest growing segment of the market, were responsible for 68% of electrical usage and

based upon the expected growth rate the electricity use by Volume servers was projected to reach 42 billion kWhrs in 2011.

(a) Typical total data center energy breakdown

(b) Typical data center cooling energy breakdown

Figure 2-1. Typical data center (a) total and (b) cooling energy breakdown [11].

As shown in Fig 2-1(a) the energy required to cool the IT equipment is roughly 25-30% of the total data center energy usage. For the projected Volumes server energy use projected growth rate this would reach 21 billion kWhrs in 2011. In the newly proposed chiller-less liquid cooled data center, the cooling energy would be reduced to 4 billion kWhrs. If an assumption is made for market penetration of a quarter of US data centers, the potential cooling energy savings would be 4.25 billion kWhrs per year.

The commercialization path for the technology is to implement warm water cooling into Volume servers for data center applications. In June, the Leibniz Superconductor Center in Germany announced the world's fastest commercially available hot-water-cooled supercomputer built with IBM System x iDataPlex Direct Water Cooled dx360 M4 Servers which included some of the technologies developed with this DOE award. The technologies are also drawing additional external interest for direct liquid cooled volume servers.

3. Background

Information Technology (IT) data centers are facilities which house numerous computer systems arranged in the form of electronic racks. Data centers vary in size and may house up to thousands of racks, with each rack typically consuming 10-30 kW of power. A study from the Lawrence Berkeley National Laboratory [14] has reported that in 2005, server driven power usage amounted to 1.2% and 0.8% of the total US and Worldwide energy consumption respectively. From 2005 to 2010, energy use by these data centers and their supporting infrastructure has increased to 2% of total US energy [1]. Cooling contributes a significant portion of this energy use. Thus, understanding and improving the energy efficiency of data center systems is important from a cost and sustainability perspective.

Figure 3-1 shows a facility level schematic for the cooling system that is used to transfer heat from the server exhaust air to the ambient outdoor air, which is the ultimate heat sink. The transfer of heat

from the IT equipment to the room level coolant flow is depicted in Figure 3-1 via the sketch labeled as the data center building. As shown in Figure 3-1, racks of IT equipment are arranged in rows to form several aisles in which two rows of racks face each other at their inlets. These racks are usually air cooled inside the servers and thus the racks require a continuous and reliable supply of cool air for their operation. This cool air is supplied by the CRACs. The cool air enters the room via floor perforated tiles, passes through the racks being heated in the process, then finds its way to the intake of the room CRACs, which cool the hot air and blow it into the under floor plenum. The chilled water from the chiller is usually pumped through a network of under floor pipes which supply and remove water to and from the CRACs. The air supplied to such equipment is typically in the 15-32°C range for allowable temperatures with 18-27°C being the nominal long term recommended band [19].

Figure 3-1. Traditional chiller plant based data center cooling loop.

As shown in Figure 2-1(a), the IT equipment consumes about 50% of the total electricity, and total cooling energy consumption is roughly 25-30% of the total energy use [4, 12, 13]. Thus, using these typical values, the cooling energy use is about 50% of the IT energy use. This is an important baseline metric with respect to the cooling inefficiencies of air-cooled data centers [11]. Subsequent results presented herein will use this baseline metric to quantify the energy savings realized through the innovative cooling designs demonstrated. Figure 2-1(b) also depicts the various cooling infrastructure components that are made up of three elements: the refrigeration chiller plant (including the cooling tower fans and condenser water pumps, in the case of water-cooled condensers), the building chilled water pumps, and the Data Center floor air-conditioning units or CRACs. About half the cooling energy is consumed at the refrigeration chiller plant and about a third is used by the room level air-conditioning units for air movement, making these the two primary contributors to the data center cooling energy use.

The cooling of a data center facility is shown in Figure 3-2. Refrigerated water leaving the refrigeration chiller plant is circulated through the CRAC units, using building chilled water pumps. This water carries heat away from the raised floor room and rejects the heat into the refrigeration

chiller plant evaporator heat exchanger. The refrigeration chiller plant operates on a vapor compression cycle and consumes compression work via a compressor. The refrigerant loop rejects the heat into a condenser water loop via the refrigeration chiller condenser heat exchanger. A condenser pump circulates water between the refrigeration chiller plant and the evaporative cooling tower. The evaporative cooling tower uses forced air movement and water evaporation to extract heat from the condenser water loop and transfer it into the outside ambient environment.

Figure 3-2. Schematic of typical data center facility cooling infrastructure

In this “standard” facility cooling design, the primary energy consumption components include [11]:

- Server fans
- Computer Room Air Conditioning (CRAC) blowers
- Building Chilled Water (BCW) pumps
- Refrigeration chiller compressors
- Condenser water pumps
- Cooling tower blowers

Several factors, which contribute to the inefficiency of current data center cooling designs and excessive energy consumption, include:

- Inefficiency of using air with low thermal conductivity and heat capacity as a coolant
- Large thermal resistance between computer chips and coolant
- Use of sub-ambient temperature air and water which requires energy intensive chillers
- Use of daisy chained loops adding inefficiencies in each heat-exchanger

This project primarily focused on an innovative server and data center cooling design that addresses the aforementioned inefficiencies by eliminating the energy usage of room air conditioning devices and the chiller compressor through the use of liquid cooling at the server and by operating at coolant temperatures that are above the ambient during the entire year.

The objective of this project was to reduce the cooling energy usage to 5% of total data center

energy usage which addresses the second area of interest, Information and Communication Technologies R&D for Energy Efficiency. In addition to significant energy usage, traditional data center cooling also results in refrigerant and make up water consumption that are eliminated in the new cooling designs discussed. In essence, the project focused on the development of two complementary novel technologies that radically reduce the energy consumption of data centers.

Firstly, a server compatible Liquid Metal Thermal Interface (LMTI) [20] was developed to improve the thermal conduction path of the hot server components to the data center ambient cooling. This liquid metal thermal interface has a thermal conductivity an order of magnitude better than state of the art materials. When integrated directly between a bare die and a water cooled heat sink, this technology achieved a significant improvement in thermal conduction and enabled the processors to operate in a much higher ambient temperature environment.

Secondly, a dual enclosure air/liquid cooling system was developed to allow direct cooling from the outside ambient environment. This Dual Enclosure Liquid Cooling (DELC) system, illustrated in Figure 3-3, uses recirculated air and water which are cooled only by heat exchange with the outside ambient air. The DELC system includes a sealed equipment rack, which was constructed to both contain and extract the heat generated from the IT equipment housed inside. A liquid loop was constructed to transport the heat from the Sealed Rack to a Outdoor Heat Exchanger, which is placed outside the data center. The system operates using only the outside air environment for cooling and offers two other key advantages. First, the system transports heat from the data center to the outside environment without exposing the equipment to the outside air that contains humidity and other contaminants which can degrade IT equipment reliability. Secondly, the Outdoor Heat Exchanger rejects the heat to the outside air environment through a radiator and fan system operating without consuming external water. The DELC system also comprised sensors and servo control algorithms which can adjust the cooling component operating parameters based upon the server rack heat load and the outdoor air temperature to minimize the cooling component energy usage.

Figure 3-3. Schematic of the Dual Enclosure Liquid Cooling (DELC) system.

The integration of LMTI and DELC system eliminated the data center refrigeration chiller plant as well as several other cooling components, thus allowing for up to 90% reduction in the cooling energy cost.

To maximize the energy impact, this project focused on the largest segment of the server market which is the Volume server. The processor power for Volume servers is currently in the range of 60-130 watts and has been increasing. The solutions demonstrated in this project are extendible to much higher power processors (> 200W) and can be applied to Mid-range and High-end servers, allowing extendibility to future server requirements. Thus far, water cooling has been based on chilled water and has been limited to High-End systems due to the cost of infrastructure, implementation and energy required to provide refrigerated chilled water. The DELC and LMTI technologies will advance the state of the art of water cooling by enabling the use of ambient cooled water to provide a cost effective solution for commercial Volume, Mid-range and High-End servers where Volume servers, the largest market segment, will use processor power of 90 watts or greater and node power of 200 watts or above.

4. Results and Discussion

The section is divided into four chapters with Chapter 1 describing the data center liquid cooling design, Chapter 2 describing the server level liquid cooling design and advanced metal interfaces utilized, Chapter 3 describing the component-level to data center-level modeling and simulations and finally, Chapter 4 describing the system characterization, operation and system performance projection.

Chapter 1: System Design

This chapter details the Volume server configuration, DELC facility design and modeling, data collection system and process, system protection measures, and server rack design, modeling and measurements.

1.1 System Overview

A dual enclosure air/liquid cooling system, illustrated in Figure 4.1-1, was developed to allow direct cooling from the outside ambient environment. On the right of Figure 4.1-1, the external dry cooler unit is depicted with a heat exchanger coil and a fan to reject the data center cooling load into the ambient air flowing through the radiator type air-to-liquid heat exchanger coils. An external pump is also shown which circulates the coolant through the dry cooler unit and buffer coolant distribution unit (also referred as a modular water unit or MWU). The recirculation valve allows part of the hot coolant to bypass the dry cooler in winter months to ensure that the liquid entering the data center is above the dew point temperature of the data center so as to prevent any condensation. The left side of Figure 4.1-1 shows the data center cooling design inside the data center which includes a buffer coolant distribution unit and the 100% liquid cooled rack. Thus, in contrast to the traditional cooling design illustrated via Figure 3-1(a), there is no refrigeration unit and only three powered cooling devices – the external dry cooler fans, the external pump, and the internal pump. The rack liquid cooling includes both air and liquid cooling devices and will be described in subsequent sections.

Figure 4.1-1. Schematic representation of the energy efficient dual enclosure data center liquid cooling test facility. IT servers are warm water cooled with the heat ultimately rejected to ambient air via a liquid-to-air heat exchanger.

Figure 4.1-2 shows photographs of the Data Center cooling loop infrastructure. Figure 4.1-2(a) shows the external dry cooler loop with five fans mounted above a large air to liquid heat exchanger coil. The air flow is drawn into the heat exchanger from below and through the sides, and the hot air expelled upwards from the fans. To the right in Figure 4.1-2(a), an auxiliary enclosure is seen which houses the external pump and recirculation valve that was described above in Figure 4.1-1. This enclosure also includes some instrumentation including a pressure meter, and temperature sensors to measure coolant temperature leaving the dry cooler as well as before and after the valve “tee”. Thus, using valve operation during winter months, it would be possible to determine the coolant temperatures before and after mixing of hot coolant that bypasses the dry cooler. Figure 4.1-2(b) shows a photograph of the piping layout inside the experimental data center. The piping coming into the data center can be seen from the outside when looking at the top right of Figure 4.1-2(a). Inside the data center there is a bypass loop to allow coolant bypass and device servicing for the 50 micron filter. Also seen in Figure 4.1-2(b) are the temperature sensors, the flow meters for the internal and external loops, pressure sensors and the buffer coolant distribution unit. This buffer unit allows separation of the internal and external coolant loops. Such separation allows the use of water inside the data center even in winter months when the external loop would require a water-glycol mixture to withstand the cold New York winters. The buffer unit also allows for the use of specially treated water on the system side, i.e. the water flowing through the rack cooling devices, which results in a greater tolerance at the data center level to less clean water in the external loop which is often the case. The buffer unit seen at the bottom of Figure 4.1-2(b) is comprised of a pump and a plate type liquid-liquid heat exchanger and if desired could be packaged into the bottom of a server rack [16].

Figure 4.1-3 provides further details of the liquid cooled server rack. The rack cooling includes front and rear covers that duct the air flow to and from the Volume servers through a side air-to-liquid heat exchanger coil (Side Car) which cools the hot server exhaust air to a satisfactory temperature for intake into the servers.

Figure 4.1-2. Photographs of data center cooling hardware (a) External dry cooler unit, (b) Internal (room) piping layout and instrumentation

Figure 4.1-3. Rack liquid cooling design, (a) Photograph of the front of the rack, (b) Plan view schematic of rack internals [18], (c) Plumbing of server node liquid cooling to rack manifolds.

In Figure 4.1-3(a), the front of the rack can be seen with the Side Car unit on the right. Figure 4.1-3(b) shows a schematic section plan view of the rack and depicts the server node as well as the side car unit. The water flow from the buffer unit first enters the Side Car heat exchanger to cool the recirculating server air flow. After flowing through the Side Car air to liquid heat exchanger, the partially heated water then enters the rack inlet manifold which distributes the water to each of the liquid cooling sub-assemblies inside the nodes. A flexible hose is attached to the inlet of the node liquid cooling assembly with a one-eared hose clamp. The other end of the hose has a quick disconnect coupler and supplies the water to the node. A similar hose returns the warm water from the node to the rack level exit manifold. Thus, there are two rack manifolds, one for liquid supply and one for return, with each manifold having 42 ports with quick disconnect couplers for connection to the node liquid cooling devices using the flexible hoses. Figure 4.1-3(c) illustrates the plumbing from the rack manifold to the node as well as the numerous ports on the manifold. The rack cooling design described above and shown in Figure 4.1-3 accommodates both air and liquid cooled devices at the server level while transferring the entire rack heat load into the water at the rack level. This is an important attribute that provides flexibility in the equipment to be housed inside the rack. However, it should be noted that while there are both air and water cooled devices in the rack, both of these sets of components must accept warm coolant whether air or water to practice the Data Center design that is presented in this report. The Side Car unit has been applied to fully air cooled racks prior to the application [17] discussed herein.

One of the key features of this server rack design is the elimination of virtually all heat loads to the data center environment. Another important feature is that the economizer based cooling isolates the rack from the outside air, thus minimizing the risk of component contamination through particulate, chemical or gaseous matter that may be present in the outdoor ambient from time to time.

Figure 4.1-4. Hybrid air-water cooled 1U server designed for intake of 45°C water and 50°C air.

Figure 4.1-4 shows a perspective view of the hybrid cooled server comprised of air cooled and water cooled devices. The server is an IBM x3550 M3 server which is 1U tall (1.75 inches or 44.45 mm) and fits in a standard 19 inch (483 mm) rack that is actually about 24 inches wide (609.5 mm). The microprocessor modules are cooled using cold plate structures. The Dual-In-Line-Memory Module (DIMM) cards are cooled by attaching them to a pair of conduction spreaders which are then bolted to water cooled cold rails. The microprocessors and DIMMs have a nominal maximum power of 130 W and 6 W each, respectively, and the server node power is about 400 W for its maximum. All the other components in the server shown in Figure 4.1-4 are air cooled. These devices include the storage disk drives, the power supply, and various surface mounted components on the Printed Circuit Board (PCB). While the air cooled version of this server had six fan-packs (two per pack), in the water cooled server three of these six fan packs were removed. The fan control algorithm was also modified to allow the server to operate up to 50 °C inlet air temperatures compared to existing commercial servers which typically enforce power down procedures if the inlet air temperature rise above 40 °C.

1.2 Temperature Excess between Outdoor Air and Server Node Coolant

The DELC system, being an ambient cooled system, is dependent on the outdoor ambient conditions. The coolant temperature leaving the Outdoor Heat Exchanger, the coolant temperature entering the rack of servers and the air and liquid coolant temperature entering the servers are driven by the outdoor ambient temperature. As a result, it is important to estimate what these coolant temperatures will be for a given location and time of the year. To do this, it is necessary to characterize the temperature difference between the entering cold fluid and the exiting hot fluid (approach temperature difference) on either side of the various heat exchanger devices in the loop including the a) dry cooler air-to-liquid heat exchanger, b) the buffer unit liquid-to-liquid heat exchanger, and the c) Side Car air-to-liquid heat exchanger. In the case of the dry cooler this will be the difference between the outdoor air temperature entering the coil and the liquid temperature leaving the coil. Figure 4.1-5 provides data based calculations for the approach temperature difference, $\Delta T_{\text{Approach}}$ for the three heat exchanger devices that are in the loop which was described in the schematic shown in Figure 4.1-1. For a sample dry cooler as shown in Figure 4.1-5(a) with a fan speed of 500 RPM and an external water flow rate of 8 GPM, the $\Delta T_{\text{Approach}}$ for the dry cooler would be 1.7 °C for a heat load of 15 kW. For a buffer heat exchanger as shown in Figure 4.1-5(b) with an external water flow rate of 7.5 GPM and an internal (rack side) water flow rate of 5 GPM, the $\Delta T_{\text{Approach}}$ for the buffer heat exchanger unit will be 4.5 °C for a heat load of 15 kW. For a side care heat exchanger as shown in Figure 4.1-5(c) with an the for an internal flow rate of 5 GPM, a rack air flow rate of 1500 cfm and an air side heat load of 5 kW, the $\Delta T_{\text{Approach}}$ would be 3.9 °C. The air heat load at the Side Car heat exchanger will be less than the total heat load because only part of the rack heat load is rejected to the air flow circulating inside the rack.

Thus, as a sample calculation, addition of all these approach temperature differences at the three heat exchange devices results in a total $\Delta T_{\text{Approach}}$ of $1.7 + 4.5 + 3.9 = 10.1$ °C for the air temperature entering the server node. Since the water supplied to the rack will absorb 5 kW in heat load at the Side Car air-to-liquid heat exchanger, its temperature for the 5 GPM flow rate will rise by 3.8 °C, and thus the total $\Delta T_{\text{Approach}}$ for water entering the node can be calculated as $1.7 + 4.5 + 3.8 = 10$ °C.

Figure 4.1-5. Approach temperature differences for the three heat exchanger devices in the loop (a) Dry cooler external air-to-liquid heat exchanger coil, (b) Buffer unit liquid to liquid heat exchanger, (c) Rack Side Car air to liquid heat exchanger

The calculations presented in Figure 4.1-5 are for a total heat load of 15 kW out of which the rack air heat load is 5 kW. In the actual system, the total load as well as the air load can vary based on the actual computational workload as well as the temperature differential between the rack coolant temperatures and the data center room ambient which will drive the heat loss into the room. There are other factors that can influence these heat loads including the heat loss in all the piping in the entire loop even though the piping in this case is insulated. In addition to these factors, the server fan speed is a function of server device temperatures and the inlet air temperature as measured by the server air inlet temperature sensor, and thus the rack air flow rate can vary from case to case. When the rack air flow rate varies, the $\Delta T_{Approach}$ for the Side Car heat exchanger will also change as shown in Figure 4.1-5(c). In winter months due to freezing weather outside and extremely low outdoor air temperatures, the coolant in the external loop will need to be a water-glycol mixture. The use of an anti-freeze in the external loop will increase the $\Delta T_{Approach}$ of both the dry cooler coil and the buffer unit liquid-to-liquid heat exchanger for the same conditions reported in Figures 4.1-5(a)-(b).

1.3 Flow Modeling and Measurement

Early in the design effort, flow modeling simulations were performed to estimate liquid flow at the external system cooling level and through the nodes within the server rack. These modeling efforts were conducted using a flow network modeling tool. Figure 4.1-6 shows an example of an early

model of the external cooling loop without an intermediate MWU. A flow element (labeled Rack Cooling) representing the overall rack impedance may be seen in the upper left corner and another element representing the dry cooler impedance may be seen in the lower right hand corner of the figure. It may also be noted that the distance in elevation (approximately 15 feet) between the server rack and the outside dry cooler and the horizontal distance (approximately 20 feet) outside the building were also taken into account. Figure 4.1-7 shows an example of the flow network model created to simulate flow distribution through the 42 servers within the rack. Data derived from experiments and modeling at the individual node level were used to determine the flow resistances of all flow elements in every node and are depicted in the middle of each leg of the “ladder” network.

Figure 4.1-6. Flow model of external system cooling loop.

Figure 4.1-8 provides an example of the results from this early model in terms of the liquid flow rate that would be supplied to each node at various rack level liquid flow rates. These results clearly illustrated the potential for a significant non-uniform distribution of flow across the rack. In the final design, this was avoided by increasing the supply and return manifold diameters to 1 inch and through the use of the parallel channel cold plate design. These design changes also reduced the overall rack liquid flow pressure drop and allowed a greater flow through each node assembly. However, the early flow model results called attention to the need to verify the uniformity of the flow distribution in the final server rack hardware. It was not practical to place a flow meter between a server node and its flow connection to the supply or return manifold, so the concept of a node flow-pressure drop simulator was employed.

Figure 4.1-7. Flow model of rack level water cooling piping.

Figure 4.1-8. Predicted nodal flow distribution from bottom to top of rack for 3/4 inch diameter supply and return manifolds for various rack flow rates.

Figure 4.1-9. Node flow-pressure drop simulator for experimentally determining flow through an individual node.

A schematic of a node flow-pressure drop simulator is illustrated in Figure 4.1-9. It is comprised of a hose assembly with a flow regulator valve in the middle and tees for pressure taps on either side of the valve. The same quick-connect fittings are provided at each end of the hose assembly as are used to connect an actual node to the supply and return manifolds. The flow regulator valve is adjusted on a test flow bench to give nearly the same pressure drop at the nominal design flow rate as an actual node. The server rack is then fully populated with node flow-pressure drop assemblies to simulate a full rack of server nodes. It may be noted that not all the test assemblies have the tees for pressure

drop sensing. It is only necessary to have one simulator hose instrumented in this fashion and move it from node position to node as desired. By reading the pressure drop across the instrumented test hose, the flow that an actual node in that position would receive may be determined. An example of the similarity between the pressure drop versus flow characteristic curve of an actual node with the parallel channel cold plate design and several test hose assemblies (i.e. hoses A, B, C and 18) is shown in Figure 4.1-10. It may be seen that a very close match is achieved at the nominal flow rate of 0.7 liters/minute. Figure 4.1-11 shows the pressure drop-flow correlation for the instrumented hose assembly used in tests. The open symbols and solid line are predicted values using the equation and the solid symbols are flow rate measured using a separate flow meter.

Figure 4.1-10. Comparison of node simulator hose assembly's flow-pressure drop characteristics with flow-pressure drop characteristics of actual node containing parallel channel cold plates.

Figure 4.1-11. Flow -pressure drop correlation for node simulator hose test assembly.

Figure 4.1-12. Node flow measurement test set-up.

Figure 4.1-13. Measured flow through simulated node as a function of node position measured from bottom of rack.

Figure 4.1-12 shows a portion of the front of the server rack with the with node flow-pressure drop assemblies in place. The instrumented node-flow pressure drop assembly connected to a differential pressure manometer is shown in the center of the picture. Tests were conducted varying the overall liquid flow rate to the rack from 11.4 to 37.5 liters/minute. The flow rates at various node positions from the bottom to the top of the rack were measured by moving the instrumented hose to each of the positions shown in Figure 4.1-13. As shown in the figure, the flow rate is nearly uniform from bottom to top. A small drop in flow within tolerable limits was observed in the topmost position.

1.4 System Monitoring and Control

To monitor the data center test facility and collect the system and facility thermal, hydraulic and power data, instrumentation was provided at various locations in the two cooling loops as shown in Figure 4.1-1 (labels indicate probes) and in two of the installed servers. Temperature, pressure, flow-rate, humidity and power measurements were collected through a programmable logic controller (PLC) unit, shown in Figure 4.1-14, and an Agilent data-logger, using a custom-built program running in DAQ Environment. The customized program allowed manual control of the data center cooling equipment: the internal loop pump, the external loop pump, the recirculation valve and the dry-cooler fan speeds. It also provided a platform to implement various automated control algorithms.

Figure 4.1-14. Photo of PLC control unit used to monitor and control the data center test facility. The PLC is connected to a remote computer with monitoring and control software.

In addition to the facility side data, thermal data from each of the servers in the rack are also collected. Figure 4.1-15 shows the schematic of the server side data collection process. The server rack houses about 40 servers, each of which is connected to the head node via an Ethernet switch. Both the head node and the Ethernet switch reside outside of the server rack. The head node is remotely accessible for starting/stopping simulated workload, data collection and server monitoring scripts. The head node can also be used to remotely power-on/off any server in the rack. Thermal data was collected from each server simultaneously at about one minute time intervals. Thermal data

collected from each server included – each CPU core temperature (6 cores per CPU, 2 CPUs per server), DIMM temperatures (12 DIMMs per server), server inlet air temperature, server power, server fans' RPM and IOH chip temperature. All of this thermal data is parsed and formatted at the head node and is made available for system monitoring and control. For example, scripts were implemented on the head node to implicitly monitor the coolant temperature entering the servers by running one of the servers at “Idle-state” and monitoring the CPU thermal data. That temperature can be utilized to provision IT heat load and/or to perform controlled shutdown of servers. Additionally, scripts were implemented on the head node to incorporate controlled temporal variation of simulated IT workload resulting in dynamically varying IT power and to test the performance of the novel dynamic servo control discussed in subsequent chapters.

The facility control algorithms are implemented both in the PLC and the DAQ programming environment. Simplified proportional-type control algorithms were implemented at the PLC for robustness. The key parameters that define the behavior of these control algorithms were made available to the DAQ program which could be used to alter the ranges of the control. The two control algorithms implemented at the PLC were i) Linearly vary the dry-cooler fan speed from a minimum to a maximum based on a specified temperature range of the water entering the buffer unit on the facility side and pump and ii) linearly vary both the dry-cooler fan speed and the external loop pump speed over their specified ranges based on a specified temperature range of the water entering the buffer unit on the facility side. These algorithms can be selected, modified and run using the DAQ control environment, but the control code itself resides on the PLC. More complex proportional-integral type control algorithms were implemented completely within the DAQ control environment due to the need for increased programming flexibility and ease of modification.

Figure 4.1-15. Server side data collection process and protection measures.

Along with the control algorithms several safety controls including leak detection were also implemented at the PLC, DAQ and server levels.

Chapter 2: Node Liquid Cooling Design and Advanced Metal Interfaces

One of the essential parts of this project was to bring liquid coolant into the server to cool the relatively high heat dissipating server components such as the processors and the memory modules. Figure 4.2-1(a) shows the air cooled version of the IBM x3550 M3 server that was retrofitted to allow liquid cooling of server components. The hybrid air/liquid cooled version is shown in Figure 4.2-1(b). Some portion of the chassis was cut in the front right section of the server and the right most server fan-pack was removed to allow for the placement of the node liquid cooling assembly.

(a)

(b)

Figure 4.2-1. Photographs of (a) Standard air –cooled version of IBM x 3550 M3 server. (b) Hybrid air/liquid-cooled version with cold-plates for CPUs and cold rails for DIMM Cooling.

Upon placement, the cooling assembly was secured in place by using chassis rework components. Moreover, as the high heat dissipating components were liquid cooled, less air flow was required to cool the other components in the server. Thus, a fan-pack was removed from each of the three cooling zones – CPU1, CPU2 and DIMM cooling zones. At the vacant fan-pack space, flow blockers were placed to prevent any air recirculation in that vacant space. In the hybrid air/liquid cooled version, the microprocessor modules are cooled using cold plate structures and the Dual-In-Line-

Memory Module (DIMM) cards are cooled by attaching each of them to a pair of conduction spreaders which are then bolted to water cooled cold rails. A detailed description of the design, modeling and performance of the node liquid cooling components and thermal interface materials used is presented in this chapter.

2.1 Overview of Liquid Cooled Server

Figure 4.2-2 illustrates several of the water cooling structures that are used to construct the hybrid air/water cooled node shown in Figure 4.2-1. The cold plate used to extract heat from the microprocessor modules is shown in Figure 4.2-2(a). It is made up of two parts – the copper core with the water flow channels that make up a parallel plate fin array, and the aluminum frame that is soldered to the copper core and provides the structural support during clamping of the cold plate to the module. The two part design allows the use of the higher thermal conductivity copper material for heat transfer while using the less expensive aluminum for structural function.

Figure 4.2-2. Liquid cooled components inside the server; (a) cold plate for the microprocessor module, (b) DIMM with conduction spreader, (c) front cold rail, (d) middle cold rail and, (e) end cold for memory liquid cooling.

Figure 4.2-2(b) displays the sub-assembly made of two copper spreader plates that are mechanically attached to a DIMM card using spring clips with a thermal interface material between the spreader and the DIMM. The spreader-DIMM sub-assembly is inserted into an electrical socket on the PCB and then the spreaders are bolted on both sides to liquid cooled cold rails which extract the heat from the DIMMs through the spreaders. There is another thermal interface material between the spreader

and the cold rail shown in Figure 4.2-2(c). Figures 4.2-2(c), (d) and (e) show three different cold rails designs that were required to meet the dimensional constraints of the server which included capacitors mounted on the PCB, cards mounted on sockets, and fans located at the front of the server. The cold rails have tapped screw holes in their solid portions. The distinct cold rail shapes shown in Figures 4.2-2(c)-(e) result from the unique geometry constraints required for each of the cold rails. The cold rails shown in Figures 4.2-2(c)-(e) terminate in hose barb connections and represent the individual prototypes made for testing and not the design used in the full node cooling assembly.

Figure 4.2-3 shows the assembled node cooling sub-assembly that is made up of two cold plates, three cold rails, and copper tubing. The structure shown in Figure 4.2-3 is lowered onto a server board and is attached at the cold plate frames to the server PCB. The water flow path is also depicted in Figure 4.2-3 with the inlet being at the copper tube shown at the bottom left of the image (left of the pair of ports). The water flow splits at the first junction resulting in parallel flow through the front and middle cold rails with the flow then joining in front of the first cold plate (upstream). After flowing through the first cold plate the water passes through the end cold rail and then the second cold plate (downstream) after which it exits the node. Figure 4.2-4 illustrates the assembly of the DIMM with the spreader attached into the server node. Similar to the drawings seen in Figures 4.2-2(c) - (e), the cold rails in Figure 4.2-4 are special prototypes manufactured for testing purposes. Only the cold rails of the cooling loop appear in Figure 4.2-4. However, viewing the complete loop in Figure 4.2-3 in conjunction with Figure 4.2-4, the reader can understand better the assembly sequence that results in the liquid cooled server displayed in Figure 4.2-1. After the loop is assembled, the DIMM and spreader sub-assembly is inserted into the electrical sockets to make good electrical contact and then the two ends of the spreader are bolted to the cold rail using screws. As mentioned earlier, thermal interface materials are used between the spreader and the DIMM chips as well as between the spreader ends and the cold rails. Details of the server liquid cooling loop design, the advanced metal interfaces used and their performances are presented in the following sections of this chapter.

Figure 4.2-3. Node cooling sub-assembly for partially liquid cooled server.

Figure 4.2-4. Illustration of DIMM-spreader assembly into liquid cooled server node

2.2 Liquid Cooled Server Design

Of the several node liquid cooling configurations explored, four main candidate node liquid cooling designs, presented in Figure 4.2-5, were selected. Figure 4.2-5(a) illustrates design A, a node liquid cooling loop design with tubed cold-plates for the cooling of processor modules and cold-rails for DIMM cooling. The tubed cold-plate is made up of two parts: a continuous copper tube with two non-flattened circular lollipop bends and three flattened sections that are soldered to an aluminum bulk frame. Figure 4.2-5(b) illustrates design B, a node liquid cooling loop design having parallel channels based cold plate for processor cooling and cold-rails for DIMM cooling. The parallel channels- based cold-plate s also made up of two parts: the copper core with the water flow channels that make up a parallel plate fin array, and the aluminum frame that is soldered to the copper core and provides the structural support during clamping of the cold plate to the processor module. Figure 4.2-5(c) illustrates design C, a node liquid cooling design having only the tubed cold plates for the cooling of processor modules. Figure 4.2-5(d) illustrates design D, a similar design that has parallel channels based cold plate for the cooling of processor modules only. If either of the two designs illustrated in Figures 4.2-5(c) and 4.2-5(d) were to be used, then the DIMMs would have to be air cooled along with other components on the board.

The parallel channels based cold plate is a relatively higher cost and higher performance cold plate while the tubed cold plate is a relatively lower cost and lower performance cold plate. Due to serial flow pattern in the tubed cold plate, the pressure drop as well as the pumping power required is higher compared to the parallel channels based cold plate. Thus, Design C had the lowest thermal performance and was also the lowest cost of the four options. Design D was slightly more expensive than design C and had improved thermal and hydrodynamic performance. Designs A and B have additional cold rails for liquid cooling of DIMMs and were more expensive relative to design C.

Liquid cooling of the DIMMs is an efficient way of cooling as it reduces the thermal path by transferring the heat to the liquid coolant at the server level. Air cooling of the DIMMs would transfer the heat dissipated by the DIMMs to the liquid coolant at the Side-car air-to-liquid heat exchanger, a less efficient and longer thermal path. Moreover, DIMM temperatures were observed to be lower when liquid cooled, potentially improving reliability and error rates.

Figure 4.2-5. Candidate node liquid cooling loop designs (a) Tubed cold plates for processor module cooling and cold rails for DIMM cooling (b) parallel channels based cold plates for processor module cooling and cold rails for DIMM cooling (c) Tubed cold plates for processor module cooling only and (d) Parallel channels based cold plates for processor module cooling only.

Figure 4.2-6(a) shows the comparison of DIMM temperatures with respect to the coolant temperature for air cooled version with that for liquid cooled version. In the air cooled DIMMs version, the DIMMs in the front bank are closer to the server fans and are cooled first by air having temperature close to the server inlet air temperature. The DIMMs in the rear bank are cooled next by the pre-heated air. Thus, the rear bank DIMMs show a much higher temperature delta from the server inlet air temperature. In the liquid cooled DIMMs version shown in Figure 4.2-6(b), the liquid coolant enters the server and flows through two parallel flow paths – one going through the front cold-rail and the other going through the middle cold-rail. The liquid coolant flow then combines and flows through the cold plate and then through the end cold rail and then through another cold plate before exiting out of the server. The middle cold rail is common to both the front and rear DIMM banks and is used to liquid cool all the DIMMs in the servers. The front cold rail and the end cold rail are used to liquid cool the DIMMs in the front bank and rear bank, respectively. Due to pre-heat from upstream cold rails and cold-plate, the liquid coolant temperature in the end cold rail is roughly 2 °C higher than that in the front and middle cold rails. However, as the rear bank of DIMMs are partially cooled by pre-heated liquid coolant, the effect of pre-heat is not as significant as in the case of air cooled DIMMs.

(a)

(b)

Figure 4.2-6(a) Comparison of DIMMs temperature delta with respect to the coolant inlet temperature for air cooled DIMMs with that for liquid cooled DIMMs. (b) Server node liquid cooling assembly with cold plates for the micro-processors cooling and cold rails for DIMM liquid cooling. 6 DIMMs in the front bank (slot numbers 2,3,5,6,8 and 9) and 6 DIMMs in the rear bank (slot numbers 11,12,14,15,17 and 18).

In Figure 4.2-6(a), it can be seen that even for the front bank of DIMMs where there is the effect of pre-heat, the liquid cooled DIMMs show a much lower temperature delta to the coolant temperature than that for air cooled DIMMs. Moreover, the temperature delta to the coolant temperature was more or less similar for all the liquid cooled DIMMs. The air cooled DIMMs showed higher

temperature variability. Thus, from a thermal and cooling efficiency perspective, it was beneficial to have liquid cooling for the DIMMs as well.

The cold rails had relatively smaller cross-section for liquid flow than the bulk of the cooling assembly, resulting in relatively higher pressure drop. Figure 4.2-7 shows the hydrodynamic performance of all four node liquid cooling designs. For a given flow rate, Design D showed the lowest pressure drop, Design A showed the largest pressure drop while Designs B and C showed similar pressure drop. As pumping power is proportional to the product of flow rate and pressure drop, cooling the processors with tubed cold plates (Design C) would have used similar pumping power to cooling both the processors and DIMMs by using a parallel channels cold plate and liquid cold rails (Design B). Thus, Design B seemed promising as it offered microprocessor and DIMM liquid cooling with moderate pressure drop.

Figure 4.2-7. Hydrodynamic performance of the node liquid cooling loop designs.

The thermal performance of the tubed cold plate was compared with the parallel channels cold plate. The thermal resistance variation as a function of flow rate for both the cold plates is presented in Figure 4.2-8(a). The temperature contours at a chip-cold plate package cross-sectional plane passing through the center of the package for tubed cold plate with a liquid coolant flow rate of 0.238 gpm are shown in Figure 4.2-8(b). The temperature contours on a similar plane for parallel channels based cold plate with the same coolant and coolant flow rate are shown in Figure 4.2-8(c). With respect to the DELC system, this plot can be interpreted in two ways. First, a target thermal resistance can be achieved by the parallel channels cold plate with a much lower flow rate which in turn results in lower pressure drop and pumping power consumption at the system level. Secondly, a lower thermal resistance at a fixed flow rate allows for a higher coolant as well as higher outdoor ambient temperature operation for the same chip junction temperature. Such dual benefit of the parallel channels based cold plate superseded its cost drawback and thus, Design B was selected as an appropriate server liquid cooling design.

Figure 4.2-8. (a) Variation of thermal resistance as a function of flow rate for the tubed cold plate and parallel channels based cold plate. (b) Temperature contours at a chip-cold plate package cross-sectional plane passing through the center of the package for tubed cold plate (c) for parallel channels based cold plate with a liquid coolant flow rate of 0.238 gpm and inlet temperature of 45 $^{\circ}\text{C}$.

As illustrated in Figure 4.2-3, the liquid coolant flow bifurcates into two parallel flow paths and passes through the front and middle cold rails. The flow distribution in the two parallel flow paths as well as the overall pressure drop in the server liquid cooling loop depends on the flow impedance along the two paths. The flow distribution also affects the liquid cooled DIMMs temperature. A comprehensive study was performed using commercial computational fluid dynamics software to figure out optimum dimensions/specifications such as tube outside diameter and wall thickness, of copper tubes in the front and middle cold rails required for sufficient flow in both the cold rails, reduced server level pressure drop in the cooling loop and efficient cooling of DIMMs.

Figure 4.2-9 shows the numerical prediction of the flow distribution between the front and middle cold rails and of the pressure drop from the cooling loop inlet to the inlet of the first cold plate for a design flow rate of 0.9 liter per minute. The pressure drop across this section is a significant contributor to and accounts for roughly 50% of the overall cooling loop pressure drop. This configuration showed a flow distribution of 26% in the front cold rail for 0.9 lpm of total flow. The flow distribution is also a function of the total flow rate through the cooling loop. Figure 4.2-10 shows the variation of flow distribution for the front cold rail (FCR) and middle cold rail (MCR) as a function of the total flow rate through the server liquid cooling loop. The numbers in the bottom two rows along the horizontal axis of Figure 4.2-10 are the percentage of flow in the corresponding cold rails and the numbers in the top row are the corresponding total flow rate. The numerical predictions suggested that there would be sufficient flow in the front cold rail even at significantly low server level flow rates. Experimental investigation of the flow distribution was also performed and the results are summarized in Figure 4.2-11. Experimental data showed higher than numerically predicted flow in the front cold rail. This was due to the fact that the front cold rail had slightly larger flow cross section than that considered for the flow modeling. As a result, the flow impedance offered by the front cold rail was roughly 50% lower than that predicted by CFD modeling. This was

beneficial from the thermal point of view as it was closer to the ideal flow distribution of 33% in the front cold rail and 66% in the middle cold rail.

Figure 4.2-9. Numerical prediction of the flow distribution between the front and middle cold rails and of the pressure drop from the cooling loop inlet to the inlet of the first cold plate for a design flow rate of 0.9 liter per minute.

Figure 4.2-10. Numerical prediction of variation of flow distribution among the front cold rail (FCR) and middle cold rail (MCR) as a function of the total flow rate through the server liquid cooling loop.

Figure 4.2-12 shows side and bottom views of the DIMM spreader sub-assembly. The sub-assembly is made of two copper spreader plates that are mechanically attached to a DIMM card using spring

clips with a thermal interface material between the spreader and the DIMM card. The spreader-DIMM sub-assembly is inserted into an electrical socket on the PCB and then the spreaders are bolted on both sides to liquid cooled cold rails which extract the heat from the DIMMs through the spreaders. A re-workable metallic thermal interface material was used at the spreader cold rail interface. Numerical modeling using commercial packages was also extensively used for geometry refinement and to predict thermal performance of various DIMM heat spreader designs. These designs were also experimentally tested using thermal test vehicles as well as tested under conditions similar to actual operating conditions. The CFD predictions and experimental data of the DIMM spreader thermal resistance were in close agreement with each other.

Figure 4.2-11. Experimental measurement of variation of flow distribution among the front cold rail (FCR) and middle cold rail (MCR) as a function of the total flow rate through the server liquid cooling loop.

Figure 4.2-12. DIMM spreader sub-assembly made up of two copper spreader plates that are mechanically attached to a DIMM card using spring clips with a thermal interface material between the spreader and the DIMM card.

2.3 Comparison with Typical Air Cooled Server

A comparison between a typical air cooled server with a server with liquid cooled processors and memory was performed to quantify the benefit of liquid cooling in terms of improvement in temperatures and server power consumption.

Tables 1 and 2 provide thermal data collected in a controlled test chamber environment to characterize both the base line air-cooled and new partially water cooled server designs. The goals of the characterization were to collect thermal and power data as well as to validate the partially water cooled server for the high air and water coolant temperatures for which it had been designed. Table 1

provides data for two air-cooled node configurations, namely for a typical 25°C air inlet as may commonly occur in a data center and a 35°C air inlet temperature as might occur in a hot spot area of a data center or for operation of these servers in a higher ambient temperature environment with the intention of realizing cooling energy savings at the facility level. Table 2 provides comparable data for two partially water cooled server configurations, namely, for a “cool” 25°C water inlet temperature and a “hot” 45°C water inlet temperature. The former would be more typical in a data center located in many parts of the world and the latter might be a worst case condition for most of the globe. It should be noted that the statements mentioned above assume the data center design described in detail in the previous section where the air and water temperatures entering the server are closely coupled to the outdoor air temperature. For the water cooled server tests, the air temperature was maintained at 5°C above the water inlet temperature in anticipation of such a difference as a typical condition in the DELC system where the water to the rack will first cool the recirculating air via the Side Car heat exchanger and then enter the node to cool the water cooled components. Figure 4.2-13 provides more detailed data for the server powers, the fan powers, the CPU lid temperatures, and the DIMM temperatures for a number of inlet coolant conditions. The two exerciser settings used for the data provided in Figure 4.2-13 are for the CPU and the DIMM memory cards, respectively. In reality, it is unlikely that a real workload would exercise only the CPU or only the memory but will be some combination of both. It should be noted that in Figure 4.2-13(b) both lines for the CPU and memory exerciser settings are nearly identical, thus appearing to be a single line. It is interesting to compare the performance of Case A for a typical “cool” air cooled node configuration representative of most air cooled data centers to Case C and Case D for the newly proposed partially liquid cooled nodes representing cool and warm water, respectively.

Table 1. Thermal test chamber data for air cooled servers.

CASE A	CASE B
<u>“Cool” air cooled node</u>	<u>“Hot” air cooled node</u>
<ul style="list-style-type: none"> ▪ 25.3°C inlet air temperature ▪ Exerciser setting at 90% ▪ 12 fans running at 7242 rpm (avg.) ▪ System power = 395 W ▪ Fan power = 19.1 W ▪ CPU lid temps. = 65.3 °C, 74 °C ▪ DIMM temperatures = 35-46 °C ▪ 12 x 8 GB DIMMS 	<ul style="list-style-type: none"> ▪ 35.4°C inlet air temperature ▪ Exerciser setting at 90% ▪ 12 fans running at 11978 rpm (avg.) ▪ System power = 423 W ▪ Fan power = 56.8 W ▪ CPU lid temps. = 68.9 °C, 71.9 °C ▪ DIMM temperatures = 35-46 °C ▪ 12 x 8 GB DIMMS

Table 2. Thermal test chamber data for partially water cooled servers

CASE C	CASE D
<u>“Hot” water cooled node</u>	<u>“Cool” water cooled node</u>
<ul style="list-style-type: none"> ▪ 49.9 °C inlet air temperature ▪ 45.2 °C inlet water temperature ▪ Exerciser setting at 90% ▪ 3 fans running at 12612 rpm (avg) ▪ System power = 411 W ▪ Fan power = 30.9 W ▪ CPU lid temps. → 62.8°C, 61.9°C ▪ DIMM temperatures → 53-56 °C ▪ 12 x 8 GB DIMMS 	<ul style="list-style-type: none"> ▪ 24.9 °C inlet air temperature ▪ 20.1 °C inlet water temperature ▪ Exerciser setting at 90% ▪ 3 fans running at 5838 rpm (avg) ▪ System power = 354 W ▪ Fan power = 8.3 W ▪ CPU lid temps. → 36.8°C, 35.9°C ▪ DIMM temperatures → 28-33 °C ▪ 12 x 8 GB DIMMS

Comparing Cases A and C (Tables 1 and 2), the typical “cool” air cooled node consumes 395 W of power while the “hot” water cooled node uses 411 W which is a difference of 16 W (4.1%) that may be explained from the difference of 11.8 W in fan power since the CPU lid temperatures are

comparable. The reason that the “hot” water cooled nodes use more fan power is because the fan speeds in this node still ramp up at the same temperatures as for the air cooled node, even though the node has 3 less fans. With a change to the fan speed algorithm, the water cooled node could use less fan power even at elevated ambient temperatures compared to an air-cooled node at typical ambient temperatures. When comparing Cases A and D, the “cool” water cooled node consumes 10.4% less power (41 W) which is partly because it has 3 fans running at low speed compared to 6 fans at low speed for Case A. Of this 41 W, 10.8 W is attributed to fan power and the remaining is likely due to reduced leakage power from the 30-40°C cooler temperatures at the CPU. While an actual test in a specific location would be needed for confirmation, it may be speculated that for many locations only a small percentage of the time would be spent at the “hot” coolant temperatures represented by Case C and that the server will likely experience the Case D conditions for most of the year. If this were the case, then in addition to the significant data center level energy savings, there can be substantial IT power related energy savings of the order of 10% due the combination of leakage power and fan power reductions. In contrast to these gains, operating the air-cooled server at elevated ambient temperatures of 35°C can result in a 7.1% (28W) increase in IT power while there will be some data center level cooling energy reductions as documented in [4].

Figure 4.2-13. Power and device temperature data for partially water cooled node, (a) Server power, (b) Fan power, (c) CPU lid temperature, (d) DIMM temperature

In a similar comparative study, the microprocessor junction temperatures and memory module temperatures were also compared. Different workloads such as CPU exerciser, Memory exerciser and Linpack were executed on the servers to provide continuous and steady heat dissipation from the processors and memory modules, and to characterize the system performance. Component information such as processor Digital Thermal Sensor (DTS) value, DIMM temperatures, system fan rpm and other information was collected using the IPMI (Intelligent Platform Management Interface) and BMC (Baseboard Management Controller) tools. For this comparative study, the water flow rate through the liquid cooled servers was maintained at 0.7 lpm.

Figure 4.2-14. Comparison of estimated junction temperature for a liquid cooled server with a typical air cooled server (a) when the CPUs are exercised at 90% and (b) when the memory modules are exercised.

Figure 4.2-14 shows the comparison of the estimated junction temperature of the two processors for a liquid cooled server with an air cooled server cooled by air at 22 °C. For the liquid cooled server, two cases were considered – one with 45 °C server inlet water temperature (and 50 °C server inlet air temperature) and the other with 25 °C server inlet water temperature (and 30 °C server inlet air temperature). Figure 4.2-14(a) summarizes the estimated junction temperature comparison when each processor was exercised at 90% while Figure 4.2-14(b) summarizes the estimated junction temperature comparison when the memory exerciser was executed. In both the cases, liquid cooled microprocessors showed much lower junction temperatures even with warm liquid coolant. Note that the 45 °C water temperature might be an extreme condition for many parts of the world and even for that condition the microprocessors were at least five DTS units cooler than the typical air cooled servers. For the memory exerciser case, although the heat dissipation from the processors is not much, the difference in the estimated junction temperature is higher. This is because the server fans are running at a lower rpm consuming lower power (see Figure 4.2-16(b)).

Figure 4.2-15 shows the comparison of the DIMM temperatures for the liquid cooled server with an air cooled server cooled by air at 22 °C. Here again, 25 °C and 45 °C server inlet water temperature cases were considered. Figure 4.2-15(a) summarizes the DIMM temperature comparison when each processor was exercised at 90% while Figure 4.2-15(b) summarizes the DIMM temperature comparison when the memory exerciser was executed. Note that the DIMMs in slots 2, 3, 5, 6, 8 and 9 are closer to the fans and are cooled by relatively cooler air while the DIMMs in slots 11, 12, 14, 15, 17 and 18 are away from the fans and are cooled by relatively warmer air due to preheat from DIMMs in the front bank. When only the processors are exercised (Figure 4.2-15(a)), the heat dissipation from the DIMMs is very small and thus the DIMM temperatures are closer to the server inlet air temperature (for air cooled server) or server inlet water temperature (for liquid cooled servers). In such cases, the benefit of going to liquid cooling for the DIMMs is negligible. However, when the memory modules are exercised, the benefit of going to liquid cooling becomes prominent. In some cases, DIMMs of a warm liquid cooled server might show lower temperatures than those shown by the DIMMs of a typical air cooled server.

Figure 4.2-15. Comparison of DIMM temperatures for a liquid cooled server with a typical air cooled server (a) when the CPUs are exercised at 90% and (b) when the memory modules are exercised.

Figure 4.2-16 shows the comparison of the server power and server fan power consumption for the liquid cooled server with an air cooled server cooled by air at 22 °C. Here again, 25 °C and 45 °C server inlet water temperature cases were considered. Figure 4.2-16(a) summarizes the server power and server fan power comparison when each processor was exercised at 90% while Figure 4.2-16(b)

summarizes the server power and server fan power consumption comparison when the memory exerciser was executed. Figure 4.2-16(a) shows that the total server power goes up when the server is cooled with a 45 °C warm water and 50 °C server inlet air. Most of this increase in the power is due to the increased power consumption by the server fans as the server sees a 50 °C inlet air temperature. If we subtract that fan power from the total power, we see that the power consumed by the server electronics is lower than that consumed by the server electronics of a typical air cooled server. This reduction in the server electronics power consumption becomes more prominent for the 25 °C water cooled server where a more than 6% reduction in power consumption was observed. This reduction in power could possibly be due to the reduction in leakage power as the liquid cooled electronics were running at much lower temperatures. In the case of the memory exerciser, this reduction in power consumption was observed to be greater than 11% where the improvement in estimated junction temperature was ~40 units.

Figure 4.2-16. Comparison of server power and fan power consumption for a liquid cooled server with a typical air cooled server (a) when the CPUs are exercised at 90% and (b) when the memory modules are exercised.

The liquid cooled volume server has roughly 65% of the heat removed by liquid cooling leaving roughly 35% of the heat to be removed by forced air and then transferred to the Side Car. However, three server fan packs were also removed from the server changing the air flow dynamics inside the server. Thus, in addition to the comparison of liquid cooled components, an evaluation of the air cooling performance in the volume server was also performed and compared to a purely air cooled server. For this purpose, an air flow bench was designed, modeled and fabricated to provide a temperature controlled environment for both air and water entering the server. Infra-red images of powered on server with its top cover removed were taken to identify the hot components, excluding the processors and memory modules, on the server board. Among other hot components on the

board, an input/output chip, referred to as an IOH chip (shown in Figure 4.2-6(b)), located downstream of CPU 1 was identified as a key component for air cooling performance comparison. In the partially liquid cooled version, this IOH chip receives reduced air flow due to one less fan pack in its zone but benefits from relatively cooler air as the upstream microprocessor is liquid cooled. Thus, the temperature of this IOH chip was compared for both air cooled and hybrid air/liquid cooled servers and the data is presented in Figure 4.2-17. For this study, the same server was first characterized for air cooling performance and then modified to the hybrid liquid cooled version and tested again for air cooling performance. For the hybrid cooled server, the air and water temperature were kept the same, that is, if water entering the servers was at 25 °C, then the temperature of the air entering the server was also matched to 25 °C. The water flow rate was fixed at 0.7 lpm.

Figure 4.2-17. Comparison of IOH temperatures for a hybrid air/liquid cooled server with a typical air cooled server (a) when the CPUs are exercised at 90% and (b) when the memory modules are exercised.

Figure 4.2-17(a) shows the comparison of the IOH temperature when the microprocessors were exercised at 90%. It can be seen that for both hybrid liquid cooled servers, the IOH temperature is much lower than that for the 22 °C air cooled server. For the 22 °C air cooled server the CPU temperature increases the server fans speed and air flow rate but also preheats the air that cools the IOH as it passes over the CPU heat sink. In the case of the 25 °C hybrid cooled server the air flow rate is reduced due to the combination of eliminating a fan pack and lower fan speed which is governed by the server inlet air temperature. However, in the hybrid liquid cooled server the IOH air is not preheated by the processor coldplate which more than compensates for the reduced air flow. In the case of the 40 °C hybrid air/liquid cooled server, the fan speed and air flow rate are higher than the 25 °C case resulting in only a ~8 °C increase in IOH temperature compared to the 25 °C case.

Figure 4.2-17(b) shows the comparison of the IOH temperature when the memory modules were exercised. It can be seen that for both the hybrid liquid cooled cases, the IOH temperature is much lower than that for the 22 °C air cooled case. In the case of the 22 °C air cooled server, the memory exerciser does not cause any significant increase in the server fans' speed resulting in lower air flow rates as compared to that for the CPU exerciser case. Although the heat dissipation from the CPU is lower when the memory modules are exercised, the thermal resistance path from the air to the IOH chip is much higher due to relatively lower air flow rate. For the hybrid liquid cooled cases, the IOH chip temperatures shows similar behavior as when the CPUs were exercised as the preheat effect is eliminated due to the liquid cooling of upstream CPU. A similar behavior was observed for other air cooled components on the server as well.

Table 3. One to one comparison of the water cooled node with its air cooled counter-part.

<u>"Cool" air cooled node</u>	<u>"Warm" water cooled node</u>	<u>"Cold " water cooled node</u>
• Inlet Air 22C	• Inlet Water & Air 40C	• Inlet Water and Air 25 C
• System power 401 W	• System Power 388 Watts	• System power 346 W
• CPU "Est Temp" 80	• CPU "Est Temp" 65	• CPU "Est Temp" 49
• DIMM Avg 40C	• DIMM Avg 46C	• DIMM Avg 33C

The Air Flow Bench also enabled the characterization of liquid and air temperature operating limits. The server inlet air temperature was observed to be the limiting component. Based on the air flow bench data, a rack inlet water temperature set-point at 38 °C was sufficient to keep the servers within safe operating limits for the long term system operation study. Additionally, a one to one comparison of air cooled and hybrid liquid/air cooled servers, summarized in Table 3, showed that the hybrid liquid/air cooled server, cooled by 40 °C water and 40 °C air, uses about 3% less IT power than that used by its air cooled counter-part cooled by 22 °C air. Further, if cooled by 25 °C water and air, the hybrid server uses ~14% less IT power which could lead to over 2 kW in IT power savings per rack of 40 servers.

In summary, liquid cooling servers provides a significant benefit in terms of lower server electronics temperatures as well as lower server electronics power consumption. By implementing liquid cooling IT power can be reduced along with the significant reduction in cooling power.

2.4 Advanced Metal Interfaces for Module Liquid Cooling

Volume servers are provided with processor chips integrated into processor modules that include a substrate, the chip, a thermal interface material (TIM, in this case often referred to as TIM1), and a lid. It is this module assembly that is ordinarily interfaced with the system thermal solution consisting of another TIM layer (TIM2) and a heat sink. In an ordinary volume server the heat sink is an air cooled heat sink. In this project the heat sink was a liquid (water) cooled heat sink. In either case, there is significant thermal resistance and thus, a relatively large temperature difference between the heat sink surface and the active semiconductor devices on the chip. As part of this project, modules and systems were cooled with a "direct attach" approach where there is only a single TIM (TIM1) between the chip and the heat sink. The difference between the conventional approach and the direct attach approach is shown in Figure 4.2-18.

One of the difficulties in implementing direct attach is handling the mechanical interface between a rigid heat sink and the semiconductor die with a TIM1 that has excellent thermal characteristics. The first solution that was investigated in this project was to utilize a liquid metal thermal interface (LMTI) composed of an indium-gallium-tin alloy which is liquid at all normal operating temperatures. This provides high performance thermal conductivity without carrying coefficient of thermal expansion (CTE) difference driven stresses between the chip and heat sink. The first step toward implementing such a solution is creating a bare die module. Volume server processors are normally lidded and are not available without a lid. We developed a "de-lidding" process to allow the use of liquid metal thermal interface material.

- **Direct Attach removes thermal resistance of Heat spreader & TIM2**
 - **Lower thermal resistance allows Higher Coolant temperature for same Chip Temp T_j**
- Figure 4.2-18. Difference between the thermal resistance path for the conventional approach and the direct attach approach.

Commercially available lidded processor modules appropriate to the set of servers utilized for the system level demonstration were procured. The module design has a heat spreader (lid) attached that needed to be removed so that the processor die would be “bare” and ready to be interfaced to a cold plate. After the heat spreader was removed, the residual thermal interface material on the die was removed. Next the die surface was prepared so that the LMTI would properly wet and contact with low thermal resistance. Finally, the seal-band material was removed from the laminate where it bridged to the heat spreader as a structural adhesive bond. This process sequence was successfully demonstrated on a number of processor packages as shown in Figure 4.2-19.

Figure 4.2-19. Chip and laminate with the residual materials removed and ready for LMTI.

A de-lidded processor was installed into one socket of a volume server and the resulting thermal performance was measured. Pictures of the resulting server are shown in Figure 4.2-20(a) and (b). The first direct attach demonstration utilized a commercially available liquid cooled heat sink. The system was operated utilizing a standard exerciser program that loads the processor at approximately 90% of its peak operating capacity. The results were compared to the case when the processor was operating in an idle state which consumes nominally 10% of the exerciser case power. Power for the two cases was approximately 100W and 10W respectively. The processor includes temperature

sensors in each of the six processor cores within the processor chip. Data is provided in Digital Temperature Sensor (DTS) units which are related to the temperature difference from a maximum usable operating temperature. In order to measure the temperature rise from the idle state to the exerciser state, DTS data was taken at a broad range of input coolant temperatures. One set of such data is shown in Figure 4.2-21.

Figure 4.2-20 (a) Server with single de-lidded module installed with a commercially available heat sink. (b) Server with single de-lidded module installed with a commercially available heat sink (close-up).

Figure 4.2-21. DTS data for varying coolant temperatures for a variety of exercise conditions on a single processor module.

This data showed that the ~100W power temperature rise with respect to the idle state was 18 °C, or a temperature rise relative to the coolant at full power of ~20 °C. Note that for a coolant temperature and air temperature of 25 °C, the processor was operating approximately 30 °C cooler for the direct attach liquid cooled case.

Based on this data, two servers were constructed with direct attach and LMTI utilizing the same liquid cooled heat sinks used with lidded modules and thermal grease TIM for the majority of the

project servers. These two servers were first built with the lidded modules and data taken across a range of coolant temperatures. The processor modules from these servers were then de-lidded and reinstalled in the servers with the same heat sinks but with direct attach LMTI. The same thermal data was then taken. Data for three of those test cases is provided in Table 4.

Table 4. Node level thermal data comparing results for lidded modules with thermal grease with results for the same modules with direct attach LMTI in the DELC server rack environment.

	Condition 1			Condition 2			Condition 3		
	Thermal Grease	Liquid Metal	Difference	Thermal Grease	Liquid Metal	Difference	Thermal Grease	Liquid Metal	Difference
Rack Level Flow Rate	4.0 gpm	4.0 gpm	-	7.2 gpm	7.2 gpm	-	9.2 gpm	9.2 gpm	-
Node Level Flow Rate	0.4 lpm	0.4 lpm	-	0.7 lpm	0.7 lpm	-	0.9 lpm	0.9 lpm	-
Rack Coolant Temp	31.3 °C	31.0 °C	0.3 °C	29.5 °C	29.6 °C	-0.1 °C	33.0 °C	33.0 °C	-0.1 °C
Rack Air Temp	36.2 °C	35.8 °C	0.4 °C	32.7 °C	31.2 °C	1.4 °C	35.0 °C	34.5 °C	0.5 °C
CPU1 ~Temp 100-Avg DTS	59.2	52.6	6.7	50.1	44.6	5.4	52.4	46.7	5.7
CPU2 ~Temp 100-Avg DTS	54.3	47.5	6.8	47.3	41.9	5.3	50.4	44.8	5.6

Figure 4.2-22. Comparison of one CPU core temperature average for standard thermal grease and direct attach LMTI implementations.

On average, the direct attach LMTI approach improved (lowered) the processor core temperature between 5 and 6 °C. Shown in Figure 4.2-22 is data for one processor taken during a coolant temperature ramp. Measured along the Coolant Temperature axis, which is an actual temperature measurement rather than a DTS estimated temperature, this CPU showed an almost exactly 5 °C improvement with the LMTI. The two LMTI nodes were installed in the rack with the other lidded, thermal grease utilizing nodes and operational thermal data was taken. This data is shown in Figure 4.2-23. The LMTI nodes clearly outperformed the “standard” nodes.

Figure 4.2-23. Core temperature variability within the 42 servers of the complete demonstration system showing the lower operating temperature of LMTI equipped nodes.

2.5 Advanced “Compliant” Heat sinks

While this direct attach LMTI performance was clearly superior to the standard approach, there remained the possibility of further thermal improvement with new advanced heat sink designs. The new heat sink designs were developed to have thermal performance superior to the channeled standard heat sinks while providing new mechanical characteristics that would allow the use of a wide range of high performance TIMs. Comparison data would be taken using LMTI however. These heat sinks consist of an array of heat conducting pins with or without links between the pins sandwiched between two thin metal sheets. The resulting heat sink is referred to as a Pin Fin Compliant (PFC) heat sink. Diagrams of this approach compared with a standard approach are shown in Figure 4.2-24.

Figure 4.2-24. Pin Fin Compliant heat sink cooled module structure compared with standard rigid heat sink liquid cooled module structure.

A heat sink constructed using this method was implemented on a single processor. The result was a 3 °C improvement over the heat sink utilized in the full node assemblies. This technology was also implemented for larger die thermal test vehicles, demonstrating performance that makes it an excellent candidate for a number of commercial applications.

Chapter 3: System Modeling and Simulations

Steady state and dynamic system models that were developed for the liquid cooled chiller-less data center design are described in this chapter. The steady state model was essentially modeled using theory, numerical simulations and manufacturer's data sheets and was used for the design selection of data center level cooling infrastructure and of server liquid cooling components and for performance estimation of the liquid cooled chiller-less data center design. The dynamic system model was essentially modeled using theory, numerical simulations and experimental characterization data of the system and was used to investigate novel cooling system servo control schemes to reduce energy consumption. The system-level models were validated against experimental data and were used to predict the system performance in different geographies.

3.1 Energy Based System Model

Figure 4.3-1 schematically represents a data center with indoor and outdoor components. The energy is exchanged between the indoor and outdoor components with the help of coolant(s) which could be either gaseous or liquid in nature. The cold coolant enters the data center and absorbs all the heat dissipated by the rack(s), (and in a more general sense, by the data center) and transports the heat outside of the data center and dissipates it to the outdoor ambient. So two of the key inputs to the system model are heat dissipation by the rack(s) (or by the data center, in a more general sense) and the outdoor ambient temperature. In general, a data center could be completely air cooled or partially air and liquid cooled or completely liquid cooled. So depending upon the cooling approach the

physical infrastructure both inside and outside of the data center is defined. Hence, other key inputs to the system model include the sub-component thermodynamic and hydrodynamic models and sub-component costs. These sub-component models physically define a particular data center design. Additionally, various algorithms/methods of operation could be implemented for a particular data center design. Thus, control algorithms/methods are also defined as the key inputs to the system model. The key inputs to the system model are: heat dissipation from the data center, outdoor ambient conditions, thermodynamic and hydrodynamic models, control methods/algorithms and sub-components cost.

Figure 4.3-1. Schematic of the system model

Figure 4.3-2 shows a flow diagram summarizing the working of the system model. First, a particular design is selected and then the sub-component information is inputted. The sub-component information for a partially liquid cooled data center may include the type of coolant in the liquid loop(s), thermal resistance curves for the cold-plates used, pressure drop curves for the liquid cooling loop(s), cost of the individual components, heat transfer characteristics of the heat exchangers used, power usage curves for the pumps, fans and blowers and other such information that make the design unique. This sub-component information can be obtained from thermodynamics theory, numerical simulations, experiments, OEM (original equipment manufacturer) data sheets or a combination of these methods. Next, a set of constraints is defined based on the anticipated working environment of the equipment such as maximum allowable junction temperatures for the processors, maximum allowable temperatures for the memory devices, hard-drives, and auxiliary board components, server inlet air temperatures, dew point temperature and other such constraints. Then, a range of operating conditions is selected such as the time variation of the outdoor air temperature and the time variation of heat generation by the rack/data center. A control algorithm or a method of operation for the system under consideration is also selected. Next, an energy balance analysis is performed using all the above information to evaluate the energy consumption of the selected design over a selected time-period at a particular geographical location. A cost analysis is also performed to estimate the data center operational costs based on the intended location of the data center. This process can be automated to explore and compare among various design choices and geographical locations.

Following is a general sequence of thermo- and hydro-dynamic calculation steps for a given outdoor ambient air temperature and rack/data center heat dissipation.

Figure 4.3-2. Flow diagram of system model operation.

- (a). Calculate the indoor and outdoor loop liquid flow rates using the corresponding pump RPM. Similarly, calculate the Outdoor Heat Exchanger fan air flow rate using the relationship between the RPM and air/liquid flow rate which can be obtained either by using OEM data sheets and relations or by using numerical simulations. In the present study, OEM data was used for the air flow rate dependence on the fan RPM. For the liquid flow rate dependence on the pump RPM, a combination of numerical simulations and OEM data were used. Analytical models as well as numerical simulations using commercially available computational fluid dynamics software were used to generate the system pressure drop curves for different cooling configurations. OEM data sheets were used to generate the pump head curves at different pump RPM settings.
- (b). Estimate the power consumption of the pumps and fans for the current RPM settings using OEM data sheets, analytical relations and/or experimental data. In the present study, the total pumping power for the pumps was calculated using the total pressure drop and volumetric flow rate in each loop. The pump electrical energy consumption was determined using the pumping power and the estimated pump efficiency based on the system pressure drop curve. For fan power consumption, an experimentally obtained relationship between RPM and power consumption was used.
- (c). Estimate the Outdoor Heat Exchanger effectiveness for the current air flow rate using OEM data sheets and analytical relations or experimental data. In the present study, the analytical

relations validated against experimental data were used to estimate the Outdoor Heat Exchanger effectiveness.

- (d). Calculate the liquid temperature entering and leaving the Outdoor Heat Exchanger and the hot air temperature leaving the Outdoor Heat Exchanger Input using energy balance, outside air temperature and Outdoor Heat Exchanger effectiveness. The IT heat load is used as the heat that is being dissipated to the outdoor ambient air.
- (e). Estimate the liquid-to-liquid heat exchanger effectiveness for the current indoor and the outdoor loop liquid flow rates using OEM data sheets, analytical relations and/or experimental data. In the present study, the OEM provided relations were used to estimate the liquid-to-liquid heat exchanger effectiveness.
- (f). Calculate liquid temperature entering and leaving the liquid-to-liquid Heat Exchanger on the indoor side and the warm liquid temperature leaving the liquid-to-liquid Heat Exchanger on the outdoor side using energy balance, the liquid temperature leaving the Outdoor Heat Exchanger and the liquid-to-liquid Heat Exchanger effectiveness. The IT heat load is used as the amount of heat that is being exchanged between the indoor and the outdoor coolant loops.
- (g). Estimate the side car air-to-liquid heat exchanger effectiveness at the current air flow rate inside the rack and the liquid flow rate in the side car heat exchanger using OEM data sheets, analytical relations and/or experimental data. In the present study, analytical relationships validated against experimental data were used to estimate the side car heat exchanger effectiveness. For the servers used in the present study, the rpm (or air flow rate) of the server fans changed predominantly based on the server inlet air temperature. The normal rpm changes due to load driven processor temperature rise were eliminated because even under full power the processors were running below the temperatures which would normally cause processor driven fan rpm increases [9].
- (h). Calculate the air temperature entering and leaving the side car and the hot liquid temperature leaving the side car using energy balance for the side car, the side car liquid inlet temperature and the side car heat exchanger effectiveness. The heat load exchanged across the side car heat exchanger is a fraction of the total IT head load. The value of the fraction depends upon the workload running on the servers. For example, for a processor intensive workload, the fraction could be 0.3 while for a memory intensive workload, the fraction could be 0.4. Since the air temperature leaving the side car is used to determine the air flow rate across the side car and this air flow rate is used to determine the side car heat exchanger effectiveness, steps (g) and (h) are iterated using the bisection method for each sample to find an equilibrium solution.
- (i). Estimate the component temperatures for the CPU chips and DIMMs using the coolant temperature leaving the side car heat exchanger, the coolant flow distribution inside the server, the fraction of heat transferred to the coolant from the DIMMs and CPUs, and thermal resistance relations obtained from server level simulations or experiments. In the present study, component level and node level simulations were performed to generate the

thermal resistance relations as functions of server level flow rates. The heat dissipation from the CPUs and DIMMs as functions of the workload were also used as input. It must be noted here that the heat dissipation from the CPUs and DIMMs is generally lower than their corresponding rated wattage and is dependent on the type of workload and the coolant and device junction temperatures.

Steps (a) through (i) were repeated for every new combination of RPM settings of the cooling system components. The values of these new RPM settings were determined based on the applied control algorithm.

3.1.1 Design Impact Study

The cost, performance and energy usage of a data center is dependent upon both the physical infrastructure and environmental conditions inside and outside the data center. While designing a system it is important to understand the cost, energy and performance impact of the design choices. Understanding and being able to quantify such impacts can help 1) guide system level design decisions; 2) quantify single component impact on the system performance and capital and operational costs; 3) relate the cooling requirements to IT load, environmental conditions, components costs and other such parameters; 4) identify possible failure locations to enable design selection; 5) explore/compare numerous design variations; and 6) identify an optimal cost-effective cooling solution within provided constraints. Two cases are presented below to emphasize the importance of a system model in guiding design decisions.

I. Single Loop vs Dual Loop

Figures 4.3-3(a) and 4.3-3(b) schematically represent two liquid cooled chiller-less data center designs. Figure 4.3-3(a) represents a single loop design while Figure 4.3-3(b) represents a dual loop design. The dual loop design is the same as that shown in Figure 4.1-1. The single loop design as the name suggests has a single liquid coolant loop transporting heat from the rack of servers to the outdoor ambient environment. In case of the single loop design, the coolant in the loop could be water or a water-glycol mixture or similar liquid coolant depending upon the location of operation of the data center. For example, if this Data Center were to be located in Poughkeepsie, NY, the coolant in the single loop design may be an anti-freeze solution to prevent damage from sub-freezing ambient conditions in winter. In the dual-loop design, the coolant in the outer coolant loop can be an anti-freeze solution and the indoor coolant loop could be water. The choice of coolants will impact the thermodynamic and hydrodynamic performance of the system could change.

Additionally, the thermodynamic and hydrodynamic performance of the node liquid cooling structures could impact the system performance. Figures 4.3-3(c) and 4.3-3(d) show two different CPU cold plate designs. Cold-plate1 shown in Figure 4.3-3(c) is made up of two parts: the copper core with the water flow channels that make up a parallel plate fin array, and the aluminum frame that is soldered to the copper core and provides the structural support during clamping of the cold plate to the processor module. Cold-plate1 is a higher-cost, high-performance cold plate. Cold-plate2 shown in Figure 4.3-3(d) is also made up of two parts: a continuous copper tube with two non-flattened circular lollipop bends and three flattened sections that are soldered to an aluminum bulk frame. Cold-plate2 is a lower-cost and lower-performance cold plate however the pressure drop and thus pumping power is relatively higher due to the serial flow pattern.

Figure 4.3-3. Design choices: (a) Single Loop, (b) Dual Loop, (c) High-cost high performance cold-plate1 and (d) Low-cost low performance cold-plate2.

Figure 4.3-4 shows the thermal resistance curves for two different cold-plates with water and with a 50% by volume water-glycol solution as the coolants. The difference in the thermal resistance curves for the two cold-plates increases when water is replaced with a water-glycol solution. The hydrodynamic performance (that is, the pressure drop variation as a function of coolant flow rate) of the cold-plates is also dependent on the coolant selection for the two cold plate designs which can impact total cooling power consumption. The system model simulator can also quantify this impact.

Figure 4.3-4. Variation of thermal resistance as a function of flow rate for different Cold-plates and different coolants.

Figure 4.3-5. Cooling power usage comparison at different outdoor ambient temperatures for different design choices.

Figure 4.3-5 shows the cooling power usage comparison for these four possible cases – (i) Single loop with cold-plate1, (ii) Single loop with cold-plate2, (iii) Dual loop with cold-plate1, and (iv) Dual loop with cold-plate2 – for 20 °C and 40 °C outdoor air temperatures. It can be seen that for 20 °C outdoor air temperature, all four cases show similar power usage. However, for a 40 °C outdoor air condition, the low thermal performance of the anti-freeze coolant causes the power consumption

of the single-loop design to be significantly higher than that for corresponding dual loop design for the same cold-plate, with the difference being greatest for cold-plate2.

It is worth mentioning here that if this data center design were to be operated at a location closer to the equator where the use of an anti-free solution is not warranted, a single loop design with water in the loop would be ideal from an energy point of view. However, concerns regarding the quality (chemistry, corrosion inhibitors, etc.) and potential environmental impact of the coolant in the liquid loop could necessitate the use of a dual loop. With a dual loop system, relatively higher quality coolant can be used in the internal loop and a relatively lower quality coolant can be used in the external loop.

II. Identifying Limiting Component

The system model can also be used to identify the “limiting” component, which is the hottest component in the system whose temperature limit is determining the cooling power consumption. An example for a single loop system for selected CPU and DIMM cold plates and control algorithm is shown in Figure 4.3-6. In this case the components are shown over a range of outdoor temperatures from 0 °C to 45 °C. The “limiting” component is the hottest CPU from 0 °C to 38 °C, then the hottest DIMM from 38 °C to 42 °C and finally the rack inlet air temperature which determines the air cooled components in the server from 42 °C to 45 °C. This enables one to study the impact of thermal cooling choices made for an individual component on the final system design.

The design choice is dependent upon the location and operating conditions of the data center.

Figure 4.3-6. Limiting components temperature variation as function of outdoor ambient air temperature.

3.2 Dynamic System Model

A representative dynamic model of the data center cooling systems was developed to study the performance of the novel cooling system and to evaluate the energy consumption of the cooling components on a transient and steady state basis. As data center server components need to be tightly regulated thermally to ensure reliability, the transient server chip temperatures can be studied and predicted by the model. To effectively keep the server chip operating temperatures within

specified limits, under varying operating conditions, an effective control scheme is needed for the cooling system components consisting of the server side buffer heat exchanger water loop pump, and the associated glycol loop pump and the dry cooler fan. The development of the dynamic model allows the design of novel data center servo control systems to closely thermally regulate servers and dynamically balance cooling components to reduce energy consumption.

Figure 4.3-7. Overall data center Cooling System & Control Simulation Model.

The data center cooling system modeling and control methodology involved obtaining the cooling system component design data from specifications. Cooling system components for which the specifications were unavailable were determined by measuring their physical parameters. The system components were developed by using a commercial design package and the dynamic models were implemented using a servo simulator. Initial set point control models were developed for the cooling system heat exchangers to establish the feasibility of closed loop control. The system component models were integrated into an overall simulation and servo control model with eleven main control system components including the server module CPU chip/Cold Plate Assemblies, the server Memory/Cold Rail Assemblies, the Side-Car liquid to air heat exchanger, buffer water/glycol heat exchanger and associated pumps, the Dry Cooler heat exchanger and fan and the servo controllers and control logic for the pumps and fans. The system was used to study various heat exchanger servo control methods and investigate novel cooling system servo control schemes to reduce energy consumption. A servo simulation model of the overall data center cooling system is shown in Figure 4.3-7. The block diagrams show the Server Rack sub-system, the MWU water/glycol buffer heat exchanger sub-system, the Dry-Cooler heat exchanger sub-system and the sub-system for the MWU, VFD pumps and dry-cooler fan servo controllers, and PLC control logic. Details of the server rack sub-system model are included in Figure 4.3-8. Figure 4.3-8 shows the incoming fluid flow to the rack which first passes through the Side Car air to liquid heat exchanger which is mainly used to cool the ancillary electronics in the server module including the disk drives.

The fluid flow then splits and separately cools two DIMM memory module cold rails and recombines to cool the first server module CPU cold plate. Subsequently, the fluid cools the third DIMM memory cold rail followed by the second CPU cold plate.

Figure 4.3-8. Rack Sub-system Model with Side Car, CPU/Cold Plates and Memory/Cold Rails.

3.2.1 Simulation Model Application Example: Exploratory Control

Figure 4.3-9 shows a schematic of the system. The servers are cooled by a water loop wherein cooled water is pumped by the MWU into the server rack and the heated rack water outflow is passed through the water to glycol buffer heat exchanger HE-A. On the glycol side of HE-A, the heated glycol is transported by the VFD pump to the Dry-Cooler glycol to air heat exchanger HE-B where it is cooled by the air flow controlled by the Dry-Cooler fan. There are three control components which include two pumps and a fan.

The simulation model was used to explore various servo control schemes to regulate server chip temperatures under different operating conditions by closed loop control of the fluid flows through the MWU and VFD pumps and the Dry-Cooler fan. One example shown below explored the feasibility of controlling the rack server chip temperatures by keeping either the rack fluid outflow temperature or the rack inflow temperatures at pre-determined set points by servo control of the pumps and fan. The model showed that for the same rack CPU regulated chip temperature under servo control, the set point control of the rack outflow temperature resulted in a much lower server chip transient temperature rise than the corresponding transient temperature rise under rack inflow temperature set point control. This is shown in Figure 4.3-10.

Controllable Data Center Cooling Components: Schematic

Advanced Control

Adjust all 3 Components to Control
either Rack Inlet or Rack Outlet
Fluid Temperature to Cool Servers

Figure 4.3-9. Data center Cooling System Control Components

Model Application:
To Maintain same
CPU Chip Temp: Use
Rack Inlet or Rack
Outlet Servo?

**Advantage of
Rack Outlet
Servo**

Sri/Arun 201-07-03

Figure 4.3-10. Comparison of rack inflow and rack outflow fluid temperature set-point control.

The comprehensive data cooling system simulation model is very flexible as all control investigations on all three available Control Components can be performed. The model allows the ability to demonstrate the performance of different control schemes and their safe operating areas without stressing the actual hardware. The system power/energy consumption due to each control technique can be assessed in detail on a transient and steady state basis. The model was also used to

develop control methods to reduce cooling power. The comprehensive model allows investigation of complex control schemes to optimize cooling system performance.

Chapter 4: System Characterization, Operation and Performance Projection

The DELC system was systematically characterized in order to obtain the thermal, hydraulic and power characteristics of the various heat exchange steps and cooling equipment. Several one day long runs were also made to determine the impact of environmental conditions on the thermal and power consumption characteristics of the system. These studies were used to develop an energy efficient PI control scheme that was utilized for a two month continuous operation of the system. In addition to the energy based static and dynamic system models discussed in the previous chapter, an empirical model based on the characterization data was also developed. These models were validated against the experimental data and were used to make projections for a typical year and for different geographical locations.

4.1 System Characterization Studies

Power, hydraulic and thermal characterization of the data center is important to determine the behavior and thermal metrics of the system as various operating conditions within the cooling loops are altered. A valuable thermal metric of interest to a data center designer is the thermal resistance across every heat-exchange stage in the cooling system. These individual thermal resistances can be characterized separately and then combined to obtain the overall thermal resistance of the data center from the component or rack inlet water to the ambient air. Measurement of power consumption by various pieces of cooling equipment under varying conditions is also important to ensure that the thermal benefit derived from a particular operating condition does not come with a significant cooling energy cost.

In this study, we characterized the impact of varying the internal and external coolant flow-rates, the dry-cooler fan speeds, the liquid-to-liquid heat exchanger arrangement and the addition of propylene glycol to the external loop. For every test that was carried out, the system was allowed to reach steady state before a new setting was tested. The average of the last two minutes of data collected from each test was used to represent the steady-state condition of the system under that particular operational condition.

IT Rack

The heat exchange at the rack occurs from the electronic components (CPUs, DIMMs) to the water as well as between the heated re-circulated air and the incoming cool water. The water enters the rack and first flows through the side mounted air-to-liquid heat exchanger to cool the rack circulating air. The water then enters the server manifold which distributes the water to each of the connected servers via flexible hoses. The water cools the thermally connected components and exits the servers into a common outlet manifold as warm water. The warm water then exits the rack and flows back to the liquid-to-liquid heat exchanger. A study of flow distribution among the servers showed that for the flow-rates of interest the distribution of flow is approximately uniform. To determine the approximate thermal resistance from the CPU coldplates and DIMM spreaders to the incoming water, two of the servers were instrumented - Node 9, which is towards the bottom quarter of the rack and Node 37 which is towards the top quarter of the rack. In each server/node, both CPU coldplates (CPU1 and CPU2) as well as the hottest DIMM spreader (attached to DIMM #18) were

instrumented with T-type thermocouples and measured using a datalogger and attached computer. The locations of the CPU coldplates and the instrumented DIMM spreader are shown in Figure 4.4-1. The maximum CPU and DIMM temperatures measured during the characterization study were 58 °C and 53 °C respectively, obtained during the same test with an outside air temperature of 29 °C, internal water flow rate of 4 GPM, IT power use of 14.5 kW and using a single liquid-to-liquid heat exchanger with water in the internal and external loop. These temperatures are well within the margins for the CPUs and DIMMs.

Figure 4.4-1. Schematic representation of the liquid cooled IBM X-3550 Server.

Figure 4.4-2(a) shows the measured CPU coldplate-to-inlet water thermal resistance (Eq. 1) and Figure 4.4-2(b) shows the DIMM-to-inlet water thermal resistance (Eq. 2) as functions of the bulk rack water flow rate (the actual flow rate at the server is approximately 40 times smaller). R_x represents the thermal resistance, T_{CP} the temperature at the base of the CPU coldplate, T_{DIMM} the temperature on the center outer surface of the DIMM spreader and Q_{IT} the power dissipated by the rack.

$$R_{CP} = \frac{(T_{CP} - T_{rack\ inlet\ water})}{Q_{IT}} \quad (1)$$

$$R_{DIMM} = \frac{(T_{DIMM} - T_{rack\ inlet\ water})}{Q_{IT}} \quad (2)$$

The results indicate a power law dependency of the thermal resistance with flow rate, with an increasing water flow rate resulting in a lower thermal resistance. Figure 4.4-2(a) shows that the CPU1 coldplate thermal resistance is higher than that for CPU2 coldplate since CPU1 is downstream from CPU2 and the DIMM spreaders and thus sees a pre-heat in the water. Variation in the coldplate assembly and the CPUs themselves result in differences in the thermal resistance between the respective CPU coldplates on Node 9 and Node 37. It is evident that for a given water flow rate there is a significant spread in the measured thermal resistance. This is due to an additional dependence on the inlet water temperature. An increase in the water temperature entering the rack results in warmer

air leaving the internal heat exchanger and entering the servers. The server fans ramp-up with the increasing server inlet air temperature and increase the air flow rate through the servers and the air-to-liquid heat exchanger. Higher fan speeds result in i) increase in the IT power consumption, ii) increase in the forced air convection heat transfer coefficient and iii) change in the effectiveness of the air-to-liquid heat exchanger. These effects collectively result in reducing the calculated component to rack inlet water thermal resistance. Similarly, when the incoming water temperature is low, the air temperatures entering the servers are lower and the server fan speeds, air flow-rate within the servers and rack, and IT power consumption all reduce, resulting in an apparent increase in the component to rack inlet water thermal resistance. Thus, the component to rack inlet water thermal resistance tends to reduce as the inlet water temperatures rises. The current data suggests an approximately linear decrease with rack inlet water temperature though a more detailed study of this dependence would be required before a firm relationship can be determined.

Figure 4.4-2 (a) CPU coldplate to rack inlet water thermal resistance and (b) DIMM spreader to rack inlet water thermal resistance. Both show a power decay trend with internal loop water flow rate and an increase in the thermal resistance as the inlet water temperature reduces.

Buffer Unit (MWU)

The buffer unit enables heat exchange between the internal water loop and the external coolant loop. The external loop coolant can be either water or water-propylene glycol mixtures for use in winter to avoid freeze damage. The impact on the thermal resistance due to flow variations in the internal and external loops, heat exchanger arrangement and addition of propylene glycol were investigated. The relevant thermal metric of interest for the buffer unit is the approach thermal resistance R_{buffer} , Eq. 3, where Q_{buffer} is heat transferred across the buffer unit, calculated from the sensible cooling of the water as it passes through the buffer unit. The thermal resistance can also be defined in terms of the heat exchanger effectiveness, ε , and the heat capacities of the minimum fluid, C_{min} , and the hot (internal) fluid, C_h .

$$R_{Buffer} = \frac{(T_{h,o} - T_{c,i})}{Q_{buffer}} = \frac{(T_{rack\ inlet\ water} - T_{buffer\ inlet\ coolant})}{Q_{buffer}} \quad (3)$$

$$R_{Buffer} = \frac{1}{\varepsilon \cdot C_{min}} - \frac{1}{C_h}$$

Figure 4.4-3 shows the variation in the thermal resistance as a function of the internal and external water flow-rates with a single heat exchanger in counter-flow configuration. The best fit surface follows a power law trend as expected by Eq. 3. The thermal resistance improves as the cold (external) fluid flow-rate is increased or if the internal hot fluid flow-rate is reduced. The reason for the reduction in resistance with reducing internal flow is due to a comparatively larger increase in the $1/C_h$ term as compared to the heat exchanger conductance term, $1/(\varepsilon \cdot C_{min})$, resulting in a reduced thermal resistance. However, this may not always be the case and is dependent on the magnitude of the conductance of the heat exchanger.

Figure 4.4-3. Approach thermal resistance surface for the buffer unit showing the thermal resistance as a function of internal and external liquid flow rates. The thermal resistance improves with increasing external (cold) loop flow rate and reducing internal (hot) loop flow rate.

Impact of Heat Exchanger Configuration

To reduce the thermal resistance across the buffer, an additional heat exchanger could be installed in series with the existing liquid-to-liquid heat exchanger. To study the impact of heat exchanger configuration, the default single counter-flow heat exchanger is compared against i) two serially connected heat exchangers in full-counter flow arrangement and ii) two serially connected heat exchanger in a counter-parallel arrangement with flow entering the first heat exchanger in counter flow and then entering the second in parallel flow. The counter-parallel arrangement can result in a lower heat exchanger conductance, $(\varepsilon \cdot C_{min})$, than a single counter flow heat exchanger. Figure 4.4-4 shows the impact on the heat exchanger conductance, Figure 4.4-4(a), and the thermal resistance, Figure 4.4-4(b), at three different internal water flow-rates as the external water flow rate is increased for the three different heat exchanger configurations.

The results show that the conductance increases with both internal and external flow rate increase for each of the configurations with the counter-parallel double heat exchanger having the lowest

overall conductance and the full-counter double heat exchanger the highest conductance. The counter-parallel heat exchanger has poorer conductance than the single heat exchanger configuration due to reverse heat flow from the external water loop back into the internal loop in the second parallel-flow heat exchanger. Figure 4.4-4(b) likewise shows that the thermal resistance for the full-counter heat exchanger is smaller than for the single counter or double counter-parallel heat exchanger configurations. The thermal resistances for all the arrangements follow a power decay relationship with the external liquid flow-rate. It can also be observed that for the single counter and double full-counter heat exchanger arrangements the thermal resistance is more sensitive to changes in the internal flow rate and reduces as the internal flow rate reduces, but for the double counter-parallel arrangement the thermal resistance is less sensitive to flow rate and increases slightly as the flow rate is reduced. This effect is due to the interplay between the heat exchanger conductance and heat capacity magnitudes in Eq. 3. For the double counter-parallel arrangement, the magnitude change in the $1/\epsilon \cdot C_{min}$ term due to the change in internal flow rate is countered by the magnitude change in the $1/C_h$ term resulting in little overall change in the resistance. The increasing internal flow rate results in a relatively smaller reduction in the $1/C_h$ component than the $1/(\epsilon \cdot C_{min})$ component resulting in an overall reduction in the resistance. For the other two configurations, the magnitudes of the terms are such that increasing the internal flow rate causes the reduction in the magnitude of the $1/(\epsilon \cdot C_{min})$ component to be somewhat smaller than the magnitude reduction in the $1/C_h$ component resulting in an overall increase in the thermal resistance.

Figure 4.4-4. (a) Heat exchanger conductance and (b) approach thermal resistance as functions of internal and external loop flow-rates and heat exchanger configuration. 1x C: single heat exchanger in counter flow, 2x C-P: double heat exchanger in counter-parallel flow and 2x C: double heat exchanger in full counter flow configuration.

Impact of Propylene Glycol Addition

Propylene Glycol (PG) was added to the external fluid loop to protect it from freezing damage during low winter temperatures. At a concentration of 20% by mass, the freezing point reduces to -8

°C and at 50% by mass to -34 °C. However, propylene glycol has poorer thermal properties with a thermal conductivity of 0.147 W/m-K and heat capacity of 2.5 kJ/kg-K as compared to water which has a thermal conductivity of 0.61 W/m-K and specific heat capacity of 4.2 kJ/kg-K. Propylene Glycol also has higher viscosity and slightly higher density than water, resulting in increased pressure drop and pump work for a given volumetric flow-rate. It is desirable to determine the change in thermal resistance due to addition of 20% and 50% by mass of PG to the external loop. Figure 4.4-5 shows the impact of adding 20% and 50% PG to the external loop on the heat exchanger conductance and approach thermal resistance. The conductance is reduced at the two higher internal flow rates and the thermal resistance increases for all internal flow rates. In both the water and water+PG tests, a full-counter double liquid-to-liquid heat exchanger configuration was used. As discussed previously, for this particular heat exchanger configuration, the approach thermal resistance is reduced as the external (cold-side) flow rate is increased and as the internal (hot-side) flow rate is reduced. At the lowest internal flow rate, the conductance is dominated by the internal loop and so addition of propylene glycol to the external loop does not have a significant effect and is slightly higher, though this is due to the slightly higher internal flow rate.

Figure 4.4-5. Approach thermal resistance as functions of internal and external loop flow-rates and external loop coolant. Addition of 20% and then 50% propylene glycol to the external loop results in an increase in the thermal resistance, with the effect growing stronger at higher internal loop flow rates.

These results indicate that the use of propylene glycol in the external loop would require comparatively higher external loop flow-rates to obtain the same approach thermal resistance as when using water, resulting in higher pumping power to obtain a given thermal resistance. However, in colder winter temperatures the system can operate at higher thermal resistances while still maintaining safe rack inlet water or rack component temperature due to the low ambient temperatures.

Dry Cooler Unit

The dry-cooler unit was used to cool the warm external coolant through heat exchange with outside ambient air that is blown across the heat exchanger fins. Similar to the buffer unit, the thermal performance of the dry-cooler is dependant on the flow-rates of liquid and air in the heat exchanger as well as the addition of propylene glycol to the external coolant. The approach resistance for the dry-cooler, R_{cooler} , is given by Eq. 4, where Q_{cooler} is the measured sensible heat loss on the liquid side and C_h is the heat capacity of the hot fluid, i.e. the external loop fluid.

$$R_{cooler} = \frac{(T_{buffer\ inlet\ water} - T_{ambient})}{Q_{cooler}} = \frac{1}{\epsilon_{cooler} \cdot C_{min}} - \frac{1}{C_h} \quad (4)$$

Figure 4.4-6 shows the variation in the approach thermal resistance as a function of the external loop water flow and the fan speed in RPM. Similar to the buffer unit, a best fit surface can be determined and follows a power trend. The thermal resistance strongly reduces with increasing fan speed (and air flow rate) up to about 500 RPM but there is little thermal benefit beyond that point. The thermal resistance also tends to increase with increasing external loop liquid flow rate. Increasing the external coolant flow rate results in a relatively smaller reduction in the magnitude of the $1/\epsilon \cdot C_{min}$ component as compared to the reduction in the magnitude of the $1/C_h$ component that results in an overall increase in the approach resistance. The fan speed rather than air flow rate was used in the analysis due to the uncertainty in the calculated air flow rate which is determined from sensible heating of the air across the dry-cooler. The relation between fan speed and air flow in cubic feet per minute is shown in Figure 4.4-7 and follows an approximately linear trend.

Figure 4.4-6. Approach thermal resistance surface for the dry-cooler unit showing the thermal resistance as a function of external liquid flow rate and dry-cooler fan speed. The thermal resistance improves strongly with increasing fan speed (and air flow) as well as with reducing external (hot) loop flow rate and follows a power trend with both parameters.

Zero fan speed operation was also tested at four different external loop flow rates (4, 6, 8 & 14 GPM). In these test conditions we rely on natural convection cooling at the dry-cooler. Due to the much lower natural convection heat transfer coefficients, the thermal resistance was found to range

from 2 to 3 °C/kW, and is over three times larger than the thermal resistance obtained at the lowest fan speed tested (120 RPM). In the current implementation, the fans could not be reliably lowered below 100 RPM due to stalling.

Figure 4.4-7. Relation between fan speed in RPM and the calculated air flow rate. The trend is almost linear (shown as a dotted line) though there is considerable spread due to the small temperature differences at higher fan speeds resulting in greater uncertainty in the determined air flow rate.

Figure 4.4-8(a) Approach thermal resistance as functions of external loop flow-rate, fan speed and external loop coolant. Addition of 20% and then 50% propylene glycol (PG) to the external loop results in an increase in the thermal resistance (b) Dry cooler conductance versus fan speed

Figure 4.4-8 shows the impact of propylene glycol addition to the approach thermal resistance. Addition of the propylene glycol shows a reduction in the dry-cooler conductance and an associated increase in the dry-cooler approach resistance. The conductance, $1/\epsilon \cdot C_{min}$, of the dry-cooler in Fig

4.4-8b shows an interesting piece-wise like behavior with an almost linear increase with fan speed up to about 400 RPM followed by an almost flat region beyond with only a slight increase in the conductance. This again emphasizes the relatively small benefit gained by increasing fan speed beyond 400-500 RPM.

Impact of Operating Conditions on Pressure Drop and Power Consumption

Changes in the flow-rates, fan speeds, heat exchanger arrangement and propylene glycol percentage have an associated impact on the liquid loop hydraulic resistance and power consumed by the different pieces of cooling equipment.

Figure 4.4-9 shows the pressure drop (a) in the internal loop (which includes the pressure drop across the rack and liquid-to-liquid heat exchanger) and the power consumed (b) by the internal loop pump as the flow rate is varied for the single and double heat exchanger configurations. Since the flow length is the same for either a full counter or counter-parallel double heat exchanger arrangement, the pressure drop and power consumed are similar. Addition of the second heat exchanger shows a clear increase in the pressure drop in the internal loop as well as pump power consumption. The increased power consumption must be weighed against the thermal benefit that is obtained by using the second heat exchanger. The pressure drop and power follow quadratic and cubic trends with the flow-rate respectively as expected from theory.

Figure 4.4-9. (a) Internal loop pressure drop and (b) internal loop pump power as a function of the flow rate. Addition of the second heat exchanger results in an increase in the pressure drop and thus energy required to provide a given flow rate.

Similarly, Figure 4.4-10 shows the pressure drop (a) in the external loop (which includes the pressure drop in the liquid-to-liquid heat exchanger, dry-cooler and piping) and the power consumed (b) by the external loop pump as the external loop flow rate is varied. Here, both the impact of the addition of a secondary buffer unit heat exchanger as well as the addition of propylene glycol to the external loop is observed. Addition of the second heat exchanger causes both the pressure drop and power consumption to rise due to the added hydraulic resistance. Addition of propylene glycol increases the pressure drop and power consumed even further due to the higher viscosity of the propylene glycol mixture. Again, the pressure drop in the external loop and power consumed follow quadratic and cubic trends with the flow-rate respectively, as expected from theory.

Figure 4.4-10. (a) External loop pressure drop and (b) external loop pump power as a function of the flow rate. Addition of the second heat exchanger as well as addition of propylene glycol results in an increase in the pressure drop and thus energy required to drive a the external coolant at a given flow rate.

Finally Figure 4.4-11 shows the power consumed by the dry-cooler fans as they are ramped from 0 to the maximum speed of 1450 RPM. The power consumed follows a clear cubic trend with the fan speed (or air flow rate). The power consumed by the dry-cooler fans is comparable to the internal and external pumps up to about 750 RPM beyond which it draws a larger and rapidly growing amount of power making it highly undesirable to operate at fan speeds beyond this point.

Figure 4.4-11. Power consumed by the dry-cooler fans as a function of fan speed follows a cubic trend. The power consumed by the fans is comparable with the pumps below 750 RPM, beyond which they progressively consume a large percentage of the total cooling power.

Total Thermal Resistance and Power Consumption

The overall impact of every significant piece of cooling equipment on the total CPU coldplate-to-

ambient thermal resistance and cooling equipment power consumption can be modeled using the characterized functions and is shown in Figures 4.4-12 and 4.4-13. In Figure 4.4-12, the internal loop flow rate is set at 4 GPM and the external loop flow and cooler fan speed are varied. The results show that the total thermal resistance is dominated by the CPU coldplate (R_{cp}) thermal resistance at the lower internal loop flow rate of 4 GPM as compared to the higher 8 GPM flow rate as shown in Figure 4.4-13. The buffer thermal resistance (R_{buff}) reduces while the dry-cooler resistance (R_{dc}) increases as external loop flow is increased. The total cooler and buffer resistance reduces with both external loop flow and cooler fan speed. The dry cooler resistance reduces with fan speed but the improvement is small after 500 RPM as described before. As shown in Fig4.4-12b and 4.4-13b the total power consumption does not change significantly until fan speeds of 500 RPM but then increases with dramatically higher power consumption, dominated by fan power, at 1450 RPM.

Figure 4.4-12. (a) Thermal resistance and (b) power consumption at a low 4 GPM internal loop while external loop flow rate and dry cooler fan speed are varied.

(a)

(b)

Figure 4.4-13. (a) Thermal resistance and (b) power consumption at a high 8 GPM internal loop while external loop flow rate and dry cooler fan speed are varied.

4.2 Day Long Operational Runs

Along with a study of the impact of flow rates, fan speeds, heat exchanger configurations and propylene glycol on the key cooling and heat exchange equipment in the test data center, it is valuable to also investigate the impact of diurnal, seasonal and weather variations on the thermal performance and cooling energy use of the test data center as a whole.

To initially study the impact of diurnal, weather and seasonal changes the test data center was operated for approximately a day in the summer, 4th August, and two consecutive days in the fall, 19th and 20th October. The summer day was clear and warm. The first fall day was rainy and cool and the second day was overcast, dry and also cool. For the purposes of this comparative test, the same CPU and memory stress tests were run which result in an approximately equivalent computational IT load. The internal and external water flow-rates were both set to 7.2 GPM with the dry-cooler fans were set to linearly vary from 170 to 500 RPM as the temperature of the water entering the buffer unit on the external cooling loop side varied from 30 to 35 °C. Below 30 °C, the fans were fixed at 170 RPM and above 35 °C they were fixed at 500 RPM. Since there is little

benefit in raising the fan speeds above 500 RPM as documented by the characterization studies, it was chosen as the upper limit of the control range. This automated control algorithm provided a simple method to increase the amount of cooling when either the IT heat load or ambient temperature rises. Figure 4.4-14 shows the key temperatures at various stages in the data center over the three test days. For the summer day, the servers were not instrumented, and for the fall days, the hottest CPU coldplate (CPU1 coldplate on Node 37) and DIMM spreader (DIMM 18 on node 37) temperatures are reported. All three days show that the data center and component temperatures generally track closely with the diurnal variations in the ambient temperature.

Figure 4.4-14. Temperature trace for a summer (a) and two fall days (b), (c) the first being rainy and the second overcast and dry. Flow rates set to 7.2 GPM both on the internal and external loop with fans set to vary from 170 to 500 RPM as the pre-buffer temperature rises from 30 °C to 35 °C. Cooling power was very similar on all three days ~430 W and is only 3.3% of the total IT power.

Due to the pre-buffer water temperature rising above 30 °C on the summer day (Figure 4.4-14(a)), the fans ramped up from 170 RPM to a maximum of 340 RPM towards the end of the test with a mean of 200 RPM during the 22+ hour run. This also resulted in a reducing temperature difference between the various probe points and the ambient as the day grows warmer. However, the much

cooler weather during the fall days results in the fan speeds remaining at a fixed 170 RPM throughout the day. The constant fan speed also resulted in an approximately constant temperature difference between the different probe points and the ambient. The impact of rain on the performance of the dry-cooler was found to be minimal being a cool fall day, with rain on a hot dry summer day expected to have a more significant effect due to the added evaporative heat transfer.

Despite the slightly higher average fan speed during the summer day the extra power consumed by the dry-cooler fans at these low speeds is small, resulting in approximately the same total cooling power of around 420-430 W being consumed on all three days. With similar IT power draw of 13.1 kW on the test days, the coefficient of performance ($COP = \text{IT heat dissipated} / \text{cooling equipment power}$) is determined to be approximately 30 and the cooling power fraction only 3.2% to 3.3% of the IT power, well below the typical 50% for a refrigerant and CRAH based data center. The measured average heat loss from the IT equipment into the room was found to be less than 4% on all three test days indicating that much of the heat generated is being absorbed by the fluid and transported away from the local data center environment. Two items not included in these efficiency and energy use calculations are the power draw by the monitoring and control system which consumes approximately 82 W and the power required to warm the pump Variable Frequency Drive (VFD)® control enclosure, located outside the building, which consumes approximately 250 W when the ambient temperature drops below 15 °C. The monitoring and control system is not included as it is a fixed energy cost that does not scale with the data center and because it is part of the experimental equipment and is thus, not indicative of a control system that would be used in an actual data center. The heating power consumed by the VFD control enclosure is not included as it is an artifact of its current location. The external pump VFD control enclosure can easily be located within the data center facility where the heater would not be necessary.

Figure 4.4-15. Temperature trace for the low cooling power test with lowered flow rates of 4 GPM on the internal and external flow loops. Fans set at a constant 170 RPM. Cooling consumes 210 W, 1.6% of the IT power.

Figure 4.4-15 shows the temperature traces for a low cooling power test that was carried out on October 5th, a clear, cool fall day. The same CPU and memory exercises were run, but the internal

and external flow rates were reduced to 4 GPM and the fans set to a constant 170 RPM. Under these operating conditions the cooling power drops to 210 W, half of the summer and fall day runs. With an IT power draw of 13.2 kW, the COP for this run is 64, with the cooling power only 1.6% of the IT power. Low flow rates help reduce cooling equipment energy use but increase the component-to-ambient thermal resistance as determined in Section 4.1. Comparing the component to ambient temperature difference between this low power fall day run against the standard runs on October 19th and 20th, the CPU coldplate temperatures are higher by approximately 6 °C and the DIMM18 spreader temperatures by approximately 4 °C due to an increase in the CPU-to-ambient and DIMM-to-ambient thermal resistance of 0.5 and 0.3 °C/kW respectively. However, the low average ambient temperature of 11 °C during this run makes this increase in thermal resistance acceptable. The average measurements and results from all four day long runs discussed are summarized in Table 5.

Table 5. Summary of test conditions during the three comparative standard tests as well as a fourth low cooling power test.

	Summer	Fall 1	Fall 2	Low Cooling Power
Date	4-Aug	19-Oct	20-Oct	5-Oct
Weather	clear	rainy	overcast	clear
Ambient Temp [C]	24.0	14.9	13.9	10.8
Ext Loop Flow [GPM]	7.1	7.1	7.1	3.9
Int Loop Flow [GPM]	7.2	7.2	7.2	4.0
Fan Speed Setting [RPM]	170-500	170-500	170-500	170
Pre Rack Liq Temp [C]	33.8	26.8	26.1	23.2
Pre Rack Air Temp [C]	36.5	30.5	30.1	29.6
IT Power [kW]	13.14	13.10	13.07	13.21
Cooling Power [kW]	0.42	0.43	0.43	0.21
Cooling / IT %	3.2	3.3	3.3	1.6
COP	31	30	30	64

A longer four day run starting on the evening of October 28th was carried out. The resulting temperature trace is shown in Figure 4.4-16. This test was initiated with 4 GPM water flow on the internal loop and 3 GPM on the external loop. The dry-cooler fans were set at 170 RPM and the recirculation valve was set to 30% open. The secondary buffer liquid-to-liquid heat exchanger was also added to the cooling loops in full counter flow arrangement. Only the bottom 21 servers in the rack were run with just a CPU exerciser as indicated by the DIMM spreader temperatures being lower than the CPU coldplate temperatures unlike in the previous day long runs. This test was unique in that it captures the impact of a freak fall snow storm that hit New York on October 29th, 2011. The sudden dip in the ambient temperatures around noon on the 29th marks the beginning of the heavy snow storm that lasted throughout the day resulting in downed trees, branches and power lines and significant power outages in the area. The snow storm was followed by sub-freezing temperatures on the 30th through the 31st (predicted to be well below -5 °C but measured to be down to -3 °C). Potential power outages leading to power loss to the IT rack and the running pumps combined with below freezing temperatures motivated the research team to add propylene glycol into the external loop while the data center test facility was running. The propylene glycol was added on the evening of the 29th and its dramatic impact is clearly seen in the step jump of 7-10 °C

in the various temperatures. As discussed in Section 4.1, the addition of propylene glycol resulted in higher approach thermal resistances across both the buffer unit and the dry-cooler. The addition of propylene glycol also resulted in a drop in the external loop flow-rate from 3 GPM to 2.6 GPM. The increased thermal resistance, due to the addition of the propylene glycol and lowered coolant flow rate, results in the observed jump in the temperatures.

Figure 4.4-16. Temperature trace during a four day period over the course of which a freak fall snow storm occurred. The trace shows the dramatic impact of adding propylene glycol to the external loop coolant to avoid freeze damage.

Server Data for Single Day Operation – August 4th, 2011

Data from the servers was simultaneously collected during the single day operations previously described. Figure 4.4-17 shows (a) the hottest core DTS values for each CPU, (b) maximum DIMM temperature for each of the 12 DIMMs and (c) the rpm of the system fans for one of the servers during the sample 22 hours run in August 2011 that began and ended in the afternoons of successive days. Observations such as variation of 5-6 °C in the DIMM temperatures and CPU 2 running relatively cooler than CPU 1, were as expected based on the computational fluid dynamic simulations of the cooling loop. The rpm of the server fans changes predominantly based on the server inlet air temperature. The more normal rpm changes due to load driven processor temperature rise were eliminated as even under full power the processors were running below the temperatures which would normally cause processor driven fan rpm increases.

Figure 4.4-18 shows the outdoor air temperature, the pre-MWU and pre-Rack coolant temperatures for the same 22- hour run. Note that because the internal and external loop coolant flow rates are kept constant through the sample run, the temperature delta between the pre-MWU and pre-Rack temperature remains constant. Also, when the pre-MWU temperature is less than 30 °C, the Outdoor Heat Exchanger fans run at constant rpm causing the temperature delta between the outdoor ambient

temperature and the pre-MWU temperature to remain constant. However, when the pre-MWU temperature exceeds 30 °C, the Outdoor Heat Exchanger fans starts to ramp up causing a drop in the temperature delta between the outdoor air temperature and the pre-MWU temperature. Hence, over the duration where the pre-MWU temperature is less than 30 °C, temperatures at all the locations of the cooling system and of the cooled electronics (that is, the pre-MWU, the pre-Rack, microprocessors junction temperature, DIMMs temperature, etc.) follow the outdoor ambient temperature profile at an essentially fixed offset.

Figure 4.4-17. Server component data showing (a) hottest core DTS numbers for CPU 1 and CPU 2, (b) DIMMs temperature for each of the 12 DIMMs and (c) system fans rpm for one of the server from a sample 22 hours run.

Figure 4.4-18 also shows the hottest DIMM temperature (DIMM 17 for this server) and the hottest core estimated temperature for each CPU. In the absence of a direct calibration between DTS values and absolute temperature, we choose to approximate the hottest CPU core temperature as 100 minus the absolute value of the DTS number. There were 38 servers in the rack with CPU exercisers and memory exercisers running on every server to provide steady heat dissipation from the processors

and from the DIMMs. Average DTS for the hottest core in CPU 1 was -43.5 with the max/min values of -36.7/-50.5. Average hottest DIMM (#17 for this server) temperature was 53 °C with the max/min values of 55 °C/50 °C. All the other servers in the rack showed similar temperatures, DTS values and fan rpm profiles.

Figure 4.4-18. Variation of temperature from the outdoor air to the server components.

From Figure 4.4-18, it can also be seen that the minimum temperature occurs around 12.4 hours and the maximum temperature occurs around 20.7 hours. Frequency distributions of the CPU DTS numbers and maximum DIMM temperatures at these time instances were evaluated and are presented in Figure 4.4-19. The mean maximum CPU1 core DTS number at time = 12.4 hrs was -50 and at 20.7 hrs was -42.1 with a standard deviation of 1.92 and 1.74 respectively. The mean maximum CPU2 core DTS number at 12.4 hrs was -51.6 and at 20.7 hrs was -43.7 with a standard deviation of 1.53 and 1.36 respectively. The variability in the DTS numbers can be attributed to the general variability in the performance of each core in a micro-processor. DTS numbers of each core of each processor were also recorded and evaluated to characterize this core-to-core and processor-to-processor variability. The mean maximum DIMM temperature at 12.4 hrs was 47.2 °C and at 20.7 hrs was 53.4 °C. Note that the variability in the DIMM temperatures is mainly due to the different types of DIMMs. All the servers that reported relatively cooler DIMMs had 8GB DDR3 DIMMs from Supplier 1 while all the servers that reported relatively warmer DIMMs had 8GB DDR3 DIMMs from Supplier 2. This is consistent with the observation that Supplier 1 DIMMs dissipate less heat than Supplier 2 DIMMs for similar performance.

(a) CPU 1 @ t = 12.4 hrs (cooler)

(b) CPU 1 @ t = 20.7 hrs (warmer)

(c) CPU 2 @ t = 12.4 hrs

(d) CPU 2 @ t = 20.7 hrs

(e) Max DIMM Temperature @ t = 12.4 hrs

(f) Max DIMM Temperature @ t = 20.7 hrs

Figure 4.4-19. Frequency distribution of CPU 1 and CPU 2 DTS numbers and maximum DIMM temperatures at t = 12.4 hrs (cooler) and t = 20.7 hrs (warmer) from the 22 hours test run.

4.3 System Servo Control

The system characterization and day long system operation data were used to develop temperature-based servo control algorithms for long term continuous operation of the DELC system. In this study, the rack inlet coolant temperature was dynamically controlled to minimize the data center cooling power consumption while under varying outdoor temperature and workload conditions.

Figure 4.4-20. Three zone control algorithm for cooling energy minimization for the DELC system (a) control flowchart (b) graphical representation of the three distinct control zones.

A graphical representation of the control is shown in Figure 4.4-20 in which the system operates at a specified minimum cooling power setting as long as the rack inlet coolant temperature being controlled (T_{Measured}), is between a Minimum and a Maximum Temperature Target.

As shown in the flow diagram of Figure 4.4-20(a), the cooling system is started at a specified minimum cooling power setting. This minimum setting need not be the global minimum for the cooling system but rather a user selectable input. At this setting, there is a certain temperature delta between the rack coolant inlet temperature and the outdoor ambient temperature referred to as ΔT_o . According to this control, if T_{Measured} approaches the $T_{\text{min target}}$ the system goes into a winter -mode operation and begins to open a recirculation valve to maintain the system above the dew point. When the recirculation valve is opened, the temperature delta becomes greater than ΔT_o to maintain rack inlet coolant temperature above the dew point or at $T_{\text{min target}}$. Holding the recirculation valve at any certain percent open setting requires a negligible fraction of total cooling energy. If the T_{Measured} increases above the $T_{\text{min target}}$, the cooling system begins to close the recirculation valves. Next, if T_{Measured} is between the $T_{\text{min target}}$ and the $T_{\text{max target}}$, the cooling system operates at its minimum cooling power setting. And if T_{Measured} is above the $T_{\text{max target}}$, the servo loop is engaged to control the cooling elements to servo T_{Measured} close to the Target temperature. For example, the external loop pump flow rate and the Outdoor Heat Exchanger fans speed could be changed proportionately to keep T_{Measured} close to the Target temperature. Thus, this approach provides three distinct zones of control as illustrated in Figure 4.4-20(b).

1) *Zone 1: Below $T_{\text{min target}}$* - In this zone, the system responds by opening the recirculation valves to keep the rack inlet coolant temperature above the dew point and/or maintain the temperature at the $T_{\text{min target}}$.

2) *Zone 2: Above $T_{\text{min target}}$ and Below $T_{\text{max target}}$* - The system operates in an energy efficient cooling mode to optimize the cooling power while letting the rack inlet coolant temperature drift between the $T_{\text{min target}}$ and $T_{\text{max target}}$.

3) *Zone 3: Above $T_{\text{max target}}$* - The system servo is initiated to control the cooling elements to maintain the rack inlet coolant temperature at $T_{\text{max target}}$.

The input to the Zone 3 PI servo control is the temperature difference or control delta between the actual and required rack inlet water temperature. The required proportional and integral gains were determined by trial-and-error by observing the dynamic behavior of the system when step changes in power or set-point were input. Figure 4.4-21 shows examples of step changes in IT power and water temperature set-point. The system is run with a P-gain of 5 and an I-gain of 0.2. Step change in the IT power (2.7 hours into the experiment) does not result in oscillatory or unstable control. The slow increase in water temperature results in the control algorithm smoothly tracking and maintaining the water temperature. At 5.4 hours, the set-point was suddenly lowered from 30 °C to 28 °C resulting in the fan and pump speeds rapidly ramping up to close the temperature gap until settling to the slightly higher operating speeds. In this case, the slightly under-damped behavior is clearly seen as the pre-rack liquid temperature oscillates before settling to within 0.1 °C of the set-point. The settling time is measured to be approx. 30 minutes. The integral control was added to eliminate a small (1-2 °C) steady-state error that was expected and observed when operating in proportional only control mode.

Figure 4.4-21. Transient response of rack inlet water temperature to changes in power and set-point.

4.4 Long Term Continuous Run with Dynamic Servo Control

To validate the dynamic servo control and to obtain a longer term measurement of the data center cooling energy use over the summer, a long 60+ day run was initiated on May 11, 2012. During the first 49 days, the IT power was kept constant between 13.5 and 14 kW as shown in Figure 4.4-22(a). The power was then intentionally varied over the next 10 days to observe the data center cooling equipment behavior and power consumption under variable workload conditions. The power was also varied during day 41 and 42 due to high ambient temperatures, shown in Figure 4.4-22(b), that required the IT load to be scaled back to avoid over temperature conditions. Note that later experimental results indicated a target of 38 °C was reasonable, which would have reduced or eliminated the need to scale back the IT load. The internal pump speed was set so as to provide approximately constant 6 GPM of water flow through the IT rack. The external loop pump and dry-cooler fans were controlled every 10 seconds using the PI based algorithm described previously with a P-gain of 3.5 and an I-gain of 0.15. The water temperature set-point was 35 °C which ensures that the air temperature entering the rack was less than 40 °C.

Figure 4.4-23 shows the key daily-averaged system temperatures during the 62 day period. The portion between the dotted lines represents the controllable ambient temperature range for a 13 kW heat load and a 35 °C set point which is Zone 3. In this Zone, the rack inlet water temperature ($T_{pre\ rack}$ shown in green) is servo controlled to the 35 °C set point. For ambient temperatures below 17 °C the equipment is set at their selectable minimum settings in Zone 2 and the rack inlet water temperature drifts with outdoor ambient temperature. When ambient temperature exceeds 31 °C the cooling equipment is set at their selectable maximum settings and the rack inlet water temperature will drift above the 35 °C set point. Thus, an ambient temperature of 33 °C would result in a pre-rack water temperature of 37 °C, 2 °C above the required set-point. On days this occurred such as days 41 and 42, the server workload was reduced to help maintain the water and air temperatures seen by the servers. The pre-server liquid and air temperatures were about 2 °C higher than the rack inlet water temperature. The implemented servo control maintained this rack water temperature to within 0.5 °C of the set-point when operating within the controllable ambient temperature band.

(a)

(b)

Figure 4.4-22 (a) IT Power and (b) ambient air temperature over the course of the 62 day run.

Figure 4.4-23. Key data center temperatures during the 62 day run with dynamic servo control implemented.

The various cooling equipment speeds (daily-averaged) as they were being servo controlled can be seen in Figure 4.4-24. As the ambient temperatures go up, the external pump and fan speeds also increase to maintain the pre-rack coolant temperature. This results in the power consumption trace as shown in Figure 4.4-25. The pump and fan power consumption increases on days where the average ambient temperature is high and is low when the temperature is low.

Figure 4.4-24. Equipment speeds as driven by the servo control in response to measured IT power and ambient air temperature.

Figure 4.4-25. Equipment power consumption over the 62 day run showing the peaks during the hotter periods and flat minimums during the cooler portions of the run window

Figure 4.4-26 shows the instantaneous IT power variation, pre-rack coolant and ambient temperature and equipment power use over a ten day period. The IT loads were varied using Linux scripts

running off the head-node (master server). This shows the impact of variable workload and ambient temperature on the coolant temperatures and cooling power use. High workloads combined with higher ambient temperatures result in high cooling power usage as the equipment speeds are maximized. This also results in the pre-rack coolant temperatures rising above the set-point as seen by the peaks. Low workloads combined with low ambient temperatures result in the temperature drifting lower and the equipment speeds and power consumption minimized as clearly seen on day 7.

Figure 4.4-26. Impact of variable workload (top) on the pre-rack water temperature (middle) and cooling equipment power use (bottom).

A key requirement of the long term study was also to quantify the energy efficiency of the data center test facility. The daily averaged cooling energy to IT energy ratio is shown in Figure 4.4-27. For the first 40 days when the IT loads are high the average energy ratio is around 3%. However, during times when the workloads are low or during times of sustained higher ambient temperatures, the daily averaged energy ratio is higher.

Statistics for the whole duration found that the energy ratio is 3.5% with an average energy use of 0.42 kW. The ambient temperature ranged from 4.7 °C to 35.8 °C with an average of 21.6 °C. The pre-rack water temperature during this same time ranged from 24.5 °C to 39.4 °C with an average of 34.6 °C. This compares favorably with the required set-point of 35 °C. Cooling power ranged from a minimum of 0.28 kW to 1.62 kW. Heat loss from the system to the room environment was found to be an average of 1%.

Figure 4.4-27. Cooling power as a fraction of IT power over the 62 day run period. Hotter days and days with lower workload result in comparatively higher cooling power fraction.

4.5 Energy Based System Model Validation

The validation study of energy based system model is presented in this section. The system model based on energy balance presented in section 3.1 was validated against six different steady state test cases as well as against day long system operation which will be described in this section.

I. Steady State Test Cases

The experimental conditions of six different steady state test cases are summarized in Table 6 below. The indoor (MWU) and outdoor (VFD) loop flow rates in Gallons per Minute (GPM) are shown in columns two and three respectively. The RPM setting of the fans for the Outdoor Heat Exchanger (OHE) appears in column 4. The IT power in column 6 was nearly constant in the range of 13.4 to 13.9 kW. The outdoor air temperature ranged from 14.7 to 27.4 °C. These values were used as the inputs to the system model and the outputs were the temperatures at different locations in the data center test facility and of the liquid cooled components. For all six cases, the coolant in the external and in the internal loop was water and processor intensive workloads were executed to provide steady heat dissipation from the servers.

Figures 4.4-28 and 4.4-29 compare the system model predictions of the facility side temperature. Figure 4.4-28 shows the pre-MWU coolant temperature prediction and its comparison with the experimental data for the six cases studied. For all the cases, the temperatures predicted by the system model were within 1 °C of the experimental data. Similarly, Figure 4.4-29 shows agreement within 1 °C between the system model prediction and the experimental data for the pre-rack coolant temperature.

Table 6. Steady state test cases for system model validation

Case #	MWU Flow Rate (GPM)	VFD Flow Rate (GPM)	OHE Fan Speed (RPM)	Recirculation Valve (%Open)	IT Power (kW)	Outdoor Air Temperature (°C)
1	7.2	8.0	500	0.0	13.4	23.2
2	9.2	8.0	500	0.0	13.4	27.4
3	4.0	8.0	500	0.0	13.7	25.4
4	7.2	7.4	400	0.0	13.5	23.0
5	9.2	6.1	225	0.0	13.7	22.8
6	4.0	3.7	105	9.5	13.9	14.7

Figure 4.4-28. Pre-MWU temperature prediction and comparison with experimental data.

Figure 4.4-29. Pre-Rack temperature prediction and comparison with experimental data.

For all these six cases, there were 40 hybrid liquid/air cooled servers inside the rack. Of the total IT heat load, roughly 67% of the heat was transferred to the liquid and 33% of the heat was transferred to air at the server level. This 33% was then transferred to the liquid at the side car air to liquid heat exchanger. Since, processor intensive workloads were executed on these servers, roughly 90% of the heat that was conducted to the liquid coolant at the server level, was generated by the processors and the remaining 10% by the DIMMs.

Figure 4.4-30(a) shows an image of one such server showing the locations of CPU1, CPU2 and the DIMM in slot number 18 (represented as DIMM 18). Figure 4.4-30(a) also shows the coolant flow path inside the server. The coolant enters the front of the server and bifurcates into two parallel flow paths passing through the front and middle cold rails, cooling the front bank of DIMMs and partially cooling the rear bank of DIMMs. The flow then recombines and passes through the CPU2, the rear cold rail and finally the CPU1. Hence, the DIMM at slot number 18 is the last DIMM to be cooled and lastly, CPU1 is cooled by preheated coolant. Figures 4.4-30(b), (c) and (d) compare the system model prediction of the liquid cooled components (CPU1, CPU2 and DIMM18) temperature with the experimental data for all the six cases. In the system model, all the servers are assumed to be similar and hence, the system model predictions are compared against the average CPUs and DIMMs temperature of all the servers in the racks. It can be seen in Figure 4.4-30 that the temperature predictions for CPUs and DIMMs are within 1.5 °C of the experimental data.

Figure 4.4-30. System model prediction of CPU and DIMM temperatures and comparison with experimental data.

Figure 4.4-31 shows the rack average CPU1 and CPU2 temperatures for the Case #1 (Table 6) with the bars showing the CPU temperature variability across all the servers. Figure 4.4-31 also compares the model prediction with the data. Figure 4.4-32 shows the rack average DIMM temperatures for Case #1 with the bars showing the temperature variability across all servers. The x-axis represents the DIMM slot numbers. The slots 2 through 9 are in the front bank while the slots 11 through 18 are in the rear bank. It can be seen that for all 12 DIMM slots, the model prediction is in very good agreement with the experimental data. Similar predictions were observed for all six test cases.

Figure 4.4-31. CPU temperature prediction and comparison with experimental data for test case 1 with the bars showing the temperature variability across all servers.

Figure 4.4-32. DIMM temperature prediction and comparison with experimental data for test case 1 with the bars showing temperature variability across all servers.

II. Temporal Day-Long Summer Operation Study

In addition to the above test cases, the day long data center test facility operation in August 2011 reported in an earlier section was also used for the system model validation. The data center test facility was continuously operated for 22 hours with varying Outdoor Heat Exchanger fan speeds and internal and external loop coolant flow rates set to 7.2 GPM and 7.1 GPM respectively. The Outdoor Heat Exchanger fans were programmed to linearly vary in speed from 170 RPM to 500 RPM as the pre-MWU temperature varied from 30 °C to 35 °C. For pre-MWU temperatures below 30 °C, the fans ran at a constant speed of 170 RPM. This control algorithm along with the IT Power and the outdoor temperature profile from this run were used as inputs to the system model. Of the total IT heat load, roughly 65% of the heat was transferred to the liquid and 35% of the heat was transferred to air at the server level. Since, processors as well as memory intensive workloads were executed simultaneously on these servers, roughly 80% of the heat that was conducted to the liquid coolant at the server level, was generated by the processors and the remaining 20% by the DIMMs. During the day long run, the data center test facility operates in a quasi-steady state mode. As long as the rate of change of the outdoor ambient temperature is slower than the intrinsic time constant of the data center test facility, the steady state system model can be used to predict the system performance.

Figure 4.4-33 compares the system model prediction of the pre-MWU and pre-rack coolant temperatures with the experimental data for the 22 hours run. Figure 4.4-34 compares the liquid cooled server component temperatures for a typical server with the experimental data for the 22 hours run. Figure 4.4-35 shows the total cooling power prediction for the day long run and its comparison with the experimental data. The total cooling power here is the sum of the indoor loop pump power, the external loop pump power and the Outdoor Heat Exchanger fan power. Figure 4.4-36 compares the prediction of the Outdoor Heat Exchanger fan rpm with the experimental data. It can be seen that the system model provides a very good prediction of the facility and server component temperatures, of the total cooling power and of the Outdoor Heat Exchanger fan speed variations.

Figure 4.4-33. Facility side temperature prediction and comparison with experimental data for a day long summer run.

Figure 4.4-34. Typical server CPU and DIMM temperature prediction and comparison with data for a day long run.

Figure 4.4-35. Cooling Power prediction and comparison with experimental data for a day long summer run.

Figure 4.4-36. Outdoor Heat Exchanger fans speed prediction and comparison with experimental data for a day long run.

4.6 Model Simulation for Simple Control Methods

The system simulator enables the study of different methods of control and impact on performance and power consumption. Figure 4.4-37 displays three different simple control methods for controlling one or all of the coolant pumping devices, i.e. the dry cooler fans, the external loop pump, and the internal loop (buffer unit) pump. For example, Figure 4.4-37(a) shows one method of controlling the dry cooler fan speed as a function of the water temperature exiting the cooler, which is also the same temperature of coolant entering the inlet of buffer unit liquid-to-liquid heat exchanger on the external side. In this simple scheme, the dry cooler fan speeds are varied between 100 rpm and 500 rpm as a linear function of the post-cooler water temperature between the 30 to 35 °C range. This scheme illustrated via Figure 4.4-37(a) is very similar to the one used during the 22 hour experimental test discussed previously.. For that 22 hour base line test, only the dry cooler fan speed was varied between 169 rpm and 500 rpm as the post-cooler water temperature changed between 30 °C and 35 °C. The external pump and internal buffer unit pump were both set to fixed speeds of 1566 rpm and 3300 rpm, respectively, for the 22 hour run. The variation of fan speed resulted in preventing the water temperature exiting the dry cooler from raising much above 30 °C, because the fans would ramp up as soon as the water temperature rose above 30 °C and the speeding fans would thus reduce the temperature differential between the outdoor air and the post-cooler coolant. When the post-cooler coolant temperature is below 30 °C, the fans are at their minimum speed and thus, for these conditions the post-cooler coolant temperature would track the outdoor air temperature resulting in cooler temperatures than were really necessary to satisfactorily cool the server rack.

The three different control methods are shown in Figures 4.4-37(a), (b), and (c), respectively, and are named as Case 1, Case 2, and Case 3, respectively. As discussed in the preceding text, the simplest one shown in Figure 4.4-37(a) involves the control of only the dry cooler fans with the external and internal pumps fixed at a constant speed, which is very similar to the control used for the 22 hour run. Figure 4.4-37(b) extends the method from Figure 4.4-37(a) in which, in addition to the control of the dry cooler fan, the speed of the external pump is also varied (900 to 1500) as a

function of the post-cooler water temperature between the same 30 °C to 35 °C range. For the method shown in Figure 4.4-37(b), the internal buffer unit pump is kept fixed at 3300 rpm. The third simple control method is depicted in Figure 4.4-37(c) and builds on the one from Figure 4.4-37(b). In this method, all three coolant pumping devices are controlled as a function of the post-cooler water temperature, i.e. the dry cooler fans are varied between 100 rpm and 500 rpm, the external pump between 900 rpm and 1500 rpm, and the internal buffer unit pump between 2300 rpm and 3300 rpm. For the control method depicted through Figure 4.4-37(c), the range of the post-cooler temperature for which the coolant pump device speeds are varied is the same as for the other two methods, i.e. 30 °C to 35 °C.

Figure 4.4-37. Control methods - Fan and pump speed as a function of post cooler water temperature, (a) Dry cooler fan control with fixed speeds for external and internal pumps, (b) Control of dry cooler fans and external pump with fixed speed for internal pump, (c) Control of dry cooler fans, external pump, and internal pump.

Figure 4.4-38(a) displays total data center cooling power consumption for the three control methods described via Figure 4.4-37 using the heat load and outdoor air temperature from the 22 hour experiment as inputs into the model. As may be expected, the three control methods each yield different cooling power consumption values in each time step over the 22 hour period, thus yielding a different average cooling power use. Case 1 is similar to the actual experiment with control only of the dry cooler fans, and results in the highest cooling power use at each time step. Case 2, which involves the control of the dry cooler fans and the external pump, and uses less power than Case 1, but more than the Case 3 in which all three coolant pumping devices are being controlled. Cases 1, 2 and 3 result in average Cooling PUE values of 1.031, 1.025, and 1.017, respectively, which

correspond to cooling power usages of 3.1%, 2.5%, and 1.7%, respectively, of the IT power.

Figure 4.4-38. Cooling power and coolant temperatures for different control methods (a) Cooling power usage by dry cooler fans, external pump, and internal pumps (b) Rack inlet water temperatures, (c) Rack inlet air temperatures.

Figure 4.4-38(b) and 4.4-38(c) show the rack inlet water and air temperatures, respectively, that corresponds to the cooling power usages discussed previously with the use of the control methods that were described via Figure 4.4-37. As expected, Case 1 which is similar to the control of the summer 22 hour experiment, results in rack inlet water and air temperatures similar (for rack inlet water) to the actual experiment in that the temperatures track with the outdoor air temperature for cooler conditions with fan control activated temperature control only at the warmest temperatures (when the post-cooler water is higher than 30 °C). For Cases 2 and 3, the rack inlet water and air temperatures are significantly more uniform presumably, because the fan and pump speed increases addressed the warmed conditions while the ramp down of fans and pumps at the cooler outdoor conditions mean that the coolant temperatures are not allowed to cool down significantly and track with the outdoor conditions. This is because for Cases 2 and 3, the fan and pump speeds never reach the minimum device speeds prescribed through the respective control method and are varied in each time step throughout the 22 hour run, thus constantly regulating the coolant temperatures to be substantially uniform. It should be noted that the rack inlet water and air temperatures of about 40 °C resulting from the use of the control method from Cases 2 and 3, are considered satisfactory but near maximum tolerable based on the design specification for the servers that were retrofitted with liquid cooling structures for this study.

4.7 Performance Prediction for Typical Year and Geographical Locations

The previous section showed that system models can be used to predict, with sufficient accuracy, the system performance for a day long operation where the outdoor ambient conditions vary from 19 °C to 32 °C. This validated system model tool can also be used to extrapolate the system performance for the entire year as well as to different geographical locations.

Figure 4.4-39. A simple graphical user interface for the system model. System model prediction of the key server components temperature, of the power consumption and annual average power consumption and annual energy and cost savings for a typical year in Poughkeepsie, NY. The typical outdoor air temperature profile was obtained from NREL database [21].

Figure 4.4-39 shows a simple graphical user interface that was developed to interactively show the system performance at different locations and to highlight the benefits of the proposed chiller-less liquid cooled data center system. The tool requires the typical outdoor ambient air temperature profile, IT rack power, electricity cost per kWhr and control algorithm as key inputs. The typical outdoor ambient air temperature profile can be obtained from national databases such as those provided by the National Renewable Energy Lab (NREL) [21]. The tool then outputs the temperature at various locations in the system such as the pre-MWU, pre-rack, rack air, CPU and DIMM temperatures. The tool also outputs the total cooling power as a function of time and also as a function of outdoor ambient air temperature. Various other plots, depending upon the need, can also be generated. The tool also calculates the annual average cooling power and represents it as a percentage of the IT power. Based on the average cooling power, the tool calculates the annual energy and operational cost savings per rack (each with 42 servers) as compared to a typical refrigeration based air cooled data center. In Figure 4.4-39, the control algorithm selected is the same as that implemented in the day long run. It can be seen that even for such a simple algorithm, the annual cooling power at Poughkeepsie, NY can be less than 3% of the IT power with up to \$6000 in annual savings in operating costs per rack of servers at an electricity rate of \$ 0.10/kWhr.

Figure 4.4-40. System model prediction for a typical year in Raleigh, NC. The typical outdoor air temperature profile was obtained from NREL database [21].

Figure 4.4-40 shows the system performance prediction for a typical year in Raleigh, NC. The control algorithm is the same as that implemented in the day long run discussed earlier. In Raleigh as well, the annual cooling power could be less than 3% of the IT power leading to significant operational cost savings. It can also be seen that Raleigh has a greater number of high temperature periods compared to Poughkeepsie resulting in relatively more hours of increased cooling power consumption. However, these periods of increased power consumption are too small a fraction of the year to have any significant impact on the annual average cooling power.

Figure 4.4-41 shows the psychometric chart illustrating the ASHRAE recommended classes for air cooled IT equipment. In Figures 4.4-39 and 4.4-40, it can be seen that during periods of high outdoor air temperature the air inlet temperature entering the servers (magenta colored curve) exceeded the ASHRAE A2 class with maximum temperature of 35 °C. While the air cooled servers retrofitted with liquid cooling in this study were tested for 50 °C inlet air temperature operation, they did not undergo long term reliability studies at these elevated temperature operations. However looking forward, the new ASHRAE A3 and A4 classes of recommended guidelines for air cooled IT equipment, servers would be qualified for 40 °C and 45 °C inlet air temperatures. Thus, future servers would be within operational range of the chiller-less liquid based cooling system approach and savings presented in these simulations could be realized.

In another similar study, the model was used to predict the system performance in nine different US cities assuming servers qualified for A3 and A4 ASHRAE classes. Figure 4.4-42 shows weather data from NREL, [21] for August 15 of a typical year for nine US cities representing different geographies and weather types including New York City (NYC), Chicago, San Francisco, Raleigh, Dallas, Phoenix, Seattle, Buffalo, and Poughkeepsie. For the data shown in Figure 4.4-42, hour 1 is

from 12 AM-1 AM. The nine cities studied comprise a wide range of climates ranging from hot (e.g. Phoenix and Dallas with maximum outdoor air temperature of 38.3 °C) to cool (e.g. Seattle with a maximum outdoor air temperature of 20.3 °C). Some cities such as San Francisco and Seattle experience very small diurnal outdoor air temperature fluctuations of less than 7 °C, while others such as Poughkeepsie see a wide temperature change of 15.6 °C for outdoor air in a single day.

Figure 4.4-41. Psychrometric chart illustrating the ASHRAE recommended guideline classes for air cooled IT equipment [22].

Model simulations were performed for the nine cities which are summarized in Tables 7 and 8 using the Case 3 control method discussed previously in section 4.7, whereby all three coolant pumping devices are controlled as a function of the post-cooler water temperature as shown in Figure 4.4-37(c). Table 7 shows the average coolant temperatures over the span of a summer day (August 15) with the loop operating with an average rack heat load of 13.1 kW. As seen from Table 7, the warmest coolant temperatures on average are experienced by Phoenix and Dallas while the coolest operation is experienced by San Francisco. Table 8 displays the average dry cooler fan speeds and the internal/external pump speeds, as well as the average cooling power usage by these devices. Table 8 also provides the average total cooling power, the average Cooling PUE, and the average percentage of the IT power that is used for data center cooling. Table 8 shows that the Phoenix and Dallas runs result in the largest cooling energy use (3.2-3.3% of IT) and the San Francisco run results in the lowest value (1.5% of IT). These values for average cooling energy usage correlate very well with average outdoor air and coolant temperature (Table 7). It should be noted that for the simulation results the IT rack power was assumed to be a constant, but in real systems the IT power

will vary with coolant temperature due to changes in the server fan power use and chip leakage power. However, it is expected that the trends presented are indicative of the data center cooling energy efficiency achievable at different geographic locations using this chiller-less liquid cooling system approach.

Figure 4.4-42. Outdoor dry bulb temperatures for nine US cities having different climates.

Table 7. Average coolant temperatures for nine US cities for model simulation using Case 3 control for August 15, NREL [21] typical year temperature data.

	Post cooler, C	Rack inlet water, C	Rack inlet air, C	Outdoor air, C
New York City	31.6	39.0	42.2	27.1
Chicago	29.4	37.9	41.2	19.0
San Francisco	27.8	36.5	39.9	15.5
Raleigh	29.2	37.6	41.0	19.1
Dallas	34.2	40.6	43.8	31.7
Phoenix	34.6	40.9	44.1	32.2
Seattle	28.7	37.3	40.7	16.7
Buffalo	30.2	38.5	41.9	20.9
Poughkeepsie	28.1	36.5	39.9	18.0

Figure 4.4-43 illustrates the hourly trends in fan and pump speeds as well as the loop coolant temperatures on which they depend on, for two cities with widely different outdoor air temperature profiles, namely, Dallas and San Francisco. While the average values for various parameters listed in Tables 7 and 8 serve to summarize the analyses on the nine US cities, the trends presented in Figure 4.4-43 provide insight on exactly how the fan and pump speeds respond to the variation in the outdoor air temperature over the day and thus the dry cooler exit water temperature on which the fan and pump speeds depend.

Table 8. Average fan and pump speeds and cooling powers from model simulation using Case 3 control
Nine US cities, August 15, Use of NREL typical year temperature data [21].

	Dry cooler fan speed, RPM	External pump speed, RPM	Internal pump speed, RPM	Dry cooler fan power, W	External pump power, W	Internal pump power, W	Total cooling power, W	Cooling PUE	Cooling power as % of IT
New York City	227.9	1091.9	2619.9	40.3	83.5	147.9	271.7	1.0207	2.07
Chicago	115.9	923.8	2339.5	18.5	64.5	119.7	202.7	1.0154	1.54
San Francisco	100.2	900.3	2300.4	18.8	62.1	116.2	197.2	1.0150	1.50
Raleigh	120.6	930.8	2351.4	18.9	65.3	120.8	204.9	1.0156	1.56
Dallas	354.0	1281.1	2935.0	118.6	108.0	186.7	413.3	1.0315	3.15
Phoenix	363.9	1295.8	2959.8	126.1	110.0	189.9	425.9	1.0325	3.25
Seattle	102.4	903.7	2306.1	18.8	62.4	116.7	197.9	1.0150	1.50
Buffalo	131.6	947.4	2378.8	19.3	67.0	123.3	209.6	1.0160	1.60
Poughkeepsie	121.3	931.9	2353.1	18.9	65.4	120.9	205.2	1.0155	1.55

Figure 4.4-43. Temperature and power variation for warm and cool climates, (a) Coolant temperatures for Dallas (warm), (b) Cooling device (fan and pump) speeds for Dallas (warm), (c) Coolant temperatures for San Francisco (cool), (d) Cooling device (fan and pump) speeds for San Francisco (cool).

Table 9 presents energy and energy cost related savings from using the chiller-less based data center liquid cooling system described in this report over a traditional chiller based data center cooling that is prevalent in the industry. The analyses shown in Table 9 assumes an IT load of 1 MW and a time period of one day, namely, the typical August 15 day that has been discussed in the preceding text. A traditional cooling system typically consumes ~50% of the IT power, which is ~500 kW for a 1 MW IT load [4]. While the use of this 50% value may not be exact for every city and for the specific time of consideration (August 15), it has been used in this study as a baseline value which may in fact be conservative for a summer day and not an annual average value. The cooling energy use that was discussed via Table 8 is used to calculate the cooling power use for this chiller-less liquid cooling configuration for a typical August 15 day. The difference between the average power usages for the single day analyzed is multiplied by 24 to calculate the energy consumption for a day and state

electricity cost estimates (2010 industrial rate [23]) are utilized to compute the energy cost savings for a single day. Since the state electricity cost does vary significantly over the US, actual realizable cost savings may not track with energy savings as illustrated by Table 9. Thus, while Seattle and San Francisco may see the same energy savings, the energy cost savings for Seattle are considerably lower than for San Francisco. However, the energy and energy cost savings presented in Table 9 for a single summer day illustrates the energy savings opportunity for this system.

Table 9. Energy/cost savings for a 1 MW IT load, for a typical August 15 day, Traditional versus DELC data center.

	Typical data center cooling, kW	DELC based data center cooling, kW	Energy Savings, kWh	Local cost of electricity, \$/kWh	Energy cost savings, \$
New York City	500	20.7	11504	0.0973	1119.3
Chicago	500	15.4	11631	0.075	872.3
San Francisco	500	15.0	11640	0.1078	1254.8
Raleigh	500	15.6	11625	0.0613	712.6
Dallas	500	31.5	11243	0.0658	739.8
Phoenix	500	32.5	11220	0.0674	756.2
Seattle	500	15.0	11640	0.0396	460.9
Buffalo	500	16.0	11617	0.0973	1130.3
Poughkeepsie	500	15.5	11627	0.0973	1131.3

5. Benefits Assessment

In 2010, there were roughly 33 million computer servers installed worldwide. In the US alone, about 12 million computer servers were installed of which 97% were volume servers and 3% were Mid-range and High-end servers [1]. Thus, to maximize the energy impact of reducing cooling energy usage this project focused on the largest segment of the server market which is the Volume server. IBM System x3550 M3 Volume server was chosen to demonstrate the new technology.

The energy usage of data centers reported by the EPA [15] was 61 billion kWhrs in 2006 which had doubled since 2000 and was projected to double again by 2011. The EPA [15] reported that Volume servers, the fastest growing segment of the market, were responsible for 68% of electrical usage and based upon the expected growth rate the electricity use by Volume servers was projected to reach 42 billion kWhrs in 2011.

The energy required to cool the IT equipment is roughly 25-30% of the total data center energy usage. For the projected growth rate for Volumes server energy use the cooling energy required could reach 21 billion kWhrs in 2011. In the newly proposed chiller-less liquid cooled data center system, the cooling energy could be reduced to 2 billion kWhrs. If an assumption is made for market penetration of a quarter of US data centers, the potential cooling energy savings could be up to 4.5 billion kWhrs per year.

6. Commercialization

The successful IBM experimental demonstration of this cooling technology is already having an impact on IBM products. In June, the Leibniz Superconductor Center in Germany announced the world's fastest commercially available hot-water-cooled supercomputer built with IBM System x iDataPlex Direct Water Cooled dx360 M4 Servers shown in Figure 6-1 below including some of the technologies developed with this award. The technologies are also drawing additional external interest for direct liquid cooled volume servers.

Figure 6-1. IBM System X iDataPlex Direct Water Cooled dx360 M4 Server.

7. Accomplishments

Technical Highlights

- Developed a liquid cooled data center system simulation model to determine the impact of engineering design choices on the projected overall system performance.
- Designed and constructed a chiller-less liquid cooled prototype data center facility in Poughkeepsie NY.
- Designed and integrated liquid cooling components for Volume Server processors and memory sub components.
- Developed an advanced metal interface for direct attach of coldplates to processor.
- Developed novel compliant coldplates for direct attach to processors.
- Created a servo control environment of the system to allow automated operation under varying IT workload and outdoor environmental conditions.
- Operated the system for a long term two month study demonstrating a cooling to IT energy ratio of 3.5%, compared to 50% for traditional air cooled data centers.
- Validated the system model with experimental data and utilized the system model to project the performance of the system in different geographies.
- The work resulted in 21 patent applications and 6 publications

Awards

- ITherm 2012 Outstanding Paper Award Thermal Management Track
- IBM Technical Exchange Conference 2012 Best Non-Confidential Oral Presentation

Publications

- [1]. “Extreme Energy Efficiency using Water Cooled Servers Inside a Chiller-less Data Center” M. Iyengar, M. David, P.Parida, V. Kamath, B. Kochuparambil, D. Graybill, M. Schultz, M. Gaynes, R. Simons, R. Schmidt and T. Chainer, ITherm 2012 May 30th -June 1st, San Diego, CA.
- [2]. “Impact of Operating Conditions on a Chiller-less Data Center Test Facility with Liquid Cooled Servers” M. David, M. Iyengar, P. Parida, R. Simons, M. Schultz, M. Gaynes, R. Schmidt and T. Chainer ITherm 2012 May 30th -June 1st, San Diego, CA. – *Outstanding Paper Award Thermal Management Track.*

- [3]. “Experimental Investigation of Water Cooled Server Microprocessors and Memory Devices in an Energy Efficient Chiller-less Data Center” M. David, M. Iyengar, P. Parida, R. Simons, M. Schultz, M. Gaynes, R. Schmidt and T. Chainer, SEMITHERM 2012 March 18-22nd, San Jose CA.
- [4]. “Server Liquid Cooling with Chiller-less Data Center Design to Enable Significant Energy Savings” M. Iyengar, M. David, P. Parida, V. Kamath, B. Kochuparambil, D. Graybill, M. Schultz, M. Gaynes, R. Simons, R. Schmidt and T. Chainer, SEMITHERM 2012 March 18-22nd, San Jose CA.
- [5]. “Experimental Characterization of an Energy Efficient Chiller-less Data Center Test Facility with Warm Water Cooled Servers”, P. Parida, M. David, M. Iyengar, M. Schultz, M. Gaynes, V. Kamath, B. Kochuparambil, and T. Chainer, SEMITHERM 2012 March 18-22nd, San Jose CA.
- [6]. “System-Level Design for Liquid Cooled Chiller-Less Data Center”, P. Parida, T. Chainer, M. Iyengar, M. David, M. Schultz, M. Gaynes, V. Kamath, B. Kochuparambil, R. Simons and R. Schmidt, ASME IMECE 2012, November 11-15th, Houston, TX.

8. Conclusions

A new chiller-less data center liquid cooling system utilizing the outside air environment has been shown to achieve up to 90% reduction in cooling energy compared to traditional chiller based data center cooling systems. The system removes heat from Volume servers inside a Sealed Rack and transports the heat using a liquid loop to an Outdoor Heat Exchanger which rejects the heat to the outdoor ambient environment. The servers in the rack are cooled using a hybrid cooling system by removing the majority of the heat generated by the processors and memory by direct thermal conduction using coldplates and the heat generated by the remaining components using forced air convection to an air- to- liquid heat exchanger inside the Sealed Rack. The system was successfully operated in New York over a two month period from May to June. The anticipated benefits of such energy-centric configurations are significant energy savings at the data center level. When compared to a traditional 10 MW data center, which typically uses 25% of its total data center energy consumption for cooling this technology could potentially enable a cost savings of roughly between \$800,000-\$2,200,000/year (assuming electricity costs of 4¢ to 11¢ per kilowatt-hour) through the reduction in electrical energy usage while also eliminating water usage of up to 240,000 gallons per day. Technologies developed under this program were used in the IBM System x iDataPlex Direct Water Cooled dx360 M4 Servers installed in a new Leibniz Supercomputer Center in Germany.

9. Recommendations

We recommend utilizing the energy efficient cooling technologies developed in this program to construct a renewable energy based net-zero data center illustrated by Figure 9-1.

The technical approach we propose is to merge a) chiller-less liquid cooling, b) renewable energy sources, c) energy storage, d) energy reuse for building heating, e) highly localized weather prediction and, f) dynamic system control and workload balancing to achieve a highly efficient renewable energy data center.

Figure 9-1. Overview of renewable energy based net-zero data center.

This approach of designing a data center for renewable energy from the ground up will provide a path to energy efficiency beyond that which could be achieved by retrofitting energy intensive refrigeration based air cooled systems with renewable energy. The liquid cooling technology provides a path to a fully self-sustainable data center.

This approach can also provide system-level architecture for mobile containerized secure grid independent data centers that can be sited in remote locations with high renewable energy content.

10. References

- [1]. J.G. Koomey, “Growth in Data Center electricity use 2005 to 2010”, Oakland, CA: Analytics Press., Aug 2011.
- [2]. Joe Kava, “Sustainable Data Centers and Water Management”, Google Data Center Efficiency Summit 2009, Mountain View, California, April 1, 2009.
- [3]. James Hamilton, “Data Center Efficiency Best Practices”, Google Data Center Efficiency Summit 2009, Mountain View, California, April 1, 2009.
- [4]. “Vision and Roadmap – Routing Telecom and Data Centers Toward Efficient Energy Use”, May 13, 2009 www1.eere.energy.gov/industry/datacenters/.../vision_and_roadmap.
- [5]. M. Iyengar, M. David, P. Parida, V. Kamath, B. Kochuparambil, D. Graybill, M. Schultz, M. Gaynes, R. Simons, R. Schmidt and T. Chainer, “Extreme Energy Efficiency using Water Cooled Servers Inside a Chiller-less Data Center”, ITHERM 2012, May 30th -June 1st, San Diego, CA.
- [6]. M. David, M. Iyengar, P. Parida, R. Simons, M. Schultz, M. Gaynes, R. Schmidt and T. Chainer, “Impact of Operating Conditions on a Chiller-less Data Center Test Facility with Liquid Cooled Servers”, ITHERM 2012, May 30th -June 1st, San Diego, CA.
- [7]. M. David, M. Iyengar, P. Parida, R. Simons, M. Schultz, M. Gaynes, R. Schmidt and T. Chainer, “Experimental Investigation of Water Cooled Server Microprocessors and Memory Devices in an Energy Efficient Chiller-less Data Center”, SEMITHERM 2012, March 18-22nd, San Jose CA.
- [8]. M. Iyengar, M. David, P. Parida, V. Kamath, B. Kochuparambil, D. Graybill, M. Schultz, M. Gaynes, R. Simons, R. Schmidt and T. Chainer, “Server Liquid Cooling with Chiller-less Data Center Design to Enable Significant Energy Savings”, SEMITHERM 2012, March 18-22nd, San Jose CA.
- [9]. P. Parida, M. David, M. Iyengar, M. Schultz, M. Gaynes, V. Kamath, B. Kochuparambil, and T. Chainer, “Experimental Characterization of an Energy Efficient Chiller-less Data Center Test Facility with Warm Water Cooled Servers”, SEMITHERM 2012, March 18-22nd, San Jose CA.
- [10]. M. David, T. Chainer, H. Dang, M. Gaynes, D. Graybill, M. Iyengar, V. Kamath, B. Kochuparambil, P. Parida, S. Rosato, R. Schmidt, M. Schultz, A. Sharma, R. Simons, “Characterization and Operational Results of the DOE Energy Efficient Data Center Test Facility”, IBM Presentation to DOE, Poughkeepsie, NY, July 19, 2012.
- [11]. M. Iyengar, R. Schmidt and J. Caricari, “Reducing energy usage in data centers through control of Room Air Conditioning units”, ITHERM 2010, June 2-5, Las Vegas, NV.
- [12]. ASHRAE book, “Best Practices for Datacom Facility Energy Efficiency”, Second Edition.
- [13]. Lawrence Berkeley National Labs, 2007, Benchmarking Data Centers – Charts”, <http://hightech.lbl.gov/benchmarking-dc-charts.html>.
- [14]. W. Tschudi, 2006, “Best Practices Identified Through Benchmarking Data Centers,” Presentation at the ASHRAE Summer Conference, Quebec City, Canada, June.
- [15]. US EPA, 2007, “Report to Congress on Server and Data Center Energy Efficiency”, Public Law 109-431, U.S. Environmental Protection Agency, ENERGY STAR Program.
- [16]. M. Ellsworth, L. Campbell, R. Simons, M. Iyengar, R. Chu, and R. Schmidt, 2008, “*The Evolution of Water Cooling for IBM Large Server Systems: Back to the Future*”, Proc. of the IEEE ITHERM Conference in Orlando, USA, May.
- [17]. R. Schmidt, M. Iyengar, D. Porter, G. Weber, D. Graybill, and J. Steffes, 2010, “*Open Side*

- Car Heat Exchanger that Removes Entire Server Heat Load Without any Added Fan Power*", Proceedings of the IEEE ITherm Conference, Las Vegas, June.
- [18]. R. Chu, M. Iyengar, V. Kamath, and R. Schmidt, 2010, "*Energy Efficient Apparatus and Method for Cooling an Electronics Rack*", US Patent 7791882 B2.
 - [19]. ASHRAE, "Thermal Guidelines for Data Processing Equipment – second edition", 2009, available from <http://tc99.ashraetcs.org/>
 - [20]. Y.Martin, T.Van Kessel, "High Performance Liquid Metal Thermal Interface for Large Volume Production", IMAPS Thermal and Power Management, San Jose CA, Nov.11-15 2007.
 - [21]. Typical year hour by hour weather data available on website of the US National Renewable Energy Lab (NREL).
 - [22]. ASHRAE TC 9.9, "2011 Thermal Guidelines for Data Processing Environments – Expanded Data Center Classes and Usage Guidance", 2011, available from http://www.eni.com/green-data-center/it_IT/static/pdf/ASHRAE_1.pdf
 - [23]. <http://www.electricchoice.com/electricity-prices-by-state.php>