

HOLISTIC AND ENERGY-EFFICIENT MANAGEMENT OF DATACENTRES

By

Yaser Talal Saber Al-Anii

**Submitted in accordance with the requirements for the degree of
Doctor of Philosophy**

Institute of ThermoFluids (iTf), School of Mechanical Engineering

The University of Leeds

March 2017

The candidate confirms that the work submitted is his own, except where work formed jointly-authored publication has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated overleaf. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

This copy has been supplied on the understanding that is copyrighted material and that no quotation from the thesis may be published without proper acknowledgment.

© 2017 The University of Leeds and Yaser Al-Anii

[This page is left blank intentionally]

WORK FORMED FROM JOINTLY AUTHORED PUBLICATIONS

The candidate has publications from work contained in this thesis. These were submitted with the support of supervisors and colleagues. The contribution of the candidate to these works is explicitly mentioned below.

1. Part of chapter 4 of this thesis presented in the proceedings of the 14th International Conference on Simulation and Experiments in Heat Transfer and its Applications (HT 2016). The paper is jointly authored by Y. Al-Anii, A. Almaneea, N. Kapur, J. Summers and H. Thompson. The title of the paper is “Effect of Processor Layout on the Thermal Performance of Fully Immersed Liquid-Cooled Microelectronics”. It is available as open access on WIT Transactions on Engineering Sciences, Vol 106, ISSN 1743-3533, doi:10.2495/HT160131.
2. Parts of chapter 3 and chapter 5 of this thesis presented in 9th IEEE/ACM International Conference on Utility and Cloud Computing (UCC 2016), Shanghai, China. The paper is jointly authored by Peter Garraghan, Yaser Al-Anii, Jon Summers, Harvey Thompson, Nik Kapur and Karim Djemame. The title of the paper is “A Unified Model for Holistic Power Usage in Cloud Datacenter Servers”. It is available: ISBN 978-1-4503-4616-0/16/12, doi:10.1145/2996890.2996896.
3. Part of chapter 3, chapter 5 and chapter 6 of this thesis presented as on March 2017 at the SEMI-THERM 33, California, USA. The paper is jointly authored by Kadhim, M., Al-Anii, Y., Summers, J., Kapur, N. and Thompson, H.. The title of the paper is “Thermal Performance of a Mixed Mode Air Handling Unit for Direct Liquid Cooled Servers”.

The candidate has collaborated with the co-authors as appears in the published papers. In addition, the supervisors provided valuable review, contribution and guidance to the candidate and his colleagues.

ACKNOWLEDGEMENT

First and foremost, all praise to Almighty Allah, for endowing me with health, patience, and knowledge to complete this stage of my life. Furthermore, I would like to acknowledge my supervisors, Dr Jonathan L. Summers, Professor Nik Kapur and Professor Harvey M. Thompson for their guidance and patience through my PhD journey. I am very grateful for their endless collaboration and support. They offered me valuable opportunities and continuously kept faith in me. I feel very honoured to have had the chance to work with them.

I am very grateful and thankful to the Higher Committee for Education Development (HCED) in Iraq for the fully funded scholarship. I also acknowledge the University of Anbar in Iraq for their financial support. Furthermore, I would like to send my thanks to Dr Kadhim Ahmed for his continuous support and encouragement.

In addition, I would like to acknowledge CoolIT Company and Mr Howard Bateman for providing the liquid cooling system and technical support. Also, I would like to acknowledge Airedale Company for supplying the air handling unit.

I would like to express my appreciation to my family back home. I don't have enough words to thank them for supporting me in all possible ways. They gave me a lot, but most especially, they gave me the strength to go on. I hope I made them proud.

Last but not least, I am very grateful and thankful to my young family. I thank my lovely daughters Balsam and Samar for making my life so fulfilled and making every day so special. I dedicate my final appreciation and thankfulness to the love of my life, my beloved wife Farah. Without her kindness, sacrifices, and patience I would not have been able to pursue my dreams. I truly appreciate her standing by my side providing me support, attention and love.

ABSTRACT

The overall power consumption of datacentres is increasing tremendously due to the high demand of digital services. Moreover, the cooling load contributes up to 50% of the power consumption due to the higher densities of newer versions of servers. However, there is an increased awareness in the operations of the sub-systems, i.e. workload, cooling load and power consumption. This awareness of the interactions between the sub-systems provides a better understanding for maintaining the datacentre as an energy-efficient infrastructure.

A direct contact liquid cooling technology is examined extensively by retrofitting to an air-cooled server. First the conventional SunFire V20z air-cooled server is benchmarked via SPECpower_ssj2008 workload to obtain some standard values. The server is placed inside a wind tunnel to ensure a controllable environment. Then an overall evaluation of the retrofitted server is presented and compared with the standard server. The retrofitted server shows a reduced cooling power consumption of 29%. In addition, the performance to power ratio increases by 10% comparing to the conventional server. The liquid cooling technology keeps the central processes units (CPUs) up to 10 °C colder than the air-cooled server. Furthermore, the new server operates in an 88% lower noise after the replacement of four fans by two pumps. However, the main restriction of using such a solution is the risk of bringing water into the microelectronics due to leakage and condensation of water.

A fully immersed encapsulated server is then investigated to assess the validity of simulating the immersed server as a porous layer. This simulation uses Darcy flow with mass, momentum and energy conservation equations. The model shows a quantitative and qualitative accuracy compared to the previous work. The model shows that the distance between processors has a strong effect on the thermal behaviour of the encapsulated server by 13.3% compared to servers' dimensions. Moreover, the model presents the optimal design and geometry of an encapsulated server with respect to the thermal performance. Although the model is simple, it can be used for an

initial prediction of the server design. This is due to the limitation of capturing the thermal behaviour of a full model.

A holistic power consumption model is presented to capture the interactive relationships between servers' sub-system. The power model relies on experimental work and is constructed based on the collected data from different cooling configurations. The model captures a detailed breakdown of the power consumption and therefore presents an accurate calculation of the partial power usage effectiveness metric. The results are limited to one microelectronic architecture within a specific IT load type. However, the results show that reducing the cooling load by 7% and increasing the performance by 5% leads to lower the partial power usage effectiveness by 1%.

Finally, the current study explores the usage of an evaporative air handling unit for energy-efficient datacentres. The air handling unit is capable of run dry and wet cooling operation. The cooling system operated successfully during July and August 2016, in Leeds. The wet cooling has a higher thermal performance than the dry cooler due to the large heat capacity of water compared to air. Therefore, the wet cooling configuration records a power usage effectiveness lower than the dry cooling by about 6.4%.

TABLE OF CONTENTS

WORK FORMED FROM JOINTLY AUTHORED PUBLICATIONS	i
ACKNOWLEDGEMENT.....	ii
ABSTRACT	iii
TABLE OF CONTENTS	v
LIST OF FIGURES.....	ix
LIST OF TABLES.....	xv
LIST OF ABBREVIATIONS	xvi
Chapter 1 INTRODUCTION.....	1
1.1 What is a Datacentre?.....	1
1.2 Thermal Management of Datacentres	2
1.3 Energy-Efficient Datacentre	3
1.4 Aim of the Current Study	5
1.5 The Contribution of the Current Study.....	8
1.6 The Terminology of the Current Study	8
1.7 The Outline of the Thesis	10
Chapter 2 LITERATURE REVIEW	12
2.1 Introduction	12
2.2 Cooling and Power Technology of Microelectronics.....	13
2.2.1 Cooling Technology of Microelectronics	14
2.2.2 Power Consumption of Microelectronics	16
2.3 Energy-Efficient Datacentre	18
2.3.1 Efficiency Metrics	19
2.3.2 Cooling Solutions of Datacentres.....	23
2.3.3 Total Datacentre Power Consumption	28
2.4 Holistic Management of Datacentre	32
2.4.1 Thermal Management.....	32
2.4.2 Datacom Solutions and Management	36
2.4.3 Integrated Management.....	38
2.5 Summary.....	41
Chapter 3 TOWARDS DIRECT LIQUID COOLING (DLC)	48
3.1 Introduction	48

3.2	Total Air Cooling (TAC) of Server.....	50
3.2.1	SunFire V20z Server.....	50
3.2.2	AMD Opteron.....	52
3.2.3	Intelligent Platform Management Interface (IPMI)	53
3.2.4	Operating System	54
3.2.5	Cooling systems.....	54
3.3	Direct Liquid Cooling (DLC) of Server	56
3.3.1	CoolIT™ Pump	57
3.3.2	The Tubing.....	60
3.3.3	The Coolant	60
3.3.4	The External Fans and Radiator	60
3.3.5	Stäubli Connectors.....	61
3.4	Preliminary Test of the Server	62
3.4.1	Fan Testing.....	62
3.4.2	Server Testing.....	65
3.4.2.1	Prepare Stress Linux Test	65
3.4.2.2	Results and Discussion of Air-Cooled and DCLC Server.....	69
3.5	Experimental Work.....	73
3.5.1	SPECpower_ssj2008	73
3.5.2	Results and Discussion.....	76
3.5.2.1	Benchmarking the SunFire V20z.....	76
3.5.2.2	Evaluation of Cooling Technologies	84
3.6	Numerical Simulation	88
3.6.1	Future Facilities – 6SigmaET	88
3.6.1.1	Governing Equations.....	90
3.6.1.2	The Computational Domain.....	90
3.6.1.3	The Turbulence Model	91
3.6.2	Building the Model	93
3.6.3	Run the Simulation.....	95
3.6.4	Results and Discussion.....	97
3.7	Summary.....	98
Chapter 4 TOWARDS TOTAL LIQUID COOLING (TLC)	100
4.1	Introduction	100
4.2	Modelling of Total Liquid Cooling	103

4.2.1	Novelty of the Model	103
4.2.2	Mathematical Formulation.....	105
4.2.3	Numerical Solution.....	106
4.2.3.1	Dimensionless Formulation	107
4.2.3.2	Solution Techniques.....	108
4.2.3.3	Parametric Criteria	109
4.2.3.4	Boundary and Initial Conditions.....	111
4.2.3.5	Mesh	113
4.2.3.6	Code Validation	114
4.3	Modelling of a 1U Encapsulated Server	117
4.3.1	The Parametric Study	117
4.3.2	Results and Discussion.....	117
4.4	Design Optimisation of the Encapsulated Server	124
4.4.1	The Design of Experiment	124
4.4.2	Optimum Design of the Server.....	124
4.4.3	Results and Discussion.....	126
4.5	Summary.....	130
Chapter 5 THE HOLISTIC POWER CONSUMPTION MODEL	132	
5.1	Introduction	132
5.2	Experimental Work.....	133
5.2.1	The Test Rig	133
5.2.2	Initial Results.....	134
5.2.3	Parametric Study	135
5.3	The Unified Power Model.....	137
5.4	The Power Usage Effectiveness PUE	140
5.5	Model Validation.....	141
5.6	Results and Discussion.....	143
5.7	Summary.....	145
Chapter 6 ENERGY-EFFICIENT DATACENTRE.....	146	
6.1	Introduction	146
6.2	Building the Rack	146
6.3	The Rack Cooling System.....	149
6.4	The Air Handling Unit (AHU)	153
6.5	Analysis of Energy-Efficient Rack-Level.....	153
6.6	Harvesting the Rejected Heat	160

6.7	Summary.....	161
Chapter 7 CONCLUSIONS AND RECOMMENDATIONS.....	162	
7.1	Introduction	162
7.2	Concluding Remarks.....	163
7.3	Recommendation for Future Work	165
References	167	
Appendix I: SPECpower_ssj2008 Reports	175	
Air-cooled SUT Report.....	175	
DCLC SUT Report.....	179	
Appendix II: Building the 6SigmaET Models	183	
Chamber Test.....	183	
The Server Chassis	183	
The Vents	184	
The Printed Circuit Board PCB	184	
The Chip Socket	185	
The Fans	186	
The Memory Modules.....	187	
The Disk Drive	187	
The Power Supply	188	
Obstructions	188	
The Cold Plates	189	

LIST OF FIGURES

Figure 1.1 Typical datacentre layout (Joshi and Kumar, 2012).....	2
Figure 1.2 Environmental classes of datacentres (ASHRAE, 2012).	3
Figure 1.3 Datacentres energy-efficiency scenarios (Shehabi et al., 2016).	4
Figure 1.4 Illustration of air-cooled datacentre (TrippLite, 2012).	5
Figure 1.5 The interactive elements of a holistic management.	6
Figure 1.6 The power consumption terminology of the current work. ..	10
Figure 1.7 The thesis outline.	11
Figure 2.1 The three levels of literature review researches.	13
Figure 2.2 Inlet temperature effect on the server's performance, P_{IT} and P_{sfan} represents the power consumption of IT and fans, respectively (Ham et al., 2015).	15
Figure 2.3 Temperature and power maps of dual core microprocessor A: Thermal and B: Power. Actual photo size is 13.2x11.6 mm ² (Hamann et al., 2006a)	16
Figure 2.4 Comparison of measured temperature vs simulation prediction (Farkhani and Mohammadi, 2010)	17
Figure 2.5 TGI vs. processes based on power and energy (Subramaniam and Feng, 2012)	21
Figure 2.6 Sankey diagram for (a) typical datacentre and (b) fully liquid-cooled datacentre (Girimella et al., 2013).....	22
Figure 2.7 Dual-enclosure-liquid cooling DELC system (Iyengar et al., 2012).	24
Figure 2.8 Air/warm water technique (Iceotope, 2013).	25
Figure 2.9 Liquid-cooled technology for rack level of datacentre (Chi et al., 2014).....	26
Figure 2.10 PUE dependency on server occupancy (Almaneea et al., 2014)	27
Figure 2.11 Model estimation power usage (a) servers level and (b) racks level (Fan et al., 2007).....	29
Figure 2.12 Proactive control strategy effect on cooling power consumption (Tarutani et al., 2015)	31
Figure 2.13 Types of the distribution system (Cho et al., 2009).....	33
Figure 2.14 Actual vs predicted power consumption (Singh et al., 2009).	37

Figure 2.15 Reduced the temperature of IT nodes when using TASA model (Wang et al., 2009)	37
Figure 2.16 Scatter-results of predicted versus real values (Moore et al., 2006).	39
Figure 2.17 Control system structure (Zhou et al., 2011).....	40
Figure 3.1 Heat load density rate (ASHRAE, 2005, Joshi and Kumar, 2012).	48
Figure 3.2 Appropriate cooling technology for various power density (Minas and Ellison, 2009).	49
Figure 3.3 Illustration of SunFire V20z server (Sun Microsystems, 2008)	51
Figure 3.4 The SunFire V20z server.....	52
Figure 3.5 Screenshot of IPMI window from the control unit.....	53
Figure 3.6 Illustration of fan locations inside the server.....	55
Figure 3.7 Cooling fans, heat sinks and baffle for CPUs.	55
Figure 3.8 Direct contact liquid cooling DCLC (CoolIT, 2013)	57
Figure 3.9 The retrofitted Sun Fire V20z Server including DCLC.	57
Figure 3.10 Schematic diagrams of the pump (CoolIT, 2012).	58
Figure 3.11 Operation curve of the pump (CoolIT, 2012).	59
Figure 3.12 The retrofitted DCLC system by CoolIT.....	59
Figure 3.13 Dimensions of the external fan (CoolIT, 2011).	61
Figure 3.14 The fan-radiator cooling system.	61
Figure 3.15 The iParaAiluRy digital tachometer.....	63
Figure 3.16 Agreements of the measured speed and captured speed by IPMI tools.....	64
Figure 3.17 Fan power measurement versus rotational speed	65
Figure 3.18 The network configuration.	66
Figure 3.19 Test rig of preliminary experimental stage.....	67
Figure 3.20 Repeated tests of the server under air and liquid cooling conditions.	67
Figure 3.21 Temperature monitoring during idle state for air-cooled server.	69
Figure 3.22 Temperature monitoring during idle state for retrofitted DCLC server.	70
Figure 3.23 Stress CPUs state for air cooled server.....	70
Figure 3.24 Stress VMs state for air cooled server.	71

Figure 3.25 Stress CPUs state for DCLC server.....	71
Figure 3.26 Stress VMs state for DCLC server.....	72
Figure 3.27 Temperature comparison of CPUs and VMs during different stress cases for air cooled server	72
Figure 3.28 Temperature comparison of CPUs and VMs during different stress cases for DCLC server.....	73
Figure 3.29 SPECpower_ssj2008 architecture (SPEC, 2012).	74
Figure 3.30 Illustration of the interval of action level (SPEC, 2012a)...	75
Figure 3.31 Illustration of wind tunnel.	76
Figure 3.32 Monitoring of power and processed transactions of SPECpower_ssj2008.....	77
Figure 3.33 Multi-level of performance of air-cooled SUT.....	78
Figure 3.34 SUT power monitoring at multi-level of utilization.	78
Figure 3.35 Average power consumption variation with utilization for air cooled SunFire V20z.	79
Figure 3.36 The matching of target and actual workload for air-cooled SunFire V20z.....	80
Figure 3.37 Performance to power consumption ratio for air cooled SunFire V20z.....	80
Figure 3.38 Monitoring of power and processed transactions of SPECpower_ssj2008 of DCLC server.....	81
Figure 3.39 Multi-level of performance of DCLC SUT.....	81
Figure 3.40 Retrofitted DCLC SUT power monitoring.....	81
Figure 3.41 Average power consumption variation with utilization of DCLC SunFire V20z.	82
Figure 3.42 The matching of target and actual workload for DCLC SunFire V20z.....	83
Figure 3.43 Performance to power consumption ratio for air cooled SunFire V20z.....	83
Figure 3.44 CPU0 temperature monitoring for air-cooled and DCLC SUT.	85
Figure 3.45 CPU1 temperature monitoring for air-cooled and DCLC SUT.	85
Figure 3.46 Monitoring of RAMs temperature of the SUTs.	86
Figure 3.47 The temperature of SP during benchmarking.	87
Figure 3.48 Outlet temperature form the SUT.	88
Figure 3.49 Differences between inlet and outlet temperatures for both SUTs.	88

Figure 3.50 Future Facilities scale of simulations (FutureFacilities, 2015a)	89
Figure 3.51 Illustration of an element in FVM (Versteeg and Malalasekera, 2007).....	91
Figure 3.52 Velocity fluctuations for turbulent regime (Versteeg and Malalasekera, 2007).....	92
Figure 3.53 The TAC server model.....	96
Figure 3.54 The DCLC server model.....	96
Figure 3.55 Comparisons of power consumption of the air-cooled SUT.....	97
Figure 3.56 Comparisons of power consumption of the DCLC SUT....	98
Figure 4.1 Different capability of heat transfer coefficients (Murshed, 2016)	100
Figure 4.2 Illustration of Iceotope module (BOSTON, 2009).	102
Figure 4.3 Schematic diagram of the physical domain of studied capsule.	111
Figure 4.4 Schematic diagram of the numerical domain of studied capsule (a) definitions and (b) boundary conditions.	112
Figure 4.5 The computational mesh (a) normal and (b) stretched toward the CPUs.	114
Figure 4.6 Mesh independency test.....	114
Figure 4.7 Validation of the governing equation used in the present code with (Bejan and Kraus, 2003)	115
Figure 4.8 Comparing the non-dimensional values of (a) stream function and (b) temperature lines, code vs. (Nithiarasu et al., 1997)	115
Figure 4.9 Compare lower and upper CPU's Nusselt numbers for the present code versus COMSOL.....	116
Figure 4.10 Effect of non-dimensional position of CPUs on ratios of non-dimensional velocity, at $S=1$ and $Ra^*=50$	118
Figure 4.11 Effect of non-dimensional position of CPUs on ratios of local Nusselt number, at $S=1$ and $Ra^*=50$	119
Figure 4.12 Effect of non-dimensional position of CPUs on ratios of non-dimensional velocity at $Ly=1$ and $Ra^*=50$	120
Figure 4.13 Effect of non-dimensional position of CPUs on ratios of local Nusselt number, at $Ly=1$ and $Ra^*=50$	120
Figure 4.14 Effect of in-between distance (S) on average Nusselt number for $Ly=1, 2$ and 3 , at $Ra^*=50$	121
Figure 4.15 Streamlines (up) and isotherms (down) for $Ly=S=1U$ at Ra^* of (a) 10, (b) 50, (c) 100 and (d) 300.	122

Figure 4.16 Average Nusselt number vs. modified Rayleigh number for all spatial ratios.....	123
Figure 4.17 Building and validation points distribution.....	125
Figure 4.18 The fluctuation of Nusselt number of both lower and upper CPUs.....	126
Figure 4.19 The fluctuation of stream function values.	126
Figure 4.20 Nusselt number of the lower CPU response surfaces for Lx with respect to (a) Ly1, (b) Ly2 and (c) S.....	128
Figure 4.21 Nusselt number of the upper CPU response surfaces for Lx with respect to (a) Ly1, (b) Ly2 and (c) S.....	129
Figure 4.22 The response surface of stream function for Lx with respect to (a) Ly1, (b) Ly2 and (c) S.....	130
 Figure 5.1 Architecture for measurement extraction.....	134
Figure 5.2 The effect of utilization level on power and CPU temperature.	135
Figure 5.3 Power consumption of the air-cooled server at various inlet air temperature.....	136
Figure 5.4 Power consumption of the DCLC server at various inlet water temperature.	137
Figure 5.5 Illustration of the breakdown of power consumption.....	140
Figure 5.6 The validation of power consumption measurements for air-cooled server.....	142
Figure 5.7 The validation of power consumption measurements for DCLC server.	142
Figure 5.8 Total power breakdown for SUT	143
Figure 5.9 Power usage of liquid cooled server sub-systems at utilization (a) 20%, (b) 60%, (c) 100%.	144
 Figure 6.1 The 45U rack during early installation stage, Lab 1.11, 2013.	147
Figure 6.2 Part of the 30 servers with the retrofitted DCLC, Lab1.11, 2013.	147
Figure 6.3 Back of the rack showing the networking cables and the coolant tubing, Lab 1.11, 2014.	148
Figure 6.4 The air heat exchanger AHU, Lab 1.11, 2015.	149
Figure 6.5 The coolant heat exchanger CHx, Lab 1.11, 2013.	150
Figure 6.6 Illustration of CHx40 (CoolIT, 2015).....	151

Figure 6.7 Effect of water inlet temperature on the cooling load (CoolIT, 2015)	151
Figure 6.8 The cloud computing rack, Lab 1.11, 2016.....	152
Figure 6.9 Schematic diagram of the connection between the rack and the AHU.....	154
Figure 6.10 Temperature records for the city of Leeds, UK (MetOffice, 2016)	155
Figure 6.11 The monitoring of temperatures at different points for a virtual workload on the rack.	155
Figure 6.12 The monitoring of temperatures at different points for a stable workload on the rack.	156
Figure 6.13 Heat transfer balance between the CHx and AHU.	157
Figure 6.14 Effectiveness calculation of the CHx.....	158
Figure 6.15 Effectiveness calculation of the AHU.....	159
Figure 6.16 Power consumption breakdown and PUE	160

LIST OF TABLES

Table 2.1 Summary of Literature Review.....	43
Table 3.1 Physical Specifications (SunMicrosystems, 2008)	50
Table 3.2 Features of SunFire V20z server (SunMicrosystems, 2008)	51
Table 3.3 The coding system of the AMD Opteron CPU (AMD, 2006).	52
Table 3.4 Specifications and coding system of the Delta FFB0412SHN fan	56
Table 3.5 Specifications of iParaAiluRy digital tachometer.....	63
Table 3.6 Parameters prediction for equation (3.2).....	64
Table 3.7 Stress command line form	68
Table 3.8 The cases used in the preliminary stage of experiments.	68
Table 3.9 Linux command line for temperature and fan speed monitoring using the IPMI.....	84
Table 3.10 Thermal design power for the main components (AMD, 2006, Pandiyan, 2013).	95
Table 4.1 The boundary conditions.....	112
Table 4.2 DoE for the first model of encapsulated server.	117
Table 4.3 DoE for the optimum design of the encapsulated server. ..	124
Table 5.1 Experimental parametric study for the DCLC server.....	135
Table 5. 2 Empirical constants for the power formula.	138

LIST OF ABBREVIATIONS

Nomenclatures

A	Area [m ²]
a ₁ , a ₂ and a ₃	Correlation parameters (used in equation (3.2))
b ₁ , b ₂ and b ₃	Correlation parameters (used in equations (5.1) and (5.2))
D	Characteristic length of CPU
Da	Darcy number
e	Effectiveness of heat exchanger
f ₁ and f ₂	Algebraic parameters (used in equation (4.16))
f ₃ , f ₄ and f ₅	Algebraic parameters (used in equation (4.17))
h	Convective heat transfer coefficient [W/m ² .°C]
i and i _{max}	Iteration identifier (used in equation (3.3))
I	Current [A]
k	Thermal conductivity [W/m.°C]
k	Turbulent kinetic energy [J/kg or m ² /s ²]
K	Permeability [m ²]
L _x	Non-dimensional width of the immersed server
L _y	Non-dimensional overall height of the immersed server
L _{y1}	Non-dimensional lower distance of the immersed server
L _{y2}	Non-dimensional upper distance of the immersed server
N	Fan speed [rpm]
Nu	Nusselt number
r _i	Convergence factor (used in equation (3.3))
Q _{conv}	Convective heat transfer [W]
Q _{cond}	Conductive heat transfer [W]
p	Local pressure or pressure drop [Pa]
P _{base}	Power consumption at active idle [W]
P _{cooling}	Power consumption of all cooling equipment [W]
P _{dynamic}	Dynamic power consumption due to utilization [W]
P _{external}	Power consumption of the cooling outside the server [W]
P _{fan}	power consumption of fans [W]
P _{IT}	Power consumption due to IT performance [W]
P _{PSU}	Power through power supply unit [W]

P_{pump}	Power consumption of pumps [W]
P_{static}	Static power consumption due to temperature rise [W]
P_{server}	Power consumption of server [W]
P_T or P_{total}	Total power consumption [W]
R	Variance factor
Ra	Rayleigh number
Ra^*	Modified Rayleigh number
RH	Relative humidity [%]
S	Non-dimensional in-between CPUs distance
t	Time [s]
T	Temperature [$^{\circ}\text{C}$]
T_{amb} or T_{∞}	Ambient temperature [$^{\circ}\text{C}$]
T_s	Surface temperature [$^{\circ}\text{C}$]
$u(t)$ and $u'(t)$	Total and turbulent instantaneous velocity [m/s]
U	Average velocity [m/s]
U	Utilisation [%]
V	Voltage [V]
x	Independent function (used in equation (3.3))

Greek Symbols

α	Thermal diffusivity [m^2/s]
β	Thermal expansion coefficient [1/K]
Γ	Stretching factor
Δ	Difference
ε	Turbulence dissipation [m^2/s^3]
ζ	Minimizer function (used in equation (3.3))
η	Efficiency [%]
θ	Non-dimensional temperature
μ	Viscosity [Pa.s]
μ_t	Turbulent viscosity [Pa.s]
ρ	Density [kg/m^3]
τ	Non-dimensional time
ψ	Stream-function
Ψ	Non-dimensional stream-function

Acronyms

1U	One rack unit (which equals to 44.45mm)
2D	Two dimensional
3D	Three dimensional
ABS	Acrylonitrile Butadiene Styrene
AC	Alternate current
AHU	Air handling unit
AHx	Air heat exchanger
AMD	Advanced micro devices
ASHRAE	American Society of Heating, Refrigeration and Air-Conditioning Engineers
BMC	Baseboard management controller
CCR	Center for Computational Research
CCS	Control and collect system
CFD	Computational fluid dynamic
CHx	Coolant heat exchanger
CMP	Chip multi-processor
COP	Coefficient of performance
CPU(s)	Central processing unit(s)
CRAC	Computer room air-conditioning
CRAH	Computer room air handling
CTFM	Compact thermal-fluid models
D	Datacentre-level (used in Table 2.1)
DC	Direct current
DC(s)	Datacentre(s)
DCEPM	Data center energy performance (productivity) metric
DCLC	Direct contact liquid cooling
DELC	Dual-enclosure liquid cooling
DHCP	Dynamic Host Configuration Protocol
DIMM(s)	Dual in-line memory module(s)
DLC	Direct liquid cooling
DoE	Design of experiment
DTM	Dynamic thermal management
ECM	Electronically commutated motor

EE	Energy efficiency
ENIAC	Electronic numerator, integrator, analyser and computer
EPA	Environmental Protection Agency
FDM	Finite difference method
FEM	Finite element method
FEP	Fluorinated Ethylene Propylene
FLOPS	Floating operation per second
FVM	Finite volume method
GB	Giga byte
HDD(s)	Hard disk drive(s)
HPC	High-performance computing
ICT	Information and communication technology
I/O	Input/output
IP	Internet protocol
IPMI	Intelligent platform management interface
IR	Infra-red
IT	Information technology
MAC	Media access control
MB	Mega byte
MIB	Management information base
MLS	Moving least square
MSEM	Multi-scale energy management
NTU	Number of transfer unit
OEM	Original equipment manufacture
OS(s)	Operating system(s)
PBT	Poly Butylene Terephthalate
PCB	Printed circuit board
PPS	Poly Phenylene Sulphides
pPUE	Partial power usage effectiveness
PSU(s)	Power supply unit(s)
PTDaemon	Power and Temperature Daemon
PUE	Power usage effectiveness
QoS	Quality of service
R	Rack-level (used in Table 2.1)
RAM	Random access memory

REE	Relative energy efficiency
RHI	Return heat index
rpm	Revolutions per minute
S	Server-level (used in Table 2.1)
SaaS	Software as a service
SERT	Server efficiency rating tool
SEPM	Server energy performance (productivity) metric
SHI	Supply heat index
SIMP	Spatially-resolved imaging of microprocessor power
SLA(s)	Service level agreement(s)
SNMP	Simple network management protocol
SOR	Successive over relaxation
SP	Service processor
SPEC	Standard performance evaluation corporation
SSH	Secure shell
SSJ	Server side Java
SSL	Secure sockets layer
SUT	Server Under Test
TAC	Total Air Cooling
TASA	Thermal aware scheduling algorithm
TASP	Thermal aware server provisioning
TAWD	Thermal aware workload distribution
TDP	Thermal design power
TGI	The green index
TIM	Thermal interface material
TLC	Total Liquid Cooling
VM(s)	Virtual machine(s)
WUE	Water utilization effectiveness

Chapter 1 INTRODUCTION

1.1 What is a Datacentre?

Datacentres represent the 21st century industrial factories requiring significant amounts of energy to give high-quality digital serviceability. The information and communication technology (ICT) demands to maximise the requesting of multi-level datasets. Nowadays, data sets are involved widely in a variety of everyday activity. For example, telecommunication, financial market, transportation control and social media. These modern digital enterprises require powerful critical mission infrastructure, i.e. datacentres (Minas and Ellison, 2009).

The datacentre requires being managed as unique infrastructure due to the interactions of response between its different principal components. The main components of each datacentre are power supply units (PSUs), information technology (IT) microelectronics and cooling systems (Joshi and Kumar, 2012) as shown in Figure 1.1. These components require a balance in operation between the performance and the energy consumption.

Furthermore, the datacentre itself needs to sustain high-performance operation through availability, reliability and resilience (Curtis, 2011). This sustainability is based on multiple heterogeneous microelectronics which play a vital role in provisioning high-quality of services. The highest performance means maximizing the number of transactions / processes / operations per unit time. For complex scientific calculations, it is best defined as Floating Operation Per Second and measured in millions or billions, MFLOPS or GFLOPS, respectively.

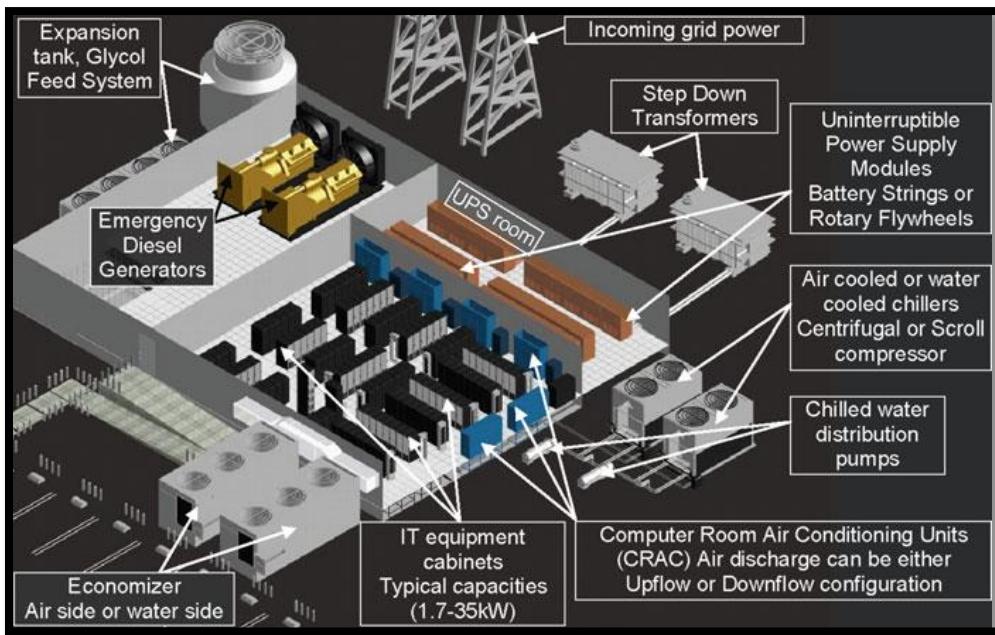


Figure 1.1 Typical datacentre layout (Joshi and Kumar, 2012)

1.2 Thermal Management of Datacentres

Running higher frequency microelectronics requires huge amounts of energy, which produces huge amounts of heat dissipation as a by-product correspondingly. It is, therefore, imperative to overcome this amount of heat by providing efficient heat sinks. The heat sink maintains the temperature at allowable operating limits (Tong, 2011). For that, many techniques are available to absorb heat and keep the temperature of the servers and their central processing units (CPUs) within the operating envelope characterised as recommended by ASHRAE standards (Hewlett-Packard, 2012, ASHRAE, 2012).

It is crucial for designers and engineers to specify the environmental envelope and define the class, or tier, of the datacentre. ASHRAE categorises four classes depending on the entire air conditioning of datacentres as shown in Figure 1.2. The narrowest range of inlet temperature to the servers is 15 to 32 °C for the A1 class. This class is applicable for the *tightly* controlled environment. However, for some control levels of the environment inside datacentres, A2, A3 and A4 have a wider range of inlet temperature. Class A2 has a wider range of inlet temperature

of 10 to 35 °C, whereas the A3 and A4 have an even wider range of 5 to 40 °C and 5 and 45 °C, respectively. The four classes share the same relative humidity (RH) and maximum dew point of 8 to 80% and 27 °C, respectively (ASHRAE, 2012).

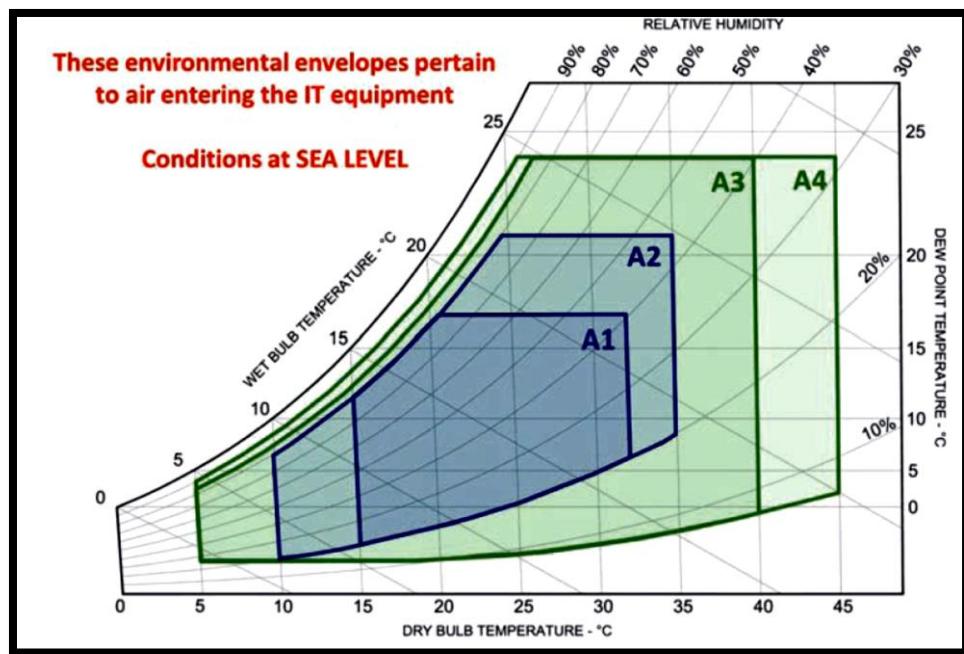


Figure 1.2 Environmental classes of datacentres (ASHRAE, 2012).

During the last decade, many investigations were interested in understanding the power consumption inside datacentres. The majority of the researches presented a breakdown of power consumption measurement to its main elements (Alger, 2009). One of the interesting conclusions is that cooling load consumed between 20 to 50% of the overall power consumption (Joshi and Kumar, 2012). Therefore, the selecting of efficient cooling technology leads to obtaining better thermal management and cuts the energy bill of the datacentre.

1.3 Energy-Efficient Datacentre

As mentioned earlier, the digital services affect, by different means, the global demand for power. For that, the multi-level components efficiency plays a very important role in this vital issue. Moreover, Moore's Law

(Moore, 1965) predicted the revolutionary progress in integrated circuits and electronic devices. Thus, more cooling load is required to overcome the huge amount of heat generated within datacentre infrastructure.

In 2010, datacentres consumed about 1.3% of electricity globally (Koomey, 2011). Although this figure represents a high percentage, it was significantly lower by 15% than predicted in the Environmental Protection Agency (EPA) report (Brown, 2008).

Recently, the power consumption is predicted to increase at 3% annual growth rate through 2020 (Shehabi et al., 2016). The reason behind this prediction coincides with the usage of efficient hyperscale datacentres and apply new saving technology at different levels as shown in Figure 1.3. However, the report did not consider many vital related issues such as; the internet of things (IoT), edge computing and micro-datacentres, etc.

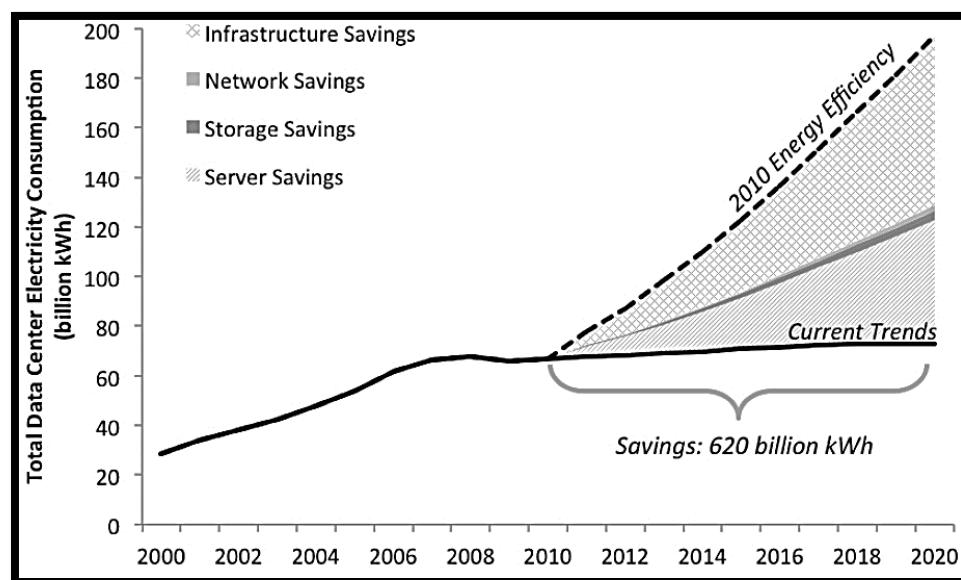


Figure 1.3 Datacentres energy-efficiency scenarios (Shehabi et al., 2016).

As a result, global interdisciplinary efforts are focusing on increasing efficiency of datacentres from many different perspectives, such as thermal management, cloud computing and computing architectures.

Despite many researches that take place to develop local enhancements in components; such as CPU, there are few that study this issue holistically (Jiang et al., 2012). To develop an overall point of view requires a broad

range of skills to penetrate through common areas of research and full awareness of interdisciplinary aspects.

1.4 Aim of the Current Study

Datacentres can be more efficient with respect to the level of energy consumption through good understanding of their conditions of operation. For instance, the reduction in the cooling load by changing cooling technology helps in obtain an efficient datacentre. Moreover, investigating the interactions between the main elements inside datacentre, as shown in Figure 1.4, plays a major role in presenting an efficient datacentre.

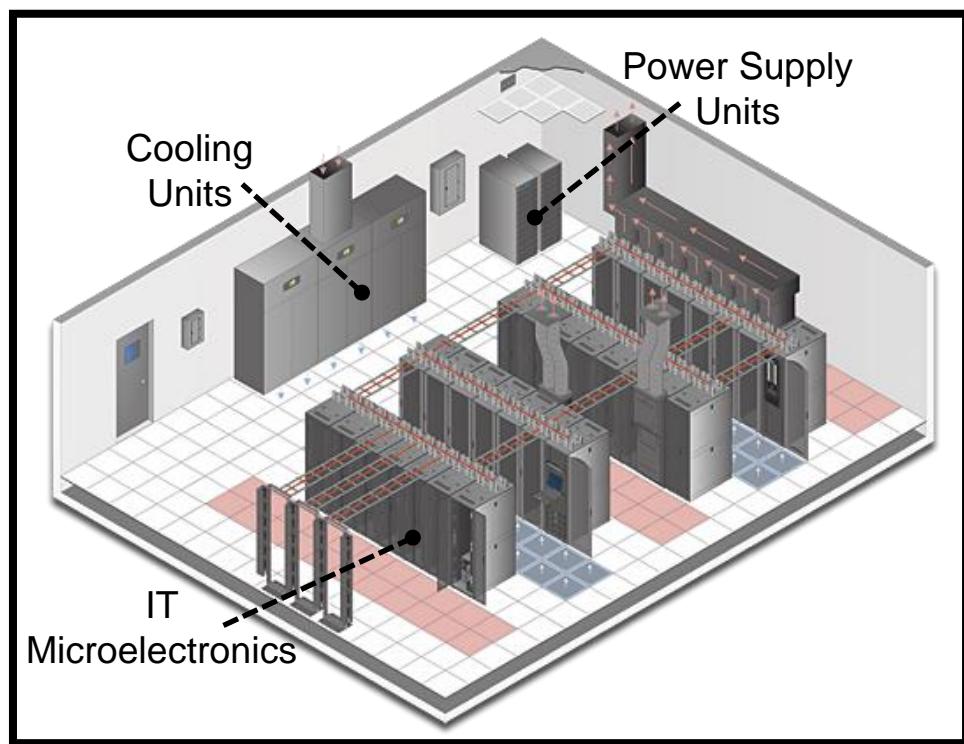


Figure 1.4 Illustration of air-cooled datacentre (TrippLite, 2012).

For that, the current study aims to investigate the impact of liquid cooling technologies, as one component of the overall datacentre, on the holistic management strategy as shown in Figure 1.5.

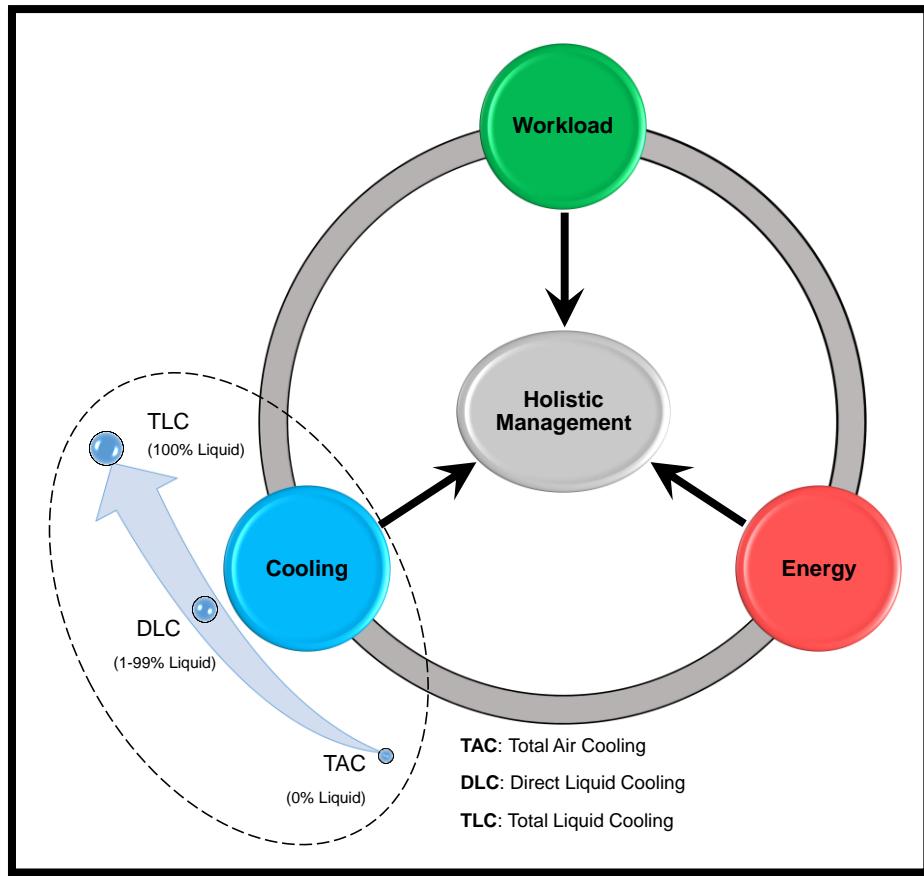


Figure 1.5 The interactive elements of a holistic management.

To achieve this aim there are six required objectives per the following stages:

1. Review the related researches, articles and patents to understand the prior work and to identify challenges with the area of research.
2. Prepare the platform:
 - a. Select the available microelectronic platform which is suitable for different digital services. It is cooled by air conventionally to play the role of server under test (SUT).
 - b. Select a benchmarking method and prepare the server to conduct benchmarking.
 - c. Evaluate the performance, power consumption and cooling effects on the SUT.
3. Based on the SUT, consider the impact of a direct cooling technology:
 - a. Retrofit the direct liquid cooling technology.

- b. Benchmark the server under similar conditions with the new cooling technology.
 - c. Compare the performance, power consumption and cooling effect on both SUTs.
4. Using immersed liquid cooling technology
 - a. Investigate the opportunities of using fully-immersed liquid cooling.
 - b. Present a novel approach by simulating the fully-immersed server as a porous cavity.
 - c. Present a parametric study and optimization for the design of the new generation of encapsulated server.
 5. Introduce a unique model for the power consumption of single server and manipulate the model to rack level.
 6. Examine the cooling load and power consumption at a rack level. This part requires a connection to an external air handling unit. The work includes an overall evaluation of energy efficient digital service.

Those objectives are achieved through experimental investigations and/or numerical simulations. The experimental parts consist of preparing the test rig and performing series of experiments. The experiments aim to monitoring the performance of the microelectronics and compare the impact of cooling technologies on the IT loads. Meanwhile, the energy and thermal balances across the units are calculated from the collected data from the sensors, such as temperature, flow rate, etc. This work is explained extensively in chapters 3, 5 and 6.

Simultaneously, the numerical simulations are based on Navier-Stokes equation which is presented in the form of a force balance as:

$$\vec{F}_I = \vec{F}_P + \vec{F}_V + \vec{F}_B \quad (1.1)$$

Where, F_I is the inertia force induced in the flow, and F_P , F_V and F_B represent the pressure, viscous and buoyancy forces, respectively.

The air and/or direct liquid cooling technologies are effective when the coolant flows at non-negligible values of velocity due to mechanical work

represented by fans and/or pumps. Therefore, the inertia force is in balance with pressure and viscous forces. This part of the numerical work is explained in chapter 3. In contrast, for the total liquid cooling, there is no use of mechanical work. The coolant movement is due to the effect of the induced buoyancy force. This buoyancy force is in balance with the viscous force only, which is illustrated extensively in chapter 4.

1.5 The Contribution of the Current Study

The current study evaluates the impact of using different cooling technologies, as shown in Figure 1.5. This evaluation helps in introduce a better awareness of the interactions between the subsystems of a datacentre. The contributions and novelty of the current study are listed below:

1. Evaluation of the performance of retrofitted liquid cooled servers.
2. A simple model of total liquid cooling.
3. A unified model of power consumption.
4. Evaluation of the power usage effectiveness of server and rack levels.

1.6 The Terminology of the Current Study

The current study deals with different level of components. Therefore, it is vital to present a clearly distinguished description of power terms of each level as shown in Figure 1.6.

1. The central processor unit (CPU) level:

The CPU represents the core of the datacentre where the IT processes/transactions/computations start to consume the required power. As a result, corresponding heat rate generated is rejected into the active coolant. The coolant is varied according to a specific cooling technology. Each CPU requires three main components of power consumption. First, the base-line power consumption (P_{base})

which is the minimum consumed power within the unit regardless of the processes/transactions/computations. Second, the dynamic power consumption (P_{dynamic}) which is related to the performed level of processes/transactions/computations, i.e. the utilisation. Third, the static power consumption (P_{static}) which is affected by the die temperature.

2. The server (or microelectronics) level:

The power consumed by the server includes all of the components of the CPU's level. In addition, there are (P_{fan}) and (P_{pump}) the fans' and the pumps' power consumption, respectively. All power components are supplied to the server through a power supply unit (PSU). Therefore, the total power (P_{total}) is based on the utilisation level, the CPU temperature, the cooling load of fans and/or pumps and the efficiency of the PSU. However, it is important to mention that the coolant, air and/or liquid, requires an external cooling power (P_{external}) to maintain the supply temperature at the threshold set point. Note that in this work, PSU refers to the power supply unit at server level, not at the level of the rack or the whole datacentre.

3. The rack level:

The rack consists of 30 servers with each server of a power (P_{total}). In addition, (P_{CHx}) and ($P_{\text{Networking}}$) represent the power consumption by coolant heat exchanger CHx and the networking switch, which are also powered through the rack.

4. The air handling unit level:

The air handling unit consists of three main components (P_F), (P_P) and (P_{spray}) which represent the power consumption of the main fan, pump and spray unit, respectively.

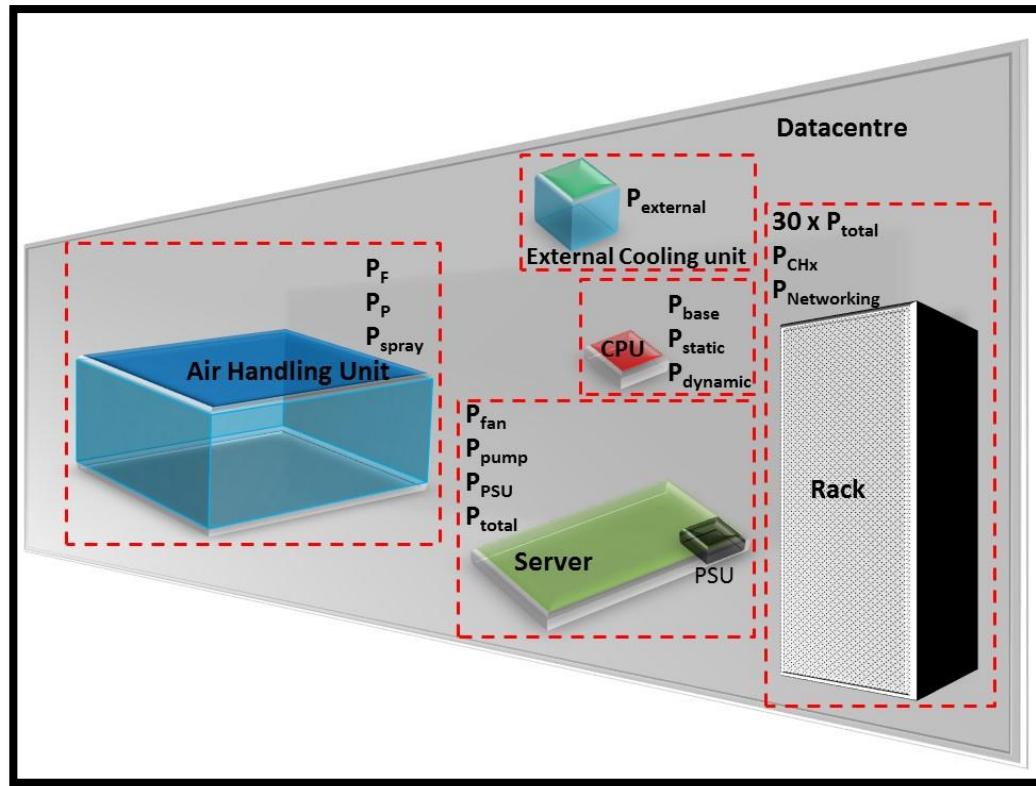


Figure 1.6 The power consumption terminology of the current work.

1.7 The Outline of the Thesis

Finally, the thesis consists of seven chapters to achieve the defined objectives. The thesis outline is presented in Figure 1.7.

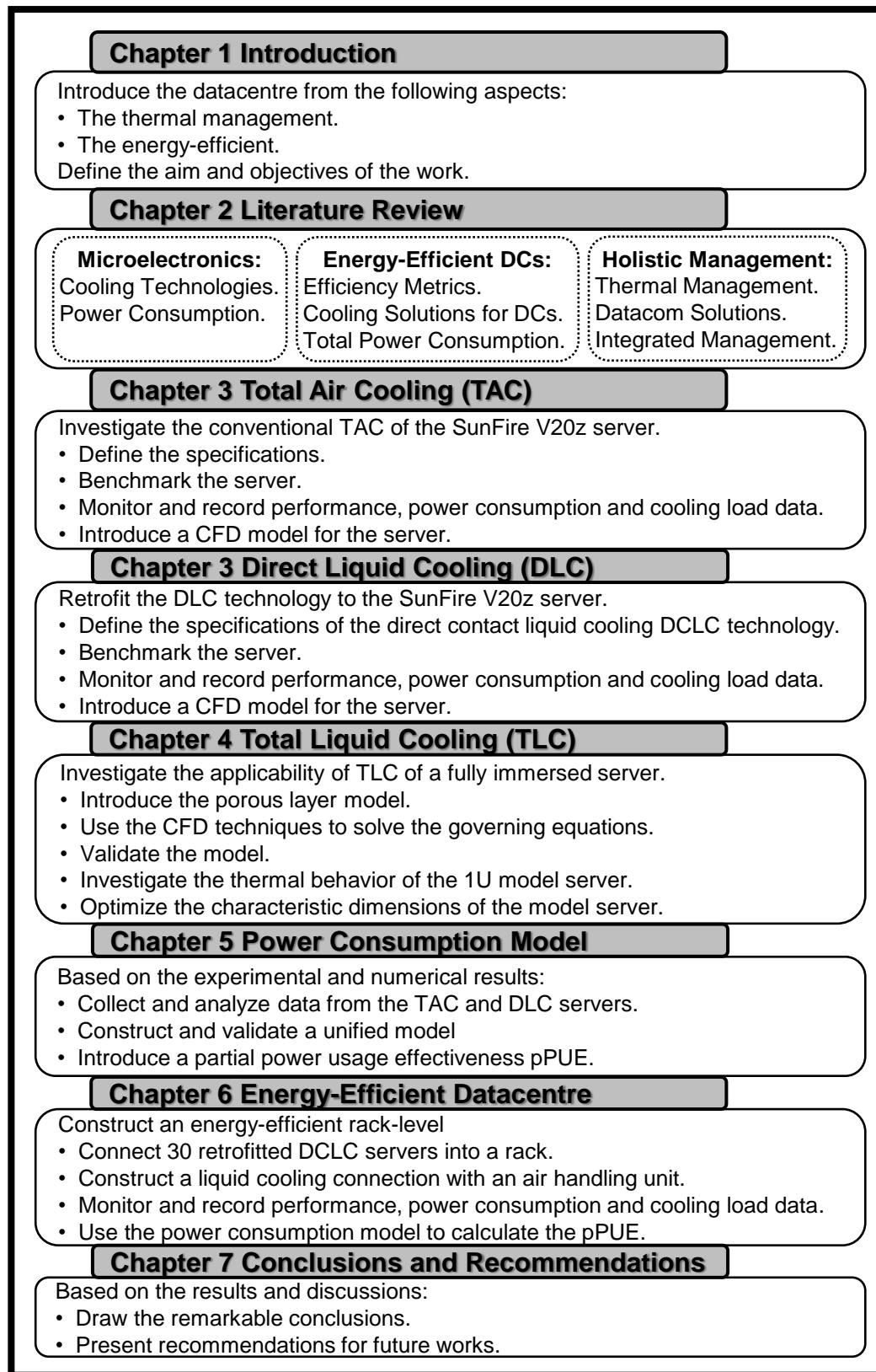


Figure 1.7 The thesis outline.

Chapter 2 LITERATURE REVIEW

2.1 Introduction

Examining the history of computing, Electronic Numerator, Integrator, Analyzer and Computer (ENIAC) was used for the first time in 1945. This massive machine was kept in a *particular room* which nowadays would be known as a *datacentre*. The first servers, or microelectronics, started to be kept in the first versions of datacentre during 1990s (Bartels, 2011). Around this time, the design, evaluation and enhancement of servers and datacentres were started. The infrastructure of datacentres received enormous interest in research both in academia and industry.

In this chapter, related works from the literature are reviewed and summarized through multiple levels of studies and this is shown in Figure 2.1. The first level deals with the server, or microelectronics, level since it represents the core of the datacentre; this level is responsible for performing the IT load. Then, multiple servers are brought together in the datacentre. The next part of this chapter reviewed datacentre level researches and especially the energy-efficient ones. Finally, although there are many investigations in the literature dealt with managements, there is only few studies applied to study the holistic perspective of management. Figure 2.1 shows the arrangement of sections based on each level.

At the end of this chapter, Table 2.1 lists each reviewed research and categorized them into a related field. Using the table offers a quick way to assess the value of each study.

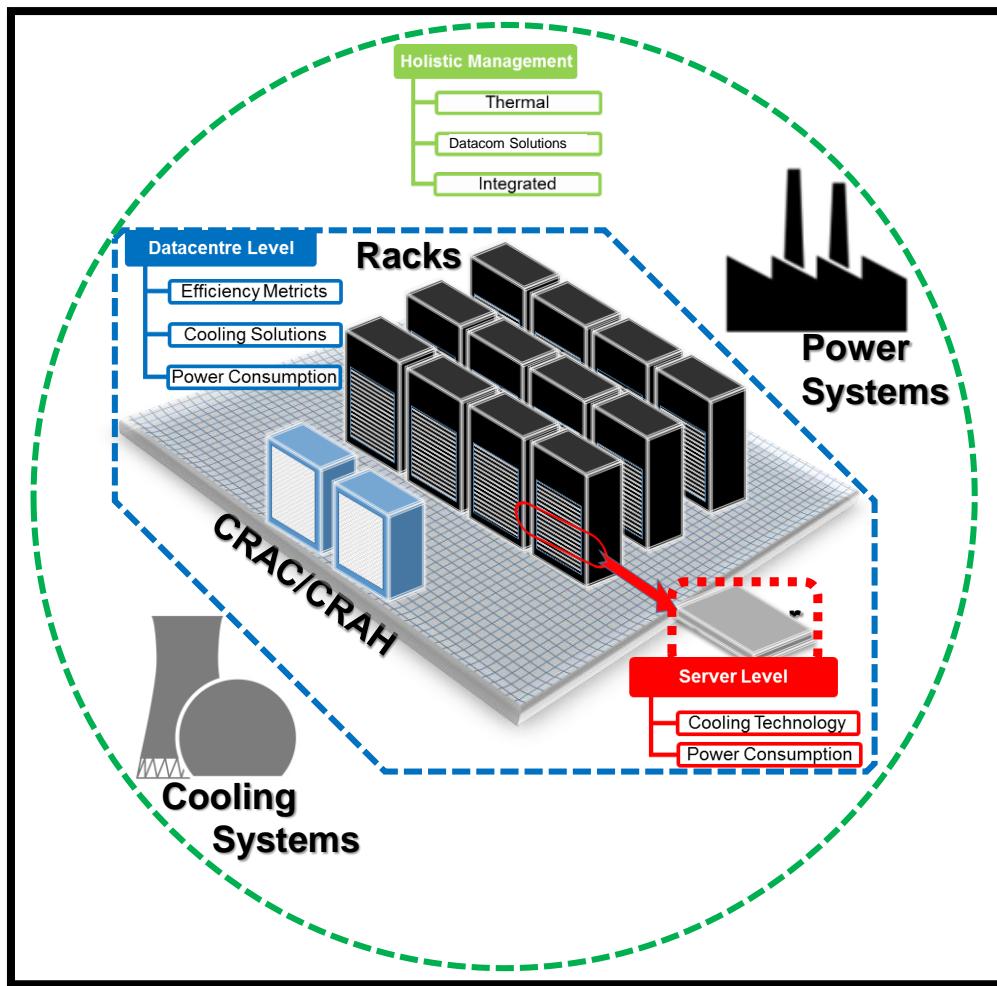


Figure 2.1 The three levels of literature review researches.

2.2 Cooling and Power Technology of Microelectronics

This part of the literature review deals with the *core* of IT which is the microelectronics. This IT load need to consume power depending on load level and rejects heat as a by-product. For this reasons, the operation of servers requires a high awareness and knowledge to better understand the relationship between the three characteristics of performance, power consumption and cooling load.

In the following sections, the performance of servers has been included either with the optimal cooling load as in section 2.2.1 or with lower power consumption as in section 2.2.2. Both sections aim to achieve energy-efficient microelectronic systems. Each section reviews key researches

over the last decade which has been selected as most related to the current study.

2.2.1 Cooling Technology of Microelectronics

Focusing on the thermal management provides a better understanding of the temperature effects on extreme power densities computing performance as presented by Choi et al. (2008). This high level of densities leads to an increase in both failure of IT and a demand for a greater cooling load. To evaluate the effects of temperature, a 3D computational fluid dynamic (CFD) tool called “ThermoStat” was used. The tool helps to understand the transient thermal behaviour of servers better. Servers are assumed to be thermally managed for reactive and proactive cooling systems. In addition, Dynamic Thermal Management (DTM) techniques are adopted to illustrate the role of several factors; such as electronic system architecture, thermal packaging and thermo-fluid characteristics. Thus, their research has helped to increase isothermal distribution of workload placement by 32.5% when compared to assuming a uniform load distribution strategy. Meanwhile, using a thermal approach reduces maximum near rack ambient temperature by 3.5 °C.

CFD techniques were used to reduce computational effort of the electronic component cooling through defining Compact Thermal-Fluid Models (CTFM) by Stafford et al. (2012). They used two different scales of difficulty to be more reliable on results. Meanwhile, they used CFD – FLUENT package to create and pre-assign the required CTFM. As a result, CTFM shows two main benefits; reducing the order of modelling and reducing CFD preliminary designs. Furthermore, through starting CTFM, thermal – fluid networks could be optimized. The results showed that there was no more than 10% of error by comparing compact and detailed models results.

A study of thermal effects of an Information and Communications Technologies (ICT) server was carried out by Sampath (2013) and Pandiyan (2013). The analyses included the measuring response of server's components to the different level of IT loads. In addition, the

researches focused on ambient temperature effects on power consumption and related cooling techniques. Simultaneously, the Computational Fluid Dynamic (CFD) models were used to investigate the possibility of predict empirical formula for power consumption. All simulation results were validated with experimental work, and the error percentages were less than 10%. The most important conclusion was that of a linear relationship between power consumption and processor utilization. This results could be used to extrapolate energy-efficient studies for datacentre level.

Recently, Ham et al. (2015) presented an investigation of a simplified microelectronics model to simulate energy efficient datacentre. The model dealt with the effects of interior thermal management techniques on cutting the energy bill. The cost of electrical power is affected by cooling load and energy consumption of IT.

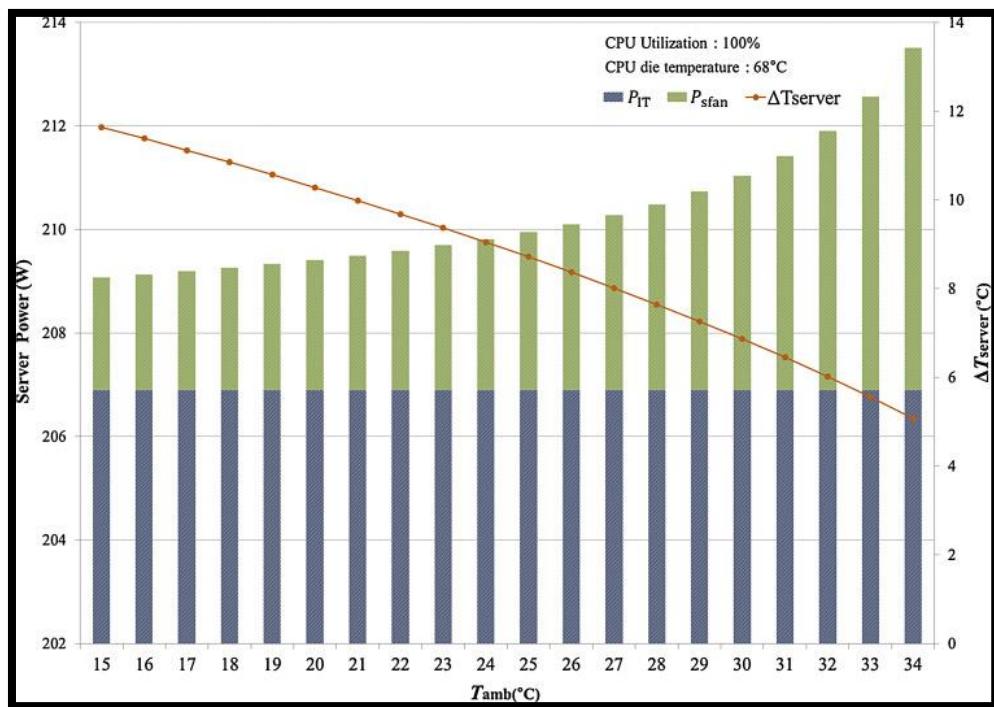


Figure 2.2 Inlet temperature effect on the server's performance, P_{IT} and P_{sfan} represents the power consumption of IT and fans, respectively (Ham et al., 2015).

For this reason, the model based on server thermal characteristics and power consumption. The results showed that increasing operation times of cooling units, i.e. economizer, will lead to more energy consumption of IT load if the thermal specification of servers neglected. In other words,

thermal criteria of each server should be considered to obtain a balance between IT high performance and energy efficient microelectronics. In practice, the power required by cooling fan inside servers is affected by environment, i.e. inlet temperature, as shown in Figure 2.2.

2.2.2 Power Consumption of Microelectronics

Power distributions within a microprocessor have been studied in some details to understand the impact of hot spots by Hamann et al. (2006b). The hot spots refer to the areas of high demanding of cooling. The research aimed to introduce power maps under various IT load. The technique used for this task was Spatially-resolved Imaging of Microprocessor Power (SIMP) (Hamann et al., 2006a). The SIMP method has two steps, firstly using Infrared IR camera to capture thermal map and secondly using the captured temperature field and define thermal resistance matrix to derive the power distribution. The pattern was obtained by two methods CFD modelling and applying power to a test chip and measure temperature field. The matrix was 30x30 and the corresponding resolution of the maps were 0.4x0.4 mm². Figure 2.3 shows thermal image under IT load and the derived power map, from the CFD model, with bright hot spot effect on high power consumption.

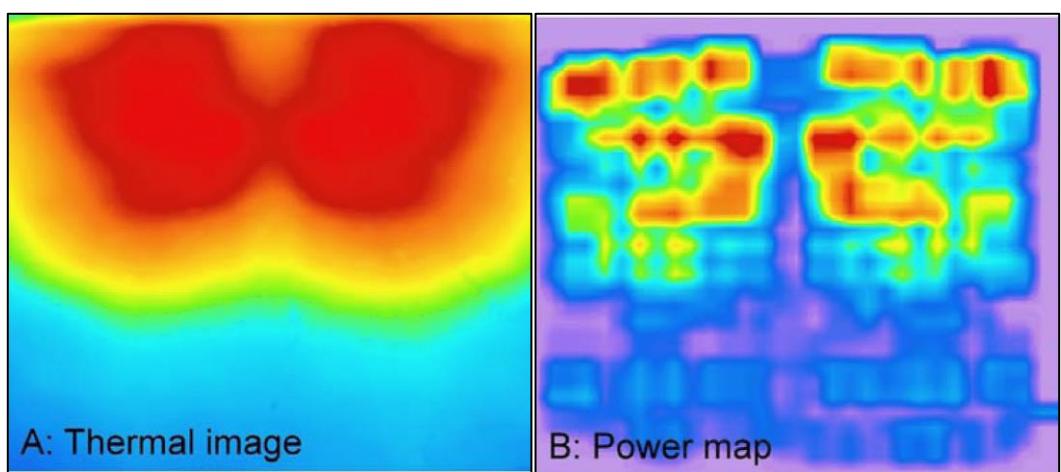


Figure 2.3 Temperature and power maps of dual core microprocessor A: Thermal and B: Power. Actual photo size is 13.2x11.6 mm² (Hamann et al., 2006a)

Using Infra-Red IR thermography to investigate Central Processing Unit (CPU) temperature distribution and corresponding power consumption has

been conducted by Farkhani and Mohammadi (2010). FLIR SC4000 camera of $10\text{ }\mu\text{m} \times 10\text{ }\mu\text{m}$ spatial resolution and 420 Hz was used to capture the thermal map of coolant. The coolant used in the setup was very high transparent mineral oil to allow tracking hot spots and high specific heat and conductivity to compensate the role of the heat sink to dissipate 100 W of heat. Standard Performance Evaluation Corporation SPEC CPU 2006 was used as a benchmark to produce the required IT load and monitor temperature for both transient and steady state. Simultaneously, a finite element method was applied to simulate the 3D core using CFD with appropriate boundary conditions. Both results, measured by IR camera and resulted from 3D model showed a linear increase in the temperature as power dissipation was increased. The agreement between experimental and simulation results was more than 97% as illustrated in Figure 2.4.

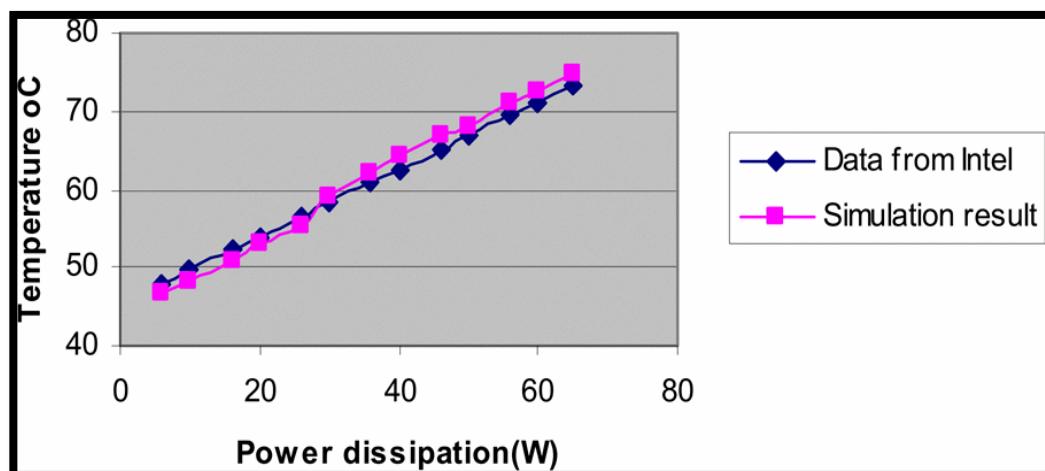


Figure 2.4 Comparison of measured temperature vs simulation prediction (Farkhani and Mohammadi, 2010)

The optimal throughput of multi core processors is very affected by core die temperature and related power consumption. For this reason, an optimised performance investigation based on thermal and power estimation has been presented by Oh et al. (2010) with reconstructed the related work formulae of power consumption. The method procedure started by estimate dynamic power profile based on temperature sensors reading, then optimize throughput related to the corresponding power consumption. Because of the estimation of power, the increase of the throughput optimization was about 4% on a temperature based method. In addition, the

reconstructed methodology reduced the hot spot prediction error to less than 3%.

Power estimation of the microelectronics level has been addressed by many researchers. They used a different definition for measurements and models to reach an estimate of the overall power consumption and cooling load. For that, Möbius et al. (2014) present his work to distinguish through a comprehensive evaluation of three groups of studies. The first group included seven approaches of CPU models of Intel servers. The second group used Intel and Advanced Micro Devices (AMD) systems to stand as virtual machines (VMs) with six approaches. Final seven types of the research were based on power consumption of the entire Intel servers. Some important points of view were concluded including the importance of presenting models based on variable workload instead of depending on available benchmarks. In addition, probabilistic parameter estimation would be a good alternative approach instead of current deterministic methods.

CPU workloads impact on power consumption of server at different benchmarks were conducted recently by Block et al. (2015). Server efficiency rating tool (SERT), which is one of the developed tools by SPEC, has been used to evaluate the energy efficiency of the server when targeting a particular IT load. The benchmarks were applied to a variety of architectures and configurations server to provide a high level of accuracy of estimated power consumption. Since SERT exercises each target by a range of activities, the results showed a full analysis range of energy use profile. This procedure was prepared to evaluate and measure server energy efficiency in *normalized* values.

2.3 Energy-Efficient Datacentre

In this section, an overview to the term energy-efficient of datacentre will be presented via three aspects. Firstly, it is important to define the metrics used to categorise energy efficiency through section 2.3.1. Therefore, selective studies which dealt with the aim of the current study will review in the first

section below. Secondly, since most researches refer that 30% or more of total power consumption inside datacentres is due to cooling, then section 2.3.2 will assess a variety of cooling technology. Thirdly, overall power consumption related papers are listed in the last part. Section 2.3.3 completes the whole point of view of mission critical computing datacentre.

2.3.1 Efficiency Metrics

The providing of digital services varies depending on the purposes of a datacentre itself. Therefore, measuring of the energy efficiency of a datacentre represents a vital issue during the evaluation process. However, there are different prospective have been used to present the *efficiency* and/or *effectiveness* of datacentre. The aim of this section is to present the different measuring techniques to evaluate the energy-efficient datacentre. For that, specific metrics which dealing with heat, power and IT performance measurements have been used to distinguish between different aims of metrics. Among all metrics, the power usage effectiveness (PUE) has the capability to capture power consumption of subsystems inside the datacentre infrastructure.

Dimensionless parameters to evaluate the thermal performance was studied by Sharma et al. (2002). It was recognised that using different criteria, i.e. inlet temperature, to address issues of the various types of the server is a complex task. For that, they derived two non-dimensional factors to evaluate servers. The authors presented for the first time such parameters which based on energy balance and fluid flow is generalized from rack to datacentre levels. The first factor is the supply heat index (SHI) which is defined as the ratio of enthalpy rise due to infiltration in cold aisle to total enthalpy rise at the rack exhaust. The second one is the return heat index (RHI), which is defined as the ratio of total heat extraction by the Computer Room Air Conditioning (CRAC) units to the total enthalpy rise at the rack exhaust. The algebraic summation of those two factors is unity since they are complementary values as shown:

$$SHI + RHI = 1 \quad (2.1)$$

Water metrics are described as efficiency management parameter and used to evaluate power consumption in datacentre by (Sharma et al., 2009). They stated ideas about simplification and correlation between power and water usage metering since they might be in the same order of effect on efficiency. Thus, determination of the water utilisation effectiveness (WUE) metric consists of a summation of the direct and indirect cooling water usage, and relates this value to energy consumption rather than power. Water efficiency is used in the sustainability index of green datacentres.

The Green Index (TGI) introduced and illustrated for high-performance computing HPC by Subramaniam and Feng (2012). In HPC, it is crucial to measure and evaluate efficiencies of both computational and non-computational parts of energy utilisation. For instance, energy usage of HPC needs to be a unique number to entering the 'TOP500' list through using LINPACK benchmark (TOP500, 2005). The procedure of calculating TGI, which is given by equation (2.2), conducting by three steps. Firstly, determination of Energy Efficiency (EE) from performance to power ratio. Secondly, obtaining relative energy efficiency (REE) based on the benchmark. Finally, calculating weighting factor W, then integrating a cross product of each REE and its weighting factor:

$$TGI = \sum_i W_i * REE_i \quad (2.2)$$

The authors conducted tests for 'SystemG', which consists of dual 2.8 GHz Quad Intel Xeon 5462 and 8 GB RAM. They managed to enrol eight values of processes as shown in Figure 2.5, where could be observed dependability of TGI on several processes.

The management of technological drivers in datacentres at different thermal, electrical and energy scales is presented by (Garimella et al., 2013). These characteristics were analysed in terms of efficiency and sustainability of datacentres in both the short-term and the long-term.

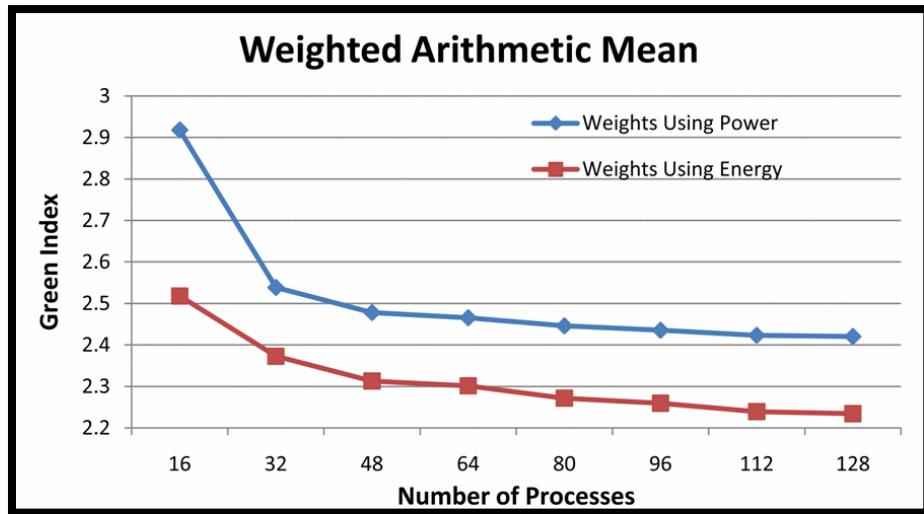


Figure 2.5 TGI vs. processes based on power and energy (Subramaniam and Feng, 2012)

Firstly, they illustrated the energy flow with Sankey diagrams, which is a particular flow diagram used to visualise energy demand based on arrow widths, as seen in Figure 2.6(a), for allocation of power drawn by the different components of a typical datacentre. In this datacentre, they presented and compared the suitable metrics and strategies to evaluate power efficiency. Secondly, they addressed the challenges by categorizing them into thermal, noise, and computation issues and electrical energy consumption as well. Finally, the authors critically evaluated a broad range of solutions for usage of cooling techniques. For example, they examined air cooling substituted by alternately advanced methods of using piezo fans, synthetic jets and ionic wind as air movers. Furthermore, liquid cooling received an excellent recommendation from the authors due to the potential for significant heat recovery based on the same computational energy consumption, Figure 2.6(b).

The most used metric is the Power Usage Effectiveness (PUE). This has been critically assessed of Facebook datacentre at Prineville by Brady et al. (2013). The data are available through the Open Compute Project. The aim of the study was to evaluate PUE metric and discuss its benefits and report its drawbacks. Moreover, exploring values of shared information to conduct precise calculation of PUE was the second aim of the study. The

study has several stages started with measuring IT, Air Handling Units (AHU) steps and all other electrical systems power consumption.

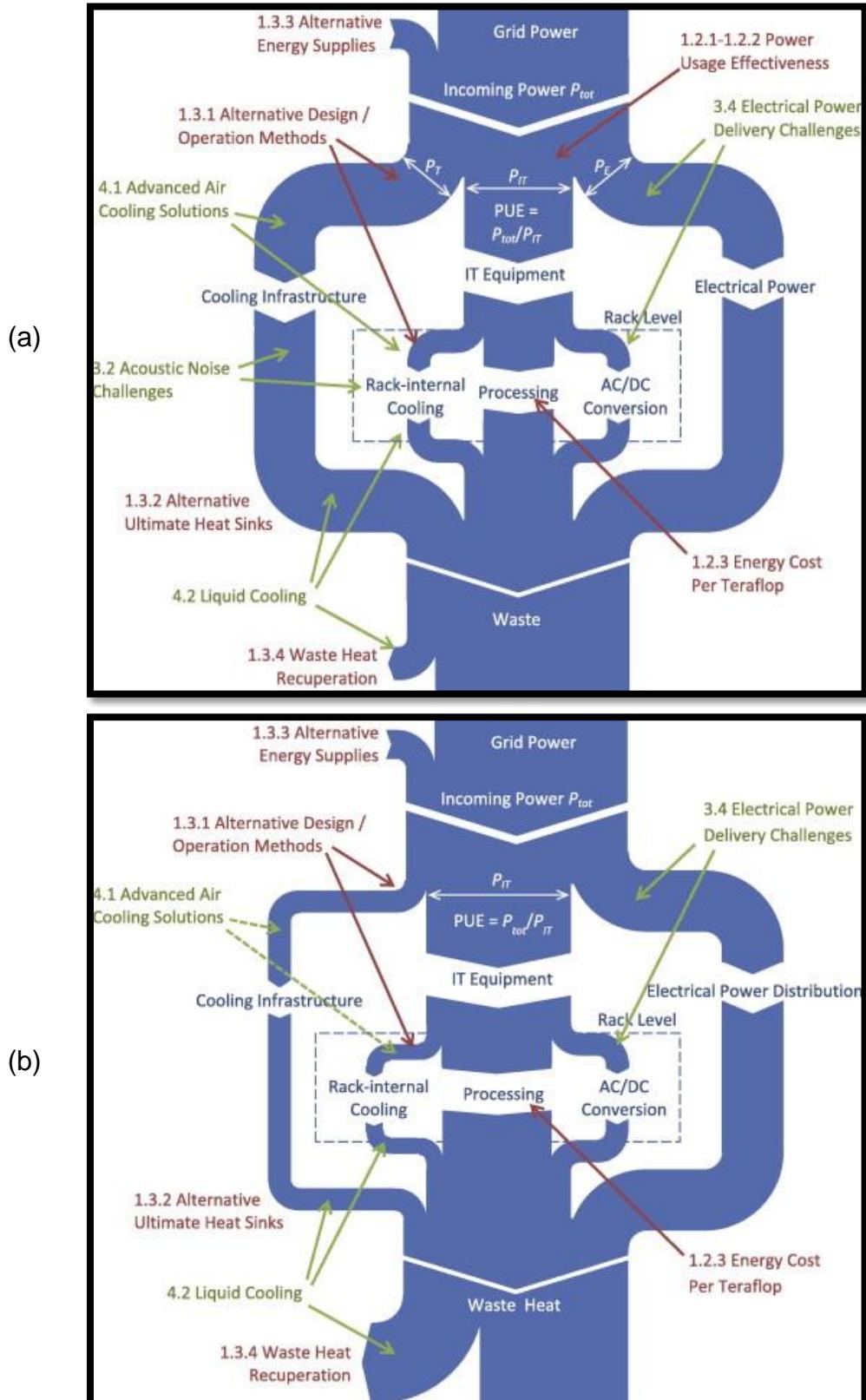


Figure 2.6 Sankey diagram for (a) typical datacentre and (b) fully liquid-cooled datacentre (Girimella et al., 2013).

A sensitivity analysis for the method of calculating PUE was also carried out. Finally, the metric subjected to critically assessment stage. Although PUE represents a useful metric to reflect energy consumption, however, it represents *effectiveness* not *efficiency*. For that, and even though many datacentres use their own calculation of PUE for marketing purposes, it still very subjective due to its limitation to express the relationship between IT load and cooling response. Furthermore, one significant drawback of this metric is its *locality*. This term refers to the impossible comparing between two datacentres in very different environments.

New IT efficiency metrics based on thermodynamic approach were presented by Beitelmal and Fabris (2014). This approach defined the ratio between IT work, which is the outcome of running the server, and the total power including cooling load. For server level measurement, the metric called server energy performance (productivity) metric (SEPM). This metric is extrapolated to datacentre level named data centre energy performance metric (DCEPM). It is defined as the ratio of average servers' power consumption to the sum of IT and cooling power consumption. These metrics restricted the efficiency of ICT, however, it may lead to redefining the datacentre standard.

2.3.2 Cooling Solutions of Datacentres

As the datacentres reject huge amount of heat increasingly, it is crucial to implement different cooling solutions to overcome this problem. The cooling solutions of datacentres varies depending on the location, size and available resources. Some researches focused on changing the layout of racks and computer room air conditioning CRAC unit. Others compare the using of air or liquid as coolant inside the server, rack and/or datacentre levels.

A CFD simulation was used by Almoli et al. (2012) to evaluate and compare traditional air-cooled datacentre versus liquid rear-door cooling technique. A 3D model was presented to simulate a three aisle of six racks datacentre. Continuity and momentum governing equations were arranged in Reynolds

Averaged Navier-Stokes (RANS) model. Three different heat load of 15, 25 and 30 kW per rack were studied for both air and liquid cooling configuration. Furthermore, there were two possible scenarios of for back-door heat exchanger. Passive back-door, which refer to circulate air with the aid of IT fans only, achieved a reduction in CRAC load of 91.9%. While active back-door, which uses its external fan, gained about 93.6%.

A concept for cooling for a highly energy efficient datacentre was presented by Iyengar et al. (2012). They named the system as ‘dual-enclosure-liquid cooling DELC’, which shown in Figure 2.7. The authors have used an IBM x3550 M3 1U server which is cooled by a hybrid air/water node. The cooling node consists of; a cold-plate to absorb heat from the CPU, spreader plates to cool Dual In-line Memory Modules (DIMMs) and cold-rails for memory chipset. The experimental results show a very significant rate of energy savings. At the datacentre level, there is a 25% energy saving, which is equivalent to minimizing the cooling-load utilization by more than 90%.

Based on previous work, David et al. (2012) investigated experimental heat rejection rates and server cooling with air / warm water techniques. They achieved a reduction in cooling energy usage by only 3.5% of IT energy and maintained the ambient air temperature within 23.8 °C.

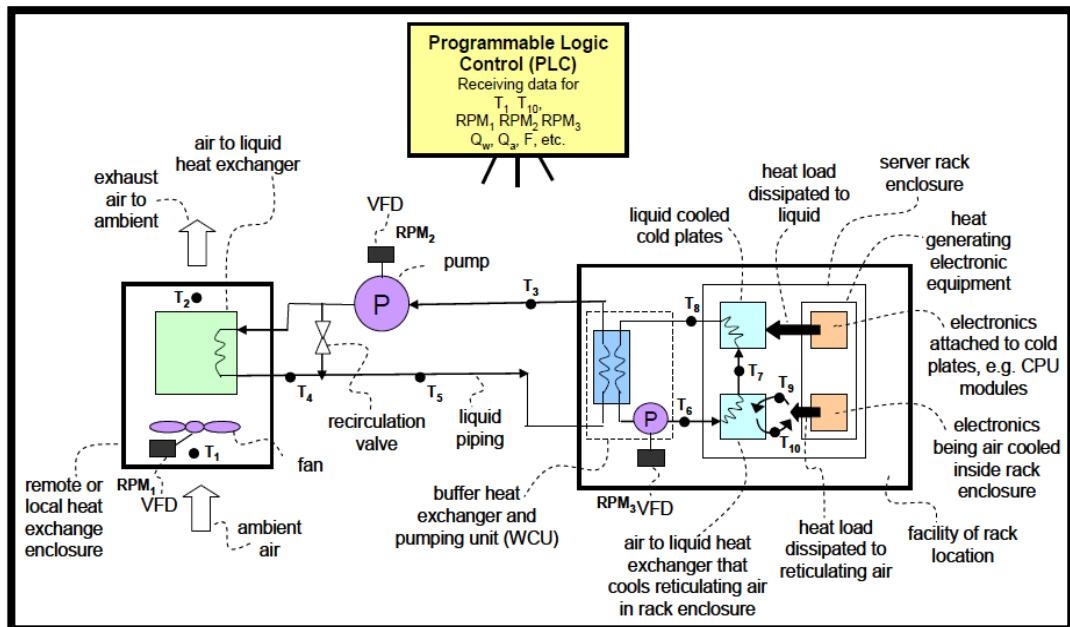


Figure 2.7 Dual-enclosure-liquid cooling DELC system (Iyengar et al., 2012).

Similar to this technique, the University of Leeds, Iceotope and 3M collaborated in a research study by using encapsulated, fully-immersed servers with 3M Novec™ liquid to improve cooling capabilities. Cooling is achieved by natural convection inside the Novec™ dielectric fluid as shown in Figure 2.8. For that, the cooling cost can be reduced by up to 97%, and for this reason, the Uptime Institute Green Enterprise IT awarded this research as 2013 winner (Iceotope, 2013).

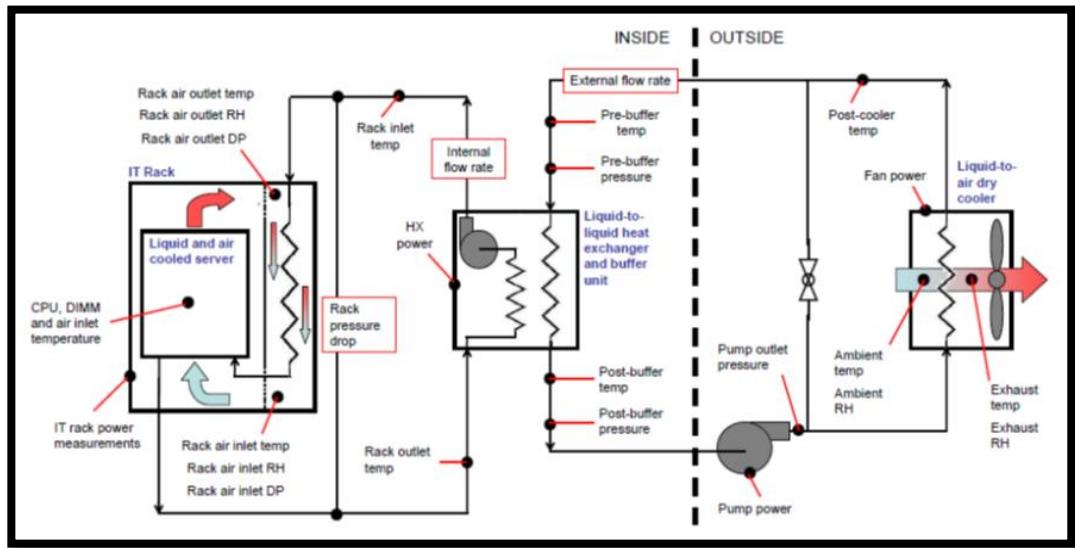


Figure 2.8 Air/warm water technique (Iceotope, 2013).

Energy savings was the aim of the CFD study presented by Cho et al. (2014). In addition, this study redefined cooling management of datacentre by enhancing thermal characteristics of 46 different configurations of the datacentre. A parametric study was included as part of the simulation and focused on the type of air distribution (location), heights of both floor and ceiling, aisle configurations and temperature of supplied coolant. The evaluation metrics used in this study were supply/return heat, rack cooling and return temperature indices. A performance index was derived from those indices to evaluate each configuration. Increasing supply air temperature from 18 to 22 °C, when using containment, represents very good status based on the performance index. Moreover, rising thermal performance of air cooling was due to increasing the height of raised floor, while ceiling height has minimal effect.

A comparison study was also considered between hybrid air-cooled and fully immersed liquid cooled datacentres from an energy perspective (Chi et al., 2014). The PUE used as an evaluation metrics due to its compatibility with the aim of the study. The hybrid air-cooled systems were the HPC at the University of Leeds, while the fully immersed solution is the same unit of Iceotope (2013) and the schematic diagram is illustrated in Figure 2.9. The results showed an enormous enhancement in energy saving, IT performance and PUE. For the fully immersed liquid cooling, power consumptions reduced by more than 97%, PUE reduced by about 23% and IT performance increased by 40%.

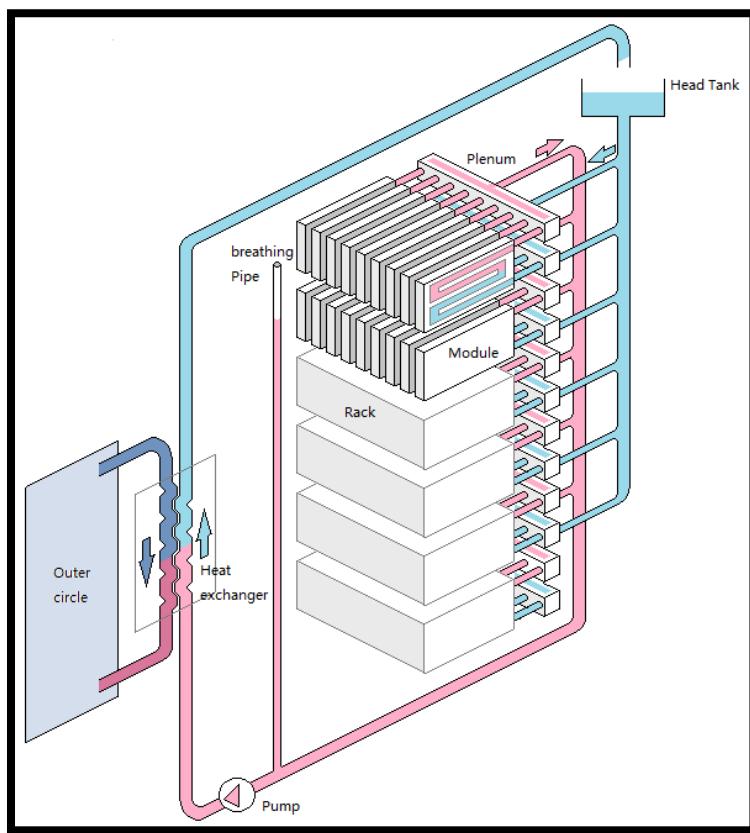


Figure 2.9 Liquid-cooled technology for rack level of datacentre (Chi et al., 2014)

CFD simulation technique was used to develop a full analysis of the liquid cooling system of datacentre (Almaneea et al., 2014). Dielectric coolant in a direct contact was used inside servers where heat is removed by buoyant flow natural convection. This convective heat is then collected by water loop to dissipate through the dry-air cooler. The server level simulation used the k- ω turbulent model with required definition of dielectric thermal properties.

In addition, an algorithm was developed for two purposes; (i) estimating the heat balances of the heat exchangers and (ii) evaluating environmental data of the cooling unit. The results showed a PUE of 1.08 which represents excellent value for the full occupied rack. However, the rack showed a lower effectiveness as a rise in PUE was observed when server occupancy dropped as shown in Figure 2.10.

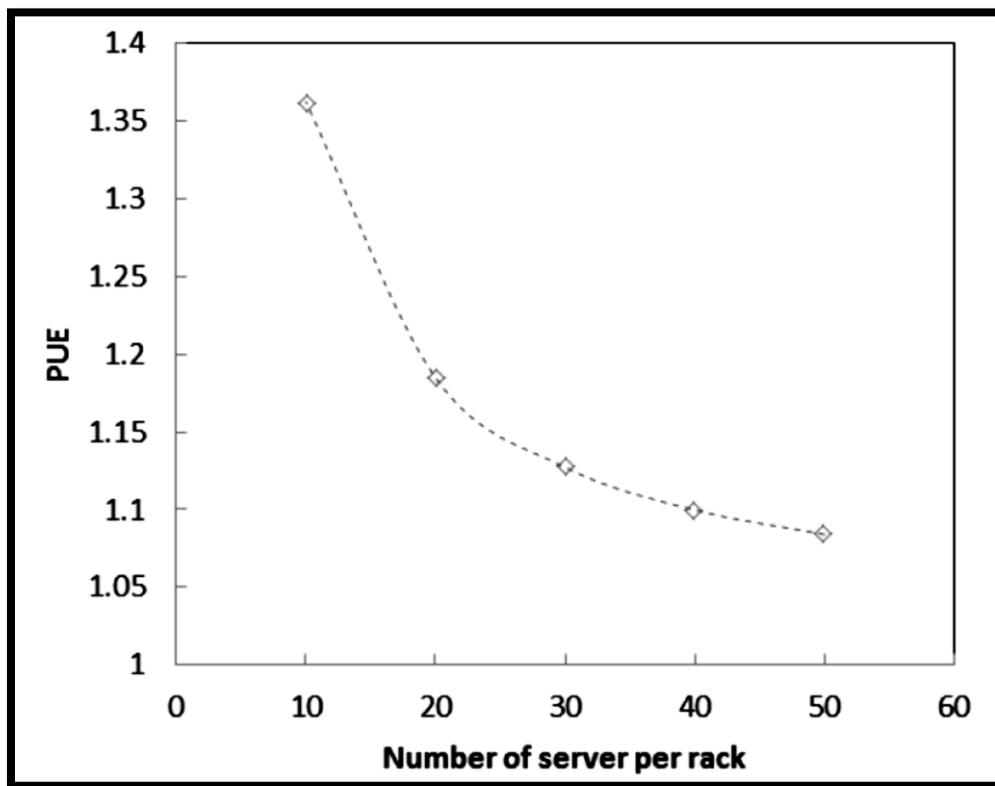


Figure 2.10 PUE dependency on server occupancy (Almaneea et al., 2014)

An overview of a variety of cooling solutions and technologies during the last decade was carried out by Li and Kandlikar (2015). Researches aimed to achieve the energy-efficient target for case studies of datacentres. The study used the concluded remarks to evaluate, compare and predict the future trend of cooling systems used in datacentres. From the cooling systems perspective, the evaluation of studies was based on the following categories; traditional air-cooled systems, single-phase and two-phase cold plate liquid-cooled systems, hybrid air and liquid-cooled systems and heat-pipe-based systems. In addition, there were opportunities of using chiller-less cooling systems and waste heat recovery systems which have been used to reduce overall energy bill. Among the conclusions drawn by this

study was liquid cooling represents more efficient method than air cooling. In addition, the cold-plate technology provided maximum transferring of heat rejected to the coolant especially when the heat needs to be used in waste recovery systems. Finally, it is predicted that a very high potential towards using the free cooling system with the aid of liquid cooling system.

Optimization of energy efficiency approach was presented to investigate 17 datacentres around the world by Cho and Kim (2016). The approach analyzed the differences of each datacentre on the climate of location and corresponding cooling load. For that, IT environments were kept within ASHRAE envelops whereas cooling energy consumption is lowered. It has been realized that optimal usage of economizer affects the cooling load required which leads to overall power saving. Moreover, the geographic location of the datacentre represents a crucial factor regarding power savings. Thus, cooling power can be reduced based on climate outside the datacentre. For example, the saving of humid-hot, dry-hot and cold regions was up to 26%, 38% and 80%, respectively, compared to the same datacentre without considering economizer. Although the research focused on those essential characteristics, however, further detailed economic investigation need to be considered when to choose and design a datacentre.

2.3.3 Total Datacentre Power Consumption

Collecting operation characteristics of up to 15,000 servers running over a period of 6 months allowed Fan et al. (2007) to analyse and assess power usage profile of datacentre. They found that even though many IT applications are well-tuned to minimal power usage, there was around 40% gap between theoretical and practical power usage trend. A modelling framework was presented to estimate peak power reduction and allow more servers to run within the same budget. The model showed excellent agreement with measured values at server level and racks level as illustrated is Figure 2.11.

The modelling power consumption of datacentre by developing analytical framework is presented by Pelley et al. (2009). The aim of this research was to abstracting clearly the power management of datacentre since there are very sophisticated interactions among entire systems. The simulation dealt with three main power consumption categories.

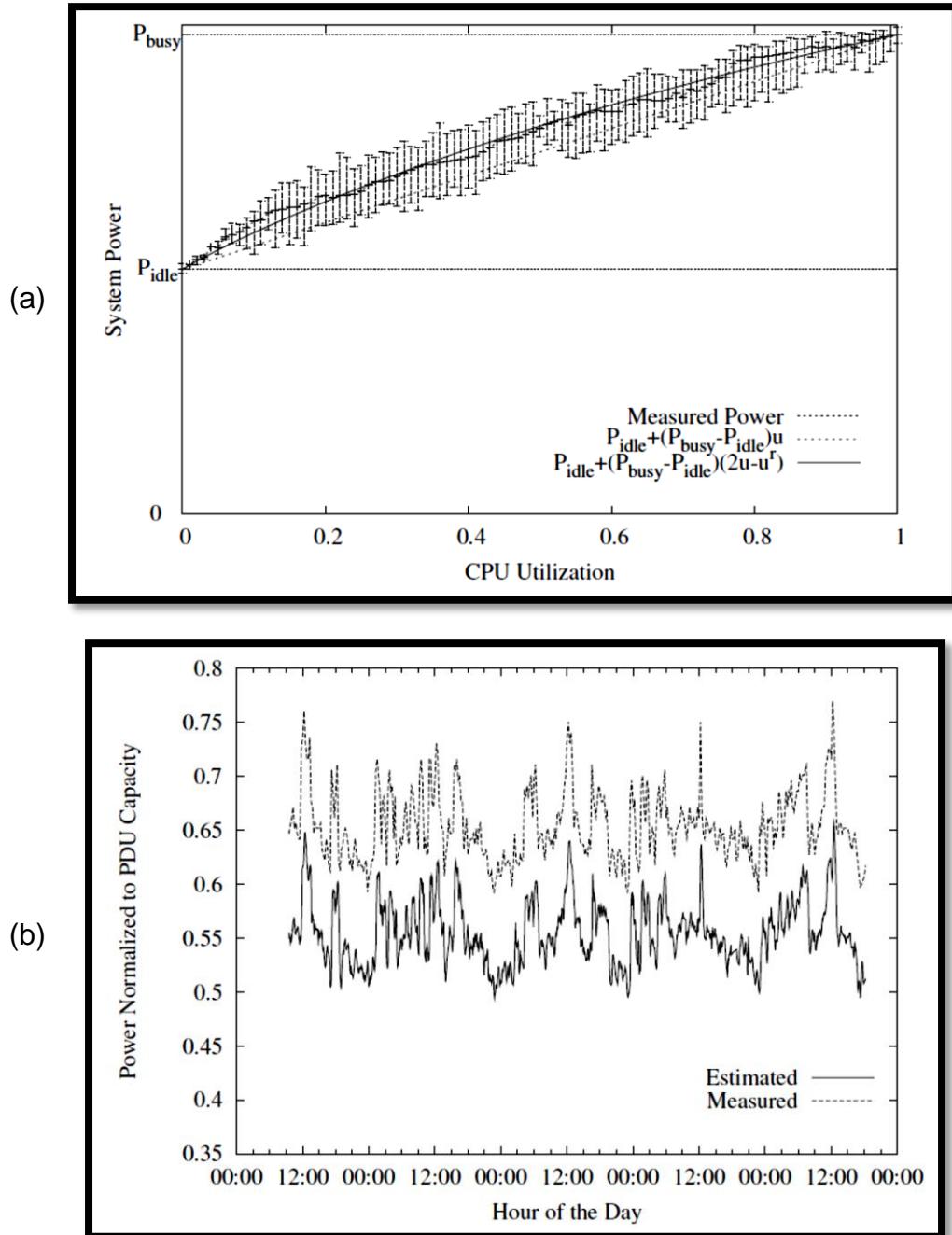


Figure 2.11 Model estimation power usage (a) servers level and (b) racks level (Fan et al., 2007).

First, servers which related to IT load awareness of tasks scheduling through available resources. Second, electrical system serviceability by

providing and distributing uninterrupted power supply. Third, the cooling systems interaction between environmental conditions and required supply conditions which share 30% of overall power consumption. Application of the model on two datacentres from last decade with legacy physical infrastructure showed a drastic reduction in power consumption. The possible scenarios applied through the model achieved power decrease of more than 60%.

Power scheme of a rack of servers' effects was studied experimentally by Nada et al. (2015). The loading of servers has been changed into four possible scenarios and they were uniform, discrete, segmented and clustered loads. The experiments have been conducted inside a one-sixth size scale room and a single rack of four servers to simulate real datacentre. SHI/RHI were used to evaluate the correlated thermal characteristics of datacentre performance. Results showed that uniformly distributed IT load consumes best energy efficient comparing thermal management to the power consumed in the rack. Moreover, clustered load provided better air circulation to cool servers down and reduce cooling load. This led to reducing overall power consumption compared to segment and discrete IT load.

Air conditioning units of datacentres are responding to the temperature threshold setting point. However, there is always a time lag between the reading of sensors and the response of CRAC/CRAH units. In order to shorten response time, Tarutani et al. (2015) presented a proactive control strategy to predict hot spots before their actually arose. The control system mimic the temperature field based on IT load assignment at the server level. The prediction of temperature has been achieved for up to 10 minutes before real time. This strategy offered highly proactive procedures to maintain the power consumption as flat as possible during experimental period as shown in Figure 2.12. One of the interesting results was power consumption has been reduced by around 30% due to a decrease of rapid fluctuation of cooling loads and keep the servers worked in the uniform thermal environment.

In an air-cooled datacentre, the effect of bypass air effects on electrical consumption was investigated by Tatchell-Evans et al. (2016). The investigation combined both experimental and modelling approaches. Experimentally, the percentage of bypass air flowing through cold aisle containment have been measured by applying differential pressure across the rack.

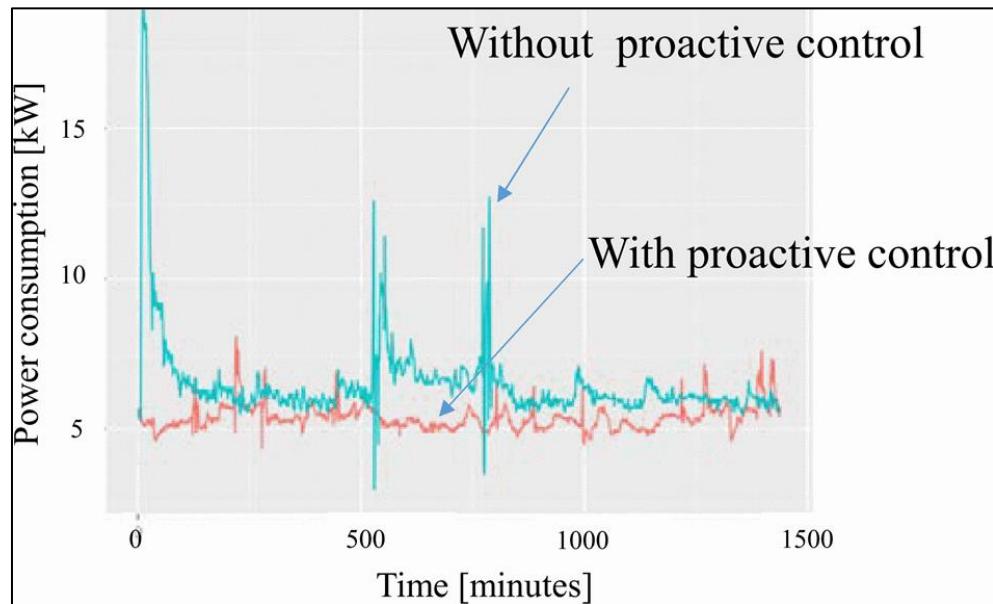


Figure 2.12 Proactive control strategy effect on cooling power consumption (Tarutani et al., 2015)

The results showed that a certain amount of cool air passed through fewer impedance paths rather than penetrate across servers. Simultaneously, a developed model has been presented based on resources information of the commercial servers and rack design. Therefore, flow measurements and predicted results from the model were used to evaluate the effect of the bypass flow on power consumption. The results showed that up to 8.8% reduction in electrical power consumption could be achieved by reducing air leaks from the cold to the hot aisle, this could be achieved by redesign of the racks. In addition, there were about 16% higher power consumption with leakage due to an excessive amount of differential pressure in the containment systems.

2.4 Holistic Management of Datacentre

The word *holistic* refers to the overall and mutual interactions among the datacentres. The criteria are IT performance, efficient cooling load and power consumption. Although each criterion has extensive opportunities for enhancement, holistic relates different aspects to one another. Thus, there are common areas between any two of them and sometimes among all of them. For this reason, this part categorised into the following sections. In section 2.4.1, the thermal management represents the starting point to evaluate holistic perspective. Furthermore, some algorithm has applied by researches in 2.4.2 to provide optimum IT performance. Finally, section 2.4.3 examines integrated management to achieve required task of management.

2.4.1 Thermal Management

The overall thermal management represents the starting point of researches in the sector of the datacentre. The reasons behind this are the effect of governing thermal environment on server's operation and the high-level costs allocated to overcome this situation. The datacentre and the Information Technology (IT) industries are affected by elevated temperatures. High temperatures in the IT systems affects both performance and reliability of the datacentre since high temperatures make up 55% of the failure factors of IT equipment (Anandan and Ramalingam, 2008).

Prediction on cooling criteria of the next generations of High-Performance Computing (HPC) facilities and datacentres was implemented by Bash et al. (2003). They addressed the sharply increasing needs of electronic cooling to overcome by-pass heat generating from IT components. Datacentres have changed in behaviour since the article was published a decade ago. However, they highlighted vital issues that require analysis, such as understanding power and effectiveness, indicating the effects of

different infrastructures and developing optimization approaches. During that period, the majority of studies assuming datacentre is a *static* infrastructure, whereas it is a real-time dynamically changing infrastructure.

A detailed review of the literature for electronic thermal management has been presented by Anandan and Ramalingam (2008). The survey covered the essential principles and guidelines for datacentre thermal management and considered cooling techniques. Furthermore, cooling methods in state-of-the-art were categorized, such as air, liquid, refrigeration, thermo-electric cooling and phase-change materials coolants. In addition, the main ideas of different characteristics of each technique and systems were explained.

Design parameters of the cooling systems for high-density usage datacentres were studied by Cho et al. (2009). Heat removal efficiencies of twelve-different arrangements of air distribution were examined and evaluated numerically using CFD simulation as shown in Figure 2.13.

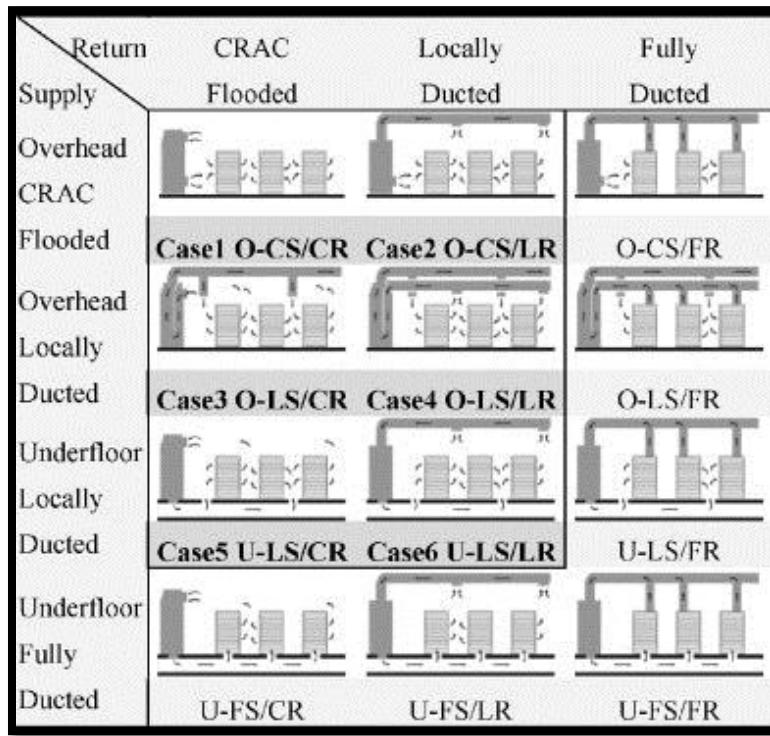


Figure 2.13 Types of the distribution system (Cho et al., 2009).

Meanwhile, to validate the numerical results, sets of experiments were conducted using a thermal infrared imaging camera and digital thermo-hygrometer in a mega-watt datacentre. Overall, the overhead-locally ducted

supply/return shows the ideal arrangement in both temperature and air flow distribution. However, the under-floor arrangements showed superior performance even though there was recirculation of the air intake temperature. The superiority of the latter is due to improvement in the Coefficient of Performance (COP). The improvement of COP is based on the supply of cold air distribution near the server intake region. This cold air reduces the cooling load of the rack and achieve overall power saving. As a conclusion, for selecting the appropriate air cooling distribution system, it is very subjectively to consider; capacity, design and condition. In addition, it is more important to determine the weight of each characteristic parameter for an efficient datacentre.

Breen and his research team studied modelling of datacentres, which started with chips and extended to cooling towers, in their recently published papers. In 2010, they investigated temperature effects on energy saving of datacentres (Breen et al., 2010). For that, the datacentre was analysed to the following levels; chip and fan, rack, Computer Room Air Conditioning (CRAC), chillers and cooling tower. The overall behaviour of the energy consumption of the datacentre was affected by the inlet air temperature to the rack and the temperature rise across it. Increased values of intake temperatures and widening temperature difference across the rack improved the energy efficiency has reduction in power consumption. In addition, free cooling assumed to play an important role assisting the primary cooling system through the dissipating massive amount of heat using the outside environment as the heat sink. Furthermore, Walsh et al. (2010) extended the work to understand the philosophy of temperature control. The first model procedure was done by assuming linear relationships between the heat sink and inlet temperatures. As temperature increased beyond 27 °C, the room temperature had the most effect on total power required since the chiller power is decreased sharply. Their second model depended on fixation heat sink temperatures while varying the intake temperature. In this case, increasing temperature to a level of 27 °C and higher required an extremely high cooling power. Cooling power at this level relates to cooling infrastructure itself and not due to IT loads, so the (COP)

reduced dramatically. Finally, Breen et al. (2012) validated their model of Multi-Scale Energy Management (MSEM) by the use of a holistic datacentre COP metric. The COP_{grand} notation proportioned inversely with overall cooling power, as shown in:

$$COP_{grand} = \frac{Q_{rack,IT}}{P_{cool}} \quad (2.3)$$

Cooling power represents the summation of power from; rack, CRAC, chilled water, chiller and cooling tower. The predictions of the MSEM model agreed to a good level of confidence with experimental results of live datacentre at the Hewlett-Packard HPL-3L.

Ouchi et al. (2012) presented cooling techniques for datacentres that have had dramatic increases in power consumption. Very high levels of power are assigned for the cooling and thermal management of datacentre environment. Unfortunately, the target temperature is not always achieved and many power failures and improper high temperature rise failures occurred. For high power demand situations, single and two phase heat exchangers, thin heat pipe, flat spreader heat pipes, Nano-fluid and plug-in connectors were discussed. Although the research was conducted on two specific techniques, the authors give summaries about five different methods. They used the following basic formulae to review the results; convective heat transfer to find heat removal and heat removal to power consumption ratio. The experimental tests on prototype server racks, which was conducted for a single-phase heat exchanger, shows higher efficiency by both saving the consumed power and lowering CPUs temperature. The ranges of power saving achieved were 44–53%, even though the temperature was kept at 27 °C. Simultaneously, the power utilisation effectiveness (PUE), which is the ratio of total facility power to IT equipment power, is about 1.2 which is an excellent value for this situation.

Hassan et al. (2013) studied cooling criteria within their university's datacentre using CFD – FLUENT code. This code is based on mass, momentum and energy governing equations to deal with air-flow, temperature and pressure distribution inside datacentre. The most important issue in this study is the detailed investigation of the rack-by-rack

inlet and exit temperature mapping, and specifying hot-spot areas within datacentres. Datacentre specifications were predefined to ensure that simulation implement real working conditions. As a result, the evaluation of 3D mapping of the governing characteristics addressed the hot-spots inside the datacentre. For that, they predict to improve and develop working conditions through certain technologies to save energy costs.

A detailed overview of advances solutions of thermal management of datacentre has presented by Fulpagare and Bhargav (2015). This review represents a guideline which gathers and analyse researches and studies dealt with thermal managements and techniques of datacentres. The first part summarizes methodologies, included experimental and numerical, and results of air-cooled datacentres and rack level layout optimization. Second part reviewed studies related to performance metrics and dynamic control and the validated models. The final parts listed the IT solutions perspective by various methodologies to enhance performance and reduce power consumptions. The research evaluated presented studies and predicted valuable future work to achieve a high level of energy efficient datacentres.

2.4.2 Datacom Solutions and Management

Linear power model was derived analytically from collected data of real-time estimation of power consumption (Singh et al., 2009). Three benchmarks suites, all running on Linux, were used to ensure the independency of the proposed model. Power envelop assumed to be degraded by a certain level. For that, a thread scheduling technique was provided to run a multi-programmed workload on chip multi-processors CMP of four cores. The model has successfully predicted power consumption levels compared to the measured values. The highest median error was 7.2% and Figure 2.14 shows the compared values of the three benchmarks.

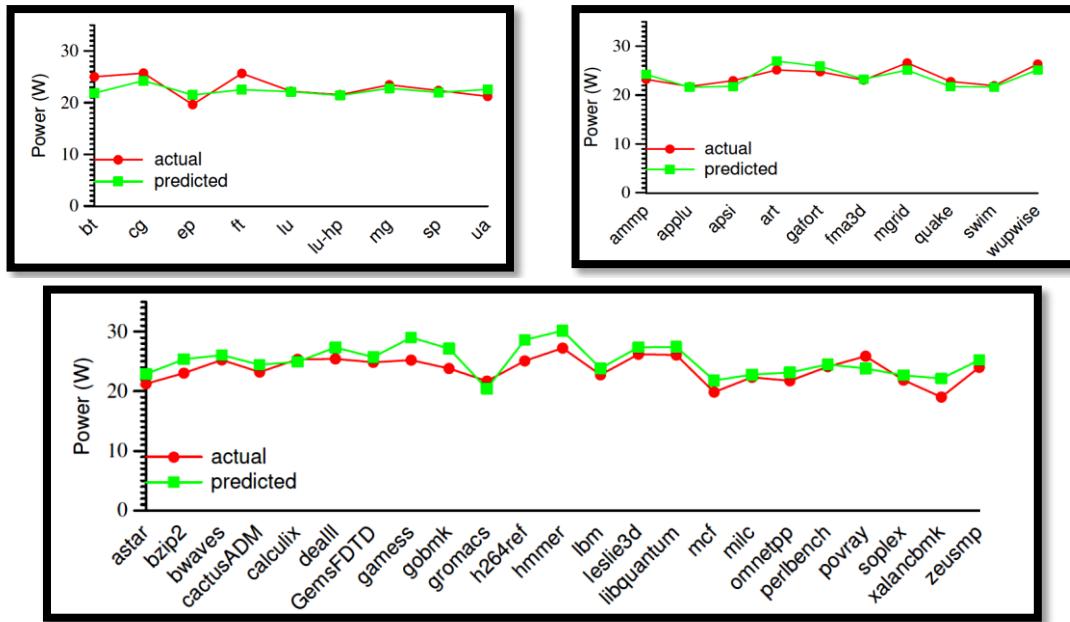


Figure 2.14 Actual vs predicted power consumption (Singh et al., 2009).

Scheduling algorithm was presented to model workload and resources of datacentres (Wang et al., 2009). All required information was based on Center for Computational Research (CCR) of State University of New York. The simulated model based on thermal mapping of the datacentre and then allocate tasks depending on the optimal location of servers. The algorithm benefit was to reduce the overall temperature of datacentre environment and thus lead to reducing cooling load of the CRAC.

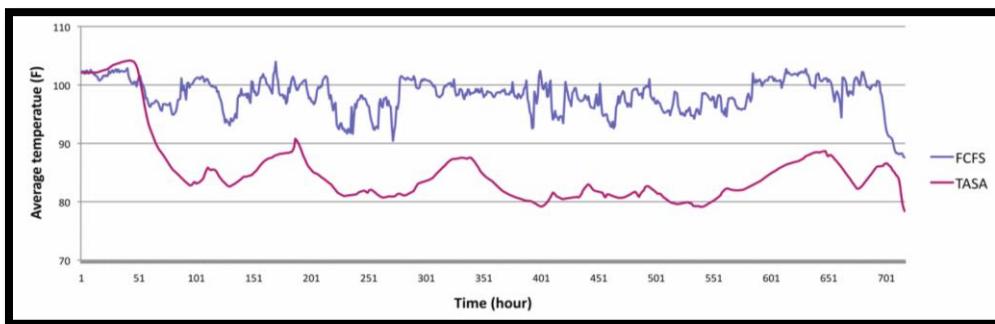


Figure 2.15 Reduced the temperature of IT nodes when using TASA model (Wang et al., 2009).

The traditional job queue policy of the CCR was First Come First Serve (FCFS), while the adopted simulation used Thermal Aware Scheduling Algorithm (TASA). The later algorithm achieved a very clear reduction in average temperature of IT node as shown in Figure 2.15.

Turning off idle servers dynamically to reduce power consumption in Internet datacentres was the aim of a research paper of (Abbasi et al., 2010). The algorithm provided to perform this mission was based on Thermal-Aware Server Provisioning (TASP) and Thermal-Aware Workload Distribution (TAWD). The energy saving, as discussed in the paper, is affected by two criteria; non-uniform traffic rate of tasks and heat recirculation between racks. Fluctuation of Internet traffic leads datacentre designers to work in *overestimation* of active servers to ensure the Quality of Service (QoS). However, TAWD model can compensate the overestimation of active server dynamically and save energy bill. In addition, TASP can provide extra energy saving when applied to large-scale Internet datacentres.

New strategies were presented to minimize power consumption of a datacentre by Cremonesi et al. (2011). The solution is based on the consideration of server's temperature and capacity as well as response time for cooling. These constraints were defined to optimize the relationship between instances and corresponding cooling power consumption.

Scheduling effects for best energy awareness were investigated by Zotkiewicz et al. (2016). An online strategy was applied for Software as a Service (SaaS) applications in cloud computing datacentre with a variety of workflow modelled services. The strategy benefited from the combination of workflow and energy aware scheduling by defining tasks based on available computing resources dynamically. To achieve better energy-efficient processes, there were four characteristics need to be examined; server's performance per power, location, utilization and potential of communication.

2.4.3 Integrated Management

The thermal mapping of datacentres with the aid of implementing different methods was studied by Moore et al. (2006). Firstly, an automated model was developed by the authors. Their code "Weatherman" succeeded to predict over 90% of results against experimental measurements to within

0.87 °C. On the other hand, online approaches, which used to capture the thermal profile of certain datacentre configuration, were valid to predict thermal mapping of real datacentres. This prediction assists to minimizing cooling cost by between 13% and 25%.

These achievements were based on a strategy that employed a number of approaches: (i) holistic activation of scheduling of IT load through different levels inside datacentre architecture, (ii) raising reliability of components, especially CPUs, by maintaining working temperature within allowable limitation, (iii) an approach that decreases both cooling cost and response time during transient intervals, and (iv) boosting efficiencies of the whole infrastructure by comparing IT load, total energy and its cost. Good agreement, to this situation, between predicted and actual temperature for 280 test with 313,600 data points, where the accuracy is ± 1.0 °C of 92% of results as shown in Figure 2.16.

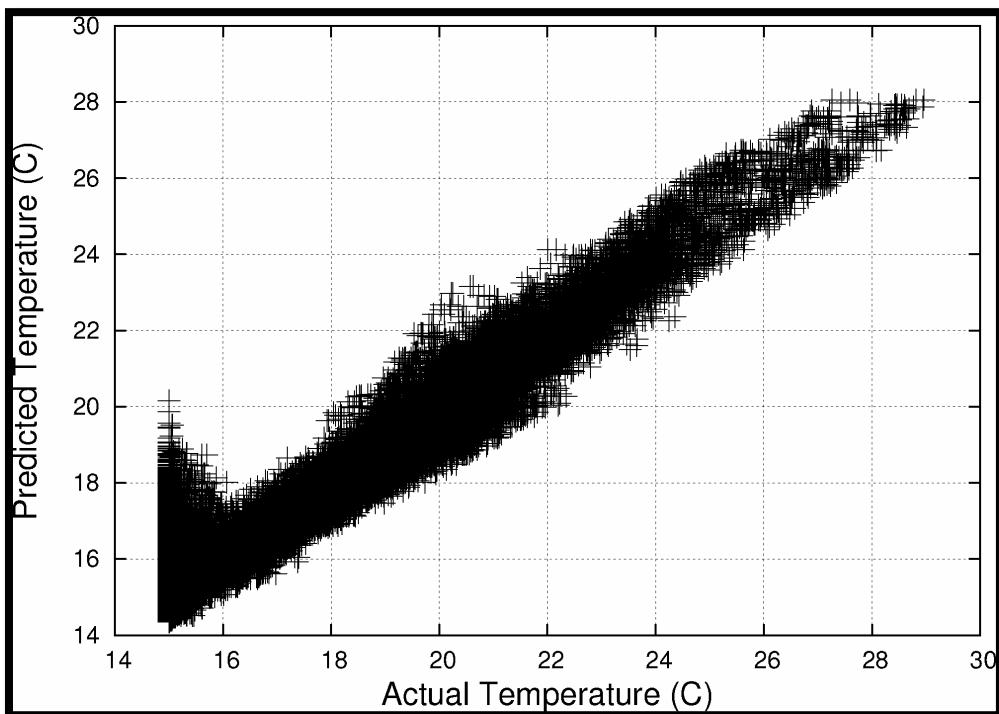


Figure 2.16 Scatter-results of predicted versus real values (Moore et al., 2006).

Performance, power and cooling holistic management for datacentres was studied as a virtualized prototype environment by Chen et al. (2010). The integrated IT management consists of three controllers; application, node and pod controllers. The application controller checks that every single or

compound application met the specified allocated target through updating response time, the control interval and utilization threshold, periodically. Meanwhile, the node controller's role is to ensure that targets are assigned to specified virtual machines (VMs) correctly. The pod controller is responsible for dedicating collections of homogenous servers. Furthermore, datacentre integrated management depends on cooling, workload and integration of IT and cooling management. Such management approaches were found to minimize IT and cooling energy consumption and by 35% and 15% respectively.

Another integrated approach was presented by (Wang et al., 2010), through awareness of dynamic control and integrated local and zonal cooling control. Results from testing 10 racks of servers showed that an integrated approach could be used to control the complicated dynamics.

Zhou et al. (2011) produced a dynamic thermal management approach for datacentres. They utilized and improved overall adaptive vent tiles for control of the cooling units starting from the fundamentals of both mass and energy conservation. The balance based on a step-by-step control at the rack level as shown in Figure 2.17, by means of supplying cold air from CRAC to rack inlet, modelling air mixture trends and controlling temperature at this level.

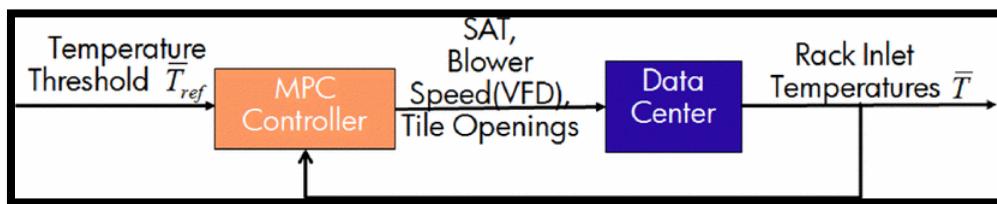


Figure 2.17 Control system structure (Zhou et al., 2011).

As a result, all these levels combined together and generated integration and optimization scheme using a MATLAB program. Meanwhile, experiments, within certain datacentre layout, were conducted to validate and evaluate controlling formulations. The results show valuable predictions for the behaviour of temperature, tile opening and blower power. The power could be minimized by 36% through efficient coordination between local need and zonal response.

Energy efficiency issues are discussed extensively by Pedram (2012) through a wide range of concepts. The research analysed system architectures by identifying the dependence of making a decision such as; processing task, power consumption and performance, thermal mapping and even Service Level Agreements (SLAs). The final step contains clarification between both power and thermal management through thermodynamic correlations. Due to the complexity of mission-critical data centres, for example; dynamic load and variety of meters, it was recommended to build-up hierarchical evaluation approaches by gathering performance versus energy at multi- level infrastructure.

2.5 Summary

At the end of reviewing all the above literature, it is obvious that the topic of energy efficient datacentres and optimum serviceability of microelectronics has been studied intensively. Four key points can be drawn from this chapter to summarise final conclusions:

1. Most studies dealt with improving cooling systems at the datacentre level. Those researches provided either new technology or enhance the available solutions. The targets for most of them were achieved through reducing the energy consumption of the cooling units.
2. Algorithms have been applied for better implementation and allocation of IT resources. Those resources could be used for higher performance and lower energy consumption if proper strategies were applied across the server and datacentre levels.
3. Since cooling solutions and IT strategies aimed to reduce energy consumption, then the overall energy consumption of the whole datacentre was reduced accordingly. For that, it is clear to conclude that nearly all conducted investigations were aimed to lower power usage and being an energy-efficient infrastructure.
4. Limited efforts have led toward holistic and energy efficient management of servers and datacentre levels. One of the important

reasons behind that is the complexity of interactions and responses between IT, cooling and power characterisations.

Finally, all researches are categorised in Table 2.1. Since cooling technologies and solutions were applied to either Server, Rack or Datacentre level, then the letters S, R and D are used correspondingly to each level, respectively. Meanwhile, IT solutions were applied either for server or datacentre level, so just the letters S and D are used to refer to each one, respectively. Furthermore, a small number of researches used certain metrics to compare and evaluate the presented solution, so it is listed in the 11th column. In addition, the next three columns specify the nature of research whether it was experimental works, numerical calculations or algorithmic solutions.

Table 2.1 Summary of Literature Review

Author(s)	Year	Cooling			Power			IT		Metrics and Benchmarks	Experimental	Numerical	Algorithm
		S	R	D	S	R	D	S	D				
Sharma et al.	2002	✓		✓						SHI/RHI		✓	
Bash et al.	2003			✓			✓			SHI/RHI		✓	
Hamann et al.	2006	✓			✓						✓	✓	✓
Hamann et al.	2006	✓			✓						✓	✓	✓
Moore et al.	2006			✓			✓		✓		✓		✓
Fan et al.	2007			✓			✓		✓				✓
Choi et al.	2008	✓						✓				✓	✓
Anandan and Ramalingam	2008			✓			✓						
Sharma et al.	2009			✓			✓			WUE		✓	✓
Pelley et al.	2009			✓			✓						✓

Cont. Table 2.1 Summary of Literature Review

Author(s)	Year	Cooling			Power			IT		Metrics and Benchmarks	Experimental	Numerical	Algorithm
		S	R	D	S	R	D	S	D				
Cho et al.	2009			✓			✓			COP		✓	
Singh et al.	2009						✓		✓		✓		✓
Wang et al.	2009						✓		✓		✓		✓
Farkhani and Mohammadi	2010	✓			✓			✓		SPEC CPU 2006	✓		✓
Oh et al	2010	✓			✓			✓			✓		✓
Breen et al.	2010	✓	✓	✓	✓	✓	✓				✓		
Walsh et al.	2010	✓	✓	✓	✓	✓	✓			COP		✓	
Abbasi et al.	2010						✓		✓		✓		✓
Chen et al.	2010			✓			✓		✓				✓
Wang et al.	2010		✓	✓					✓		✓		✓

Cont. Table 2.1 Summary of Literature Review

Author(s)	Year	Cooling			Power			IT		Metrics and Benchmarks	Experimental	Numerical	Algorithm
		S	R	D	S	R	D	S	D				
Cremonesi et al.	2011	✓					✓		✓				
Zhou et al.	2011			✓			✓		✓		✓		✓
Breen et al.	2012		✓			✓				COP	✓		✓
Subramaniam and Feng	2012				✓			✓		TGI, EE & REE		✓	✓
Almoli et al.	2012		✓	✓								✓	
Iyengar et al.	2012	✓	✓	✓							✓		
David et al.	2012		✓	✓							✓		
Ouchi et al.	2012			✓			✓			PUE	✓	✓	
Stafford et al.	2012			✓								✓	✓
Pedram	2012			✓			✓		✓			✓	✓

Cont. Table 2.1 Summary of Literature Review

Author(s)	Year	Cooling			Power			IT		Metrics and Benchmarks	Experimental	Numerical	Algorithm
		S	R	D	S	R	D	S	D				
Hopton and Summers	2013	✓	✓		✓	✓					✓		
Garimella et al.	2013	✓		✓	✓			✓	✓			✓	
Brady et al.	2013			✓			✓			PUE		✓	✓
Sampath	2013	✓			✓						✓	✓	
Pandiyan	2013	✓			✓						✓	✓	
Hassan et al.	2013			✓			✓					✓	
Almaneea et al.	2014	✓	✓							PUE		✓	
Chi et al.	2014	✓	✓	✓	✓	✓	✓	✓	✓	PUE	✓	✓	
Beitelmal and Fabris	2014	✓		✓	✓			✓	✓	SEPM & DCEPM			✓
Möbius et al.	2014	✓			✓			✓					

Cont. Table 2.1 Summary of Literature Review

Author(s)	Year	Cooling			Power			IT		Metrics and Benchmarks	Experimental	Numerical	Algorithm
		S	R	D	S	R	D	S	D				
Cho et al.	2014			✓						SHI/RHI		✓	
Ham et al.	2015	✓		✓	✓		✓					✓	
Block et al.	2015	✓			✓			✓		SERT - SPEC	✓		✓
Li and Kandlikar	2015	✓	✓	✓									
Nada et al.	2015	✓	✓		✓	✓		✓		SHI/RHI	✓		
Tarutani et al.	2015	✓		✓			✓	✓				✓	✓
Fulpagare and Bhargav	2015			✓									
Cho and Kim	2016			✓			✓		✓			✓	
Tatchell-Evans et al.	2016	✓	✓		✓	✓					✓	✓	
Zotkiewicz et al.	2016			✓			✓		✓		✓		✓

Chapter 3 TOWARDS DIRECT LIQUID COOLING (DLC)

3.1 Introduction

During the last decade, the heat load density within datacentres has increased dramatically as shown in Figure 3.1 (ASHRAE, 2005, Joshi and Kumar, 2012). This increase has led to a production of a huge amount of waste heat from IT industry. Therefore, it is crucial to improve cooling solutions to manage the waste heat adequately.

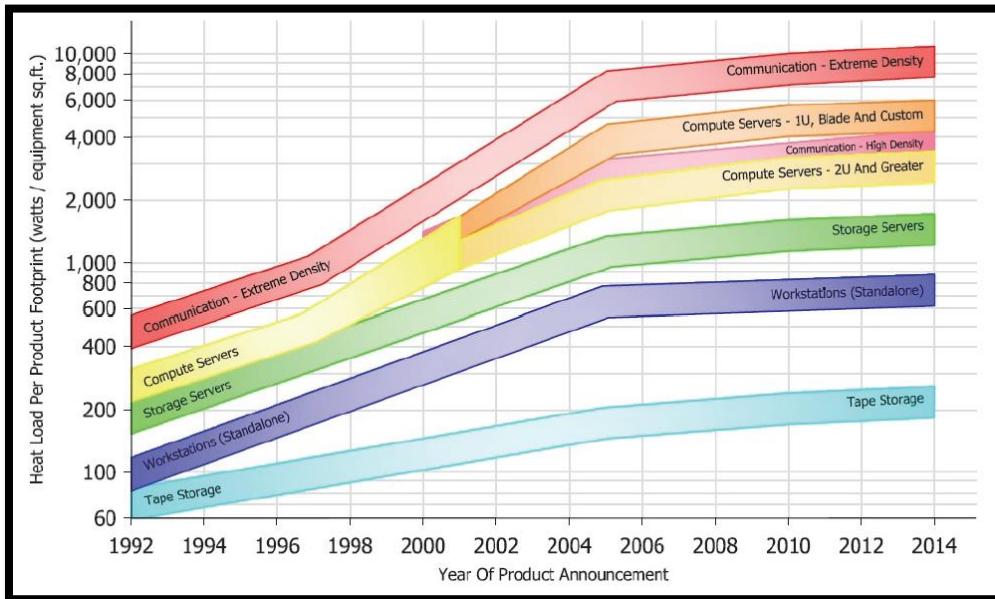


Figure 3.1 Heat load density rate (ASHRAE, 2005, Joshi and Kumar, 2012).

For that, various investigations focused on two main research categories of evaluation of cooling methods. The first aspect is dealing with the enhancement of traditional cooling technologies. For example, reducing air

leakage from hot to cold aisles has seen a large improvement toward reducing computer room air conditioning (CRAC) load (Niemann et al., 2011). The second aspect is introducing new cooling technology to sustain the higher power density of the new generation of microelectronics. The newer microelectronics produces a large amount of waste heat and also have much higher level of heat density (ASHRAE, 2005, Joshi and Kumar, 2012).

Figure 3.2 illustrates the well-known cooling methods which recommended to the corresponding Watt per square feet (Minas and Ellison, 2009). Therefore, the liquid cooling technologies are most appropriate beyond 5.5 kW/m² power density.

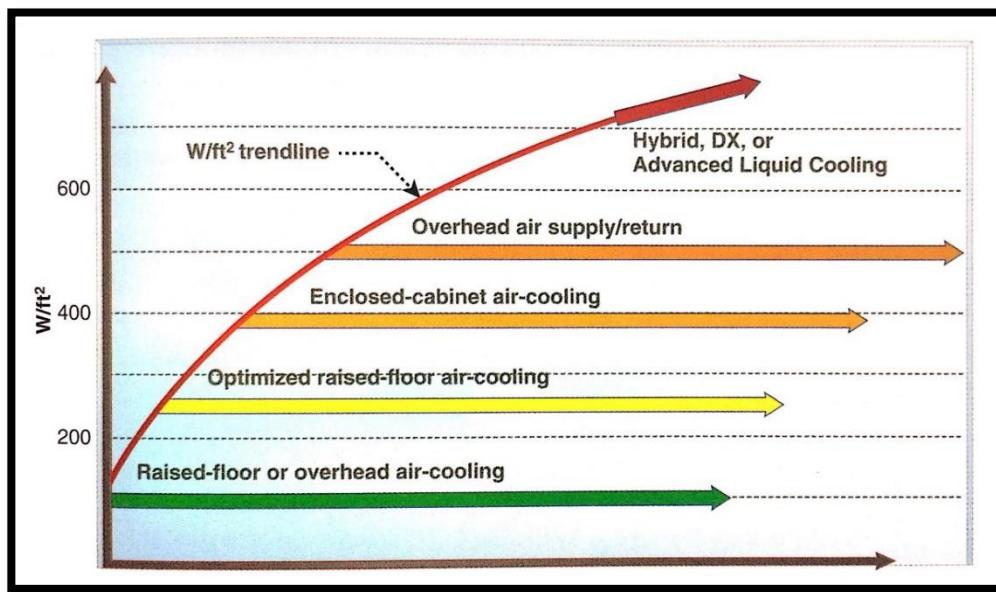


Figure 3.2 Appropriate cooling technology for various power density (Minas and Ellison, 2009).

The scope of this chapter is to present an experimental investigations and numerical simulation to compare and evaluate air and direct liquid cooling technologies. Therefore, the total air cooled server is benchmarked experimentally. Then, the same microelectronics architecture is retrofitted using direct contact liquid cooling. The retrofitted servers are benchmarked using a similar scenario. Meanwhile, both servers are simulated numerically in order to better understand the interaction between the cooling, energy and IT load.

3.2 Total Air Cooling (TAC) of Server

3.2.1 SunFire V20z Server

Sun Microsystems started its very first components within the workstation wars in the 1980s. At that time, Sun Microsystems presented their products as part of minimized cost to play an important role in the technical workstations industry. After two decades, in 2003, Sun Microsystems started an alliance with Advanced Micro Devices (AMD). This partnership aimed to use AMD's Opteron processor to present x86/x64 servers. During the next year, SunFire V20z server was introduced from Sun Microsystem as an Opteron-based server (SUNNYVALE, 2006). The SunFire V20z is a one rack unit (1U •) and has the physical specifications listed in Table 3.1.

Table 3.1 Physical Specifications (SunMicrosystems, 2008)

Parameter	Value
Height	42.93 mm ~ 1U
Width	430.276 mm ~ 10U
Depth	711.2 mm ~ 16U
Weight	15.88 kg

The location of the components within the server and the specifications are illustrated in Figure 3.3 and Table 3.2, respectively (SunMicrosystems, 2008).

The service processor (SP) is embedded in the server to give required management tools. These tools are essential to increase the availability of server and provide an independent operating system.

The operating system of the service processor is Linux and Secure Sockets Layer (SSL) encryption to increase the security of the connection. The connection is available via two dedicated 10/100 Megabits per second (Mbps) to ensure the connectivity of the server to all machines in the lab.

• The 1U is a rack unit height of 44.45 mm.

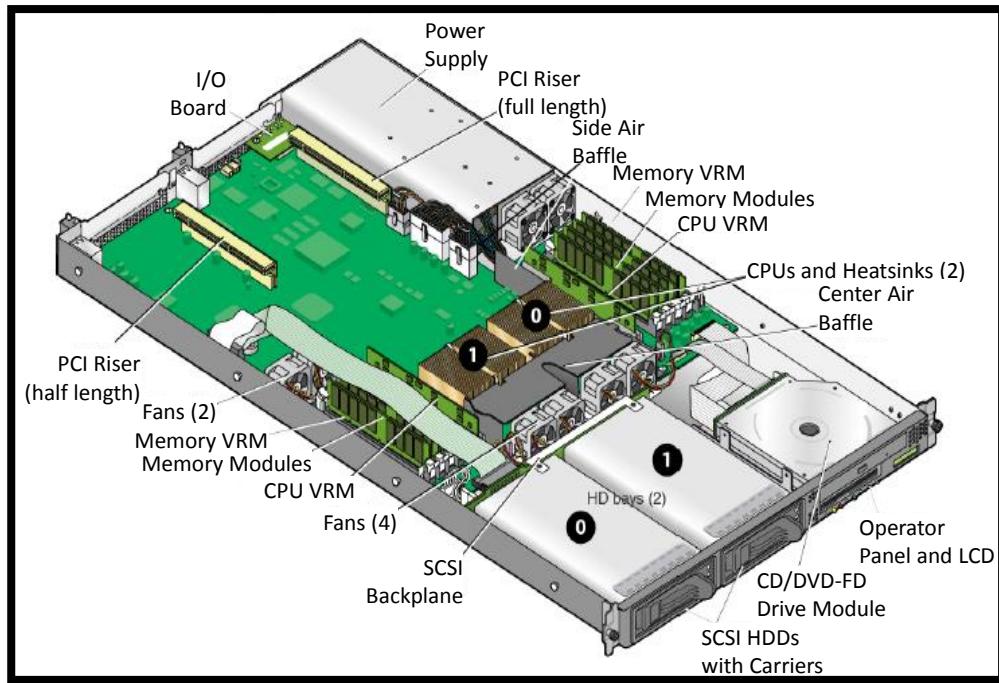


Figure 3.3 Illustration of SunFire V20z server (SunMicrosystems, 2008)

The SunFire V20z server is suitable for different information technology (IT) roles due to its capabilities. Those capabilities offer the required environments for hosting application platform and database workload. Furthermore, in high-performance compute (HPC) clusters and corporate datacentre are in need to such servers to perform the IT side of operations. The server has primary (or master) and secondary (or slave) central processing units (CPUs), named CPU0 and CPU1, respectively. Each CPU has its own banks of random access memory (RAM). The bank of four slots of a capacity range between 256 MB to 8 GB. Figure 3.4 shows the photo of SunFire 20z server.

Table 3.2 Features of SunFire V20z server (SunMicrosystems, 2008)

Component	Description
CPU	2 AMD Opteron processors, 64-bit x86 architecture.
RAM	8 slots, 512 MB to 16 GB ECC, registered DDR.
Hard disk drive(s)	1 or 2 U320 SCSI HDD(s).
Network I/O	Dual embedded Gigabit Ethernet.
Other I/O	Internal CD-ROM (or DVD-ROM), Diskette drive, Embedded SVGA Video, Keyboard and Mouse.

3.2.2 AMD Opteron

The central processing unit CPU used in SunFire V20z server is AMD Opteron, as mentioned earlier. It is an original equipment manufacture OEM microprocessor and part number details are given in Table 3.3 (AMD, 2006). The CPU is a dual core, 2 GHz frequency and 64-bit data width. The maximum operation temperature is 71 °C, and the thermal design power (TDP) of this CPU is 85.3 W. This TDP is dissipated through a copper heat sink attached to CPU. A thin layer of thermal paste is spread between CPU and heat sink to ensure perfect contact.



Figure 3.4 The SunFire V20z server.

Table 3.3 The coding system of the AMD Opteron CPU (AMD, 2006).

OSA270FAA6CB							
OS	A	270	F	A	A	6	CB
Brand AMD Opteron Server	Power limit is standard, 85.3 W	Model number of 2 GHz	Package 940 Pin Lidded OpPGA	Operation voltage is variable	Case temp is variable, specified 71 °C	Cashe size of 2 MB	Dual processor

3.2.3 Intelligent Platform Management Interface (IPMI)

The intelligent platform management interface (IPMI) represents a standardized computer system interface used for managing and monitoring operation conditions of server (Minyard, 2004). The interface is responsible for conducting communications between hardware and management subsystems. These communications are exchanged via controllers known as baseboard management controller (BMC) (ORACLE, 2007). The IPMI is connected to the control unit, which is a TOSHIBA laptop, to access the local host for both monitoring and logging data from sensors. Those sensors logging data from different components inside the server and report any failure in operation due to various reasons. The locations of sensors are marked in the IPMI window as shown in Figure 3.5.



Figure 3.5 Screenshot of IPMI window from the control unit.

The 8 components provided with temperature sensors are; ambient, RAM0, CPU0, CPU1, HDD, RAM1, Ethernet and service processor. In addition, the sets of cooling fans are provided with tachometer output sensor. Each sensor detects the desired values and compares it with the defined thresholds to ensure that the components are working properly.

3.2.4 Operating System

Ubuntu, in brief, is a Debian-based environment which is a Linux operating system (OS). Ubuntu 14.04 long term support (LTS) is chosen to withstand as the operating system of all servers. This operating system, Ubuntu, is selected due to its compatibility with cloud application and networking, and it is available for free. These two characteristics make this OS lead other OSs. According to the monitoring of *The Cloud Market*, the number of platforms running Ubuntu was more than seven times that using Windows during October 2016 (EC2, 2016).

As each OS has different sets of software, it is important to capture the operating system idle power. This type of idle power defines as the power consumed by the server after turning it on and complete the initialisation. In addition, the term baseline power refers, in this work, to the value of the power consumed by the server after completing processes/transactions/workloads. This power consumption can be used as more representative measurements since it reflects the change from utilisation level to active idle. The baseline power consumption is affected by various operating systems due to different configurations used in each OSs (Gough et al., 2015).

3.2.5 Cooling systems

Each server is cooled using a set of 8 fans distributed inside the server as shown in Figure 3.6.

1. 2 fans attached to the power supply unit (PSU). These fans are running even with pre-operation status in the low rotational speed of 4150 rpm. This speed is enough to maintain air circulation inside the PSU. However, the speed is increase to 11600 rpm during normal operation of the server. This allows the PSU to be kept cool within service threshold. In addition to this role, these 2 fans help in cooling down the primary bank of RAMs.

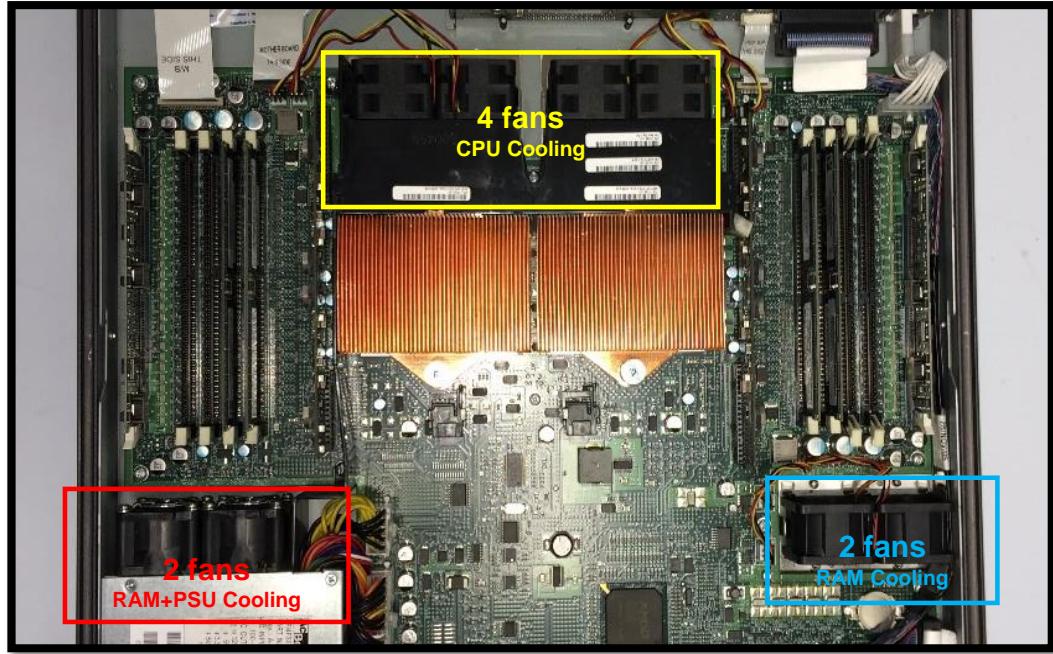


Figure 3.6 Illustration of fan locations inside the server.

2. 4 fans are responsible for cooling the central processing unit (CPU) down (2 fans per CPU). The air stream is sucked by these set of fans and passed over the heat sink. To ensure the maximum air stream is forced to pass through the fins of the heat sink, plastic baffles are used as shown in Figure 3.7.

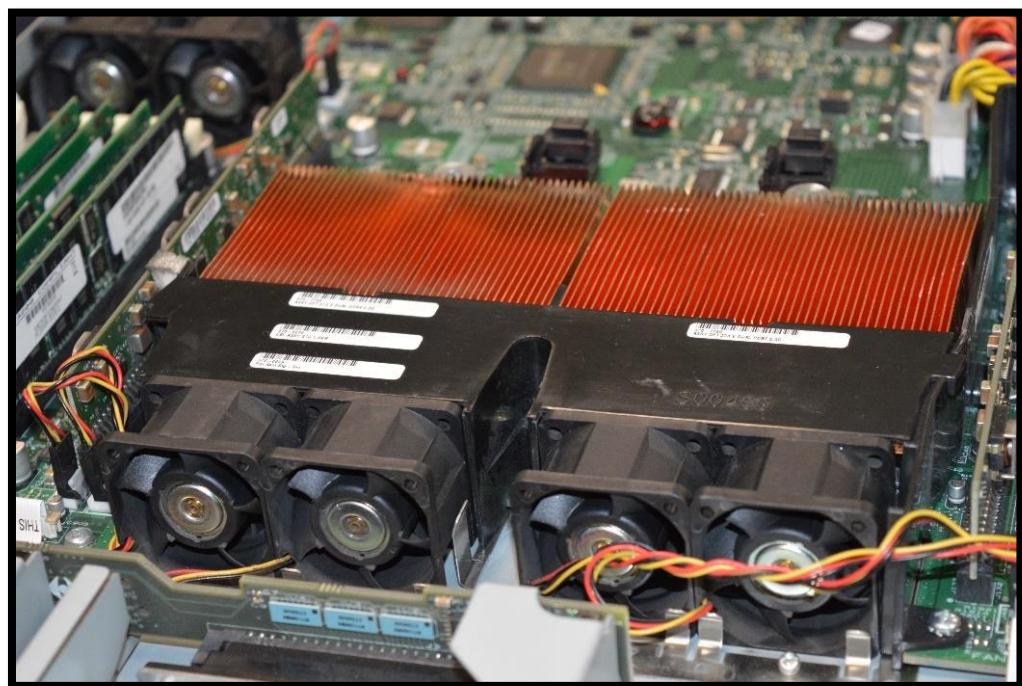


Figure 3.7 Cooling fans, heat sinks and baffle for CPUs.

3. 2 fans are used to drive a certain amount of cold air to pass through the secondary bank of RAMs.

The fan is one of DELTA series under the model FFB0412SHN. It has the specification listed in Table 3.4 (DELTA, 2015).

Table 3.4 Specifications and coding system of the Delta FFB0412SHN fan

Rated Voltage	Operating Voltage	Input Current	Rated Speed	Max Air Flow	Noise (dBA)
12 V	4.5 – 13.8 V	0.5 A	13,000 rpm	0.68 m ³ /min	54.5*
FFB0412SHN					(R00/F00)
FFB	04	12	SH	N	R00
Series name	Square frame of 40mm edge	Rated voltage	Super high speed	28mm Frame thickness	Failure detector
					Speed sensor

3.3 Direct Liquid Cooling (DLC) of Server

The SunFire V20z server was retrofitted with a direct liquid cooling (DLC) system. The DLC system means to bring the liquid, as working fluid, inside the server to cool down the electronic components. The electronic components play the role of “heat source” and produce heat due to power consumption to perform IT load. Then, this heat is rejected away into specified “heat sink” depending on cooling method. For the air-cooled server, the heat is rejected to a combination of heatsink and set of fans. Whereas for liquid-cooled server, the heat rejects to the cold plate.

CoolIT, which is a company founded in 2001 in Calgary, Canada, presents a direct contact liquid cooling (DCLC) technology. The DCLC represents a particular case of DLC in which there is a “physical contact” between the “source” and the “working fluid” via cold-plate. The CoolIT DCLC system consists of a pump, cold-plate, reservoir, magnetic multi-platform mounting

* This level is equivalent to clear conversation or air conditioner according to BRUNEAU, M. 2013. *Fundamentals of acoustics*, John Wiley & Sons.

kit, tubes and Stäubli fluid quick connector (CoolIT, 2011, CoolIT, 2012). Figure 3.8 shows the schematic diagram for the DCLC system. The retrofitted server with the illustration of liquid loop is presented in Figure 3.9.

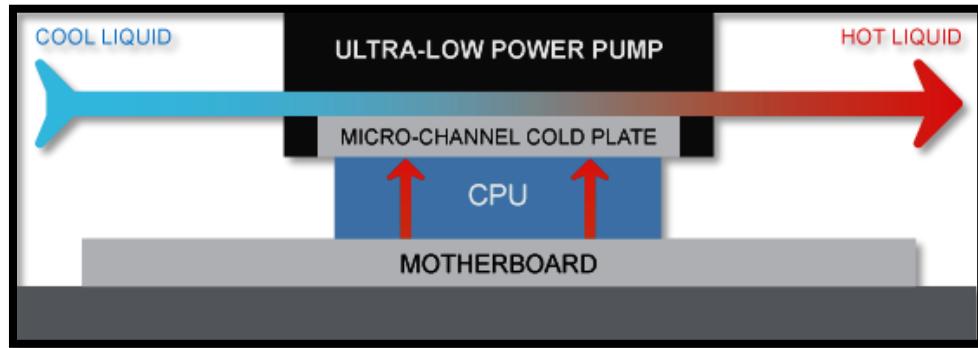


Figure 3.8 Direct contact liquid cooling DCLC (CoolIT, 2013)

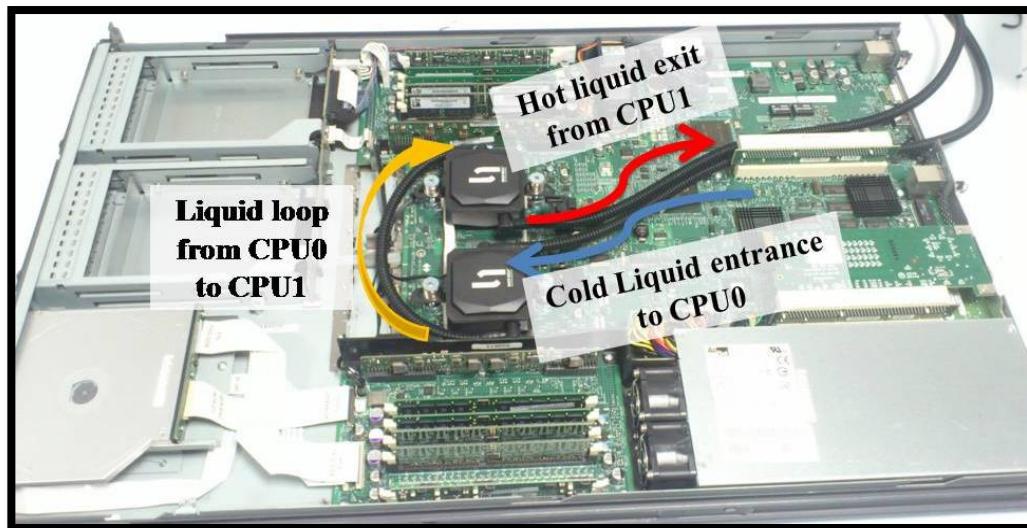


Figure 3.9 The retrofitted Sun Fire V20z Server including DCLC.

3.3.1 CoolIT™ Pump

The cold plate technology uses a brushless electronically commutated motor (ECM). The ECM converts the direct current (DC) via integrated inverter power supply into alternate current (AC) (Yedamale, 2003). The inverted AC has high efficiency performance and imply square current wave instead of sinusoidal wave. The nominal DC voltage of the pump is 12 V and the nominal current is 0.12 A, therefore, the power consumed by the pump is 1.44 W. The nominal and maximum flow rate are of 0.57 and 0.90

litre/minute, respectively. The acoustic noise of the pump is less than 23 dBA* measured at nominal speed of 2200 rpm.

The pump is made of semi-crystalline engineering thermoplastic called Polyphenylene Sulfide PPS reinforced with 30% glass filler. While the impeller is made from Poly Butylene Terephthalate PBT which is synthetic thermoplastic with 30% glass filler. The case, which represents the reservoir, is made of Acrylonitrile Butadiene Styrene ABS due to its high toughness and easy formation properties.

Moreover, all components of the pump are made of UL94V-0 material which refers to flammability rating. This rating describes the material as one that will not burn with direct flame combustion for 50 seconds nor drip flaming particles. So, the working condition of this material provides a high rate of safety inside the server. Figure 3.10 shows the schematic diagram and dimensions of the pump, and Figure 3.11 presents the operation curve of the pump at nominal and maximum conditions.

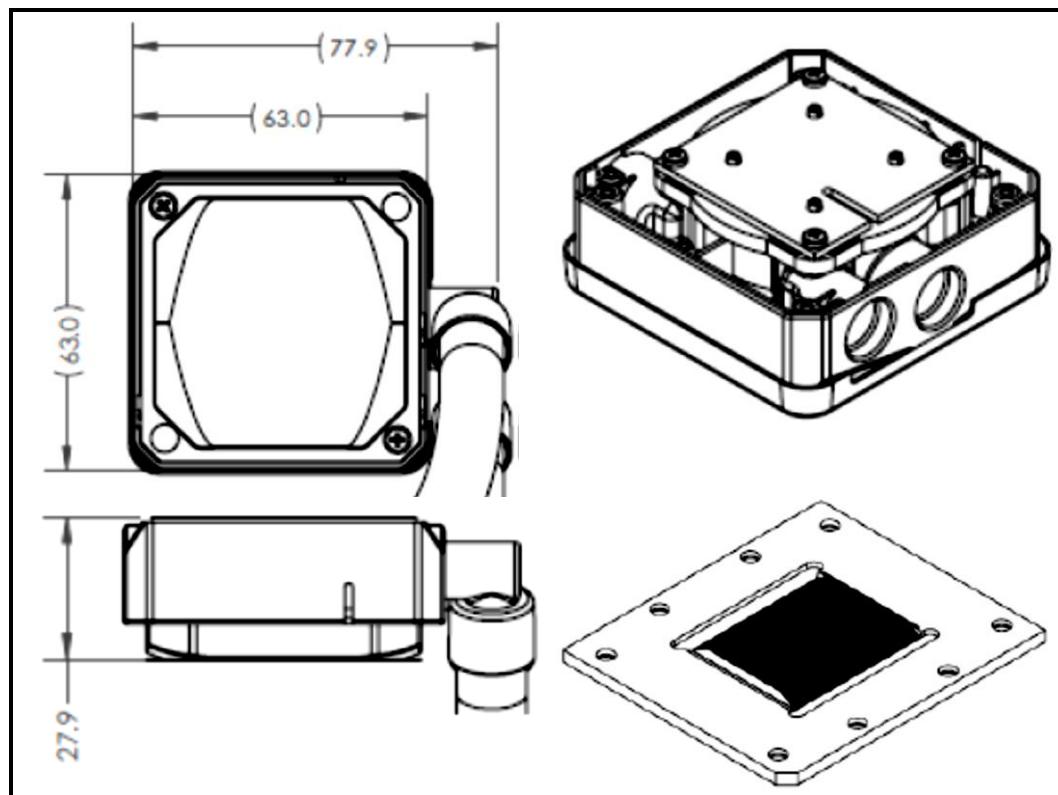


Figure 3.10 Schematic diagrams of the pump (CoolIT, 2012).

* This level is equivalent to whispering at 5 feet.

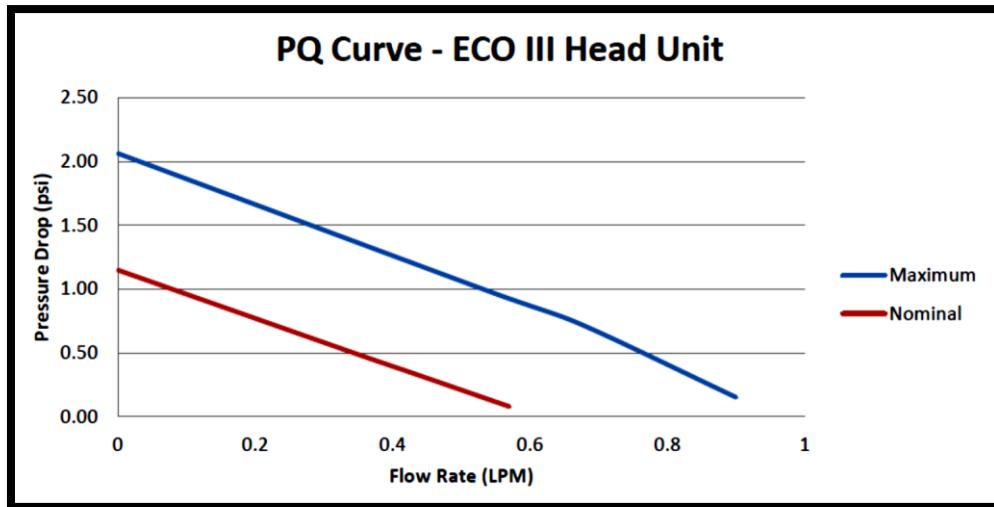


Figure 3.11 Operation curve of the pump (CoolIT, 2012).

The pump encapsulated inside the reservoir where the coolant circulated in contact with the copper cold plate through a microchannel. Two pumps are retrofitted instead of the heatsink and fan cooling system as presented in Figure 3.12.

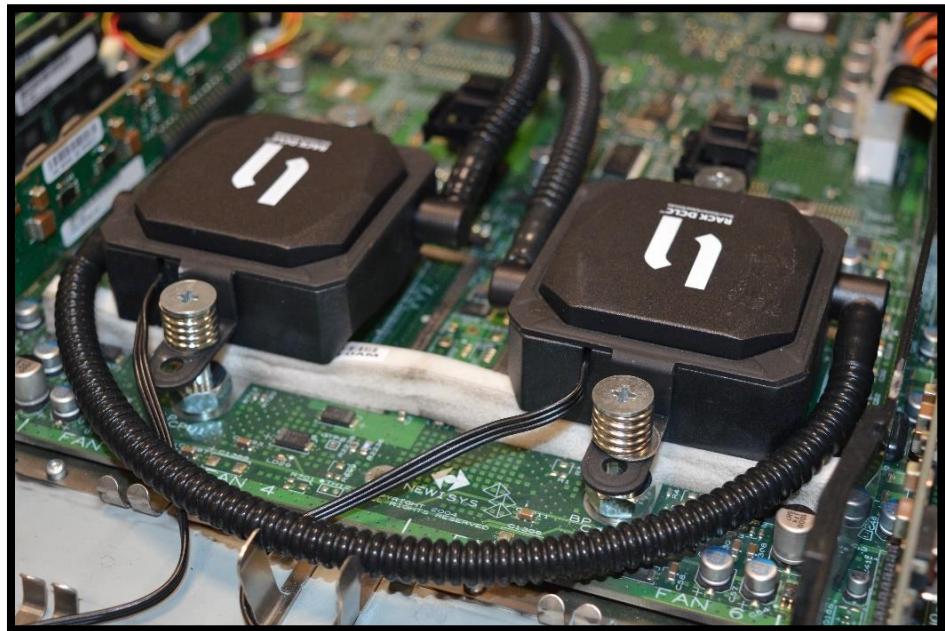


Figure 3.12 The retrofitted DCLC system by CoolIT.

Finally, it is important to mention that the pumps connect in series. This connection provides $n+1$ redundancy for the cooling system. So, even though failure happens to one of the pumps, the other one still circulates the flow. The flow is reduced to 25% when failure occurs to the pump.

3.3.2 The Tubing

The three parts of tubes are made of flexible corrugated Fluorinated Ethylene Propylene (FEP) material. The FEP is chemically inert to various industrial liquids, and offers non-flammability and thermal stability. The first tube connects the first pump, which cools the master CPU, with the intake of cold liquid. The second tube connects both pumps, whereas the third tube carries the hot liquid from the second pump to the heat rejection system. The internal and external diameters of the tube are 6 mm and 8 mm, respectively.

3.3.3 The Coolant

The working fluid in the DLC and DCLC systems is a liquid mixture. This liquid mixture consists of the following:

1. 86% water.
2. 13% low toxicity Propylene Glycol.
3. 1% anticorrosion and antifungal additions.

The freezing point of the coolant is -4.7°C , whereas the boiling temperature is 100°C . The mixture has specific gravity and specific heat of 1.011 and 4.075 kJ/kgK , respectively.

3.3.4 The External Fans and Radiator

Cooling of single server unit requires a heat rejection system. This system allows the coolant to dissipate the heat and cool down before flowing back to the cold plate again through the tubing connection. The system consists of six fans, with separate powering control, and a radiator. The fan is made of Poly Butylene Terephthalate PBT with 30% glass filler. The rated voltage and current of the fan are 12 V and 0.2 A, respectively, so the nominal power required is 2.4 W at speed of 1800 rpm. The fan schematic diagram

and configurations are illustrated in Figure 3.13 and Figure 3.14, respectively.

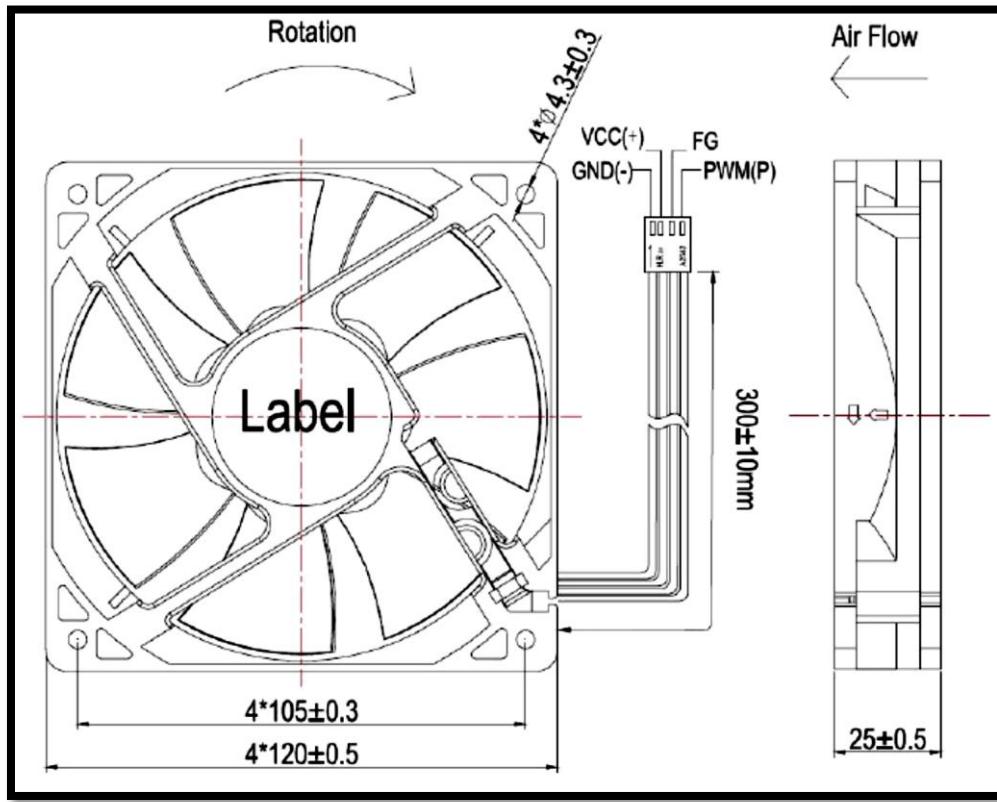


Figure 3.13 Dimensions of the external fan (CoolIT, 2011).



Figure 3.14 The fan-radiator cooling system.

3.3.5 Stäubli Connectors

Stäubli was founded in 1892 in Horgen, Switzerland. This company serves a wide variety of industrial connectors from pharmaceutical to aeronautical

engineering. These connectors provide an innovative customised connection (Stäubli, 2016). It is a versatile connector used in dry quick-connect technology. This technology provides essential connection for liquid cooling of microelectronics, where there is no tolerance for any liquid leakage.

In 2013, CoolIT announce their partnership with Stäubli to present direct liquid cooling industry standards (Massolin, 2013). The two companies put their potential together to produce consistent design of connectors to ensure reliable passages for coolant inside the servers.

3.4 Preliminary Test of the Server

3.4.1 Fan Testing

Each fan has a speed sensor that gives tachometer output through the IPMI. Although the operation curve for the fan is available in (DELTA 2015), no specific details are available about the correlation between fan speed and power consumption. Therefore, an experimental test includes running the fan using a controllable power supply within the ranges of 4.48 – 13.83 V and measuring the power consumption at each level.

Simultaneously, a digital tachometer, iParaAiluRy shown in Figure 3.15, is used to capture the rotational speed of the fan. This digital tachometer sends a laser beam to the blade of the fan and capture the reflection of the laser beam. One reflection occurs on a silver light-reflective sticker per revolution. The silver sticker attaches to one fan blade. Depending on the number of laser signal detected, an indicator of the number of revolution per minute (rpm) is recorded by the device. The resolution of the tachometer 1 rpm is capture with an error rate of $\pm 0.05\%$. The specifications of the tachometer is listed in Table 3.5 (iParaAiluRy).



Figure 3.15 The iParaAiluRy digital tachometer.

First, the measured rotational speed N using tachometer compares with the values captured by the IPMI. Utilization level increases for four different levels of 25 %, 50%, 75% and 100% to ensure the validation of measurements. Figure 3.16 shows high agreement between the two values which is 99.95%.

Table 3.5 Specifications of iParaAiluRy digital tachometer.

Parameter	Value	
Size	131 x 70 x 29 mm	
Display	5 digits, 18 mm LCD	
Range	2.5 to 9,999 rpm	Over 10,000 rpm
Resolution	0.1 rpm	1 rpm
Accuracy	$\pm(0.05\% + 1 \text{ digit})$	
Detecting Distance	50 to 500 mm	
Sampling Time	0.8 sec	
Operation Temperature	0 to 50 °C	

Next step includes collecting data of the running fans. Three random fans are selected to perform this test to ensure the reliability of results. The results gather, as illustrated in Figure 3.17, and represent a polynomial

relationship between the fan speed N and corresponding power consumption P_{fan} as shown:

$$P_{fan} = a_1 N + a_2 N^2 + a_3 N^3 \quad (3.2)$$

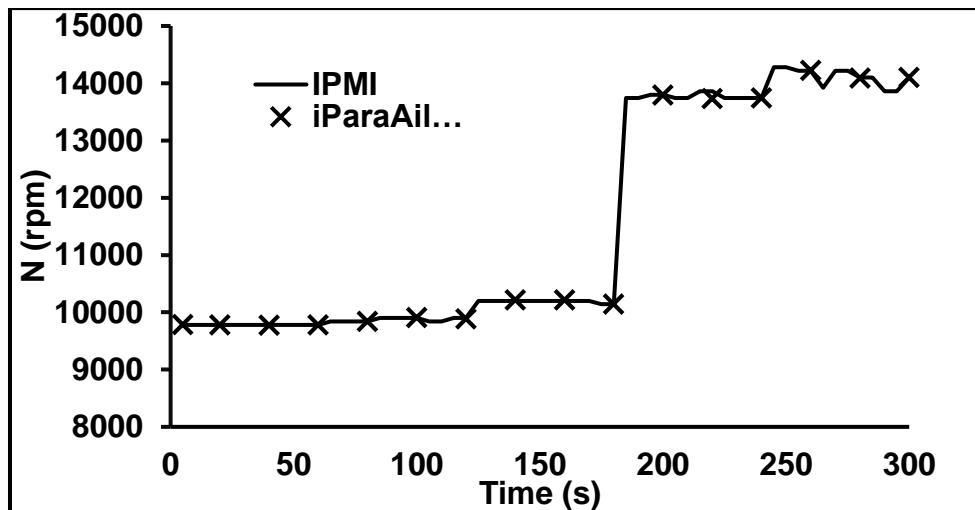


Figure 3.16 Agreements of the measured speed and captured speed by IPMI tools.

To find out the parameters a_1 , a_2 and a_3 , the collected data was analysed using non-linear regression. The estimation is based on Gauss-Newton method. This method, in brief, assumes an initial approximation of values and then performs linearization around this selected value. This requires a minimizer function ζ to achieve convergence as:

$$\zeta = \sum_{i=0}^{i_{max}} \left(r_i(x^{(k)}) + \nabla r_i(x^{(k)})^T (x - x^{(k)}) \right)^2 \quad (3.3)$$

Where k is the number of current iteration, x is the values of current approximation, and r is the evaluated function in matrix form at the assumed guess values. Once convergence occurs, after several iterations, the empirical constraints are ready to compute as shown in Table 3.6, with a proportion of variance ($R = 0.993$).

Table 3.6 Parameters prediction for equation (3.2).

Parameters	a_1	a_2	a_3
Values	2.46×10^{-4}	-3.70×10^{-8}	3.11×10^{-12}

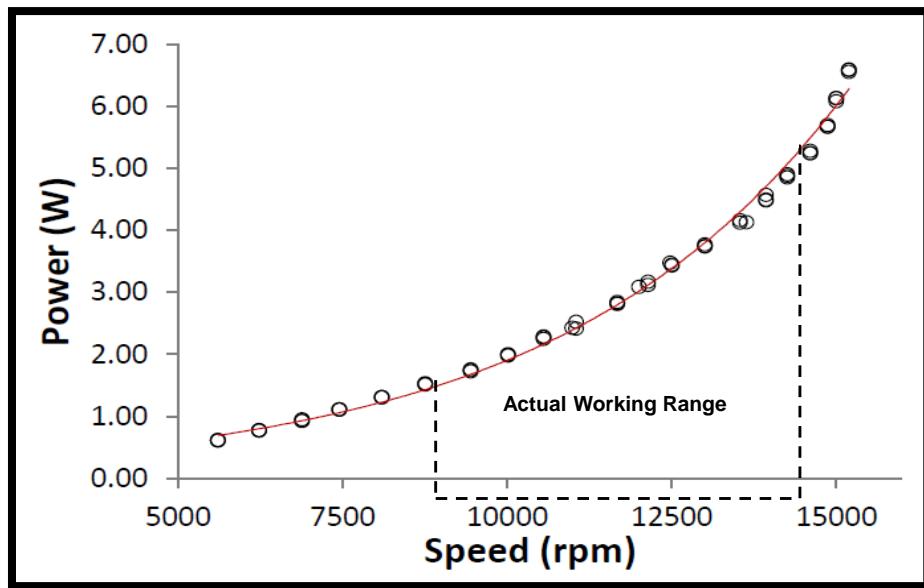


Figure 3.17 Fan power measurement versus rotational speed

3.4.2 Server Testing

3.4.2.1 Prepare Stress Linux Test

The Stress Linux is used to provide the server under test SUT with a range of loading levels. The compatibility of Stress Linux with the command line makes the writing of scripts fast, easy and flexible to achieve the target loading. For example, the CPUs and virtual machines VMs loading is illustrated in Table 3.8. Loading the CPUs and VMs in such a combination provides results about the effect of IT performance on power and cooling loads. Composing configurable command line options in this way provides a mechanism of the response of the SUT to parameter change.

The control server connects to the SUT via local network according to the configuration illustrated in Figure 3.18. Since the control server connects remotely, the Stress Linux script is performed through a secure channel. This secure channel provides a secure shell (SSH) to send a command-line remotely. Meanwhile, the control server connects to the IPMI of the SUT for monitoring purposes. The IPMI exports temperature of eight different

locations, for example those of the CPU. In addition, the tachometer sensors record the rotational speed for all the fans.

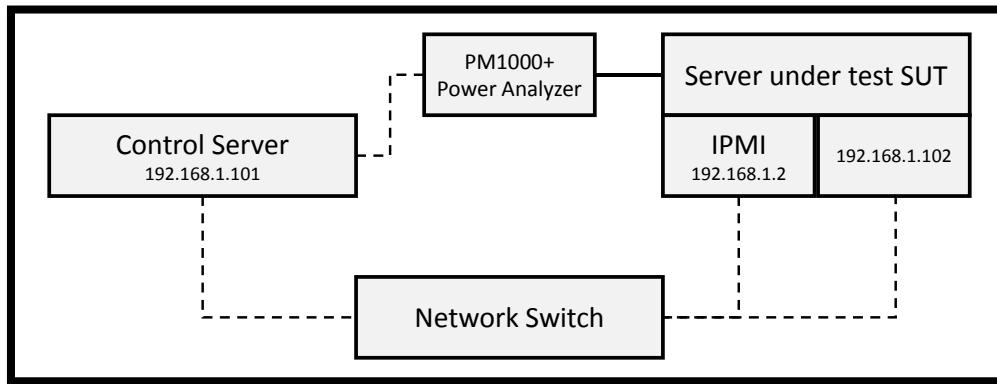


Figure 3.18 The network configuration.

Preliminary tests represent the first phase of experimental work. The results collection is based from about 40 hours of live monitoring. In this phase, there are sets of tests to the SunFire V20z as described in section 3.2. This server run in full load status using an IT loading program, a burn in test approach. In this work Stress Linux is used to perform the server loading tests, due to its faster response time, even though it has some problems dealing with error codes (Kanoun et al., 2005). The test rig of this stage consists of test server, control and monitoring server system, power supply and power measurement system and external radiator cooling system as shown in Figure 3.19.

During these tests, it was planned to measure both temperature and load. Therefore, temperature is measured by the built-in sensors and reported by Intelligent Platform Management Interface IPMI. However, to ensure the reliability of measurements, each experiment was repeated three times. The repeated results indicate the temperature differences at each level of utilisation as shown in Figure 3.20. Meanwhile, the server power load is measured using a direct power meter.

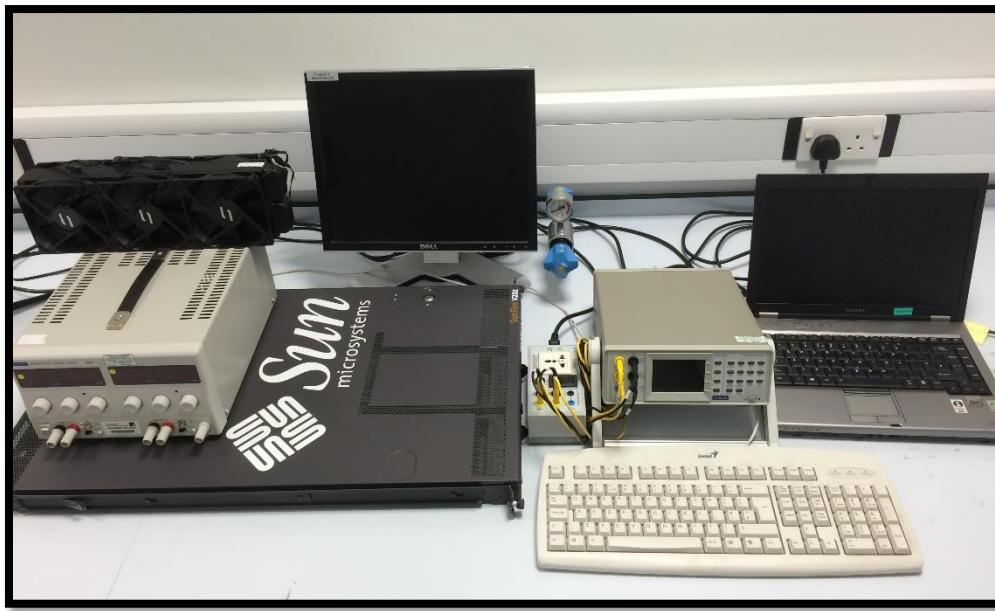


Figure 3.19 Test rig of preliminary experimental stage.

The second step measures the temperature of the retrofitted DCLC system. There is an external fan and heat exchanger to give a closed loop with the pump. These external fans help to quantify the effect of heat exchanger cooling rate by controlling the number of running fans, from natural to forced heat transfer, on the performance of system.

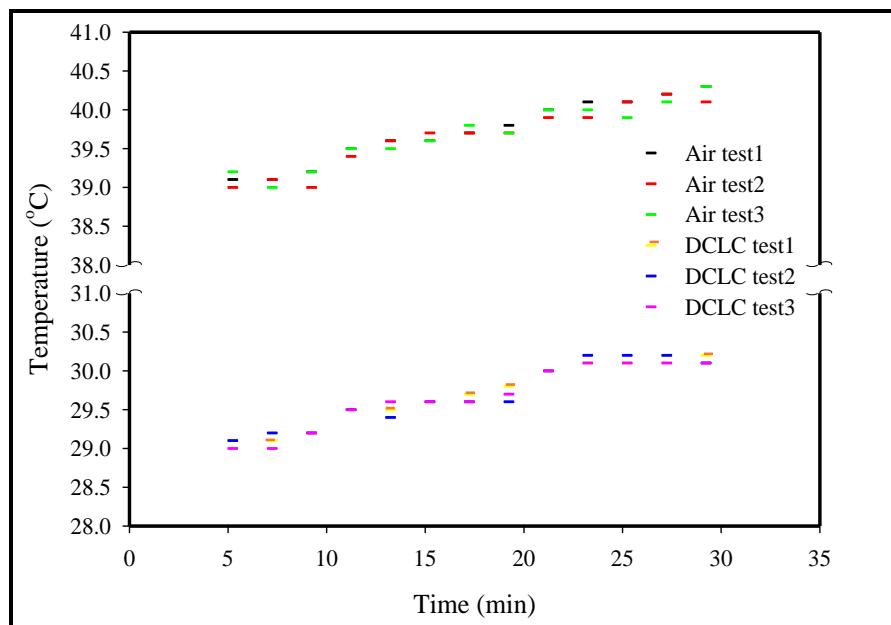


Figure 3.20 Repeated tests of the server under air and liquid cooling conditions.

It is possible to compare results obtained from air and liquid cooling based on a variety of parameters and configurations, i.e. temperature, fan speed and power. The testing procedure of the servers are;

1. Get connection to the intelligent platform management interface (IPMI) of the server from control unit.
2. Run the server in the idle condition to capture baseline power.
3. Monitor temperature of the eight points, ambient, RAM0, CPU0, CPU1, HDD, RAM1, Ethernet and Service Processor (SP).
4. Running the server at two cases of loading by using the command stress which is a LINUX based command. Cases are setup that stress the CPU and stress the virtual memory (VM). Each case lasts 40 minutes, the first 30 minutes under stress from the idle state and the last 10 minutes after stress are used to monitor the temperature drop behaviour. The command lines used are listed in Table 3.7.
5. Run the server at different cases of loading by using StressLinux also. The cases include stressing the CPU and VM separately for three levels. Then, two more cases include stressing both CPU and VM for two combinations. The cases last 6 minutes each and the commands are listed in Table 3.8.

Table 3.7 Stress command line form

Unit	Stress Command Line
CPU	stress --cpu 4 --timeout 30m
VM	stress --vm 2 -vm-bytes 1500M --timeout 30m

Table 3.8 The cases used in the preliminary stage of experiments.

Case	Stress Command Line
1	stress --cpu 4 --timeout 6m
2	stress --vm 2 -vm-bytes 1500M --timeout 6m
3	stress --vm 4 -vm-bytes 1500M --timeout 6m
4	stress --vm 8 -vm-bytes 1500M --timeout 6m
5	stress --cpu 4 --vm 2 -vm-bytes 1500M --timeout 6m
6	stress --cpu 4 --vm 4 -vm-bytes 1500M --timeout 6m

3.4.2.2 Results and Discussion of Air-Cooled and DCLC Server

Each monitoring case lasted 30 minutes idle and followed by 10 minutes at power off status. The overall behaviour is quite similar, rapidly increasing in temperature following application of the load, followed by a relatively constant phase after approximately 5 minutes. Once the load is removed there is a sharp drop in the temperature. The CPU0 recorded the highest temperature in both cases and that is very logically due to responsibility of this CPU to carry-out the “no-load” during idle status. For the air-cooled server, the temperature reached 40.3 °C as shown in Figure 3.21. Whereas in the retrofitted DCLC server, the temperature is just 30.2 °C as shown in Figure 3.22. From a related point of view, the temperature of Ethernet and SP recorded much higher temperatures in the air-cooled server. The level of differences is 5-6 °C, and that is due to the impact of hot air, which leaving heat sink, on Ethernet and SP component.

RAM1 reached higher temperature in the air-cooled case also, and this because the fans of CPUs add more suction and reduce the air flow towards RAM1. RAM0 and HDD do not affect by changing the cooling techniques.

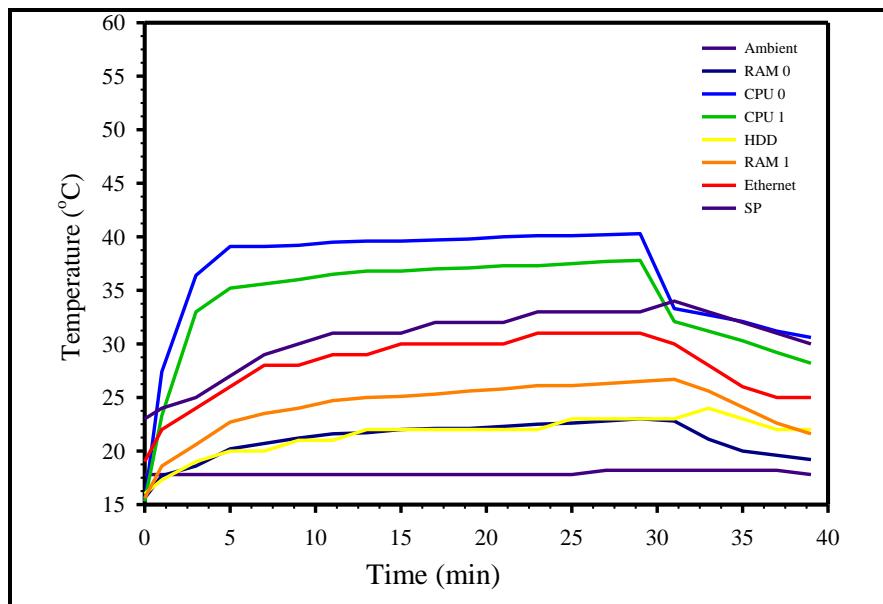


Figure 3.21 Temperature monitoring during idle state for air-cooled server.

During stressing the CPUs, the temperature goes higher to 50 °C level because of the load on both CPU0 and CPU1 Figure 3.23. Meanwhile,

RAM0 and RAM1 sharing part of load, and their temperature being higher also. RAM0 recorded maximum temperature of 36.8 °C.

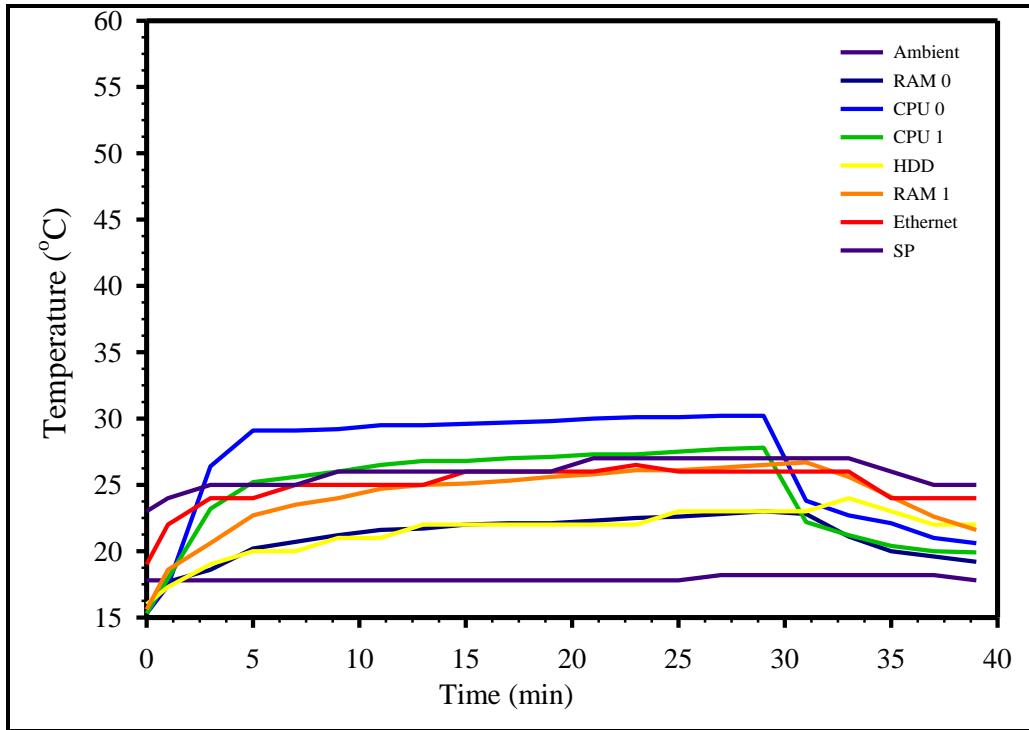


Figure 3.22 Temperature monitoring during idle state for retrofitted DCLC server.

From the other hand, during stressing the VMs, CPUs do not reach the previous level. Since the test does not employ CPU's resources, as shown in Figure 3.24.

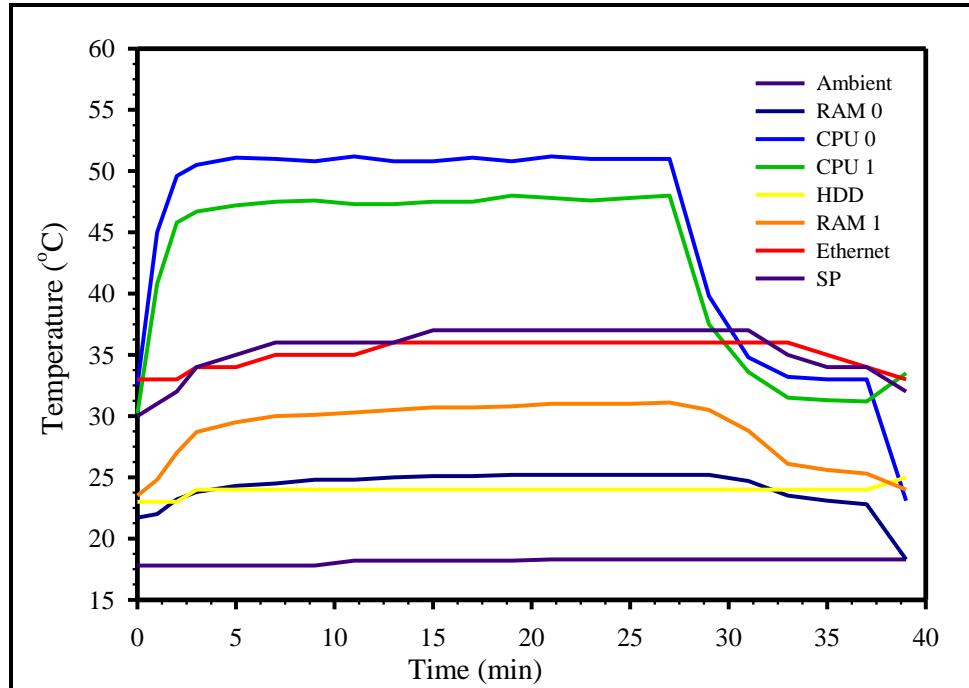


Figure 3.23 Stress CPUs state for air cooled server

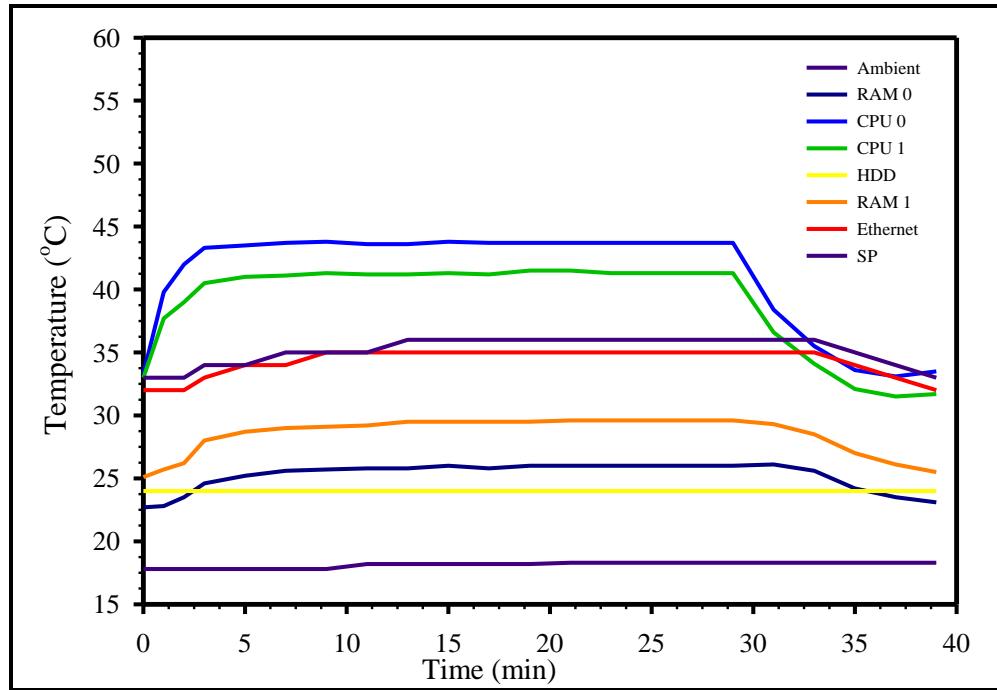


Figure 3.24 Stress VMs state for air cooled server.

By adopting the DCLC cooling, the CPUs and VMs temperatures are still within 29.4 °C by different loading status. Figure 3.25 and Figure 3.26 present temperature monitoring during CPUs and VMs stresses, respectively.

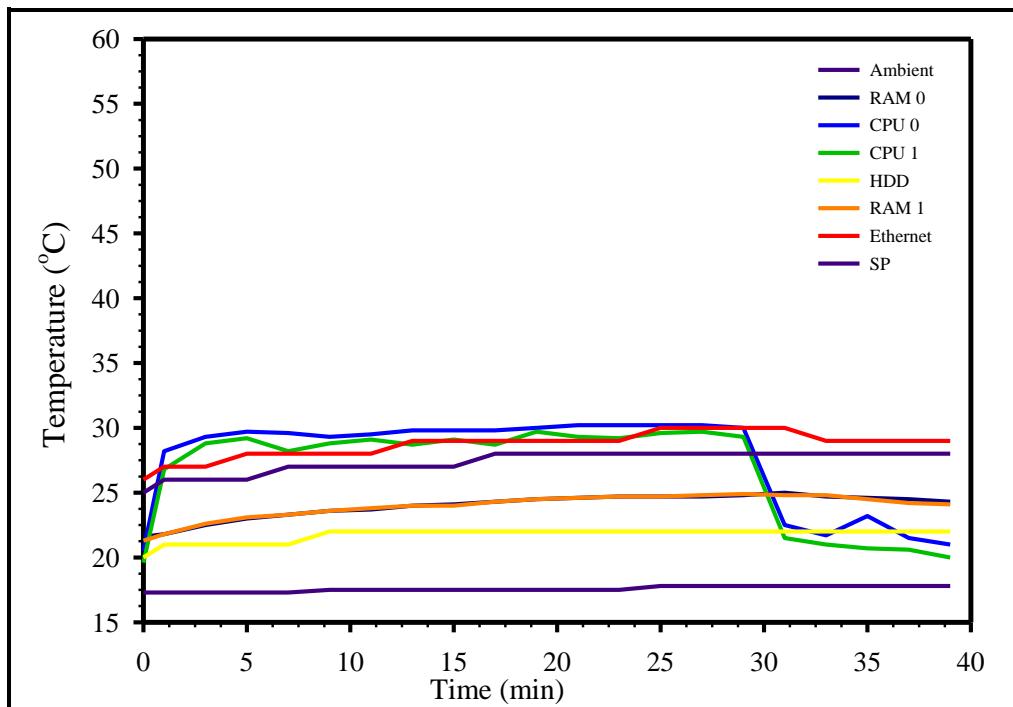


Figure 3.25 Stress CPUs state for DCLC server.

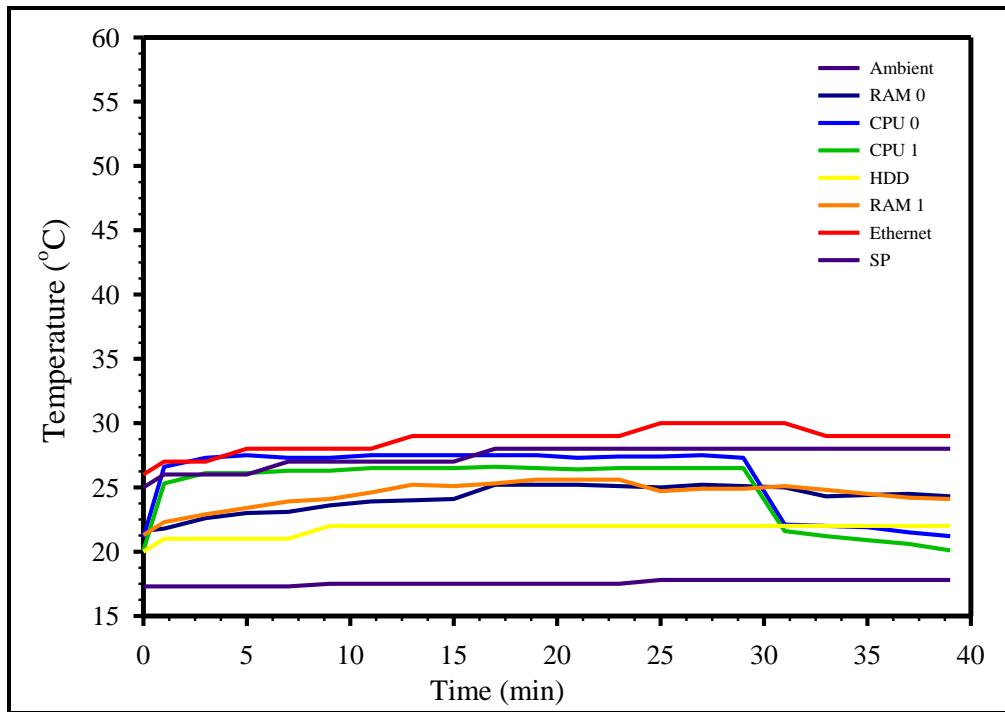


Figure 3.26 Stress VMs state for DCLC server.

Finally, Figure 3.27 and Figure 3.28 compare the highest recorded temperature for RAM0, CPU0, CPU1 and RAM1 during different type of loading. The setting of the stress test as per Table 3.8 is conducted to present the impact of cooling methods on the temperatures of the component. The DCLC maintains the temperature within 23.3-40.1 °C. This is very important in case of thermal design and packing of electronics (Remsburg, 2011, Remsburg, 2000).

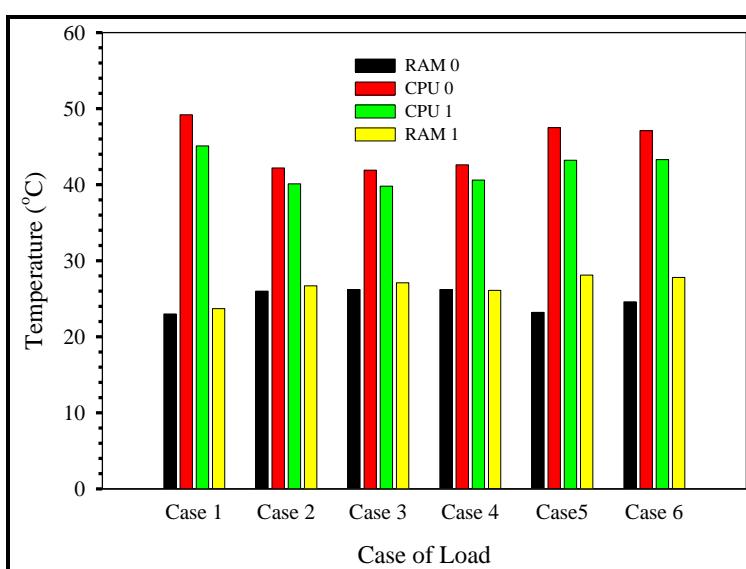


Figure 3.27 Temperature comparison of CPUs and VMs during different stress cases for air cooled server

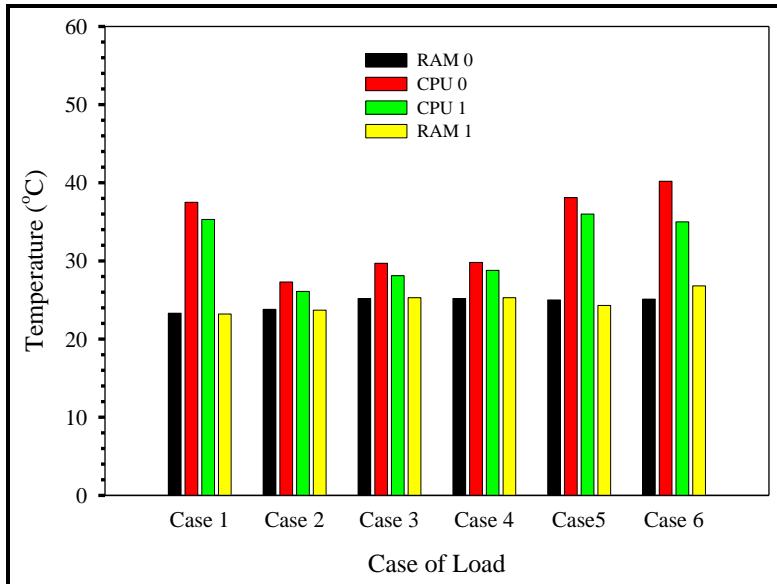


Figure 3.28 Temperature comparison of CPUs and VMs during different stress cases for DCLC server

3.5 Experimental Work

In this part, the air-cooled server is being under test to create a standard benchmark. This benchmark is used to evaluate the effect of the DCLC retrofitted server. The SPECpower_ssj2008 is selected to perform the benchmarking due to its availability and configurability.

3.5.1 SPECpower_ssj2008

The Standard Performance Evaluation Corporation (SPEC) is a non-profit group dealing with power consumption of servers. This corporation includes member from a wide variety of researcher and contributors among many academic institutions and industry. The SPEC has produced Java based scripts as a benchmark tool. This framework called SPECpower_ssj2008 and it is used to measure power consumption and performance of servers in 2007. SPECpower_ssj2008 consists of three software, which are installed on server under test SUT and the control server, run together to perform the required measurements (SPEC, 2012b, SPEC, 2016):

1. Server Side Java (SSJ): The SSJ workload tests different aspect of servers via Java designed script. Using Java is preferable due to its scalability and portability across various operation systems OSs. This workload can target CPUs, memories and virtual machines VMs. This SSJ exercises the required components with range of configurable characteristics such as load level, increments and time intervals. It is important to mention that the SSJ is installed on the server under test SUT unit end.
2. Power and Temperature Daemon (PTDaemon): The PTDaemon initializes, connects and validates the SUT with the SPECpower_ssj2008 benchmark. In addition, it controls both power measurement device and temperature sensor. Those controlled values were reported at the end of the test. The PTDaemon installed on a control server not the SUT.
3. Control and Collect System (CCS): The CCS application is a multi-threaded Java workload. It coordinates the networking between the control server and SUT.

The architecture of performing SPECpower_ssj2008 benchmark is shown in Figure 3.29.

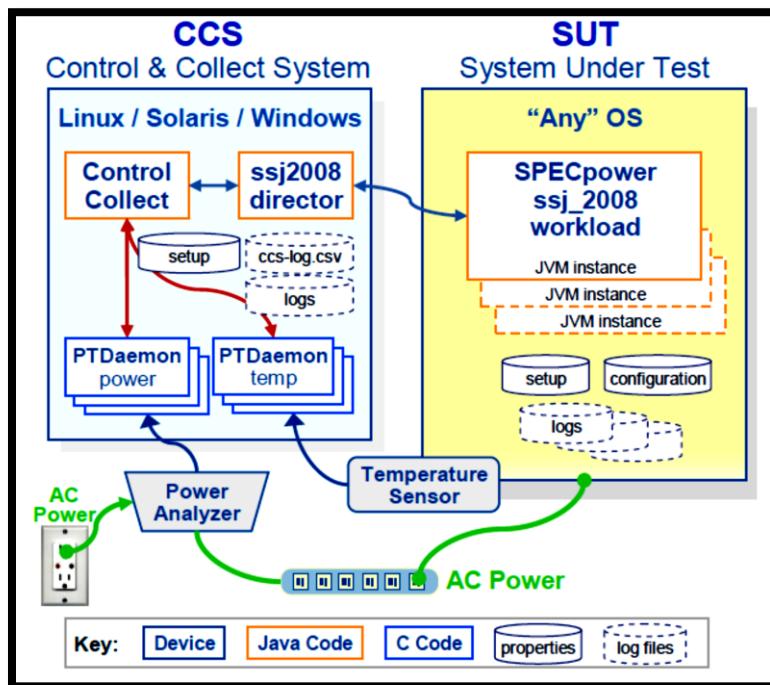


Figure 3.29 SPECpower_ssj2008 architecture (SPEC, 2012).

The SPECpower_ssj2008 workload utilises three main phases. The first one is the calibration phase when three calibrations are conducted. The second phase is the 10 load levels which starts from 100% load and decreases at 10% decrement. The final phase represents an active idle phase. Each load lasts 240 seconds with up to 30 seconds for ramp-up and ramp down periods and up to 10 seconds in-between each intervals as shown in Figure 3.30 (SPEC, 2012a). So, the total testing time of each benchmark workload lasts approximately 75 minutes.

The power measurement during these experiments is logged and monitors via a Voltech PM1000+. The PM1000+ analyser powers the SUT and is connected to the CCS via null serial cable (Burdett, 2014). The reported power measured by the PM1000+ is measured using the definition of the true power as:

$$\text{Power} = \frac{1}{T} \int_0^T V_i I_i dt \quad (3.4)$$

Where, T is the total time of measurement, V_i and I_i are the instantaneous voltage and current, respectively. The AC input voltage and frequency ranges of the monitor are 85-264 V and 45-65 Hz, respectively (Voltech, 2011), so is well suited to the 240V, 50Hz conditions under which the server was run.

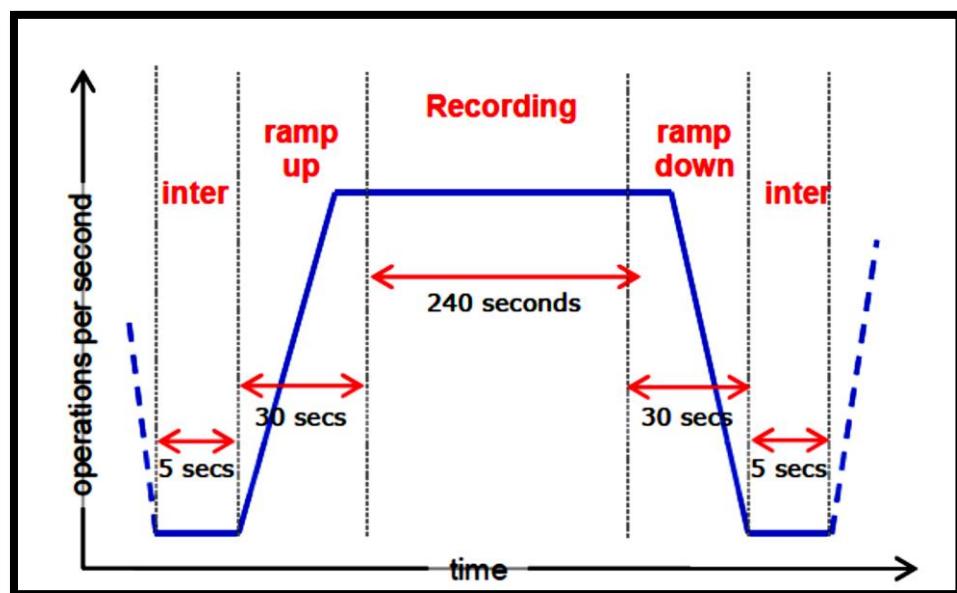


Figure 3.30 Illustration of the interval of action level (SPEC, 2012a).

The SUT is located inside a wind tunnel to maintain a controllable environment as shown in Figure 3.31. The wind tunnel consists of three main sections; entrance, test-section and outlet regions. The entrance region is connected to the air cooling supply duct. This region has an adjustable blade which slides to accommodate different sizes of servers. The test section accommodates 1U to 4U servers. There is another adjustable blade inside the outlet region to provide a close fit to servers of different cross sectional areas. These adjustable blades ensure no over or reverse flow occur at entrance or outlet regions, respectively.

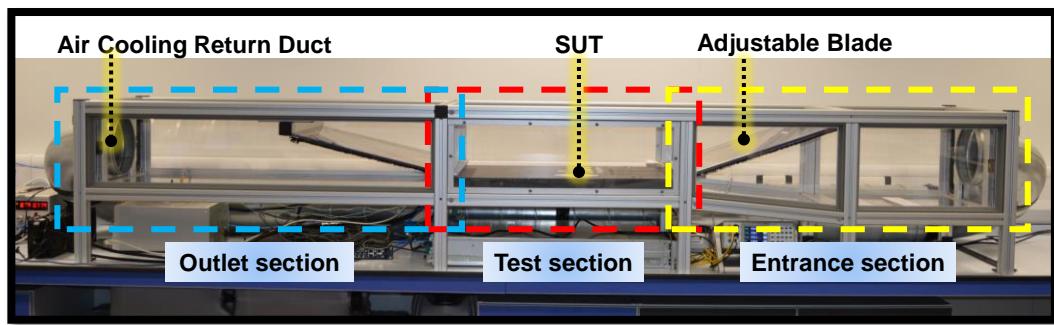


Figure 3.31 Illustration of wind tunnel.

3.5.2 Results and Discussion

3.5.2.1 Benchmarking the SunFire V20z

The SunFire V20z server for power consumption and cooling load investigation needs to first be benchmarked. This benchmarking provides required power and performance values against which it is possible to evaluate any changes with respect to the original status and measure enhancement ratio. The SPECpower_ssj2008 performs a series of transactions which capture the performance of the processor. The SPECpower_ssj2008 exports all quantitative and qualitative values into different form of data presentations. There are raw data files which contain all monitoring values of transactions and power measurements. In addition, there are report files which illustrate performance and power in summary form. The reports of the SPECpower_ssj2008 are in Appendix I.

First, the traditional air-cooled server was run as server under test (SUT) to obtain the benchmark of an original server. Figure 3.32 illustrates the simultaneous monitoring of processed transactions and the corresponding power consumption of the server at each level of the SPECpower_ssj2008.

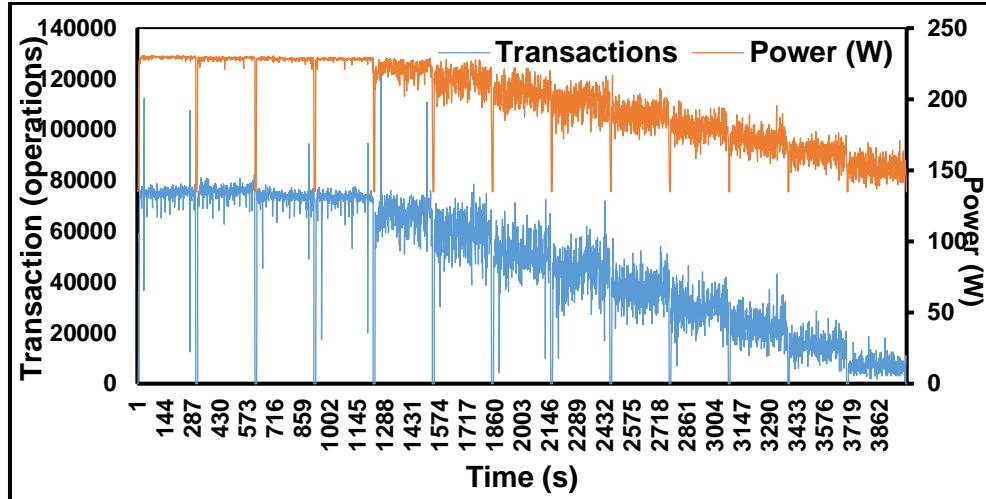


Figure 3.32 Monitoring of power and processed transactions of SPECpower_ssj2008.

The Figure 3.32 presents the decrement of power consumption of the server with the decrease of number of operations processed in the CPUs. The three stages of calibrations inspect the available computational resources in the SUT. The measuring of computational resources provides the distribution of the performance in the next stage of evaluation. During the 100% utilization, the CPUs show the highest level of power consumption because the processes use all resources to perform the calculation. This means the CPUs reach their limit of allocated processes. However, more fluctuations are presenting as the transactions number reduced. The reason behind this fluctuation is due to the reduction in the number of operations which can be conducting through the same resources.

To capture the estimated performance in the air-cooled server, Figure 3.33 shows the multi-level of processed transactions. In addition, the corresponding power consumption is illustrated in Figure 3.34. As mention earlier, the 100% utilization, in Figure 3.34, uses the available resources completely. Whereas when the available resource is greater than that required for transactions, there is an increase in the fluctuations due to loading and completion of tasks.

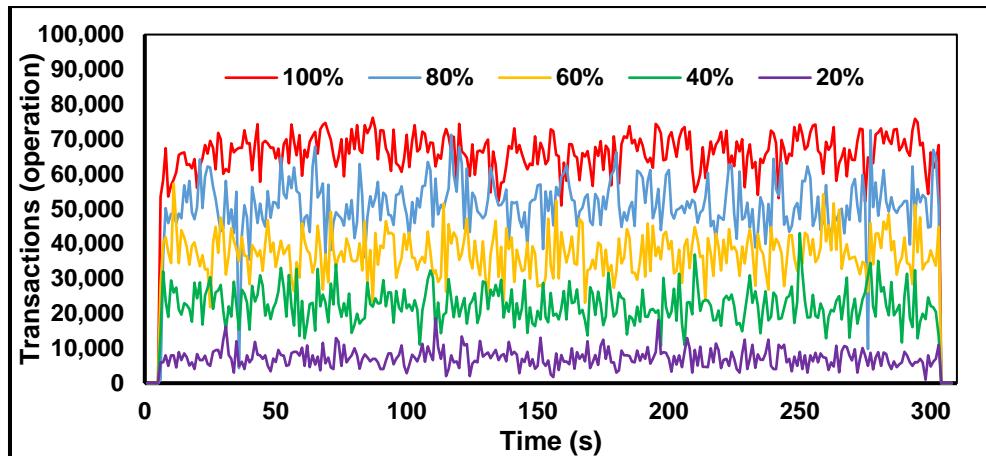


Figure 3.33 Multi-level of performance of air-cooled SUT.

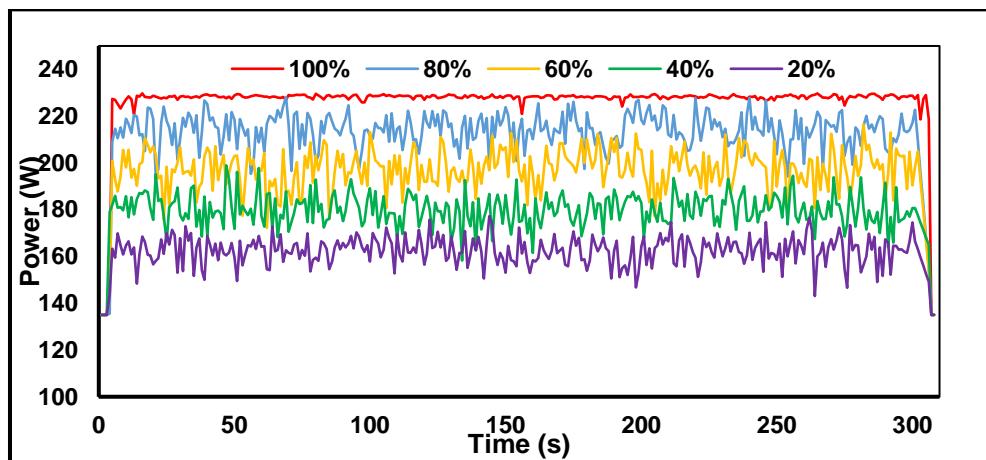


Figure 3.34 SUT power monitoring at multi-level of utilization.

The transcript of SPECpower_ssj2008 workload is designed to record the power consumption during the 240 seconds and then reports the average power consumption. It is obvious that the power consumption decreases linearly with the decrement of utilization level. The power consumption is 230 W at 100% utilization, then reduces by a slope of -0.9 with respect to the utilization decrement as shown in Figure 3.35.

The SPECpower_ssj2008 sets a target operational load at 100% which is concludes from the calibration phase. The SSJ operations at 100% utilization phase is 73,709 (Appendix I) measured in server side Java operations (ssj_ops). Then the workload decreases uniformly by 10% decrement until complete the benchmarking at active idle of zero ssj_ops. The active idle represents no actual compute performance, however, the CPU still consume power. The benchmark reports the matching between

the actual load and the target load as shown in Figure 3.36. The average accuracy of convenience of the actual load to the target load is 98.8%.

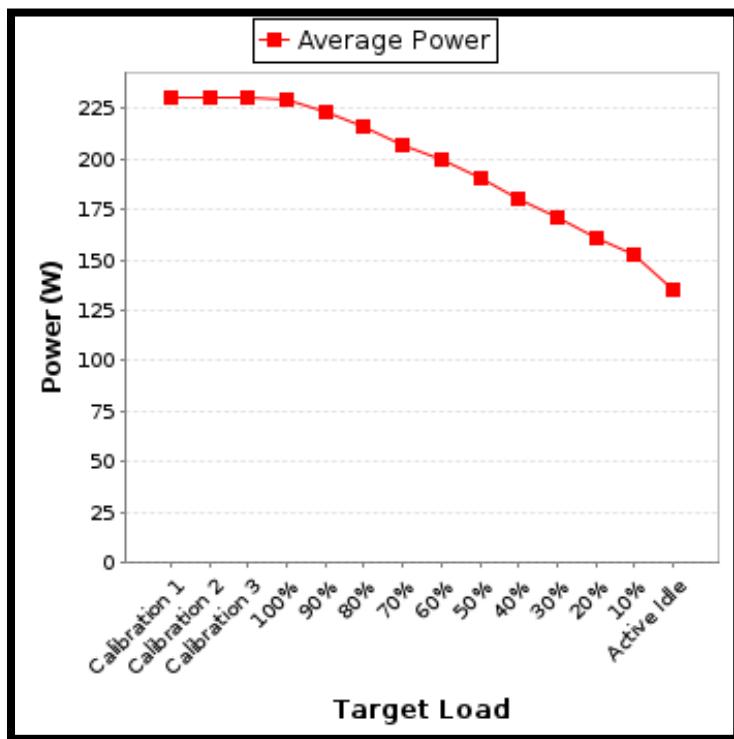


Figure 3.35 Average power consumption variation with utilization for air cooled SunFire V20z.

In addition, the benchmark generalizes all measurements to an important value by calculating the performance to power ratio. This ratio makes the benchmarking and comparing IT servers more beneficial. The reason for using such a ratio is due to both the difference of server's ability to perform computing and the required power consumption. Figure 3.37 presents the behavior of performance to power ratio and calculation of average ssj_ops per Watt. The benchmark reports the maximum and overall performance to power ratios of 321 and 198 ssj_ops/W, respectively.

The retrofitted DCLC server has benchmarked using the same scenario. Figure 3.50 shows qualitatively similar behaviour of the simultaneous monitoring of processed transactions and the corresponding power consumption for liquid cooled server. However, the DCLC server consumes less overall power and achieve nearly the same transactions as the air-cooled server.

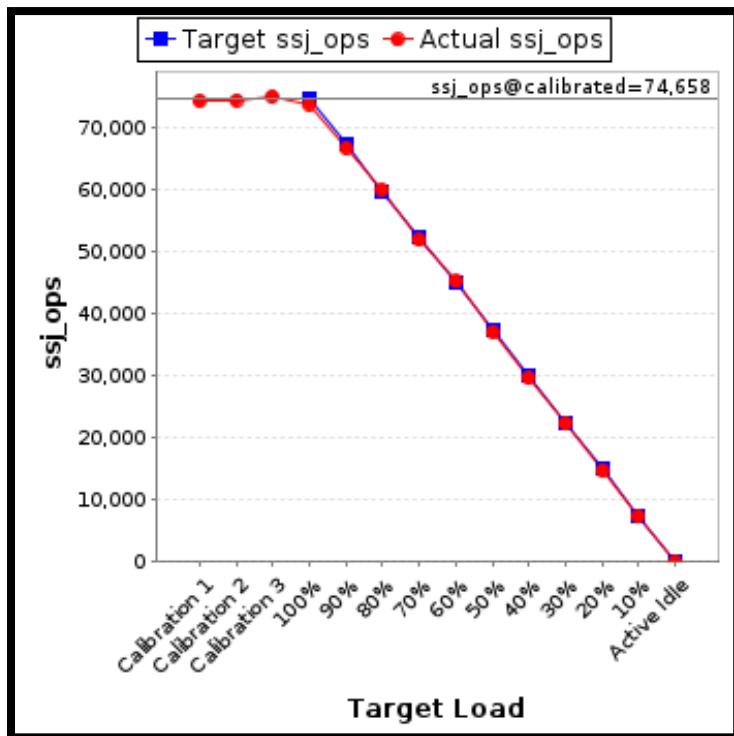


Figure 3.36 The matching of target and actual workload for air-cooled SunFire V20z.

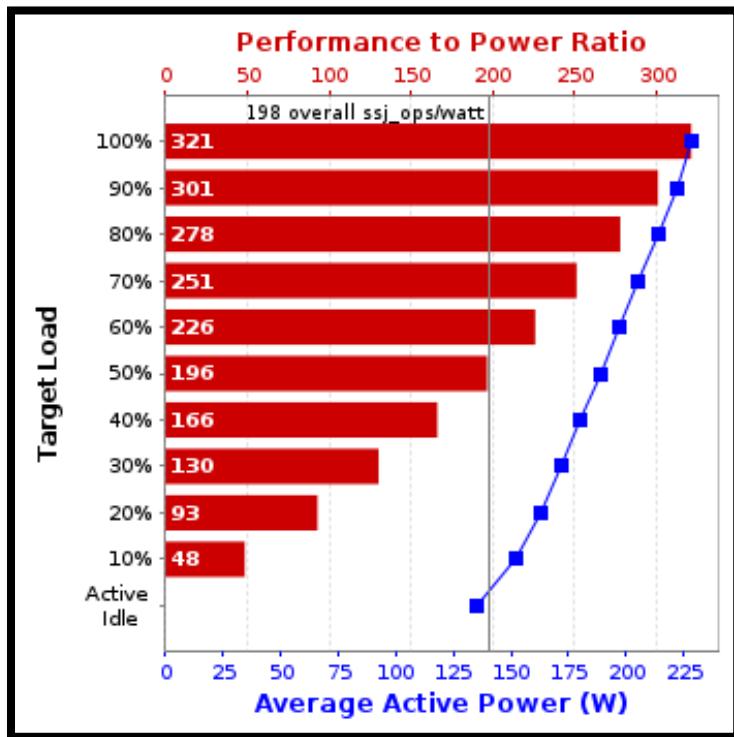


Figure 3.37 Performance to power consumption ratio for air cooled SunFire V20z.

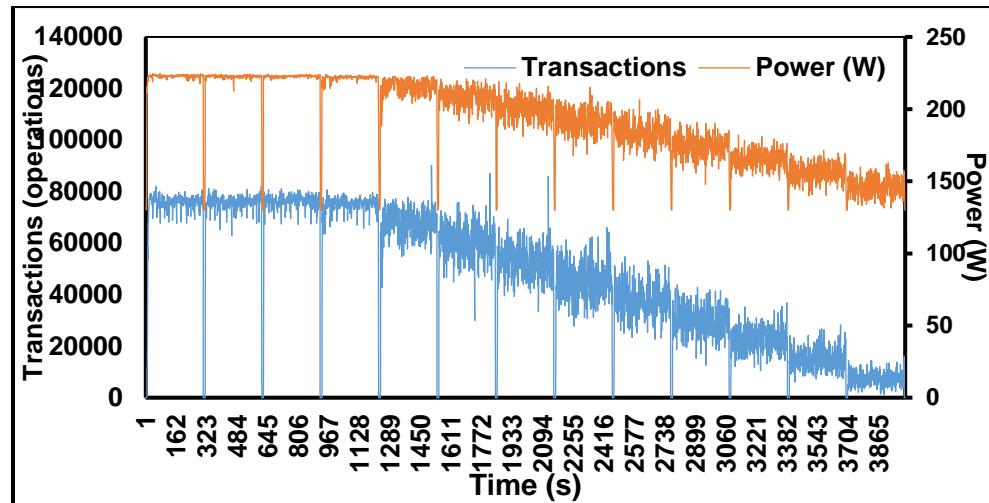


Figure 3.38 Monitoring of power and processed transactions of SPECpower_ssj2008 of DCLC server.

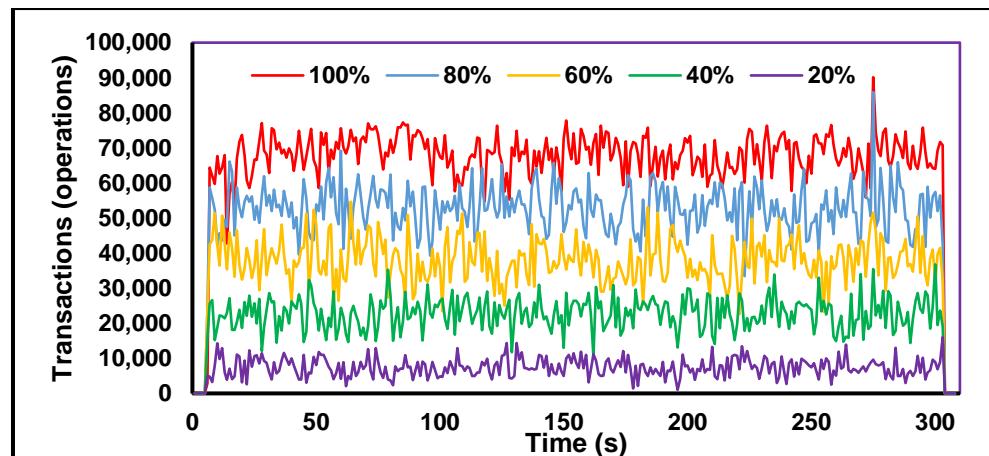


Figure 3.39 Multi-level of performance of DCLC SUT.

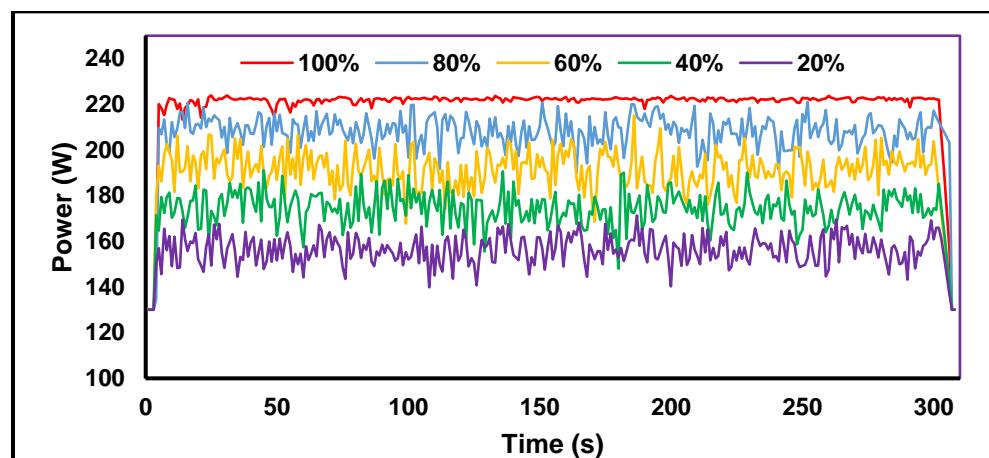


Figure 3.40 Retrofitted DCLC SUT power monitoring.

The performance in the retrofitted DCLC server increases as shown in Figure 3.39. However, the corresponding power consumption is reduced as illustrated in Figure 3.40.

The DCLC server has a lower power consumption, which is 212 W, and it decreases linearly with the decrement of utilization level. The retrofitted server reduces in a slope of -0.88 with the corresponding utilization as shown in Figure 3.41.

The SSJ operations at calibration phase has a negligible reduction comparing to the air-cooled server. In addition, the benchmark workload reports that the matching of the actual load and the target load as shown in Figure 3.42. The average accuracy of convenience between both actual load and the target load is 99.4%.

Furthermore, the performance to power ratio increases by 10% when using the DCLC server. Figure 3.43 shows a linear behavior of performance to power ratio and reports an overall 208 ssj_ops per Watt. The maximum and overall performance to power ratios of 350 and 218 ssj_ops/W, respectively.

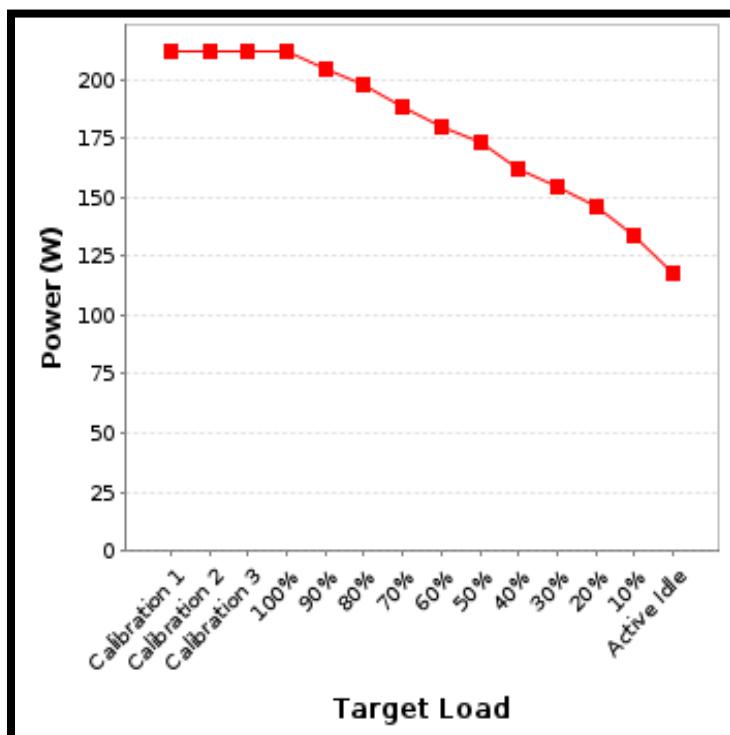


Figure 3.41 Average power consumption variation with utilization of DCLC SunFire V20z.

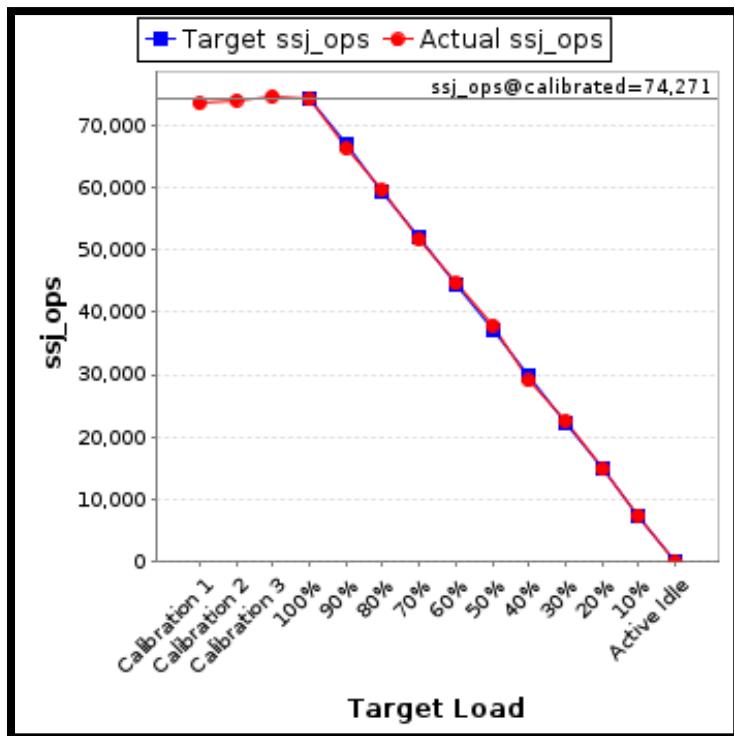


Figure 3.42 The matching of target and actual workload for DCLC SunFire V20z.

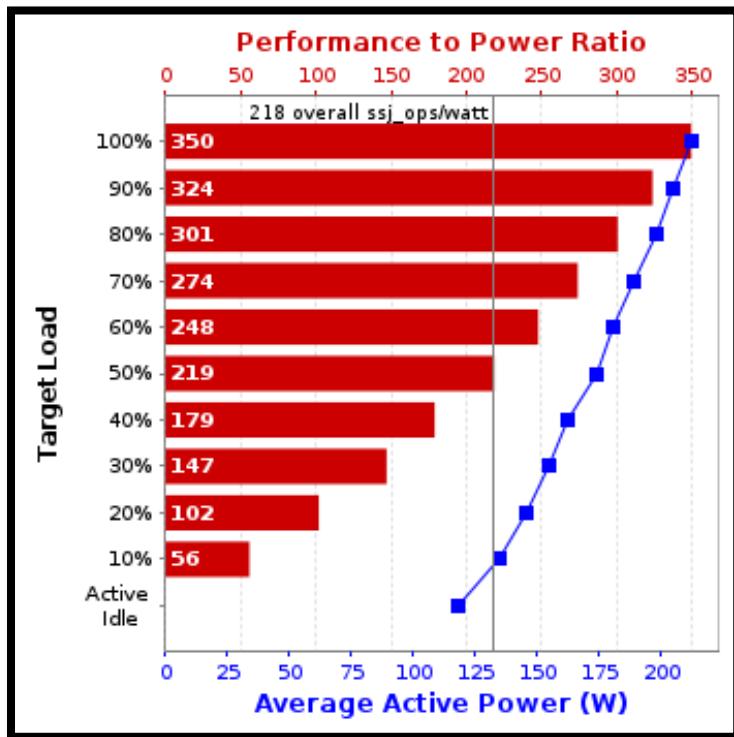


Figure 3.43 Performance to power consumption ratio for air cooled SunFire V20z.

3.5.2.2 Evaluation of Cooling Technologies

The enhancement of the DCLC cooling technology over the air-cooled server is calculated through the thermal balance. For this reason, an IPMI command line is used to log the monitoring of CPUs temperatures and fans speed. The script is shown in Table 3.9, allows the IPMI to pull data into accessible files.

The command line sets manually to run simultaneously with the SPECpower_ssj2008 workload. Normally, the connection with the IPMI starts before the benchmarking workload. This is done to ensure time matching between both files, since they record time step.

First, it is important to reflect the temperature comparisons of both air-cooled and DCLC server under the same benchmark workload. Figure 3.44 presents the temperature monitoring of both SUT. It is obvious to capture that both servers behave in a very similar way with respect to each level of utilization, or transaction's processes. However, the DCLC reduces the primary CPU's temperature, CPU0, by around 6.5 °C during the benchmarking time.

Table 3.9 Linux command line for temperature and fan speed monitoring using the IPMI.

Temperature Logging
While true; do echo \$(date +%s) \$(sensor get -t temp awk -F'temp' '{print \$2}' ORS=',') >> temp.txt; sleep 1; done
Fan speed Logging
While true; do echo \$(date +%s) \$(sensor get -t fan awk -F'tach' '{print \$2}' ORS=',') >> rpm.txt; sleep 1; done

Meanwhile, the temperature of the secondary CPU reduces by about 10 °C due to the use of DCLC technology. Figure 3.45 illustrates the temperature monitoring of both cooling technologies.

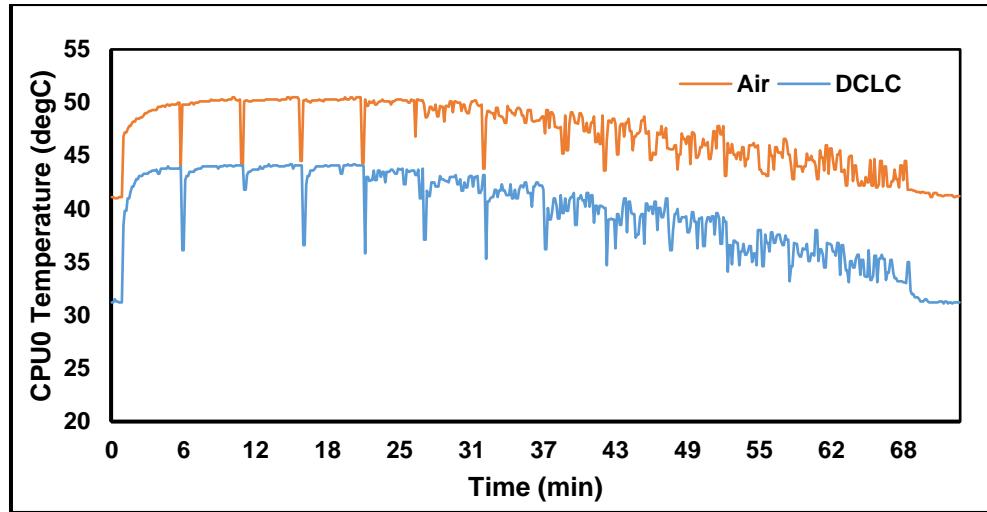


Figure 3.44 CPU0 temperature monitoring for air-cooled and DCLC SUT.

Since the workload targets all available resources, the RAMs are involved as well. Although the server is cooled either by air or liquid, the RAMs still cooled by air only. However, it is noticeable that the DCLC affects the temperature of both banks of RAMs.

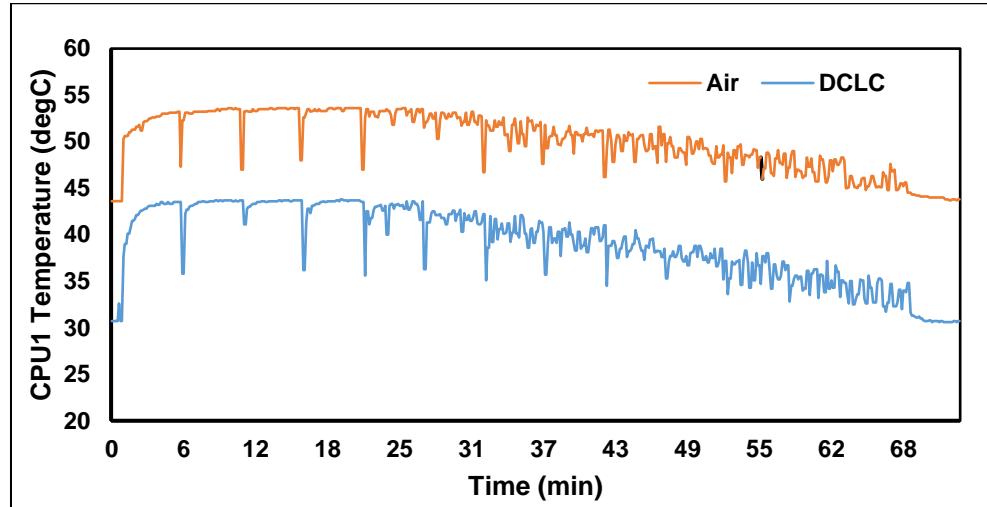


Figure 3.45 CPU1 temperature monitoring for air-cooled and DCLC SUT.

Figure 3.46 presents the monitoring of both banks of RAMs during the workload. RAM0 represents the bank related to the primary CPU, and cooled by the two fans. Those two fans are allocated primarily for the cooling of the power supply unit (PSU). Therefore, this is one reasons for the temperature of the RAM0 to be higher than RAM1. In addition, RAM0 is linked to the primary CPU, so this group of RAMs have a priority to share available resources and perform required operations. The other bank of

RAMs is cooled by two fans. At the higher levels of utilization, the temperature of RAM1 is the same for both cooling technologies. However, as the utilization reduces below 50%, the DCLC maintains the temperature below the air cooled one by more than 1 °C. The explanations related to the correlation of speed fan and the computational load. Since the fan speed reduces as the utilization decrease, so the cooling capacity is reduced. In addition, the number of transactions conducted through both CPUs increases when use DCLC. This means that the computational processes which allocated for RAMs is reduces as well.

The service processor SP is in the downstream of air flow as illustrated earlier in Figure 3.4. This component is affected by the hot air stream, as well as other components in close proximity such as the Ethernet chipset. Figure 3.47 shows the effect of retrofitted the DCLC technology for the SunFire V20z server on the SP's temperature. The SP's temperature rise is a combination of SP's own temperature and the rise due to the hot stream. During the high level of utilization, the air-cooled server records 38 °C for the SP. Its temperature keep constant due to CPUs downstream temperature. However, the DCLC collects the whole amount of heat generated in the CPUs, so there is no heat rejected to the back of the server. Therefore, the temperature of the SP reduced from 37 °C to 35 °C with the reduction of utilization.

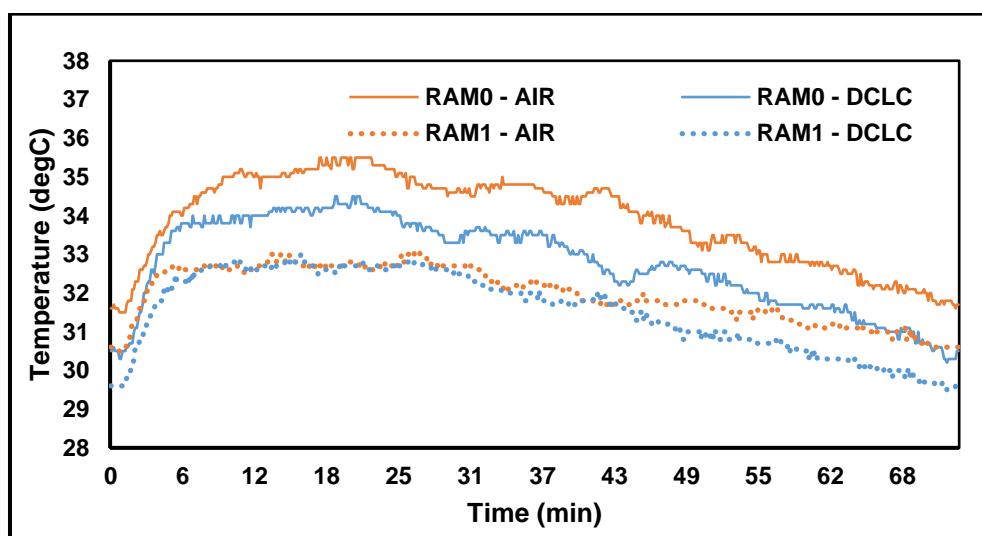


Figure 3.46 Monitoring of RAMs temperature of the SUTs.

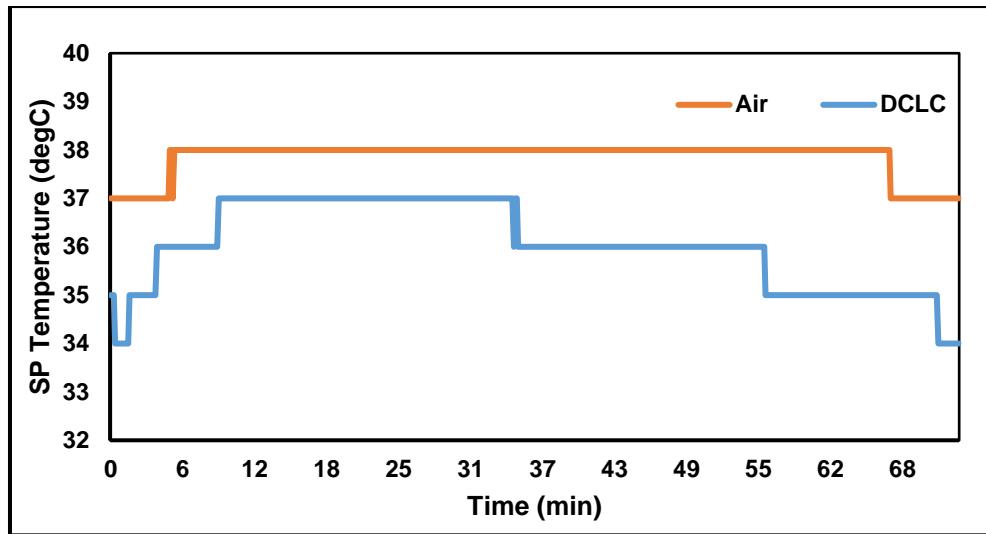


Figure 3.47 The temperature of SP during benchmarking.

The environmental impact of the DCLC technology can be examined through measuring the temperature across the SUT. The outlet temperature from the SUT is logged using three K type thermocouples. The K type thermocouple is a Nickel–Chromium / Nickel–Alumel. These are commonly used within the laboratory as they are both inexpensive and reliable. The temperature range is -270 to 1260 $^{\circ}\text{C}$, with accuracy of ± 1.1 $^{\circ}\text{C}$. The three thermocouples recorded the temperature and the average is presented in Figure 3.48. The air-cooled server rejects air stream hotter than the DCLC by about 1.5 $^{\circ}\text{C}$. In addition, the DCLC provides more stable temperature than the air-cooled server. The air-cooled server relies on the fan speed which is very sensitive to any change in temperature, so the fan speed keep changing. The inlet and outlet temperature differences for both SUTs are shown in Figure 3.49.

Finally, it is crucial to mention that the DCLC offers a high potential to optimize the energy consumption due to two important reasons. First, the direct contact of the liquid with the heat sources, i.e. the CPUs, removes the whole amount of heat away. This removal occurs via liquids which is much higher heat capacity than air and keep the outlet temperature in lower level. Whereas the air flow, for the air-cooled server, is rejected to the room and increase the outlet temperature. The second reason is related to the energy harvesting technology. The hot liquid offers higher opportunities to re-use the energy than the hot air flow. It is very important to make a

balance between those two reasons due to the extra power consumed by the external cooling mechanism. Although it is obvious to record power saving in the DCLC technology over the TAC, the external cooling power should be considered when calculating the enhancement ratio.

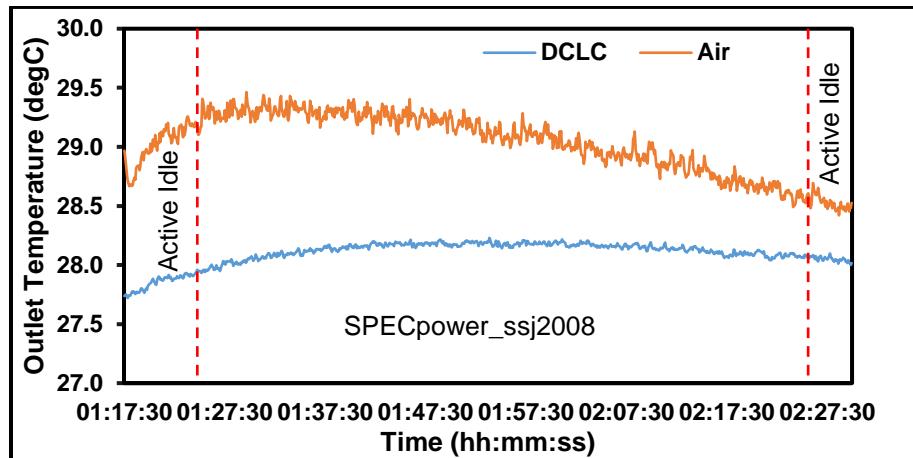


Figure 3.48 Outlet temperature form the SUT.

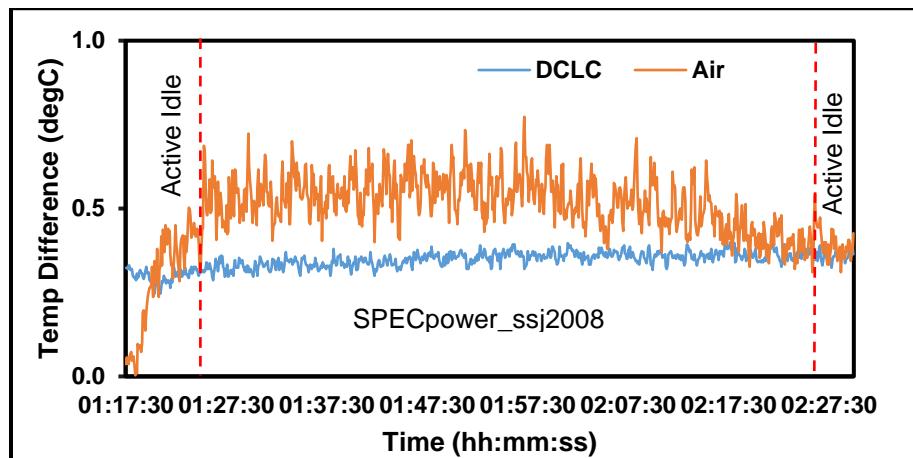


Figure 3.49 Differences between inlet and outlet temperatures for both SUTs.

3.6 Numerical Simulation

3.6.1 Future Facilities – 6SigmaET

Future Facilities Company, founded in 2004, presents a range of different packages for modelling within the information and communication technology ICT industries. The solutions are varying from simulation at a microelectronics level through to a datacentre levels, see Figure 3.50. One

of their CFD package is 6SigmaET which provides thermal simulation for microelectronics. The 6SigmaET deals with complex geometries and physics using very robust and fast parallelised solver. Therefore, this parallelised solver can perform a simulation of more than 3×10^8 mesh cells and scale up to 128 cores for HPC clusters (FutureFacilities, 2015a).

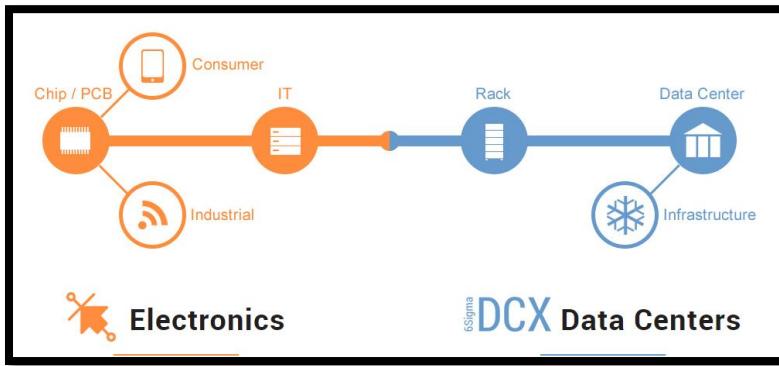


Figure 3.50 Future Facilities scale of simulations (FutureFacilities, 2015a).

Solving simulation problems through CFD can be performed by one of three methods (Eymard et al., 2000, Droniou et al., 2010). These methods are:

1. The Finite Element Method FEM: This method, in brief, divides the whole body into small elements. The solution based on the inter-relations between those elements. The FEM is widely used for solid structural problems and can be used with fluid flow as well.
2. The Finite Difference Method FDM: In this method, the variables of the governing equations are discretised over series of grid points. The FDM is commonly used to solve transport phenomena of fluid flow and heat transfer problems.
3. The Finite Volume Method FVM: This method requires an integration of the governing equations around the control volume. The control volume defines all required parameters and boundary conditions. Then, the integration presents the final distribution of functions over the mesh elements.

3.6.1.1 Governing Equations

Solving any CFD problems consists of three principal steps; select the governing equations, discretise of the volume and apply the appropriate boundary conditions. The present case study is solved by selecting Navier-Stoke equations for fluid flow as transport phenomena coupled with the energy equations to allow heat transfer measurements to be made. The transport phenomena are governed by conservation of mass, momentum and energy. The common formulae of those equation are listed in different forms (Patankar, 1980):

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho u) = 0 \quad (3.5)$$

$$\frac{\partial}{\partial t}(\rho u) + \nabla \cdot (\rho uu) = \nabla \cdot (\mu \nabla u) - \left(\frac{\partial p}{\partial x} \right) + B_x + V_x \quad (3.6)$$

$$\nabla \cdot (\rho uh) = \nabla \cdot (k \nabla T) + S_h \quad (3.7)$$

Where; B_x is the body force per unit volume, V_x is the viscous term and S_h is the volumetric rate of heat generation.

3.6.1.2 The Computational Domain

The computational domain represents the space that enclosed by known boundaries and suitable to apply the governing equations. This enclosed system is called control volume (Anderson and Wendt, 1995). The control volume discretises by converting differential form of the governing equations into algebraic form.

The 6SigmaET deals with the thermal problem of the server level using FVM. The FVM is preferable to use due to its applicability with any arbitrary geometrical shape. Each element has control volume and its boundaries, as well as the nodal points as shown in Figure 3.51 (Versteeg and Malalasekera, 2007). Therefore, the governing equations are solved by integrating the parameters with respect to the centre of the control volume, i.e. centroid of the element.

The 6SigmaET creates the best appropriate mesh based on the geometrical shapes and distribution of the entire components.

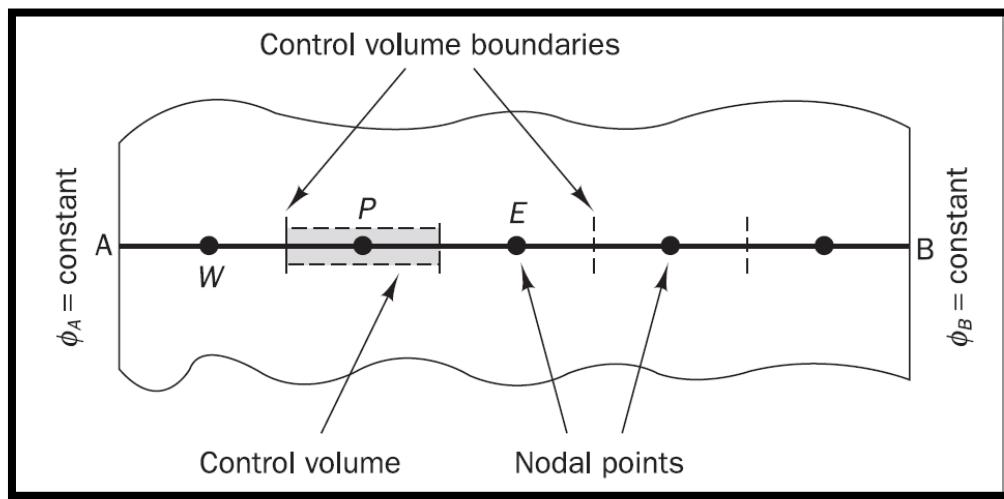


Figure 3.51 Illustration of an element in FVM (Versteeg and Malalasekera, 2007).

3.6.1.3 The Turbulence Model

Fluid flows in one of the following regimes; laminar, transitional or turbulent. Laminar flow occurs when ‘lamella’ of fluid layers flows over another generally there is no disturbance on such a flow. The turbulent regime is recognised by the irregular disturbances and fluctuations as well as a main directionality of flow there is a secondary disturbance on the top of this flow. For that, the velocity value, as shown in Figure 3.52, at any time is a function of mean and fluctuated velocity as shown:

$$u(t) = U + \dot{u}(t) \quad (3.8)$$

There are many models simulate the turbulent flow regime, however, 6SigmaET uses $k-\epsilon$ model. The $k-\epsilon$ model is widely used in industrial applications and especially the datacentre simulation. This preferability is due to lower computational expense and an increased solution stability (Bhopte et al., 2006, Joshi and Kumar, 2012). This model defines two parameters; k and ϵ which are the kinetic energy and energy dissipation rate, respectively.

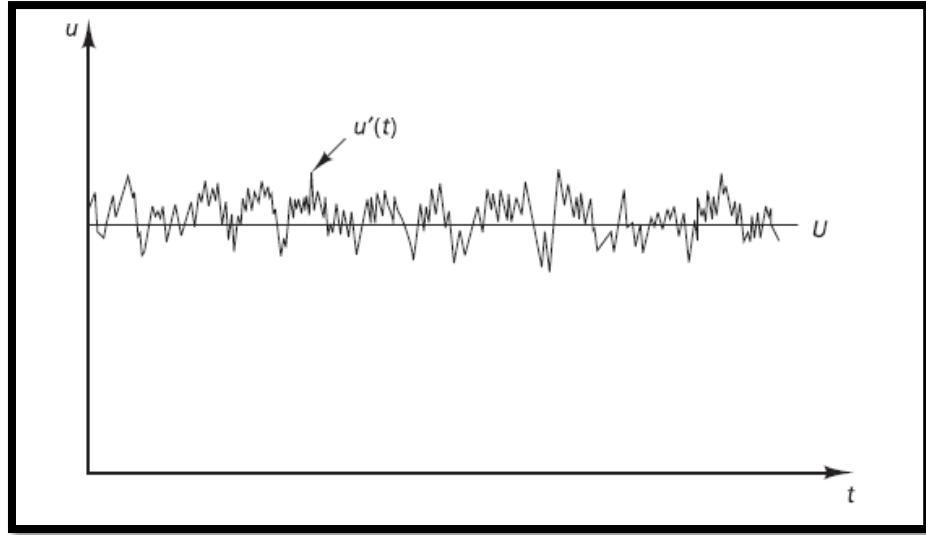


Figure 3.52 Velocity fluctuations for turbulent regime (Versteeg and Malalasekera, 2007).

The dynamic viscosity is defined as:

$$\mu_t = \rho v_t \quad (3.9)$$

Where v_t is the kinematic viscosity and given by;

$$v_t = C_\mu \frac{k^2}{\epsilon} \quad (3.10)$$

C_μ is a constant, whereas k and ϵ can be determined by solving the following equations:

$$\frac{\partial \rho \bar{U}_i k}{\partial X_j} = \frac{\partial}{\partial X_j} \left(\left(\mu + \frac{\mu_t}{\sigma_k} \right) \frac{\partial k}{\partial X_i} \right) + P + G - \rho \epsilon \quad (3.11)$$

$$\frac{\partial \rho \bar{U}_i \epsilon}{\partial X_j} = \frac{\partial}{\partial X_j} \left(\left(\mu + \frac{\mu_t}{\sigma_k} \right) \frac{\partial \epsilon}{\partial X_i} \right) + C_1 \frac{\epsilon}{k} (P + C_3 G) - C_2 \rho \frac{\epsilon^2}{k} \quad (3.12)$$

P is the function of shear production, while G is the turbulent kinetic energy due to buoyancy, and both given by:

$$P = \mu_{eff} \frac{\partial \bar{U}_i}{\partial x_j} \left(\frac{\partial \bar{U}_i}{\partial X_j} + \frac{\partial \bar{U}_j}{\partial X_i} \right) \quad (3.13)$$

$$G = \frac{\mu_{eff}}{\sigma_T} \beta \frac{\partial T}{\partial x_j} \quad (3.14)$$

The parameters appear in equations (3.10) to (3.14) are listed below (Joshi and Kumar, 2012):

$$C_\mu = 0.09, C_1 = 1.44, C_2 = 1.92 \text{ and } C_3 = 1.0 \quad (3.15)$$

3.6.2 Building the Model

The model is built step by step as illustrated in Appendix II using the tutorial guides (FutureFacilities, 2015b, FutureFacilities, 2015c). It is important to notice that every “bold” word represents an option from a drop list or a choice in a dialogue box. For example, **New Test Chamber** is one of the options within **New** drop list and so on. It is important to list the thermal design power TDP for main components in Table 3.10 (AMD, 2006, Pandiyan, 2013).

The model tree includes the following main parts:

1. Chamber Test

The chamber test simulates the wind tunnel in the experimental test rig. The chamber provides a control volume for the coolant entering the server. Both height and width are similar to the corresponding values in Table 3.1. However, the depth varies per the wind tunnel shape.

2. Chassis of the Server

The geometry of the chassis is like the actual SunFire V20z unit. The dimensions are listed in Table 3.1

3. The Vents

The SunFire V20z has perforations at three sides; the front, the rear and the upper sides. The available options for the perforations is vent areas on the selected sides.

4. The Printed Circuit Board PCB

The SunFire V20z has two possible cooling systems as mentioned earlier; the air cooled with extruded fins heatsink and the liquid cooled cold plate. Moreover, there are two CPUs and eight small daughter board for memory cards.

5. The Chip Socket

The two CPUs of the studied server consist of chip socket, heatsink and thermal interface material.

6. The Fans

The studied server has 3 sets of fans as mentioned earlier in Section 3.2.5. Both RAMs and CPUs six fans are attached to the PCB since they are powered when the server is running. However, the rest two fans are designed to be powered separately with the power supply unit PSU.

7. The Memory Modules

The memory modules consist of socket and daughter board

8. The Disk Drive

The disk drives are added to the chassis level. The addition contains creating the drive bay first then adding the drives.

9. The Power Supply

Servers have power supply unit PSU or multiple units. The PSU is a complex unit and consists of many components. However, 6SigmaET presents a simple model of PSU. Although the simplicity of the model, it reflects the correct characteristics of heat and fluid flow of the actual PSU. The only restriction is the prediction of hot spot inside PSU block, which is not one of interest of current study.

10. Obstructions

Each server has different type of components which are not categorized above. So, those components are called obstructions. It could be wire extension, power joint or even mounting kits. The three generalized types of obstructions are:

- Solid Obstruction: This type of components is a 3D solid body of any shape. These obstructions block the air flow totally.

- Perforated Obstruction: It is either a flat plate or polyline plates of any shapes. These obstructions contain series of perforations which reduce the fluid flow.
- Porous Obstruction: This obstruction is either 3D or plate of any shapes. In addition, the flow resistance in all direction can be adjusted independently.

11. The Cold Plates

The cold-plate define under the obstruction blocks since there is no flow across. However, the design either performed within the geometrical facilities or imported from specific design, i.e. computer aided design CAD file.

Table 3.10 Thermal design power for the main components (AMD, 2006, Pandiyan, 2013).

Part	Number	TDP (W)
CPU	2	85.3
RAM	8	0.81
PSU	1	4.5% of total load
HDD	2	7.8
Fan*	4 or 8	Equation (3.1)

3.6.3 Run the Simulation

After complete and build the model, the two servers are ready for running the simulation. The first server represents TAC server with aid of 8 fans, where 4 of them cool down the CPUs, used for cooling process Figure 3.53. Whereas the second server uses cold plate attached to the CPUs for cooling instead of 4 fans and heatsinks as shown in Figure 3.54.

Running the simulation requires definition of the mesh size and count which is automatically generated by the software. The generation is based on the complexity of the case study. However, it is easy to modify the number of the mesh and compare results for mesh independency tests.

* The number of fans depends on cooling technology, 4 for DCLC and 8 for TAC.

1. From the **Solver**, click **Verify Model**. This verification looks for any intersection or overlapping between components.
2. Click **Run** to start the simulation.

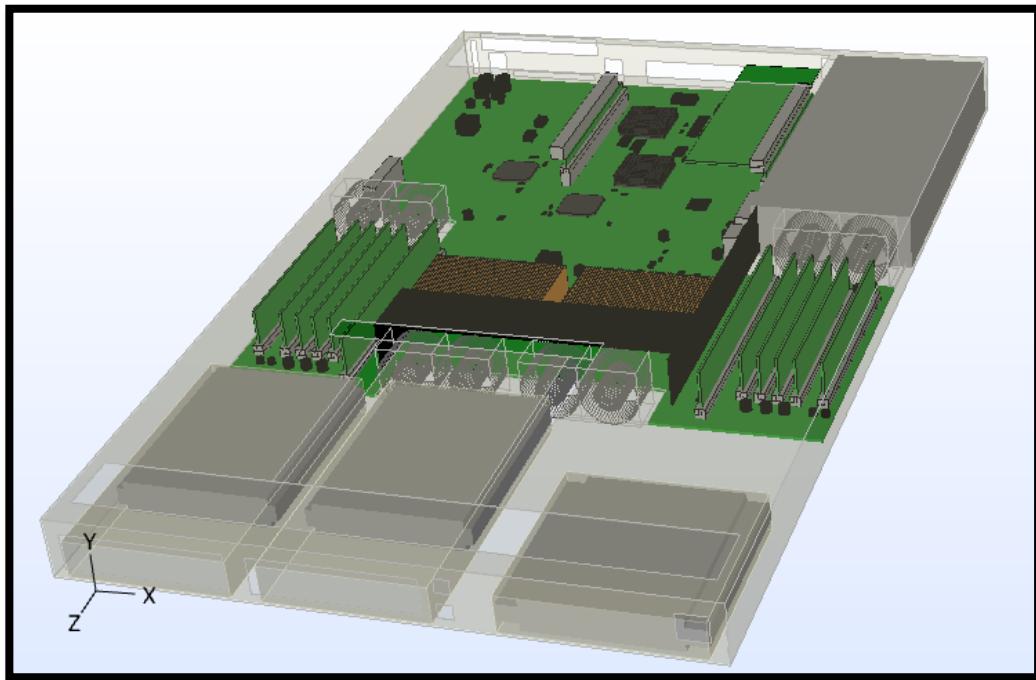


Figure 3.53 The TAC server model.

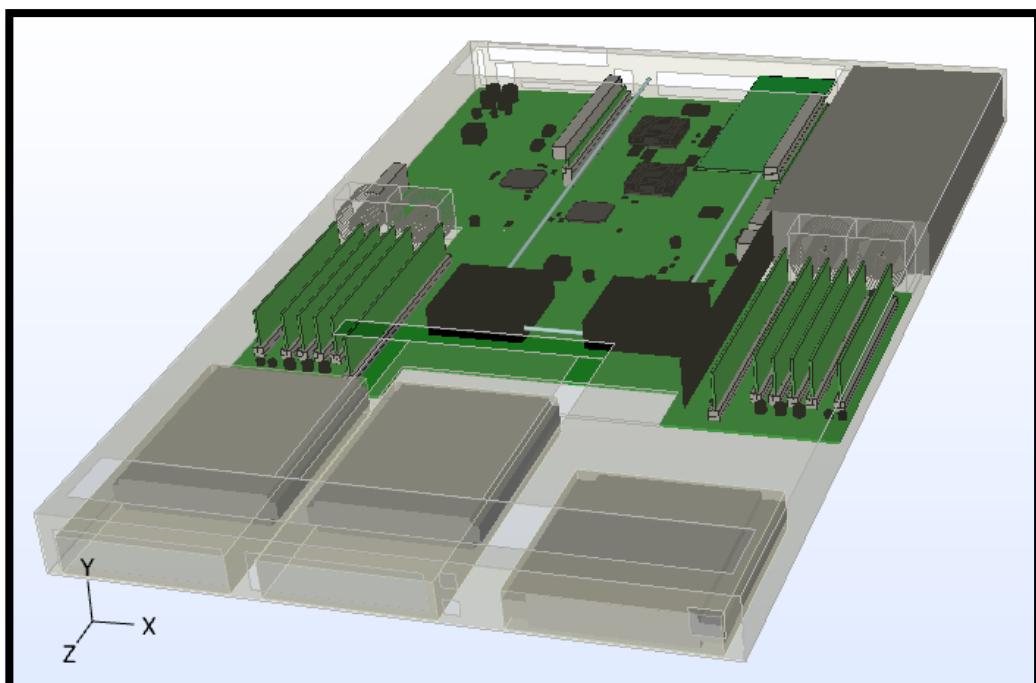


Figure 3.54 The DCLC server model.

3.6.4 Results and Discussion

The 6SigmaET performs a CFD turbulent simulation with capability of a parametric study. The parametric study includes different ranges of geometry and placements of components, the utilization level, cooling technology and the environmental aspects.

In this simulation, the results are validated with benchmark values for similar conditions. Therefore, the first test uses air-cooled server and assumes the utilization of the CPUs is 100%. The utilization reduces at 10% decrement until 10% to mimic the workload of the SPECpower_ssj2008. The predicted active idle is 140 W, whereas the actual idle load is 136 W. From Figure 3.55, as the utilization level decreases the accuracy of the simulation increases. The maximum difference is 6.1%, whereas more than 80% of results have an average difference of 2.7%.

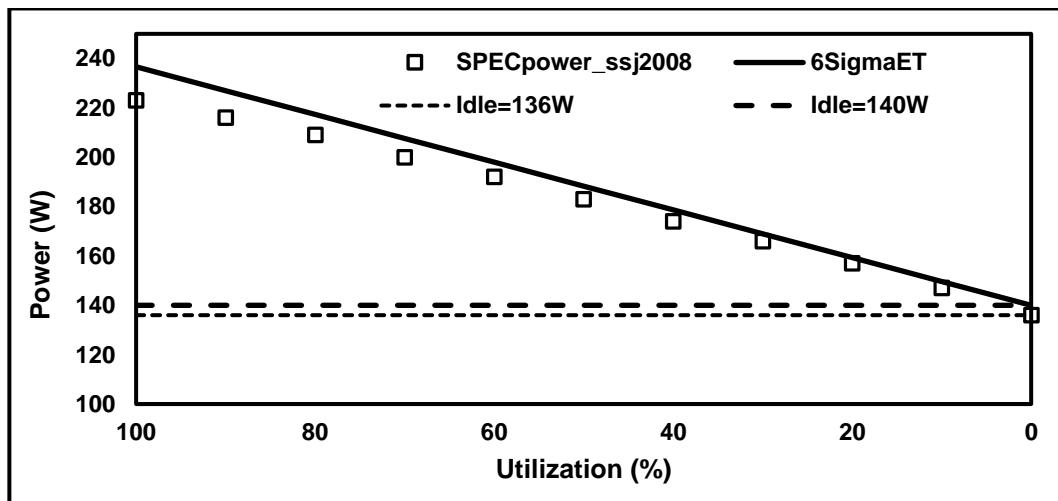


Figure 3.55 Comparisons of power consumption of the air-cooled SUT.

The DCLC technology shows slightly higher inaccuracy in the simulation predictions. Figure 3.56 illustrates the power consumption of the DCLC server from SPECpower_ssj2008 and 6SigmaET. Both results show similar overall trend of power linear dependency on utilization. However, the simulation results are less accurate than the air-cooled server. One possibility of this conclusion is because the DCLC is not a predefined item in the software library. Although the maximum difference between both

results is 13%, majority of differences is about 7%. This differences is still acceptable as an initial prediction for such a complex problem.

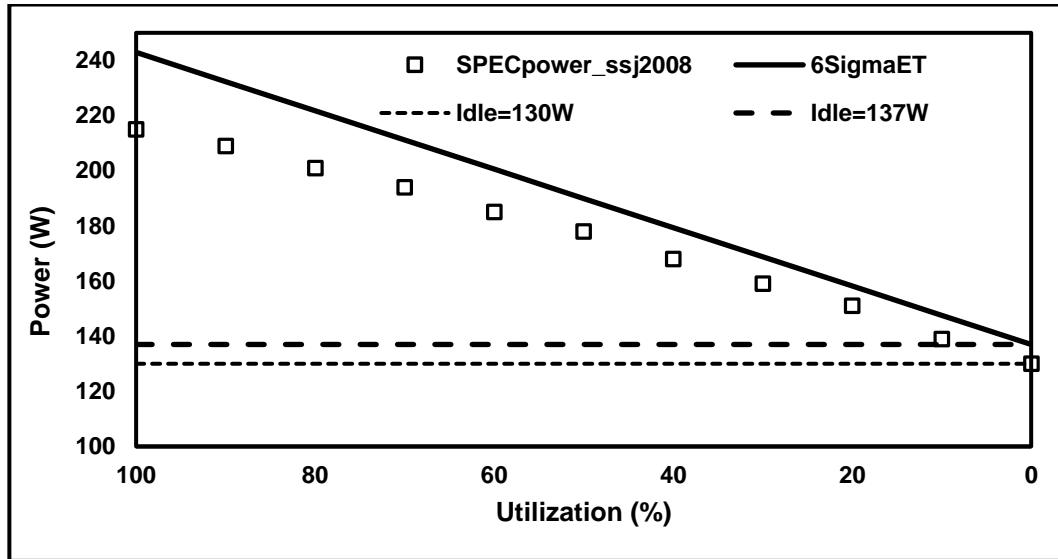


Figure 3.56 Comparisons of power consumption of the DCLC SUT.

3.7 Summary

The current chapter presents the ability of retrofitting the SunFire V20z server with DCLC. This server's CPUs are cooled by air with aid of set of fans traditionally. After describe the server and inspect its characteristics, the cooling system is modified to direct contact liquid cooling DCLC provided by CoolIT. Both conventional and retrofitted servers are tested to evaluate their performance. The performance measurements have conducted using Stress Linux and SPECpower_ssj2008. Moreover, there are some results related to the environmental effect on power consumption from both experimental work and numerical simulation. Those results are included in the following chapter.

It is crucial to mention that using the direct contact liquid cooling has the following limitations. Firstly, although using a well-engineered tubing system, there is a risk of microelectronics damage due to water droplets. The droplets can be either from a leakage or from condensation if the inside coolant temperature dropped below the dew point of the environment. Secondly, the applicability of such a cooling solution requires an external

heat removal unit. This unit need to be designed efficiently for a wide range of IT loads and datacentre's serviceability.

However, the important concluding remarks can be summarised as below:

1. The maximum power consumption of the DCLC server reduces by 18 W. This saving represents 7.8 % of the total power consumption.
2. For the inside-server cooling, the DCLC saves 88% at chipset scale and 29% at whole server scale.
3. Using the liquid cooling technology increases the overall performance to power ratio by 10% over the air-cooled technology.
4. The DCLC reduces the temperature of primary and secondary CPU noticeably by 6.5 °C and 10 °C, respectively.
5. Retrofitting the DCLC to cool the CPUs reduces the noise by 88.1% comparing to the air-cooled CPUs.

Chapter 4 TOWARDS TOTAL LIQUID COOLING (TLC)

4.1 Introduction

Total liquid cooling (TLC) has been used in the *macro* scale of electronic and electric components and exploited for after the World War II. However, the TLC of the microelectronics was only beginning to apply within the mid-1980 (Bar-Cohen et al., 2006). At that time, IBM presented the indirectly water-cooled mainframe computers. Despite a relatively primitive implementation of liquid cooling, these liquid-cooled computers offered essential improvements in the capability of heat removal (Chu et al., 2004). Heat transfer coefficients in liquids are very high in comparative with air in both free and forced convection as shown in Figure 4.1.

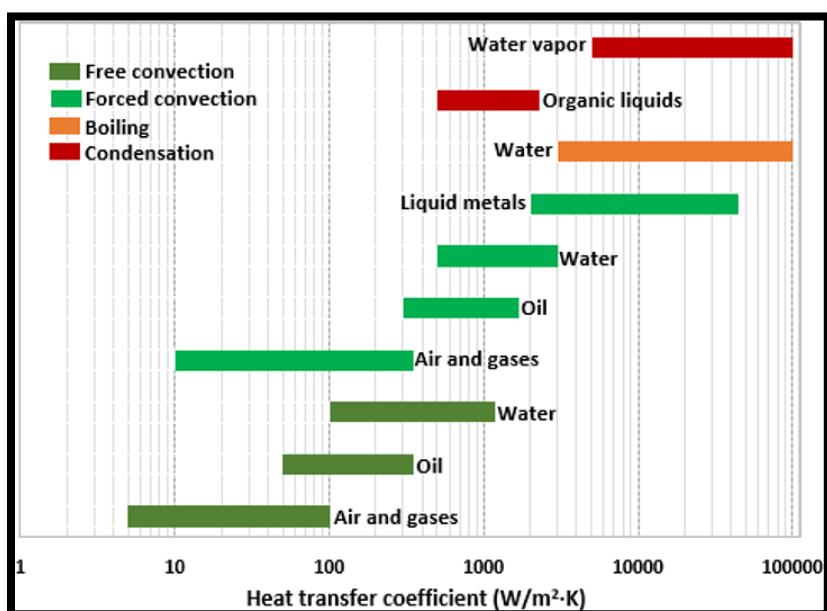


Figure 4.1 Different capability of heat transfer coefficients (Murshed, 2016)

In liquid immersion (Chester et al., 2013, Hopton and Summers, 2013), a dielectric liquid is used in direct contact with all of the microelectronics. It is encapsulated vertically, as shown in Figure 4.2, on one of the capsule's walls and the heat is transferred to the opposite, cooler wall via a natural convection current. In this study, the geometrical investigation of an immersed server is carried out by assuming the fluid flow in the capsule, which is the container of server and cooling technology, is governed by creeping flow, where inertial forces are small compared to viscous forces, in a saturated porous media in the presence of buoyancy forces. The use of such a reduced order model enables a wide-ranging parameter study to be carried out in feasible timescales. Similar to the enclosed server capsule discussed earlier, (Nithiarasu et al., 1997) studied a square cavity filled with fluid in a saturated porous medium where the left wall is hot and the right wall is cold while the top and bottom walls are insulated. The study demonstrated that both Rayleigh number and porosity value affect the wall Nusselt number dramatically.

One of the important investigations of rectangular cavity enclosure with fluid and porous media was studied numerically (Chan et al., 1970). The top and bottom walls are insulated and the heat is transferred from the left wall and the enclosure is cooled by constant temperature on the right wall. The paper showed that the convective flow distorts temperature field with increasing Rayleigh number. The paper also presented the correlation of heat transfer enhancement with respect to Rayleigh number, Darcy number, and the cavity's geometric aspect ratio. Also (Poulikakos and Bejan, 1984), investigated an enclosure with a fluid saturated porous media with hot and cold sources on each side of the enclosure. A part of the left wall was hot and the other part was cold while the remaining walls were insulated. The investigation found that in the situation when the hot part was above the cold part the natural circulation was either incomplete vertical penetration or incomplete horizontal penetration. However, as expected, in the case when the cold part was above the hot part the flow spreads throughout the porous medium.

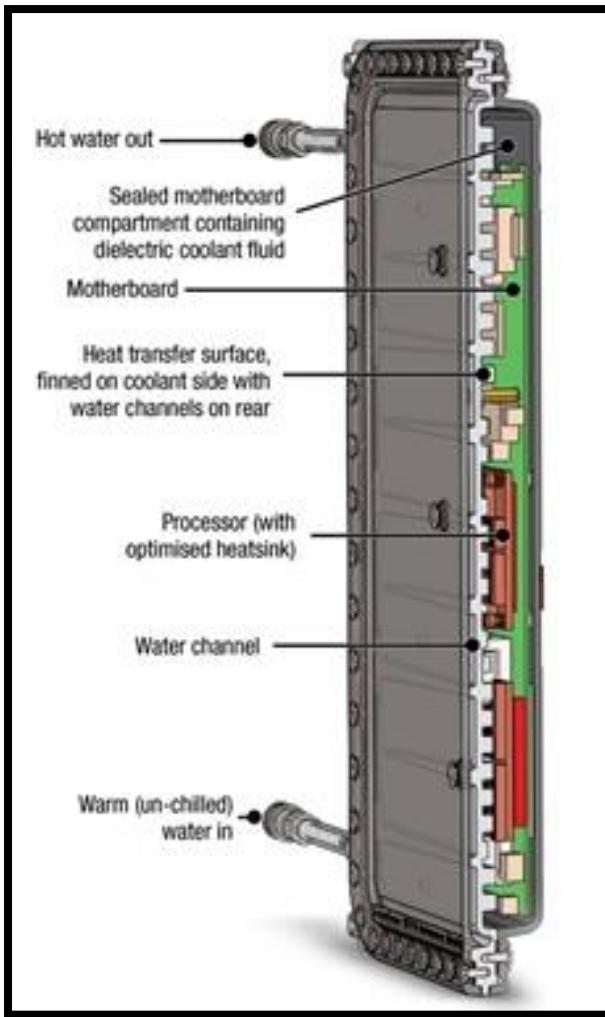


Figure 4.2 Illustration of Iceotope module (BOSTON, 2009).

An enclosure model with a hot patch on one side and a cold patch on the other side was studied by (Ismaeel, 2011). The study demonstrated that the best heat transfer performance was obtained when the hot patch was at the bottom and the cold patch at the top. In contrast, the lowest heat transfer performance occurred when the hot patch was placed at the top and the cold patch at the bottom. Other investigations analysed the situation when the heat was from below for enclosures that had fluid saturated porous media with different parameters (Kladias and Prasad, 1990, Prasad, 1991). Prandtl and Darcy numbers effects on fluid flow in a porous media has studied by (Kladias and Prasad, 1989). The heat was generated from the bottom of the enclosure and the study showed that the convective heat transfer regime increased with increasing Rayleigh number, Darcy number and Prandtl number.

In this chapter, a novel approach studies the total liquid cooled TLC server which immersed in dielectric fluid. The novelty of this work is the modelling of the server and coolant as a porous layer.

4.2 Modelling of Total Liquid Cooling

Fully immersed liquid-cooled server is modelled as a cavity where convective currents flow inside it. This flow could be in three different modes; forced, mixed or natural convective flow. Forced convective flow requires external or mechanical work to circulate the coolant. Natural convective flow depends on buoyancy forces to circulate the fluid. Finally, mixed convective flow occurs when a combination of both forced and natural in either aid or oppose flow regime exists. The natural convection assumed to be the only type applicable to the current study.

4.2.1 Novelty of the Model

The present model deals with natural convective flow due to buoyancy. The density change is modelled by Boussinesq's approximation, as shown in equation (4.1), and it describes the density as a function of the temperature differences:

$$\rho = \rho_{\infty} [1 - \beta(T - T_{\infty})] \quad (4.1)$$

Where ρ and ρ_{∞} represent the local and free stream density, which depends on local temperature T and free stream temperature T_{∞} , respectively. In addition, β represents the thermal expansion coefficient.

The server consists of different electronic elements which produce heat waste with their IT productivity. So, those electronic elements provide heat sources inside the servers which can generate a tremendous amount of heat. Furthermore, this heat needs to be rejected into a cooling path which can be represented by the heat sink. The amount of heat generated is large enough to damage the internal components of the server. Therefore, the

heat should be transferred and then rejected outside the server. Moreover, the coolant should have some important characteristics to play the essential cooling role. Since the coolant in contact with the electronics, then it should have dielectric properties to protect the server electrically. In addition, the higher the thermal conductivity, the more effective the heat removal. Moreover, the coolant fluid needs to be in a reasonable range of thermal expansion to drive the natural convection. For example, liquid-cooled fully immersed rack system uses the engineered dielectric fluid Novec (Iceotope, 2013). Novec has a relatively high thermal expansion ratio of 0.00145 K^{-1} compared to other liquids (Chi et al., 2014, Chi, 2013). The Iceotope module cross section is shown in Figure 4.2 and illustrate processors, heat sinks, heat transfer surface and water flow paths (BOSTON, 2009).

The present study deals with the immersion of a dual CPU server inside a vertically aligned capsule. The capsule is filled with dielectric liquid, which removes heat from the CPUs by a naturally driven convective circulation. Convective flow can move freely inside the capsule from the heat source on one vertical wall of the capsule to the opposite vertical wall, which is acting as a heat sink. Each microelectronics system, or server, contains many components such as power supply, heat sinks and spreaders, RAM, VRMs etc... These elements impede the coolant flow and can inhibit the natural convection inside the encapsulated server, which reduces the heat transfer capability.

The two dominant forces acting on the fluid flow are in competition, namely coolant buoyancy forces of the natural convection and the viscous forces due to the microelectronic components and capsule walls. Flow is therefore modelled by natural convection inside a saturated porous medium. The coolant represents the saturation fluid, whereas the microelectronic elements represent the solid structure of the porous medium.

The need for such a simple model is helpful for a rapid prediction of initial design of immersed servers. In addition, initial prediction helps in better awareness of thermal behaviour in the encapsulated server to avoid hot spot. The hot spot has worse effect especially in HPC application because

the high range of produced heat within small areas, e.g. up to 300 W/cm² (Tong, 2016).

4.2.2 Mathematical Formulation

In the present study, natural convection in the server capsule is modelled as an isotropic porous vertical layer. This layer is assumed to fulfil the following assumptions (Al-Anii, 2005):

1. 2-dimensional (2D) domain.
2. Single-phase flow.
3. Thermodynamic equilibrium between the porous components and the saturated fluid.
4. Boussinesq's buoyancy approximation is valid.
5. Flow is slow so that fluid inertia can be neglected.
6. Flow in the porous media is based on Darcy's law.
7. Any chemical reactions and viscous heating are neglected.

The governing equations based on the conservation of mass, momentum and energy are listed respectively using vector notation (Cheng, 1979, Bejan and Kraus, 2003):

$$\nabla \cdot \vec{v} = 0 \quad (4.2)$$

$$\vec{v} = - \left(\frac{K}{\mu} \right) (\nabla p - \rho \vec{g}) \quad (4.3)$$

$$\sigma \left(\frac{\partial T}{\partial t} \right) + \vec{v} \cdot \nabla T = \alpha \nabla^2 T \quad (4.4)$$

where; K is the permeability of the medium, σ the volumetric heat capacity and α is the thermal diffusivity.

The density of the body force term of equation (4.3) varies according to the Boussinesq's approximation as mentioned earlier in equation (4.1).

After the definition of the density, equation (4.3) can, in fact, be simplified to one dependent variable instead of two velocity components. The dependant

variable is known as stream-function and it represents each velocity component. Using the stream-function makes the mathematical problem simpler and facilitates flow visualisation. Physically, both horizontal and vertical components of velocity can be represented by stream-function derivatives as shown as:

$$u = \frac{\partial \psi}{\partial y} \quad (4.5)$$

$$v = -\frac{\partial \psi}{\partial x} \quad (4.6)$$

The reduction of the momentum equation to a single component equation is achieved by taking the curl of equation (4.3), resulting in equation (4.7) as follows:

$$\nabla \times \vec{v} = \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right) \vec{k} = \frac{K}{\mu} (\nabla \times \nabla p) + \left(\frac{Kg}{\mu} \frac{\partial \rho}{\partial x} \right) \vec{k} \quad (4.7)$$

Now, after reducing velocities to stream-function, mass, momentum and energy conservation governing equations are in the following form:

$$\nabla^2 \psi = -\frac{Kg\rho\beta}{\mu} \frac{\partial T}{\partial x} \quad (4.8)$$

$$\alpha \nabla^2 T = \frac{\partial T}{\partial t} + \left(\frac{\partial T}{\partial x} \frac{\partial \psi}{\partial y} - \frac{\partial T}{\partial y} \frac{\partial \psi}{\partial x} \right) \quad (4.9)$$

4.2.3 Numerical Solution

The solution procedure of equations (4.8) and (4.9) requires three steps, which are explained in detail in this section. The steps are non-dimensionalisation, selecting proper solution procedure and choosing appropriate mesh.

4.2.3.1 Dimensionless Formulation

In order to non-dimensionalise equations (4.8) and (4.9), all independent (x, y, t) and dependent (ψ, T) variables need to be re-defined to introduce the relevant non-dimensional parameters. The new definitions are in terms of characteristic values, this is achieved by the following non-dimensional procedure (Al-Anii, 2005). First, X and Y are the new spatial parameters and them re-produced with respect to CPU length D :

$$X = \frac{x}{D}, \quad Y = \frac{y}{D}, \quad \tau = \frac{t\alpha}{D^2}, \quad \Psi = \frac{\psi}{\alpha}, \quad \theta = \frac{(T - T_\infty)}{(T_s - T_\infty)} \quad (4.10)$$

Substituting the parameters of equation (4.10) into equations (4.8) and (4.9), and then grouping dimensional terms to obtain the non-dimensional characteristics numbers called Rayleigh number, Ra , and Darcy number, Da , respectively:

$$Ra = \frac{g\rho\beta D^3 \Delta T}{\alpha\mu} \quad (4.11)$$

$$Da = \frac{K}{D^2} \quad (4.12)$$

These two non-dimensional numbers can be combined to define the modified Rayleigh number as:

$$Ra^* = Ra \cdot Da = \frac{g\rho\beta DK\Delta T}{\alpha\mu} \quad (4.13)$$

Re-arranging the non-dimensional governing equations yields equations to be used in subsequent steps. Therefore, the momentum and energy equation in the non-dimensional form are in the form of:

$$\nabla^2 \Psi = -Ra^* \theta_X \quad (4.14)$$

$$\theta_\tau = \nabla^2 \theta - (\theta_X \Psi_Y - \theta_Y \Psi_X) \quad (4.15)$$

4.2.3.2 Solution Techniques

Firstly, the resulting equations (4.14) and (4.15) should be transformed from partial differential equations (PDEs) forms to linear algebraic forms using the finite differences method FDM. The discrete equations contain spatial and time dependent terms, rather than first and second order discretisation terms. The final form of the momentum equation (4.14) results in an elliptic partial differential equation, Poisson type, which is solved using successive over relaxation (SOR). In brief, this method solves for the entire nodal point with respect to the values of neighbourhood nodal points. For each nodal point, iterations repeated for a particular time until convergence occurs. The convergence is based on an error less than 10^{-4} which enables the desired final values of the entire distribution of stream function to be obtained (Fletcher, 1991, Fletcher, 1988).

On the other hand, the energy equation (4.15) is a parabolic partial differential equation. A time marching technique is employed as the optimal method of solution, where the growth of the non-dimensional temperature is based on a convergence based on an error of the order of 10^{-6} . Since a steady state solution will emerge, the dependency of the energy equation upon a time will vanish after a period of successive time intervals and convergence is reached (Fletcher, 1991, Fletcher, 1988). The reduced and formulated finite difference forms of the two governing equations are written as follows:

$$\Psi_{i,j} = f_1(\Psi_{i-1,j}, \Psi_{i,j+1}, \Psi_{i,j-1}, \Psi_{i+1,j}) + Ra^*[f_2(\theta_{i,j}, \theta_{i+1,j})] \quad (4.16)$$

$$\begin{aligned} \theta_{i,j}^{\tau+1} = & f_3(\theta_{i-1,j}^\tau, \theta_{i,j-1}^\tau, \theta_{i,j}^\tau, \theta_{i+1,j}^\tau, \theta_{i,j+1}^\tau) \\ & - [f_4(\theta_{i-1,j}^\tau, \Psi_{i,j-1}^\tau, \theta_{i+1,j}^\tau, \Psi_{i,j+1}^\tau) \\ & - f_5(\theta_{i,j-1}^\tau, \Psi_{i-1,j}^\tau, \theta_{i,j+1}^\tau, \Psi_{i+1,j}^\tau)] \end{aligned} \quad (4.17)$$

Where, f_1 and f_3 are the second order discretization factor of stream and temperature, respectively. Whereas, f_2 , f_4 and f_5 are the first order discretization factor. In heat transfer and fluid flow problems, it is preferable

to use the non-dimensional parameters to express the physical meaning of transport phenomena. For this reason, the current code uses the values of non-dimensional temperature to evaluate the Nusselt number. Nusselt number is a non-dimensional value which is defined as the ratio of convective to conductive heat transfer. Mathematically, Nusselt number derived to form compatible with non-dimensional temperature function as illustrated:

$$\begin{aligned}
 Q_{conv} = Q_{cond} &\Rightarrow hA(T_S - T_\infty) = -\frac{kA(T - T_\infty)}{x} \Rightarrow \\
 \therefore h &= -\frac{k(T - T_\infty)}{XD(T_S - T_\infty)} \\
 Nu_{local} &= \frac{hD}{k} = -\left. \frac{d\theta}{dX} \right|_{normal}
 \end{aligned} \tag{4.18}$$

Equation (4.20) gives the local Nusselt number along the required length which is the length of the CPUs. However, the average value can be obtained by using Simpson's rule of numerical integration along a specified length as given by:

$$\begin{aligned}
 \overline{Nu} &= \frac{\int Nu_{local} ds}{\int ds} \tag{4.19} \\
 &= \frac{\Delta x}{3} [f(x_0) + 4f(x_1) + 2f(x_2) \\
 &\quad + 4f(x_3) + 2f(x_4) + \dots + f(x_n)]
 \end{aligned}$$

4.2.3.3 Parametric Criteria

The present study investigates the overall thermal behaviour of the 2D, fully immersed in the dielectric fluid and encapsulated server. The server, have different geometrical criteria, however, the four non-dimensional parameters are defined as shown in Figure 4.3 and listed below:

1. The height of each CPU is D.
2. The width of the capsule is Lx.

3. The distance from the bottom edge of the capsule to the bottom edge of the lower CPU is Ly1.
4. The in-between distance from the higher edge of the lower CPU to the bottom edge of the upper CPU is S.
5. The distance from the top edge of the capsule to the top edge of the upper CPU is Ly2.

It is important to mention that each dimension of the capsule is based on the CPU length which is D. In addition, the CPU length is *assumed* to be equal to the server unit length, i.e. D = 1U. Therefore, each dimension is a multiplication of D which in turns mean the multiplication of 1U. This assumption is used to help in simplification of fully immersed, liquid cooled capsules and servers.

There are two geometrical aspects for the current study to better understand the opportunities and limitations of the dimensions.

1. The first stage, section 4.3, investigates the interactive thermal effects between CPUs. In addition, this stage deals with specified width Lx of 1U to easily compare the results with the available information of the 1U SunFire server. Although the server has a width of 1U, the other dimensions can be varied. For symmetry purposes, this stage assumes that Ly1 equals Ly2, and have the values of 1.0U, 1.5U, 2.0U and 3.0U. However, the in-between distance between CPUs, which is S, has the following values 0.5, 0.75, 1.0, 1.25, 1.5, 2.0, 2.5 and 3.0 all with respect to U. This wide variety of S helps in the better capture of thermal interactions between CPUs.
2. The second stage, section 4.4, presents a detailed study of selecting the optimum geometry of the server. All parameters Lx, Ly1, Ly2 and S have a range of 1U to 5U to examine the thermal behavior of each set of configurations.

Furthermore, for both stages above, the modified Rayleigh number is also varied independently in the range $Ra^*=0.5, 1, 5, 10, 20, 50, 100, 150, 200$

and 300. This range represents the utilization levels explained in the previous chapter.

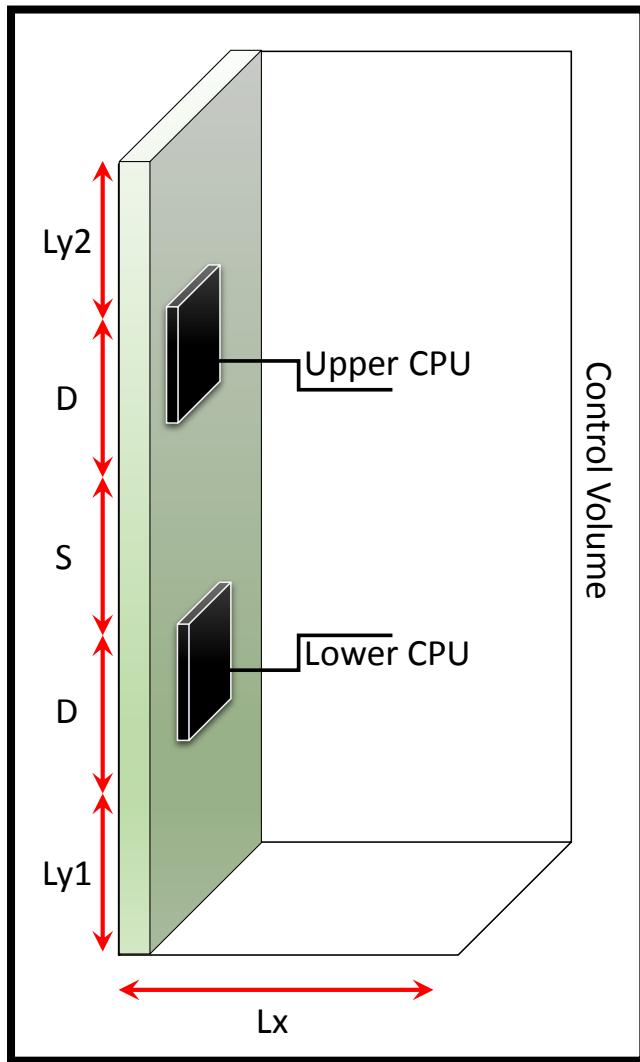


Figure 4.3 Schematic diagram of the physical domain of studied capsule.

4.2.3.4 Boundary and Initial Conditions

The liquid in the capsule is heated by two heat sources, namely the lower and upper CPUs. In the meantime, upper and lower edges of the capsule are insulated, to ensure that all the heat captured by the fluid is transferred to the opposite side. This side modelled as a heat sink where the heat is rejected to the primary water circuit. Therefore, the boundary condition is given as a constant zero non-dimensional temperature condition. In addition, the left edge of the capsule is insulated from three regions; below

the lower CPU, above the lower CPU and the in-between the CPUs. The geometrical dimensions, nodal mesh and boundary conditions are illustrated in Figure 4.4.

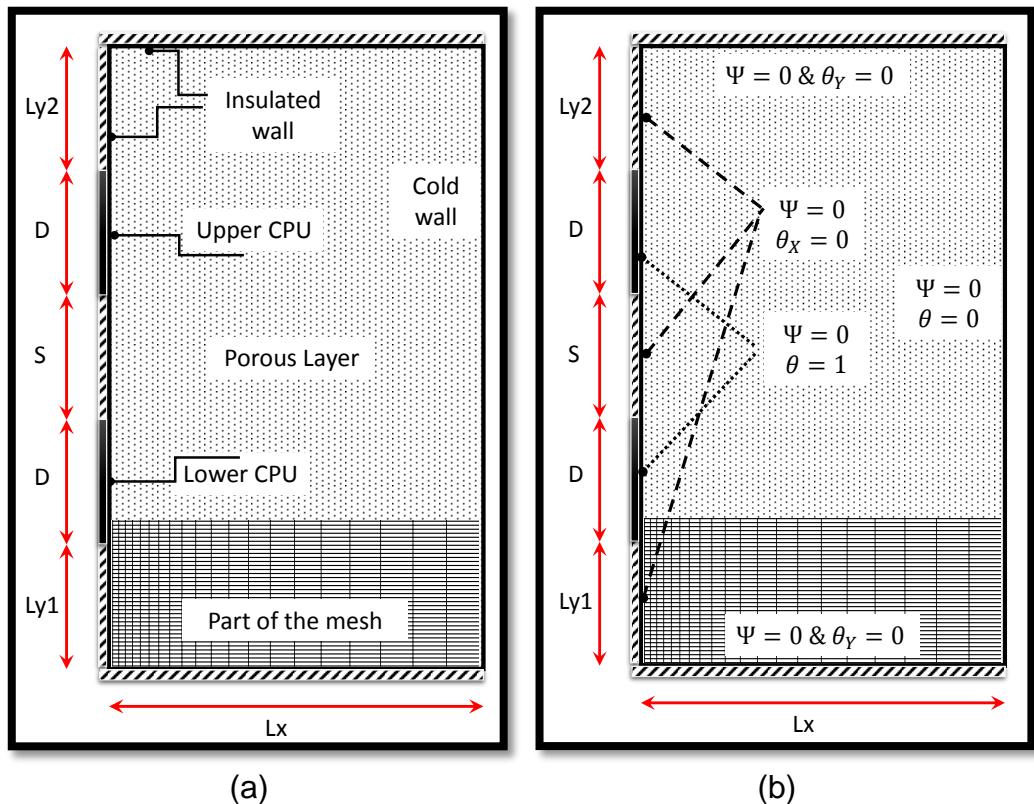


Figure 4.4 Schematic diagram of the numerical domain of studied capsule (a) definitions and (b) boundary conditions.

Table 4.1 The boundary conditions.

Part	BCs	Numeral BCs
CPU	Heat Source Hot wall	$\theta = 1$
Right side	Heat Sink Cold wall	$\theta = 0$
Bottom edge	Insulated wall	$\frac{\partial \theta}{\partial Y} = \theta_Y = 0$
Below the lower CPU	Insulated wall	$\frac{\partial \theta}{\partial X} = \theta_X = 0$
In-between the CPUs	Insulated wall	$\frac{\partial \theta}{\partial X} = \theta_X = 0$
Above the upper CPU	Insulated wall	$\frac{\partial \theta}{\partial X} = \theta_X = 0$
Top edge	Insulated wall	$\frac{\partial \theta}{\partial Y} = \theta_Y = 0$
ALL boundaries	Non-Permeable	$\Psi = 0$

The parameters are presented in the non-dimensional form in Table 4.1. In addition, the table includes the numeral value of each boundary condition and the definition as well. Finally, all non-dimensional parameters are given a zero-initial condition.

4.2.3.5 Mesh

The physical domain of the present study has a rectangular cross section. This cross section matches the Cartesian coordinate system in Figure 4.5 (a). However, the mesh of the domain is stretched and clustered toward the CPUs to investigate the thermal interactions more accurately Figure 4.5 (b). For the mesh stretching process, the x-axis nodal points are re-distributed by defining a slope of distribution of 0.2, and damping factor of 2 (Fletcher, 1991, Fletcher, 1988):

$$x = x_{i=1} - \Gamma(x_{i=1} - x_{i=max}) \quad (4.20)$$

$$\Gamma = 0.2\Lambda + 0.8 \left(1 - \frac{\tanh(2 - 2\Lambda)}{\tanh(2)} \right) \quad (4.21)$$

Equations (4.20) and (4.21) give the criteria of these values and the new position of the nodal point. To check the mesh independency, different mesh sizes selected. The range of the mesh size is from 11x101 to 61x101 nodes. The measured Nusselt number against that range varies as shown in Figure 4.6. It is obvious that the values of Nusselt number being independent. Therefore, the selected mesh size is 46x101 with a change of less than 0.25% from 61x101 mesh.

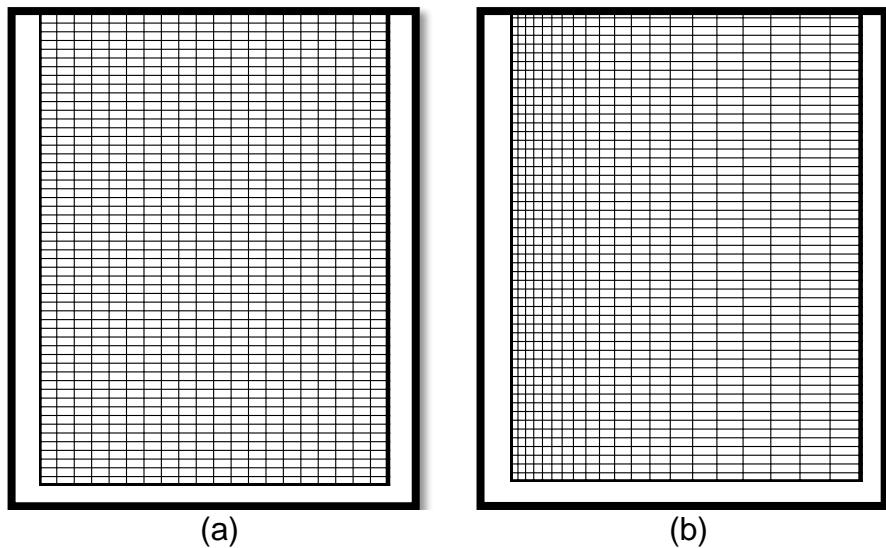


Figure 4.5 The computational mesh (a) normal and (b) stretched toward the CPUs.

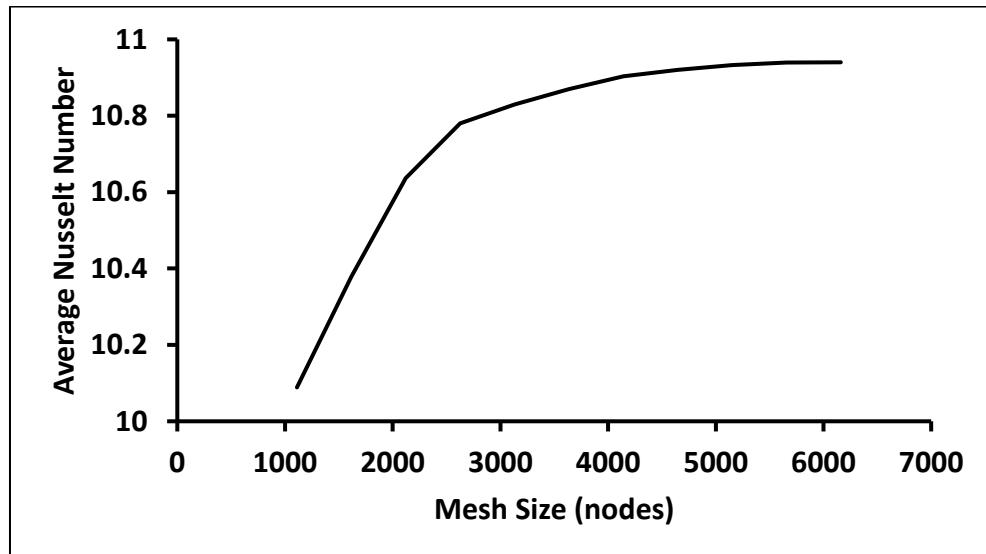


Figure 4.6 Mesh independency test.

4.2.3.6 Code Validation

The validation of the initial results is conducted in three steps. The first step tends to validate the presenting of the governing equations. To perform this step, the calculated Nusselt number requires a validation with a generalised formula. The natural convective inside a porous layer has been studied extensively by Bejan and Kraus (2003). So, the code prepared to mimic the conditions of (Bejan and Kraus, 2003) and the results shown in Figure 4.7.

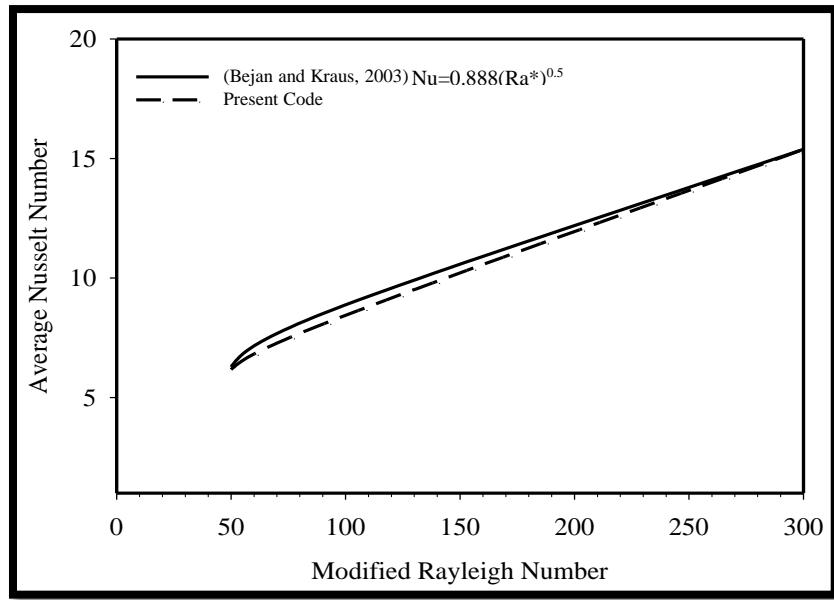


Figure 4.7 Validation of the governing equation used in the present code with (Bejan and Kraus, 2003)

The next step includes comparing the results of non-dimensional temperature and stream function with a case study. The selected case study investigated the natural convective heat transfer inside a porous cavity (Nithiarasu et al., 1997). To perform this step, the code geometry and boundary conditions are modified while the governing equations kept as in equations (4.16) and (4.17). Figure 4.8 shows good qualitative agreement between present code results and (Nithiarasu et al., 1997) for both non-dimensional stream function and temperature.

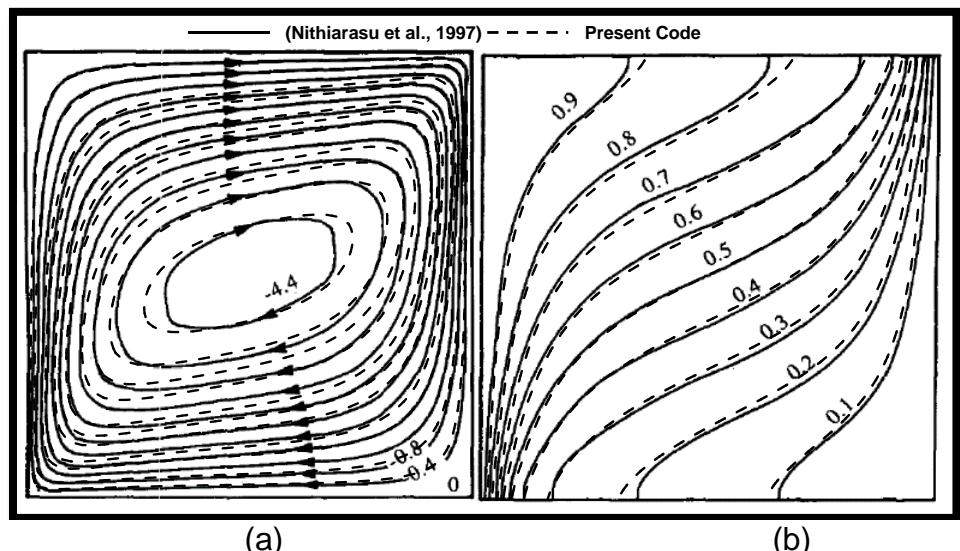


Figure 4.8 Comparing the non-dimensional values of (a) stream function and (b) temperature lines, code vs. (Nithiarasu et al., 1997)

The final step compares current simple model with the corresponding predictions from solutions of the full two-dimensional Navier-Stokes for natural convection using COMSOL. A collaboration with Almaneea et al. (2014) to validate the current model is carried out. With six cases of same geometry and boundary conditions and a Rayleigh number of 50,000, the solutions of the two models are compared. For the simplified model the modified Rayleigh number, Ra^* , and Darcy number are set to 50 and 10^{-3} , respectively, to simulate the same level of convective heat transfer. The Nusselt number is calculated in both cases for a range of CPU configurations and is presented in Figure 4.9. The maximum differences between results for all cases of lower and upper CPUs is 5.9% and 7.3%, respectively. These agreements in results suggest that the porous medium modelling assumptions are valid.

From the three steps mentioned earlier, it is obvious that derived governing equations, geometry and boundary conditions settings and the application of the present code is valid. This validity gives the code ability to use in design and analyse the thermal performance of encapsulated server.

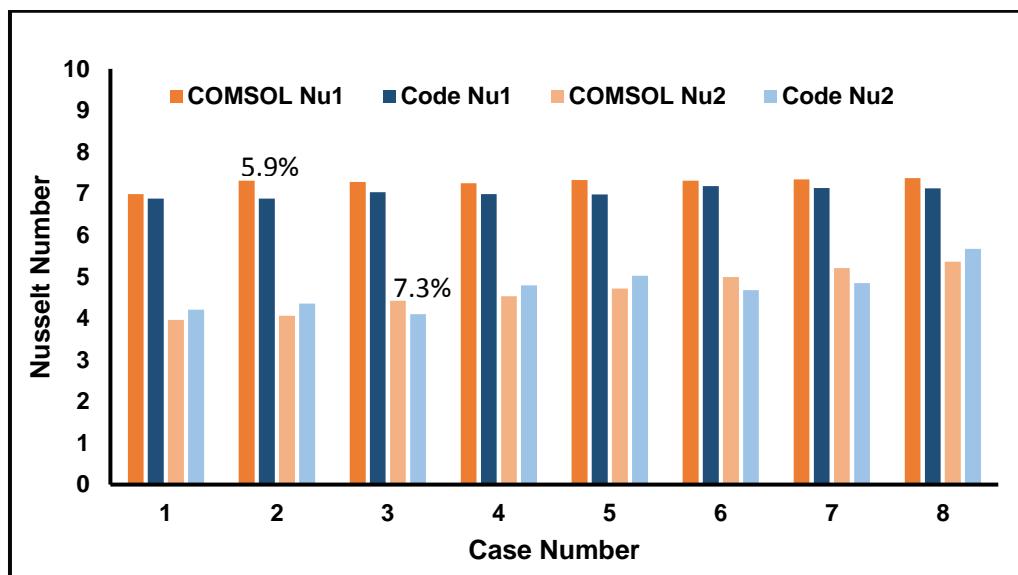


Figure 4.9 Compare lower and upper CPU's Nusselt numbers for the present code versus COMSOL.

4.3 Modelling of a 1U Encapsulated Server

4.3.1 The Parametric Study

As mentioned in the parametric criteria, section 4.2.3.3, this section includes the study of thermal interaction between CPUs. This section deals with server width L_x of 1U to ease the comparison with results from previous chapter. The parametric study applies to vary the geometrical parameters as shown in Table 4.2.

Table 4.2 DoE for the first model of encapsulated server.

L_x [U]	$Ly_1=Ly_2$ [U]	S [U]
1.0	1.0, 1.5, 2.0 & 3.0	0.5, 0.75, 1.0, 1.25, 1.5, 2.0, 2.5 & 3.0
$Ra^* = 0.5, 1, 5, 10, 20, 50, 100, 150, 200 \& 300$		

4.3.2 Results and Discussion

The positions of the CPUs are represented using two spatial parameters, namely Ly , the vertical distance of the first CPU from the bottom of the capsule, and S , the vertical spacing between the CPUs. The heating fluxes, which are related to the workload of the CPUs, are represented by varying the Rayleigh number.

Figure 4.10 presents the ratio of the fluid velocity at grid points adjacent to the lower CPU to corresponding velocities along the CPUs to show interactions between the thermal plumes associated with each CPU. Ly has an influence on the patterns of velocity ratio and as Ly increases the velocity ratio slightly decreases. For example, for $Ly=1U$, the velocity at the leading edge of lower CPU is smaller than the similar position of the upper one. The heating effect of the lower CPU is initially more abrupt than at the upper one since in the latter case the cooling fluid is already carrying more thermal energy. However, as the liquid moves up the upper CPU the velocity

increases more quickly due the thicker thermal boundary layer and this effect is particularly strong for the upper quarter of the upper CPU. Furthermore, when Ly is set to 3U, the mean velocity above the upper CPU is significantly bigger than over the lower CPU due to the larger distance to the upper edge of the capsule. Figure 4.10 suggests that $Ly=1.5$ represents a critical value since the trend of decreasing velocity above the CPUs (for $Ly < 1.5$) changes to that of increasing velocity above the CPU for $Ly > 1.5$.

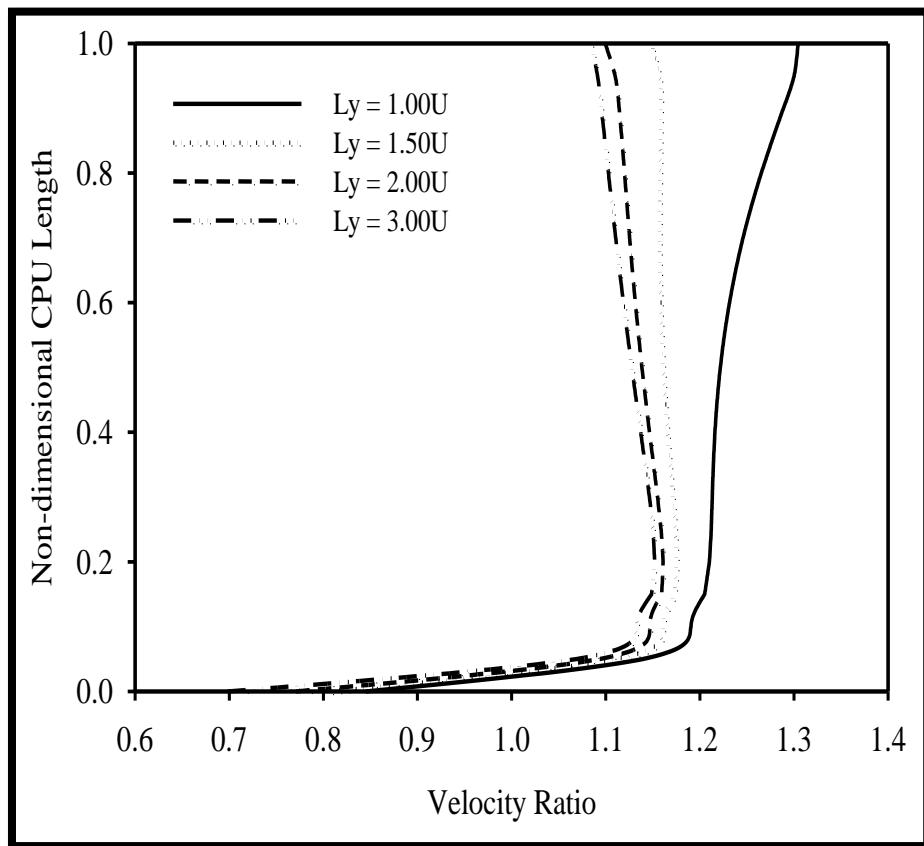


Figure 4.10 Effect of non-dimensional position of CPUs on ratios of non-dimensional velocity, at $S=1$ and $Ra^*=50$.

From a heat transfer perspective, the boundary layer thickness has a direct effect on the heat removal from the CPUs. This is investigated via the local Nusselt number being calculated along unit each CPU. As with the velocity ratios shown in Figure 4.10, the ratio of local Nusselt numbers for the lower to upper CPUs is plotted in Figure 4.11. For $Ly=1$, the lower CPU has a local Nusselt number that is on average twice that of the upper CPU. The upper CPU has a thicker boundary layer since the fluid adjacent to it is hotter because it is located closer to the lower CPU, i.e. $S=1U$. Therefore,

the ability of the dielectric fluid to remove heat from the upper CPU is reduced. As Ly increases the ratio of the two local Nusselt numbers decreases, since the boundary layer thicknesses over both CPUs are similar. It can also be seen that the influence of Ly on the local Nusselt number ratio reduces since Figure 4.10 and Figure 4.11 show marginal differences between $Ly=2$ and $Ly=3$.

The effect of the vertical CPU separation, S , is now investigated and the results are shown in Figure 4.12 and Figure 4.13. As S increases, the velocity ratio decreases incrementally as can be seen in Figure 4.12 and the opportunity to re-accelerate the convection current is increased, resulting in a decrease of the velocity ratio.

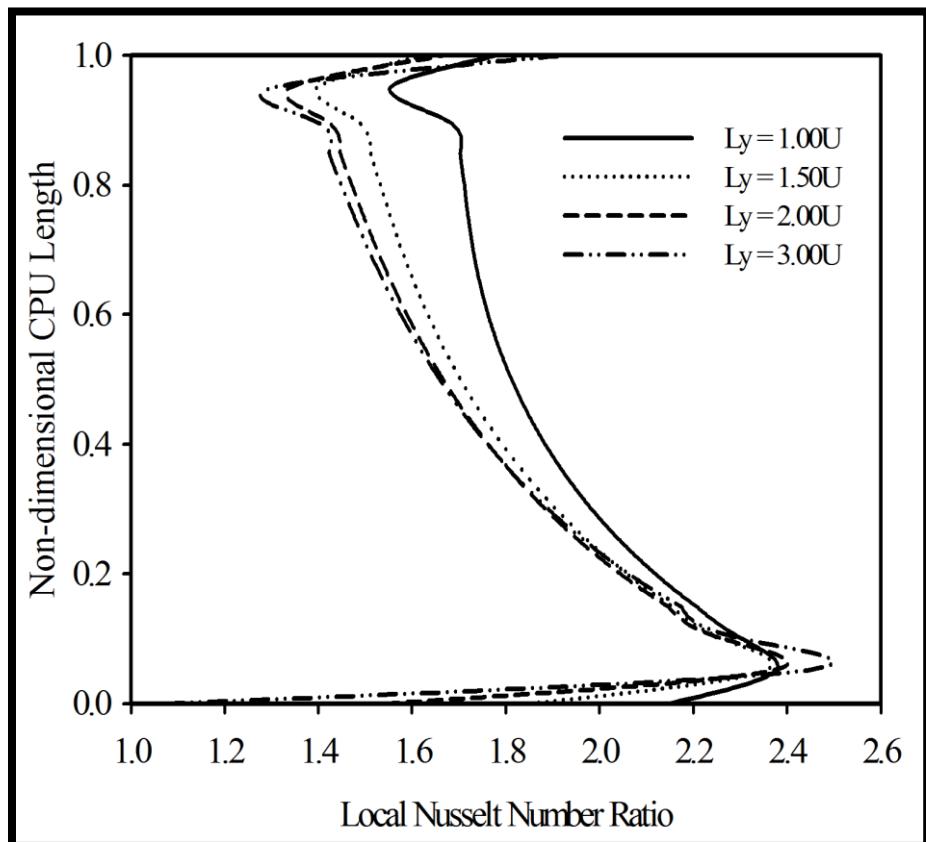


Figure 4.11 Effect of non-dimensional position of CPUs on ratios of local Nusselt number, at $S=1$ and $Ra^*=50$.

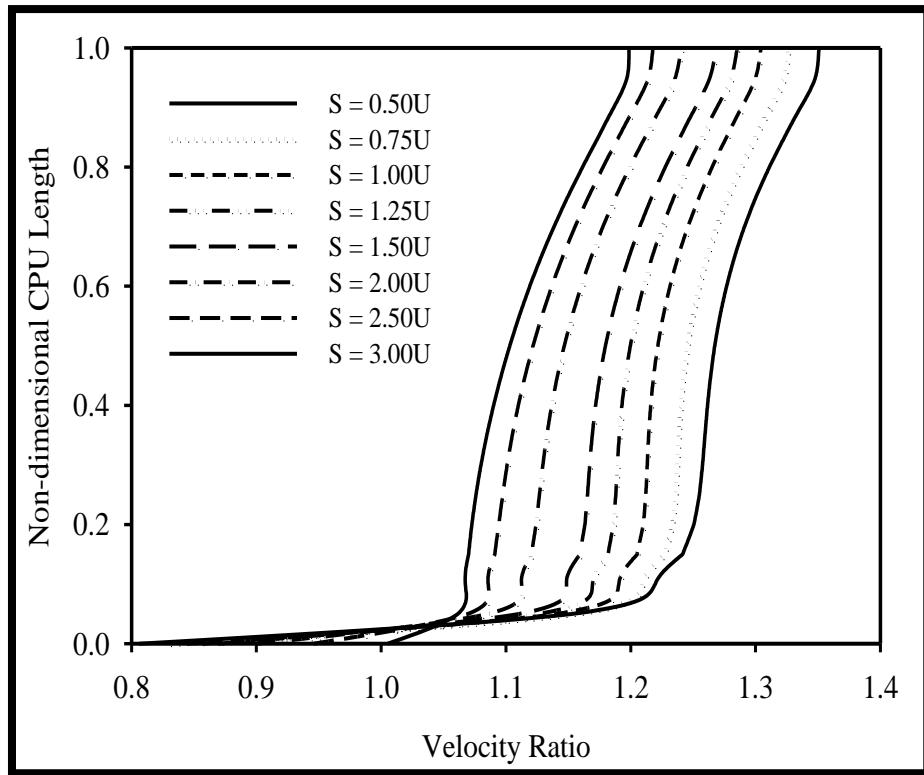


Figure 4.12 Effect of non-dimensional position of CPUs on ratios of non-dimensional velocity at $Ly=1$ and $Ra^*=50$.

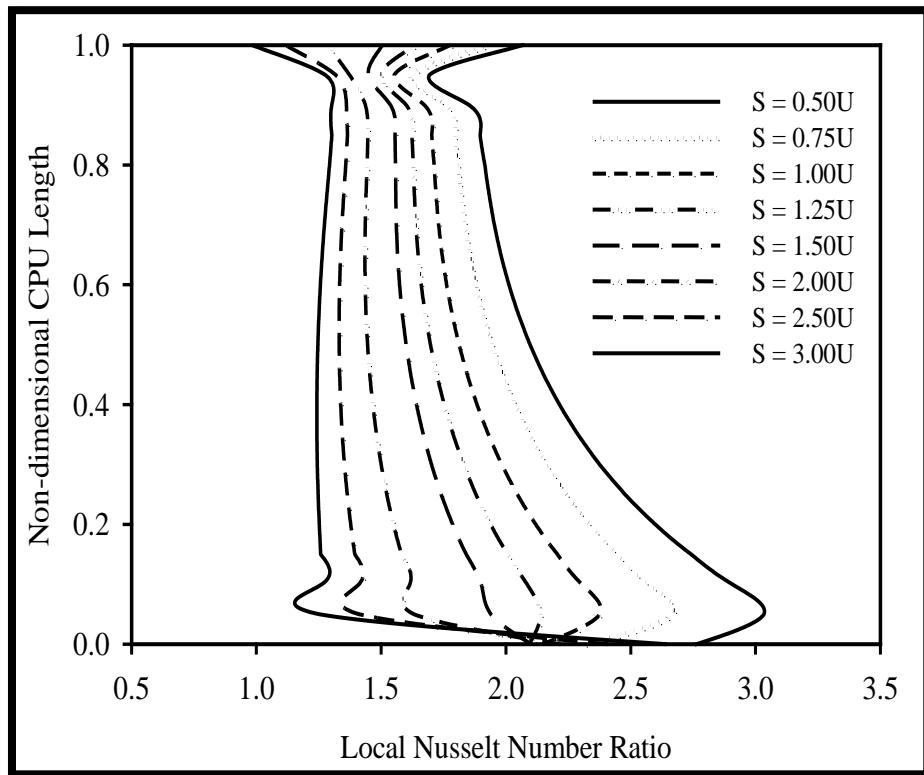


Figure 4.13 Effect of non-dimensional position of CPUs on ratios of local Nusselt number, at $Ly=1$ and $Ra^*=50$.

The effect of increasing S on the Nusselt number is more interesting, as shown in Figure 4.13. When S is small, there is a strong interaction between the convection currents over the CPUs and the lower current influences the upper one over the first quarter of upper CPU's length, leading to a lower local Nusselt number that is three times greater than the upper one. As S increases the Nusselt number over the lower CPU is still higher than that over the upper CPU, however the latter is increasing towards the former. For S values greater than $2U$, the Nusselt number over the lower CPU is approximately constant while over the upper CPU it increases. This occurs since the greater vertical separation of the CPUs enables the dielectric liquid to cool before it reaches the upper CPU, which results in a greater acceleration of the convection current over the upper CPU.

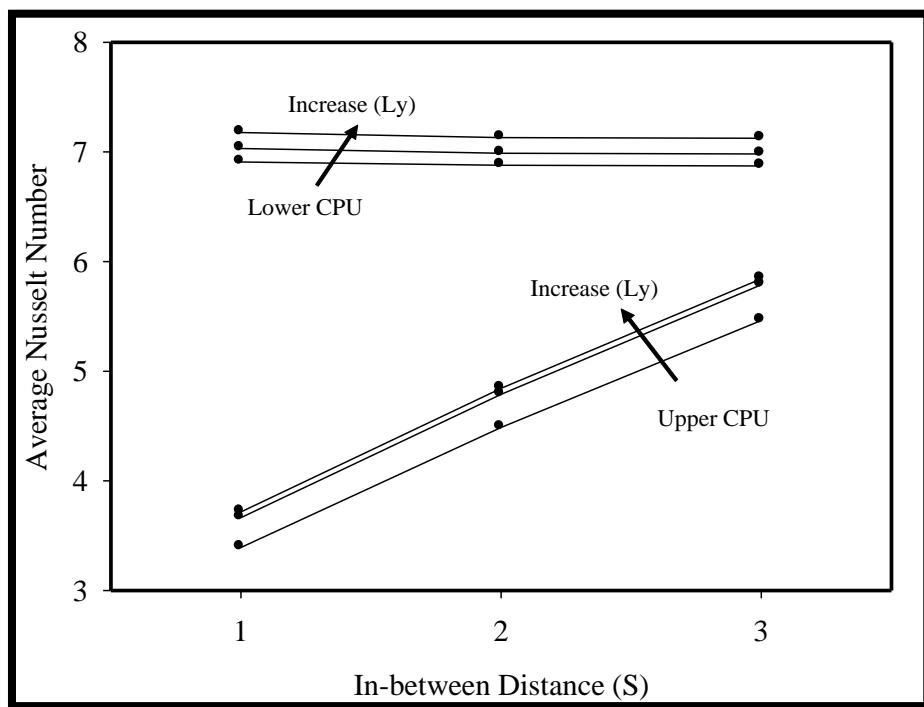


Figure 4.14 Effect of in-between distance (S) on average Nusselt number for $Ly=1, 2$ and 3 , at $Ra^*=50$.

Figure 4.14 illustrates the effect of the in-between distance S , with combination of increasing Ly_1 , on average Nusselt number. It is obvious that both lower and upper CPUs show an increase in Nusselt number with increasing Ly . Although the in-between distance does not affect Nusselt

number of the lower CPU, the Nusselt number of the upper one increases significantly.

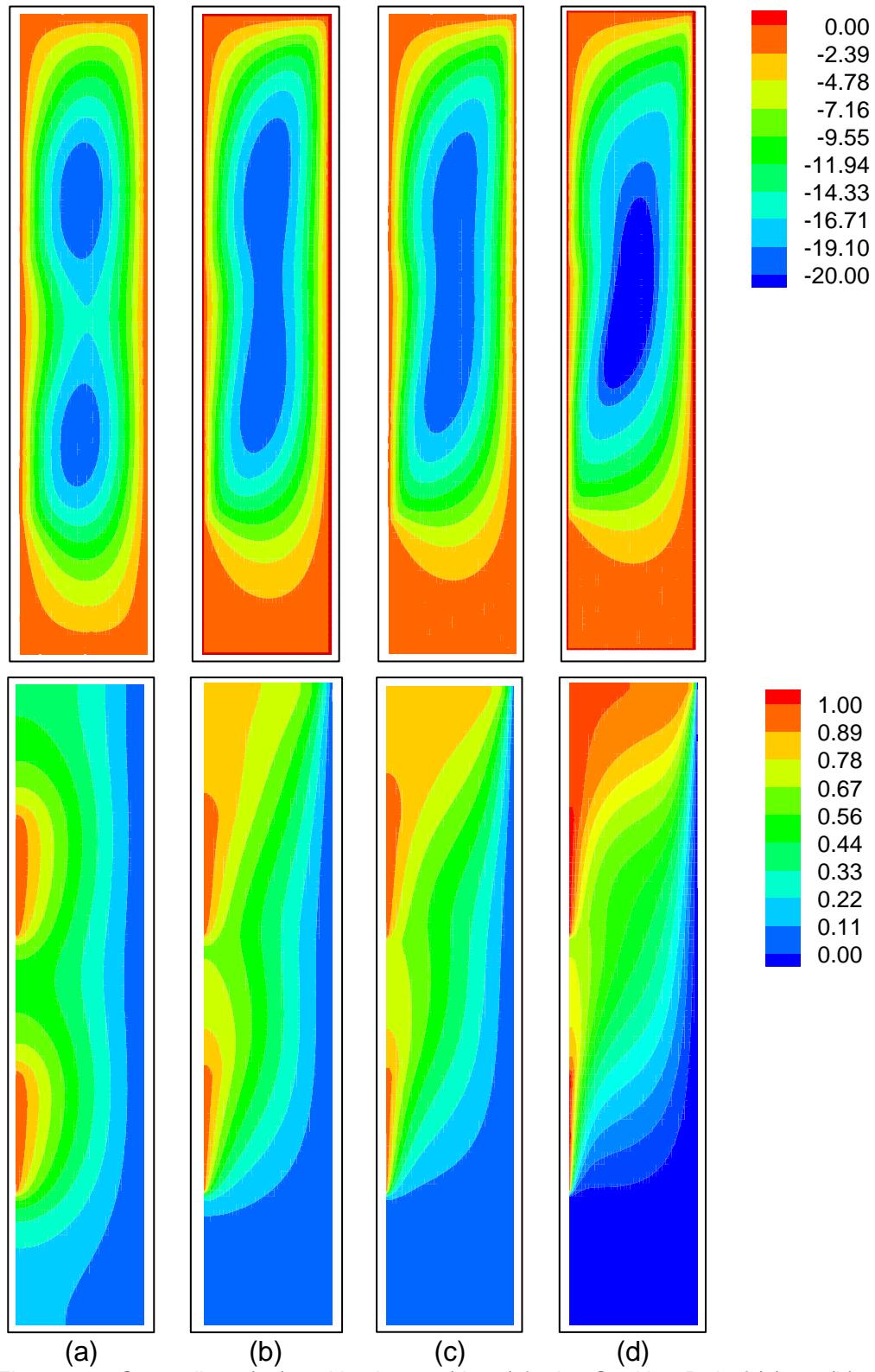


Figure 4.15 Streamlines (up) and isotherms (down) for $Ly=S=1U$ at Ra^* of (a) 10, (b) 50, (c) 100 and (d) 300.

The streamlines and isotherms are related since the fluid circulation is a function of density change, which is also dependent on the local temperature gradient. Figure 4.15 shows stream-lines for different values of the modified Rayleigh number and as the buoyancy forces are increased with increasing modified Rayleigh number, the stream-lines are displaced upwards and the convection current strengthens due to the larger local temperature differences shown in Figure 4.15. As the modified Rayleigh number increases the maximum stream function value increases (implying a faster natural convection current) resulting in a larger average Nusselt number.

This increased heat transfer is shown more clearly by the relationship between average Nusselt number and modified Rayleigh number, presented in Figure 4.15. The data can be described by a single correlation equation, which highlights the influence Ra^* , Ly and S on the average Nusselt number. A Gauss-Newton technique is employed in obtaining this correlation and the result is shown in equation (4.22). The correlation coefficient between the curve (4.22) and the computational data is 0.968.

$$Nu = 0.801 \text{ Ra}^{*0.5} \text{ Ly}^{0.075} S^{0.085} \quad (4.22)$$

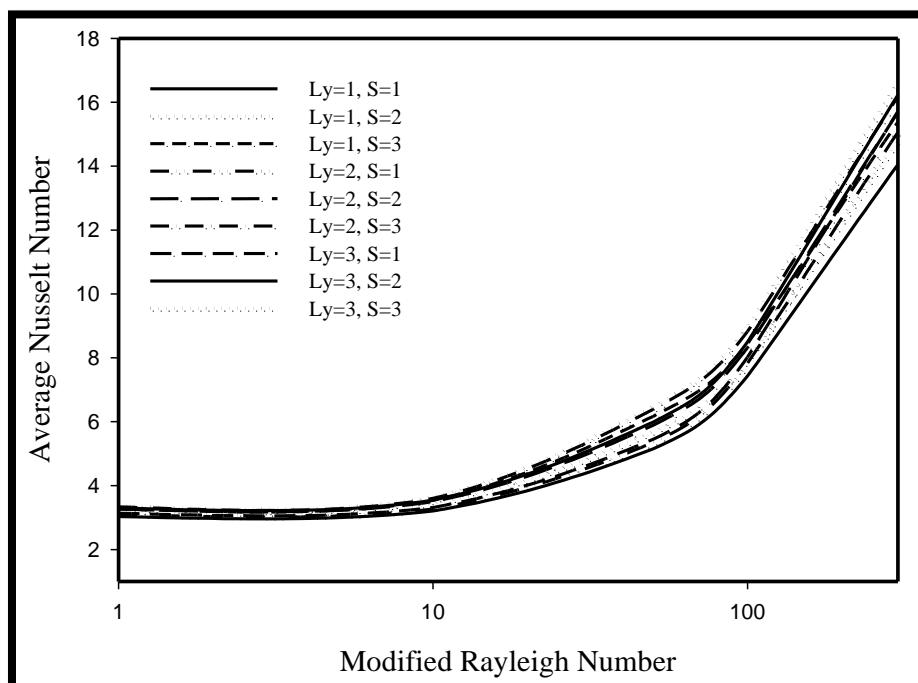


Figure 4.16 Average Nusselt number vs. modified Rayleigh number for all spatial ratios.

4.4 Design Optimisation of the Encapsulated Server

4.4.1 The Design of Experiment

Refer to the parametric criteria, section 4.2.3.3, this section studies the optimum design of the encapsulated immersed server. The server's characteristics L_x , Ly_1 , Ly_2 and S is varied from 1U to 5U. Although all vertical characteristics are within the range 1U to 5U, however, the maximum overall width of the server is restricted to 10U. The overall width Ly is given by equation (4.23):

$$Ly = Ly_1 + Ly_2 + S + 2D \quad (4.23)$$

Therefore, there are 47 possible combinations of Ly_1 , Ly_2 and S to obtain the value of 10U with presence of 2U of the CPUs. The Design of Experiment DoE applies to vary the geometrical parameters as shown in Table 4.3.

Table 4.3 DoE for the optimum design of the encapsulated server.

L_x [U]	Ly_1 [U]	Ly_2 [U]	S [U]
1.0, 2.0, 3.0, 4.0 & 5.0	1.0, 2.0, 3.0, 4.0 & 5.0	1.0, 2.0, 3.0, 4.0 & 5.0	1.0, 2.0, 3.0, 4.0 & 5.0
$Ra^* = 0.5, 1, 5, 10, 20, 50, 100, 150, 200 \text{ & } 300$			

4.4.2 Optimum Design of the Server

The current stage performs an optimisation procedure via a multi-objective problem which aims to maximise average Nusselt number and stream function for $1U \leq L_x \leq 5U$ and $5U \leq Ly \leq 10U$ for a different levels of modified Rayleigh number as shown in Table 4.3. The optimisation problem can be defined as follows:

Objective function: Maximise Nu and Ψ .

Subject to: $1U \leq L_x \leq 5U$ & $5U \leq L_y \leq 10U$.

The Moving Least Squares (MLS) method is used to build the response surface functions which are based on specified DoE points. The points obtained using an Optimal Latin Hypercube approach, that uses a permutation genetic algorithm to produce uniform points spread within the design space (Narayanan et al., 2007). The algorithm generates 30 design points that are sufficient to provide the response surface functions for both Nusselt number and stream function accurately. These 30 points are divided into two groups and they are:

1. Building points: They are 20 points and used to determine the response surface.
2. Validation points: They are 10 points and utilised to validate the response surface.

The 30 points are distributed the two design variables in the spacing design uniformly as illustrated in Figure 4.17.

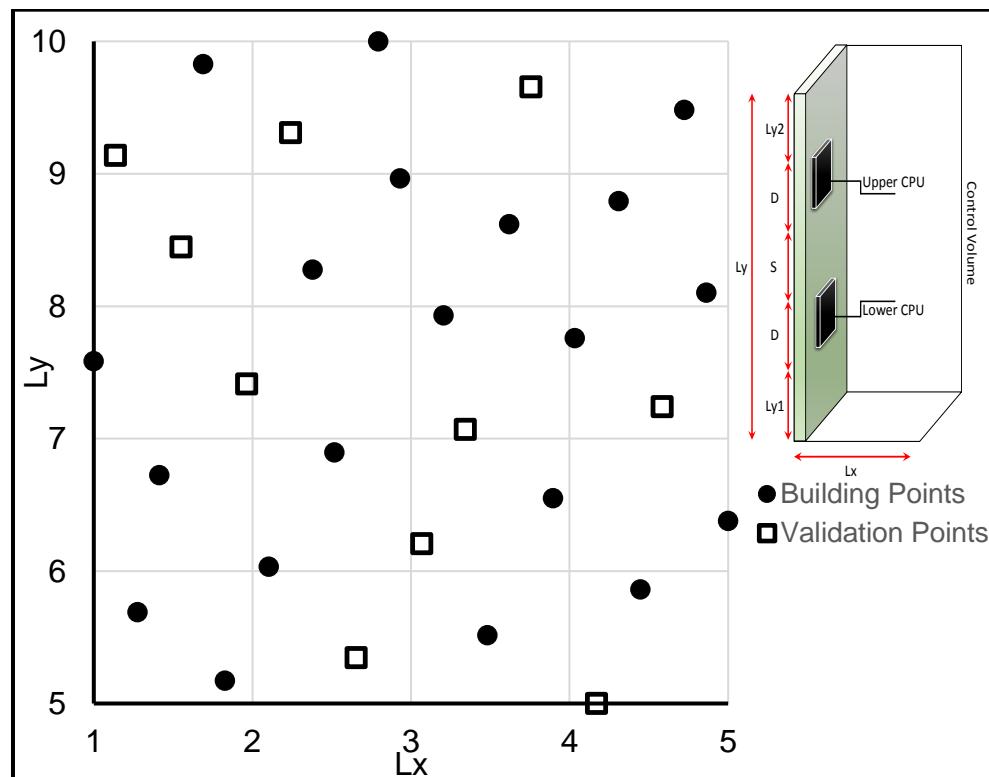


Figure 4.17 Building and validation points distribution.

4.4.3 Results and Discussion

It is very important to study the effects of each characteristic length of sever, however, this situation is not easy as expected. For this reasons, the optimization procedure is chosen to predict the optimum dimension of the desired server. The non-dimensional parameters, Nusselt number and stream function, show a fluctuation when tested with respect to the cases as shown in Figure 4.18. Meanwhile, stream function behaves very fluctuating as shown in Figure 4.19. This behavior predicts the effect of characteristic lengths on the stream values inside the encapsulated server.

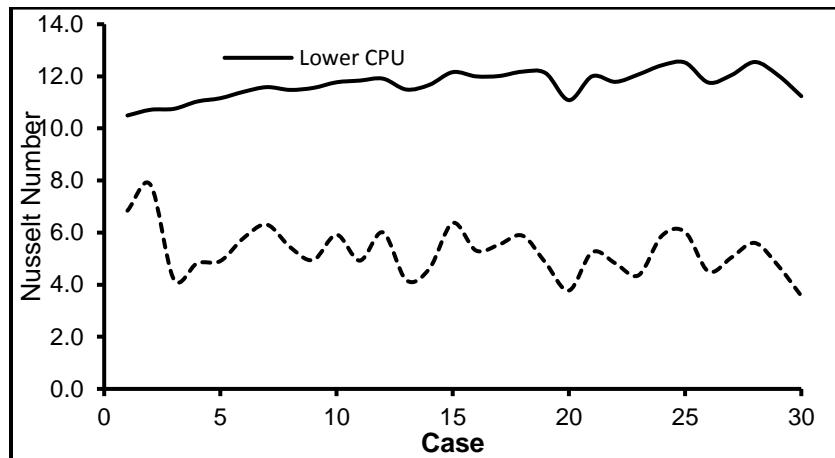


Figure 4.18 The fluctuation of Nusselt number of both lower and upper CPUs.

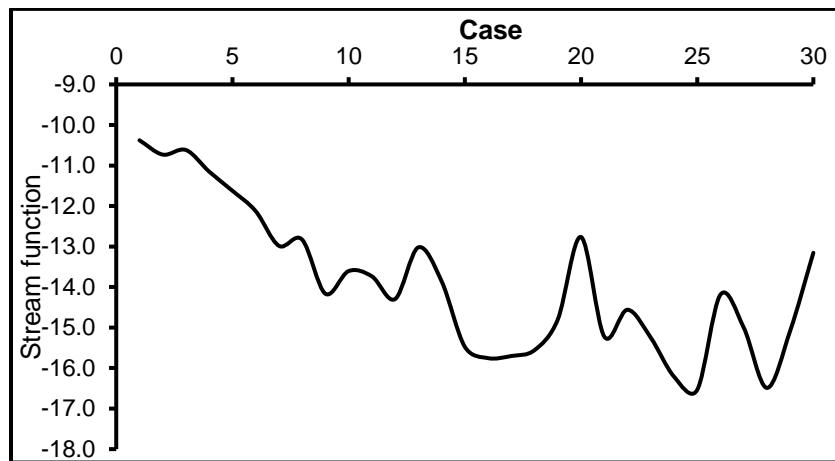


Figure 4.19 The fluctuation of stream function values.

Since the initial test predicts the fluctuation of non-dimensional parameters, then it is crucial to specify the response surfaces. Average Nusselt number

for lower and upper CPUs show different response surfaces regarding to L_x with respect to Ly_1 , Ly_2 and S . For the lower CPU, the average Nusselt number shows a maximum value near boundaries as shown in Figure 4.20 (a), (b) and (c) respectively. The maximum values occur at higher level of Ly_1 , Ly_2 and S with moderate to high width of server, i.e. more than 3.5 U. The reason behind that is due to large distances between the heat sources. So, the coolant rejects part of his heat, which gained from lower CPU, before pass across the upper CPU.

In contrast, the Nusselt number of the upper CPU be at higher values at L_x equals to 1U and high Ly_1 , Ly_2 and S . The opposite behavior exists because the upper CPU is very affected by the upper insulated edge which increase the stratification of hot fluid.

However, as the server being thinner, i.e. $L_x=1U$, the circulation of fluid being faster. So, too much higher rate of heat transfer occurs correspondingly. Figure 4.21 shows the response surfaces of Nusselt number regarding each set of characteristic lengths.

The stream function plays an important role in the thermal behavior of the encapsulated server. This important role is due to the variety of shape and values of the circulation corresponding to each design configuration.

The server of L_x more than 4U shows the highest stream function values especially for high values of Ly_2 and S . However, increasing of the lower distance Ly_1 has no significant change on the stream function values comparing to Ly_2 and S . As the server being wider, the circulation has more space to carry the heat to the opposite side of the server. This wide space provides wider range of density gradient inside the capsule which leads in turn to dissipate more heat from source to sink.

The optimised MLS method gives excellent agreement with merged DoE_m at value of $R^2=0.942$ for average Nusselt number and stream function.

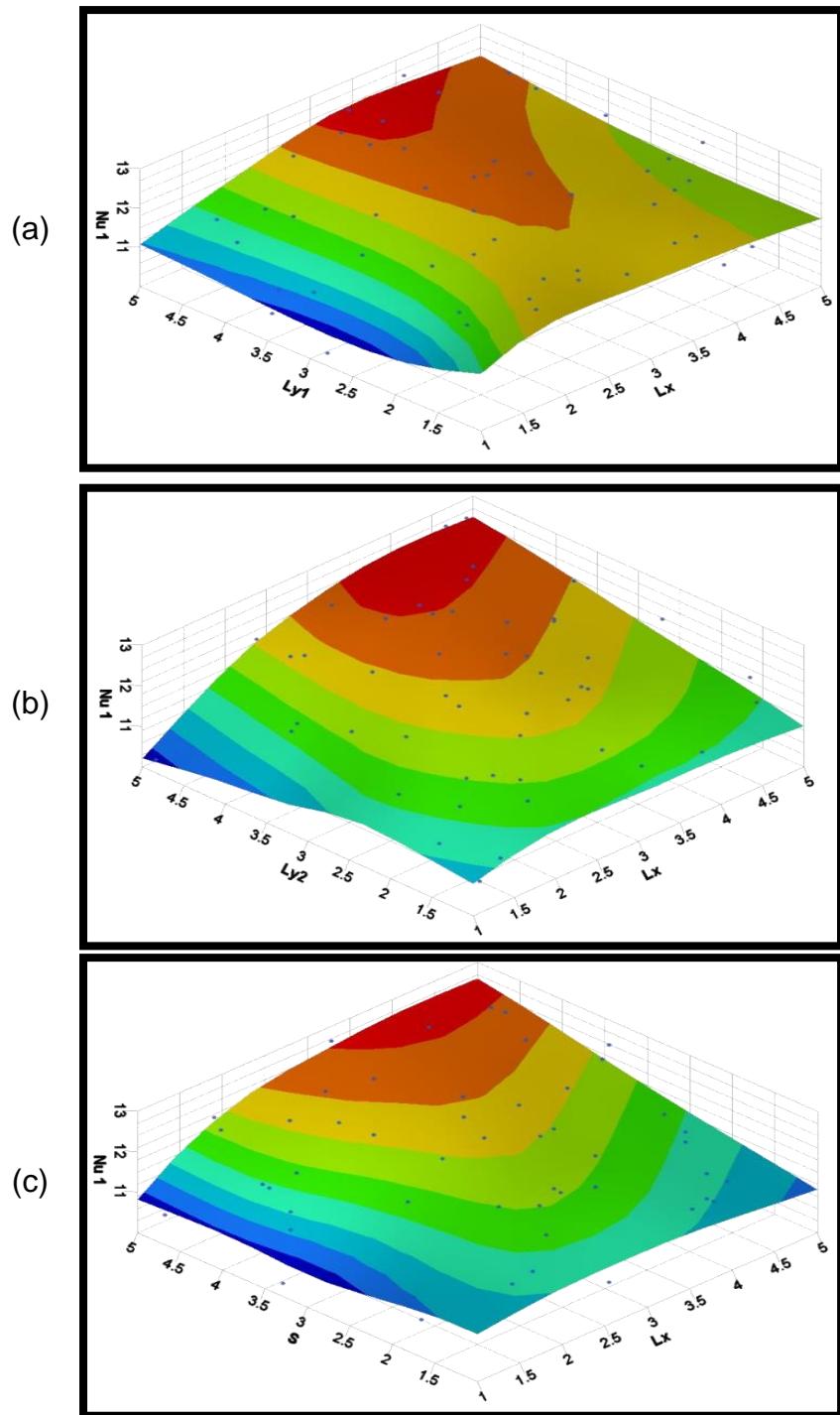


Figure 4.20 Nusselt number of the lower CPU response surfaces for L_x with respect to
(a) Ly_1 , (b) Ly_2 and (c) S .

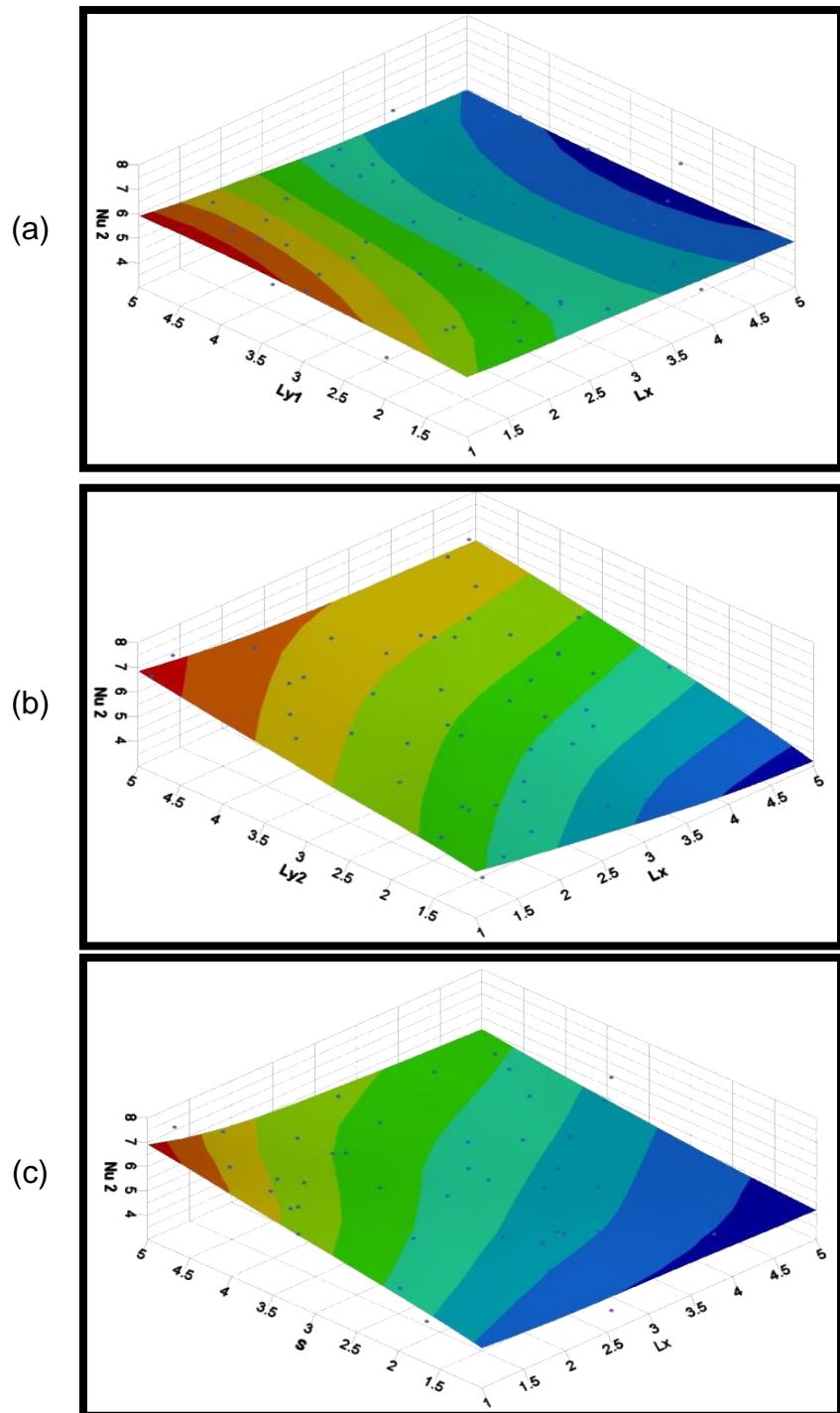


Figure 4.21 Nusselt number of the upper CPU response surfaces for Lx with respect to
(a) Ly_1 , (b) Ly_2 and (c) S .

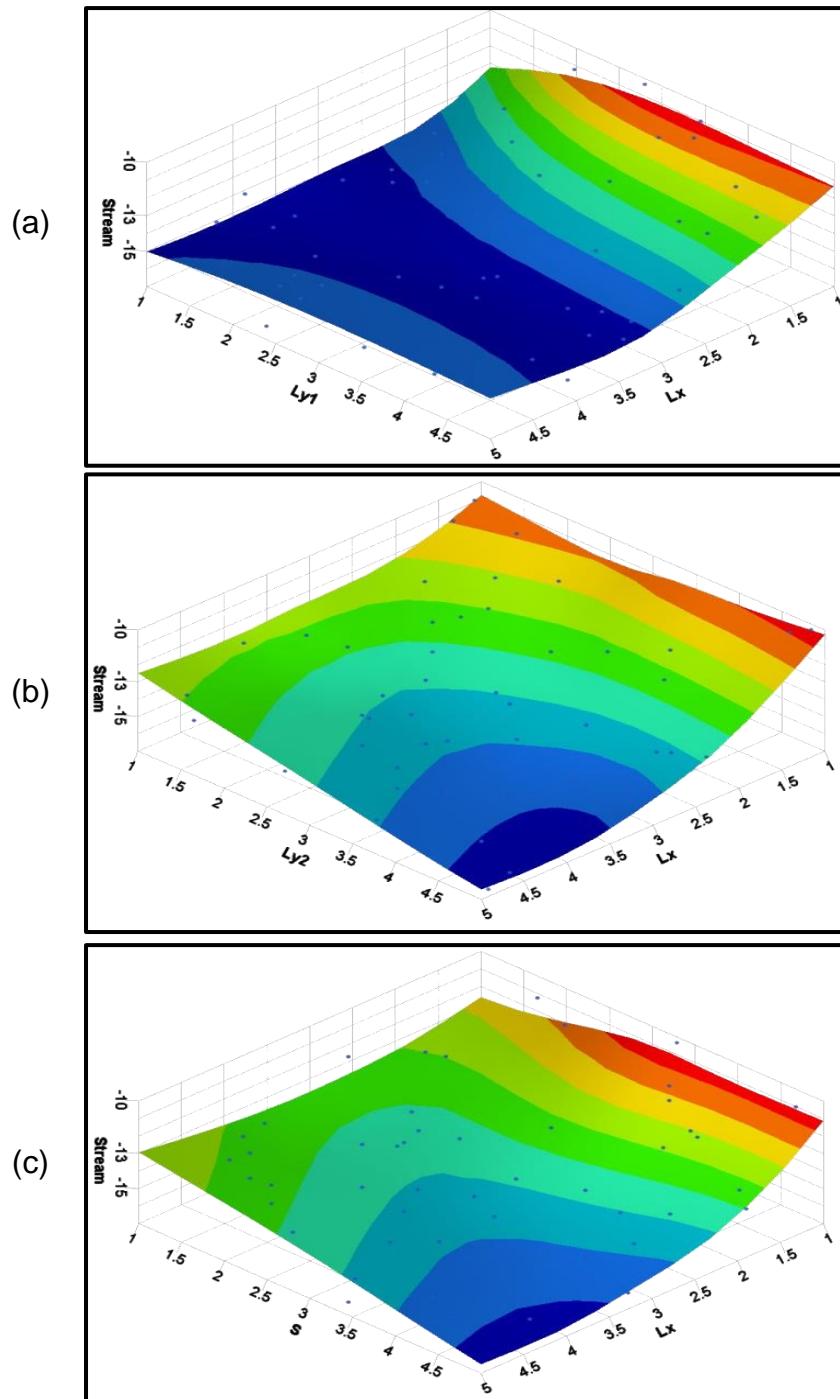


Figure 4.22 The response surface of stream function for L_x with respect to (a) Ly_1 , (b) Ly_2 and (c) S .

4.5 Summary

In this chapter, the modeling of total liquid cooling TLC technology presented. The presented model uses a novel approach to deal with the encapsulated server as a porous layer. The governing equations re-

arranged to include Darcy flow effects. The study involves non-dimensionalisation of characteristic parameters to generalize the solution. The written code uses successive over-relaxation and time marching techniques to solve momentum and energy equations, respectively.

The first stage of the investigation deals with 1U server to better understand the thermal interactions between the CPUs and the coolant. Distances from each CPU to the nearside edge and in-between the CPUs are studied extensively. Results show that the in-between distance affects the heat transfer rate more than the distance from CPU to nearside edge.

Furthermore, the model presents the optimum design of fully immersed server. The design is based on geometrical aspects and thermal behaviour for enhanced heat transfer conditions. Optimal Latin Hypercube approach is used to obtain building and validation points, whereas MLS method is used to build the response surface functions. Results evaluate different possible scenarios based on the priority of each parameter.

The presented model has limited to a 2D solution, which in turn means a limited capability of capturing the thermal performance of actual staggered CPUs layouts. However, the model can be used for initial design purposes to predict the thermal behaviour and allocate the hot spots inside the immersed servers.

Chapter 5 THE HOLISTIC POWER CONSUMPTION MODEL

5.1 Introduction

The microelectronics, and hence the datacentre, have vital interactions between the main components. These components are, as mentioned earlier in Figure 1.5, the IT workload, cooling load and energy consumption. The interactions between those sub-systems exhibit implicit and explicit relationships with respect to their operational characteristics. Thus, a logical deduction is that changes to one sub-system will impact other sub-systems within the server.

Although it is possible to directly measure the power consumption of each sub-system, the total power consumption does not equate the sum of these sub-systems values. Calculating the overall power consumption of the whole system is more complicated and challenging (Moreno et al., 2014).

For example, if the transaction / operation / process increases, it is not easy to predict the effect on the cooling load or power consumption. This is crucial when it is assumed that the interactions between the sub-systems are not correlated linearly (Sampath, 2013). Furthermore, selecting the coolant, air or liquid, reflects different performance and power consumption profiles. This becomes even more challenging outside the boundary of a single server and within the context of the entire facility. Thus, it is important that not only is the holistic power usage of a datacentre quantified, but how alterations within a specific sub-system impact such parameters.

5.2 Experimental Work

5.2.1 The Test Rig

The experimental set up was explained earlier, sections 3.4 and 3.5, and the test rig is shown in Figure 3.19. The experiments were conducted in the laboratory to investigate the sub-system operation of a SunFire V20z. Stressing the server to different levels enables to vary the IT performance creating increased server CPU utilization.

Meanwhile, measurements collected using the method illustrated in Figure 5.1. Specifically, application performance was collected from the generation of system log files detailing application performance at fixed time intervals. Server resource usage was collected through the `top` command within Linux. Server temperature was determined by using the `PCMI` command to `ls-sensors`, and server electrical power collected using a Vol-Tech 9000 power meter.

The power consumption of the cooling fans is collected from the corresponding fan speed per equation (3.2) and Table 3.6. The fan speed is captured through tachometer sensors via the IPMI platform. The power consumption of the DCLC system is given as 1.44 W per pump (CoolIT, 2012).

To create different operational conditions for experiments, the `cpu-limit` that enables the ability to control the CPU utilization of the server has been used. Each experiment was conducted 10 times for a period of 5 minutes, resulting in a total of 70 experiment cases. All sub-system metrics after experiments were automatically transferred to a Windows machine via `bash` scripts and `scp` for conducting analysis in RStudio.

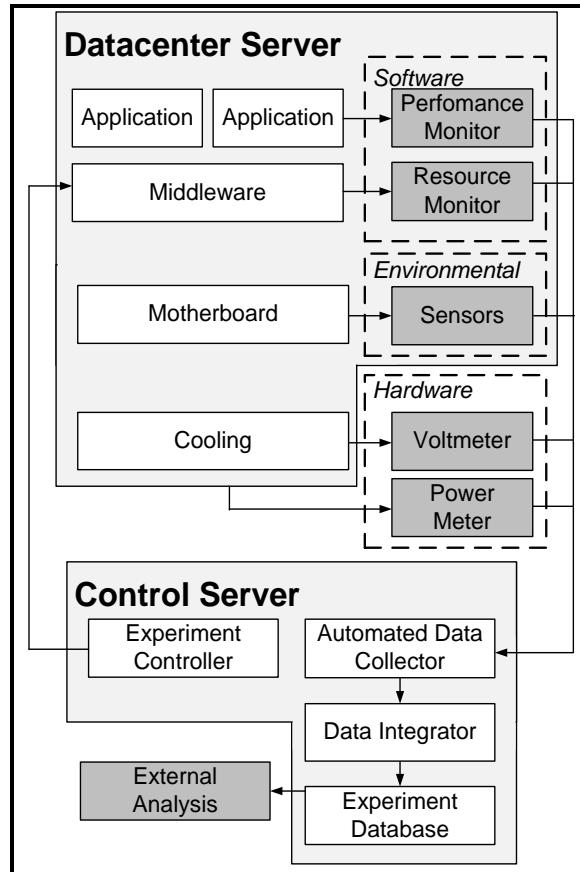


Figure 5.1 Architecture for measurement extraction.

5.2.2 Initial Results

The first stage of experiment is stressing the DCLC server and monitoring both temperatures and power consumption. As the utilization level increases the power consumption increases as well. This relation reflects the interaction between workload and power. Meanwhile, using the CPU at high utilization level to perform specific transactions / operations leads to increases the die temperature of the CPU as shown in Figure 5.2. The most interesting point is the relation between all three components.

To address this question, it is required to implement a parametric study consisting of all three parameters, i.e. utilization, power and temperature. This aspect to investigate the effect of inlet temperature on the power consumption and IT performance.

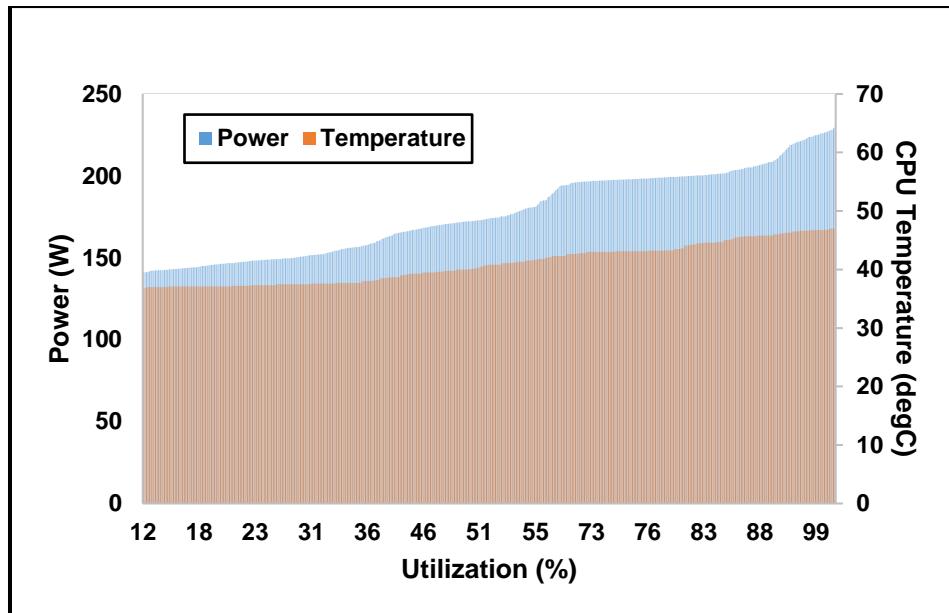


Figure 5.2 The effect of utilization level on power and CPU temperature.

5.2.3 Parametric Study

First the parametric study of this part is implemented on the DCLC server, and it includes changing the utilization level and inlet water temperature, as in Table 5.1. There was a challenge regarding the variation of water inlet temperature due to the closed circuit from the cold plate / pump end and the external fan / heat exchanger from the other end. For this configuration, it was not possible easily to get the temperature of inlet and outlet and the flow rate of water as well. Therefore, the closed circuit was replaced by an open circuit to maintain the inlet temperature while the server is at the specified utilization level. At each level, the power and CPU temperature were recorded for further analysis.

Table 5.1 Experimental parametric study for the DCLC server

Utilization (%)	Inlet water temperature (°C)
Active idle, 25, 50, 75, 100	15.7, 24.1, 35.0, 42.2, 51.3

It is worth explaining the reasons behind the selection of the lower and upper temperatures. The dew point, the point that water vapor begin to condense, in the lab was 15 °C. So, the 15.7 °C was the lowest temperature

limit to keep the server safe. While the upper limit was constrained by 51.3 °C to avoid overheating of the CPU at high utilization level. It is known from Table 3.3 that the maximum temperature of the CPU is 71 °C. However, the lower limit of warning threshold of the server was set to 62 °C.

Regarding the air-cooled server, the air conditioning system is set to provide the laboratory environment with a temperature of about 21 °C. This temperature required to be varied to capture the inlet temperature. It was possible to increase the room temperature by turning on some servers, which share the same environment. However, this scenario offered more 4 °C which still not wide range.

To overcome this problem, it is beneficial to use the numerical predictions from the 6SigmaET simulation, section 3.6. The simulation provides the sufficient data to implement more accurate calculations of the power consumption. Figure 5.3 shows the effect of air inlet temperature on the power consumption of the air-cooled server. The *Exp* refers to the experimental measurements, whereas *Num* refers to the numerical results of the 6SigmaET simulation. In addition, the effect of the water inlet temperature on the power consumption of the DCLC server is presented in Figure 5.4.

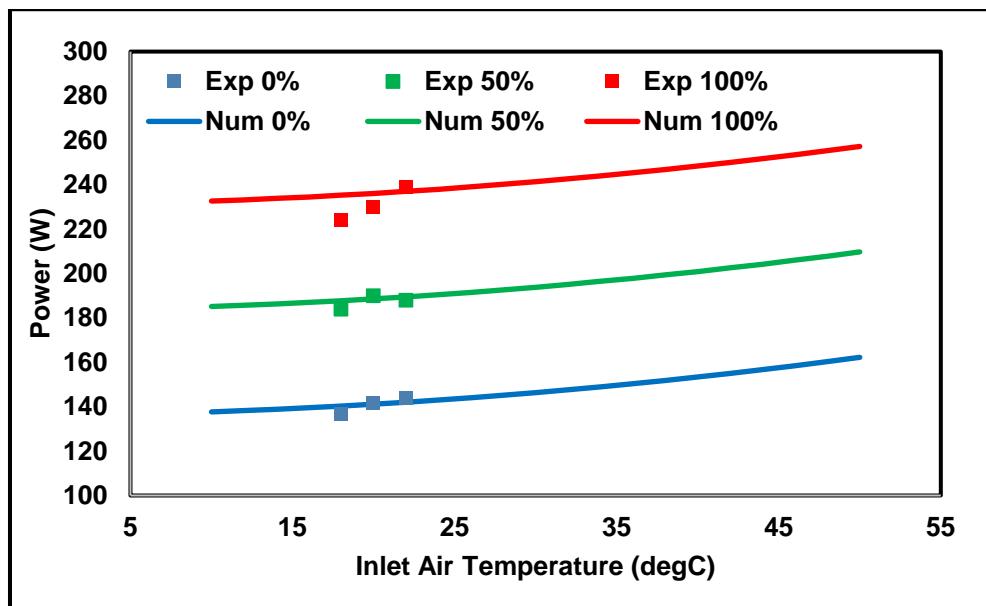


Figure 5.3 Power consumption of the air-cooled server at various inlet air temperature.

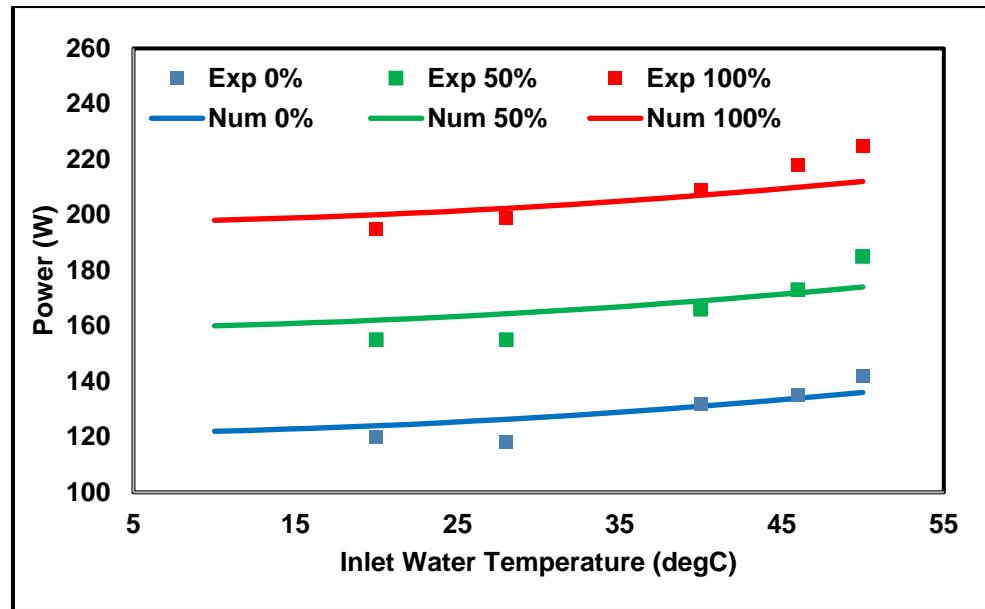


Figure 5.4 Power consumption of the DCLC server at various inlet water temperature.

5.3 The Unified Power Model

The power consumption, as stated earlier, affected by both IT workload and cooling level. Therefore, it is useful to present a unified model for capturing the relationship between resource utilization, microelectronics processor power, and cooling comprising electrical and non-electrical power.

The research focuses on the relationship between thermal characteristics of servers and the corresponding effect on the cooling system. The first part of model construction entails sectoring microelectronics processor power into three categories: baseline power, dynamic power, and static power. Baseline power (P_{base}) represents the power consumed by the motherboard and server processor(s) when it is idle (i.e. ~ 0% CPU utilization). The value of P_{base} is obtained from the SPECpowe_ssj2008 benchmark. Dynamic power $P_{dynamic}$ is the server power consumption driven by an increase in CPU resource utilization u . The results of the benchmarking for air-cooled and DCLC servers show a linear relationship between the utilization and the corresponding power consumption as expressed in equation (5.1). Static power P_{static} is the power consumed by the CPU driven by its die temperature T (i.e. physical temperature of the core). The parametric study

results in a second order polynomial relationship between power consumption and the temperature of CPUs as in equation (5.2). Each of these are interrelated to one another, and although their respective relationship appears intuitive due to experimental results. The relationship is derived from complex electronics with static leakage power of CPU consumption (solely dependent on temperature and voltage).

Each of these categories together form the total power consumption consumed by IT processes P_{IT} as shown in equation (5.3)

$$P_{dynamic} = b_1 u \quad (5.1)$$

$$P_{static} = b_2 T_{CPU} + b_3 T_{CPU}^2 \quad (5.2)$$

$$P_{IT} = P_{base} + P_{dynamic} + P_{static} \quad (5.3)$$

The constants b_1 , b_2 and b_3 for dynamic and static power are calculated based on Gauss-Newton method, as explained in section 3.4.1. The resulting parameters are listed in Table 4 with an R value of 0.983 indicating high accuracy.

Table 5. 2 Empirical constants for the power formula.

Parameters	b_1	b_2	b_3
Values	0.706	0.05	0.005

The power consumption of the SUTs additionally comprises the operation of pumps P_{pump} and/or fans P_{fan} for cooling. This results in the total power consumption of the entire server P_{Server} is defined as in equation (5.4).

$$P_{Server} = P_{IT} + \sum_{i=1}^m (P_{pump})_i + \sum_{j=1}^n (P_{fan})_j \quad (5.4)$$

Where m and n represent the total number of pumps and fan within the server, respectively. This summation of cooling components allows for the model to capture the power profiles of server cooling comprising liquid, air or a combination of both.

The total power drawn by a server P_{total} includes both power consumed by the microelectronics components, P_{Server} , and power losses in the PSUs, P_{PSU} . It is possible for PSUs to exhibit varied operational efficiencies, i.e. the ratio between input and output of electricity.

$$P_{PSU} = \sum_{k=1}^l ((1 - \eta) P_{PSU})_k \quad (5.5)$$

Where η represents the power supply unit efficiency percentage and l represents the number of PSU used to power the server.

While P_{total} represents the consumed power by microelectronics components, such components require cooling power internal and external of the server to function. Internally, there exist two types of cooling; (i) the DCLC pumps, and (ii) the fans which move colder air to pass over microelectronic components. The hot working fluid requires mechanical work to reject heat and back to the supply condition. This is performed by the heat exchanger and fans, and represents the external power $P_{external}$. As shown in Figure 5.5, $P_{external}$ represents the amount of power expended to achieve temperature homeostasis between the input and output temperature of the server.

Thus, the summation of internal and external power consumption of cooling forms the cooling load $P_{cooling}$:

$$P_{cooling} = \sum_{i=1}^m (P_{pump})_i + \sum_{j=1}^n (P_{fan})_j + P_{external} \quad (5.6)$$

The total power consumption of the SUT when considering all sub-systems holistically is represented in equation (5.7).

$$P_{total} = P_{IT} + P_{PSU} + P_{cooling} \quad (5.7)$$

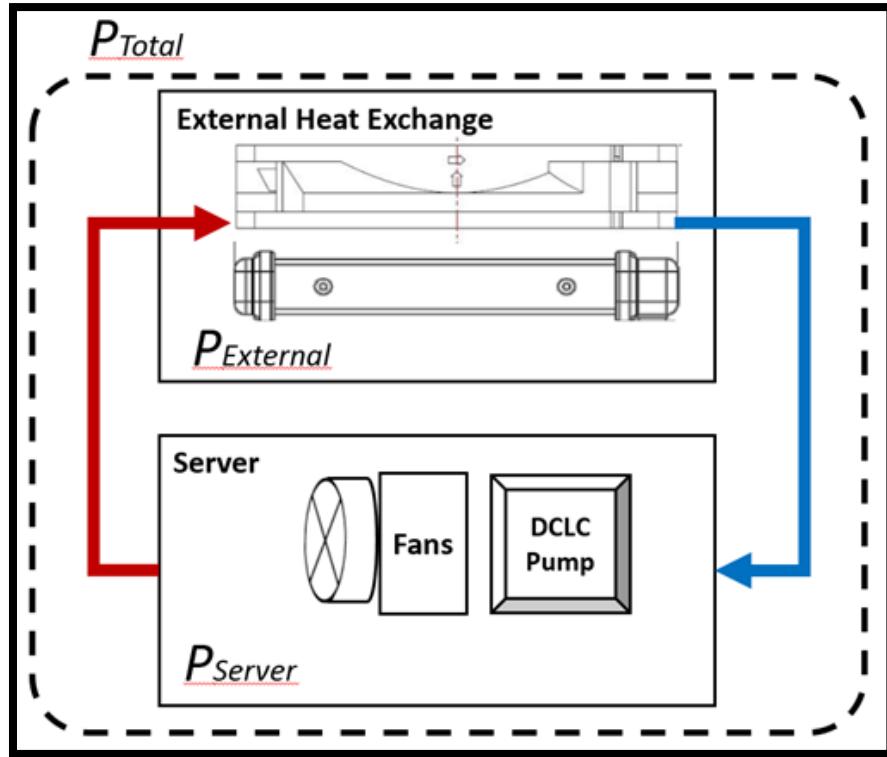


Figure 5.5 Illustration of the breakdown of power consumption.

5.4 The Power Usage Effectiveness PUE

The power usage effectiveness (PUE) is a widely-used metric in datacentre industry. In this part, the partial power usage effectiveness pPUE is defined to capture a precise measurement for the current SUT calculations.

From equation (5.7) and its respective parts, it is possible to calculate the pPUE of the server. The pPUE is defined as the ratio of power consumed within the IT and cooling, and aims to determine the effectiveness of server power usage accurately. There are two methods to measure the pPUE by exclusion and inclusion of internal cooling and the scale of the control volume as shown in equations, respectively.

$$pPUE_I = \frac{P_{IT} + P_{cooling}}{P_{IT}} \quad (5.8)$$

$$pPUE_{II} = \frac{P_{Server} + P_{external}}{P_{Server}} \quad (5.9)$$

The reason for providing two measurements for pPUE is due to limitations of knowledge pertaining to server operation. Feasibly, if the specification for fan and pump characteristics are unavailable or immeasurable it is not possible to fully construct a unified model. Thus, pPUE_I assumes full knowledge of sub-system operation and provides a higher degree of accuracy. The pPUE_{II} instead measures the temperature difference across the server to evaluate P_{external}, treating the internal cooling architecture as a black-box.

5.5 Model Validation

In order to validate the proposed unifying model for server power consumption, numerous experiments are conducted to study model accuracy. Number of SunFire V20z servers are used to validate the results with conventional air-cooled and DCLC technologies. This is important as it drives different cooling behaviours within the system (for example, the operation temperature directly affects the power consumption of the CPU cores). The temperature of supply and return water from the external heat exchanger were recorded at each level of the experiment. The recorded water temperature assists in determining the precise calculation of cooling load and leads to predicted overall power consumption. In addition, it is required to measure the flow rate of the pumps. The flow rate captures by timing the collection period of 300ml of water and dividing the volume by time.

SPECpower_ssj2008 benchmark is used for workload imposing to control the server utilization between 0 and 100% at 10% increments. At each utilization level, the server is monitored and both the CPU die temperature and fan speed are recorded. The recorded data is analysed, averaged and fitted into the proposed model to evaluate server power consumption.

Figure 5.6 and Figure 5.7 contrasts the modelled server power profile against real measured outputs for air-cooled and DCLC servers, respectively. The model can successfully capture server power consumption with an error ranging between –5.60 to 3.92% for air, and –0.97 to 2.94% for DCLC. The weakness of the model arises for idle server power usage, where the error rate deviates up to 7% for liquid cooling. The reason for this deviation is due to parameter selection for equation (5.3) to represent $P_{dynamic}$, i.e. a zero value for utilization with the applied constant results in a large discrepancy. While the model accurately models the power consumption at various utilization level.

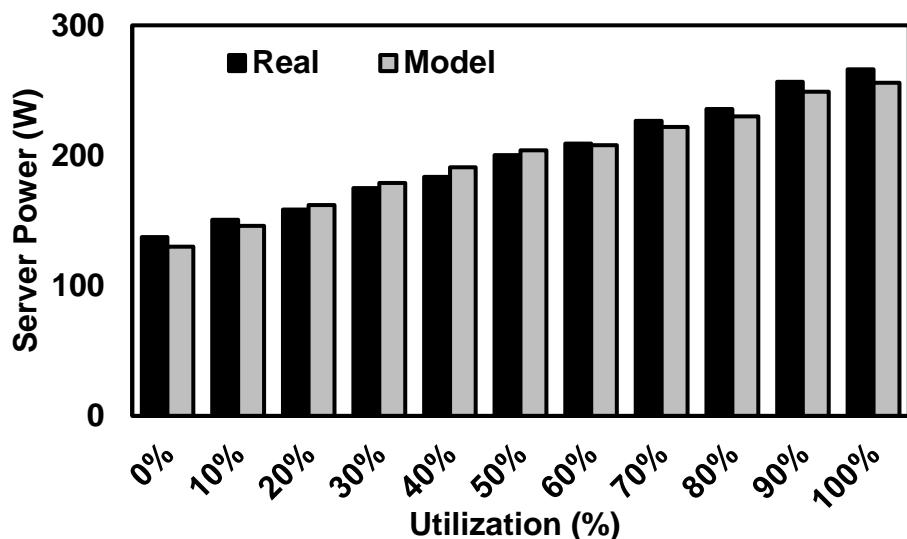


Figure 5.6 The validation of power consumption measurements for air-cooled server.

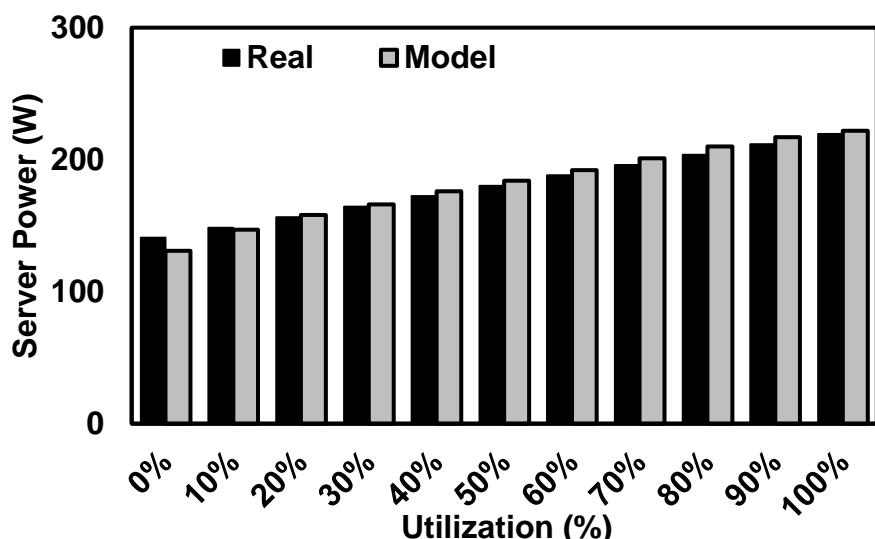


Figure 5.7 The validation of power consumption measurements for DCLC server.

5.6 Results and Discussion

The breakdown of the total power consumption is demonstrated in Figure 5.8. It is observable that increasing the resource utilization results in power usage within each sub-system increasing at linear rate. Therefore, the $P_{dynamic}$ experiences the largest growth from 1% to 33%, driven by increased resource utilization of the server. Furthermore, it is indicated that the total cooling constitutes an additional 5.9% to 10% of power for the server to operate, increasing with higher levels of resource utilization as shown in Figure 5.9. In contrast, air cooling results in 3.44% to 13.44% additional power consumption, with similar breakdown of sub-system power usage observed within liquid cooling. Air cooling consumes on average 7% additional power in comparison to liquid, with the largest deviation of 13% at 100% utilization.

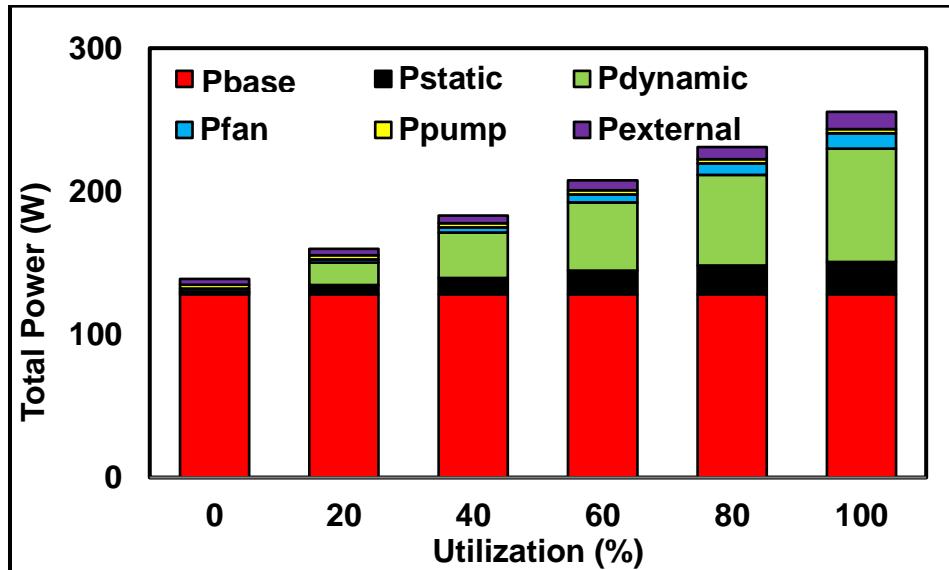


Figure 5.8 Total power breakdown for SUT

The pPUE metrics detailed in equations (5.8) and (5.9) applied to the DCLC analysis results in values of 1.096 and 1.062 for $pPUE_I$ and $pPUE_{II}$, respectively. While this difference is small, around 3.2%, it is worth highlighting that $pPUE_I$ provides increased accuracy for measuring power effectiveness, and will be magnified within the context of thousands of servers. Achieving the difference between the two metrics requires more

efficient server of a reduced cooling load by 25% or increased the IT load by 18%.

Moreover, the external heat rejection is very close to the IT equipment, whereas in larger scale systems the liquid/air must travel significant distance to reject heat. Furthermore, in many scenarios it is not practical or economically feasible to collect data from every sub-system for energy-aware decision making. Therefore, the metrics combined with the unified model allow for providers to determine and select which metric is most suitable for their requirements.

When applying pPUE_I to servers with air and liquid cooling, it is observed a value on average of 1.16 and 1.096, respectively. This is due to the capability of liquid to remove heat is greater than air and sustain a lower temperature due to its thermal capacity.

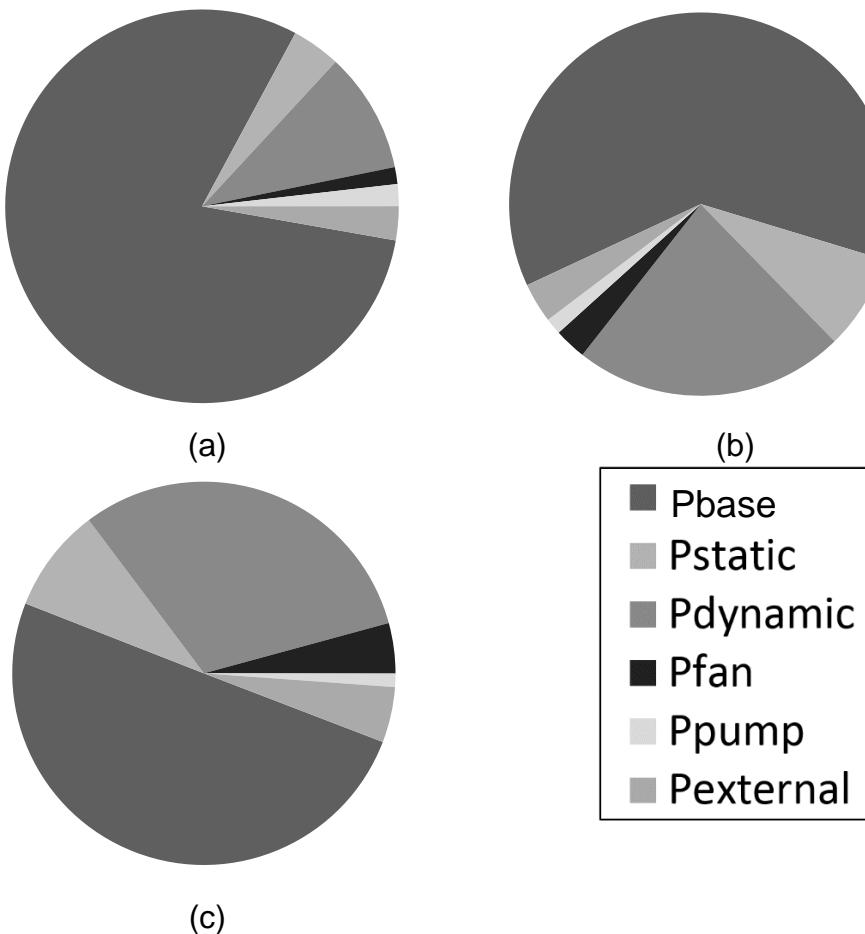


Figure 5.9 Power usage of liquid cooled server sub-systems at utilization (a) 20%, (b) 60%, (c) 100%.

5.7 Summary

The presented holistic power model can produce detailed power profiles of all server sub-systems. More importantly, the model can quantify the cooling power expended under different utilization levels and cooling technologies. The model demonstrates that up to 10% and 13% of the total server power usage is driven by cooling for air and liquid, which is a non-detectable by server power meters. Therefore, it is envisioning the enhancement of numerous energy-efficient scheduling algorithms capable of capturing this new power profile.

There are intuitive relationships between sub-systems parameters from application performance, microprocessor temperature, fan speed and server power. This relationship is not strictly linear and is dependent on utilization levels and cooling type.

Finally, according to the definition of the sub-systems, the calculation of the power pPUE depends on the detailed measurements of power consumption. In addition, the pPUE_{II} presents an interactive relationship between the IT and cooling loads. At the server level, for example, the reducing the cooling load by 7% and increasing the IT load by 5% leads to result in 1% lower pPUE_{II}.

Chapter 6 ENERGY-EFFICIENT DATACENTRE

6.1 Introduction

This chapter presents an energy-efficient datacentre via conducting experiments to measure the cooling load and power consumption. First, the building procedure of the rack will be explained in details. Second, a brief description about the air handling unit AHU will be presented. Later, there will be a discussion about the use of the produced unified power model and the applicability of the partial power usage effectiveness.

6.2 Building the Rack

During 2013, the School of Mechanical Engineering, the University of Leeds started a collaboration with CoolIT. The aim of this collaboration was to investigate the use of DLC technology on a rack-level to present a model of energy-efficient cloud computing. Therefore, a 45U rack, as shown in Figure 6.1, was installed in the laboratory 1.11 in the first floor, the Energy Building.

The next stage of building the rack was provided the IT servers. The SunFire V20z servers have been used. Firstly, 30 servers were retrofitted with the DCLC system as shown in Figure 6.2. Then, all the 30 servers were fitted inside the rack, powered and ready for installing the OS. The OS is the same version of Ubuntu, which presented in section 3.2.4. The very top server called the *head node* which play the control server to all servers in the lab. The head node was given an Internet Protocol (IP) of

192.168.1.101 and defined as eng01. The access to eng01 is by its default username and password of mainuser and AdminAdmin, respectively. Meanwhile, the access to the IPMI was via 192.168.1.1 with the default word admin for both username and password.



Figure 6.1 The 45U rack during early installation stage, Lab 1.11, 2013.



Figure 6.2 Part of the 30 servers with the retrofitted DCLC, Lab1.11, 2013.

As the OSs installed, each server connected to a network switch with activation of the Dynamic Host Configuration Protocol (DHCP). There is a script contains the media access control (MAC) address, which is a unique identifier, for all servers. This script is located in eng01 and responsible for generating the name and IP address based on activation of the DHCP and connect to the MAC address. This results in defined the second server from the top as eng02 with IP of 192.168.1.102 and so on. Simultaneously, all servers connected through eth0 in series to ensure an access to the IPMI. In Figure 6.3, the yellow cables connect the servers to the switch for internet and/or local hosting networking. Whereas the blue cables are in daisy chain for the IPMI access.

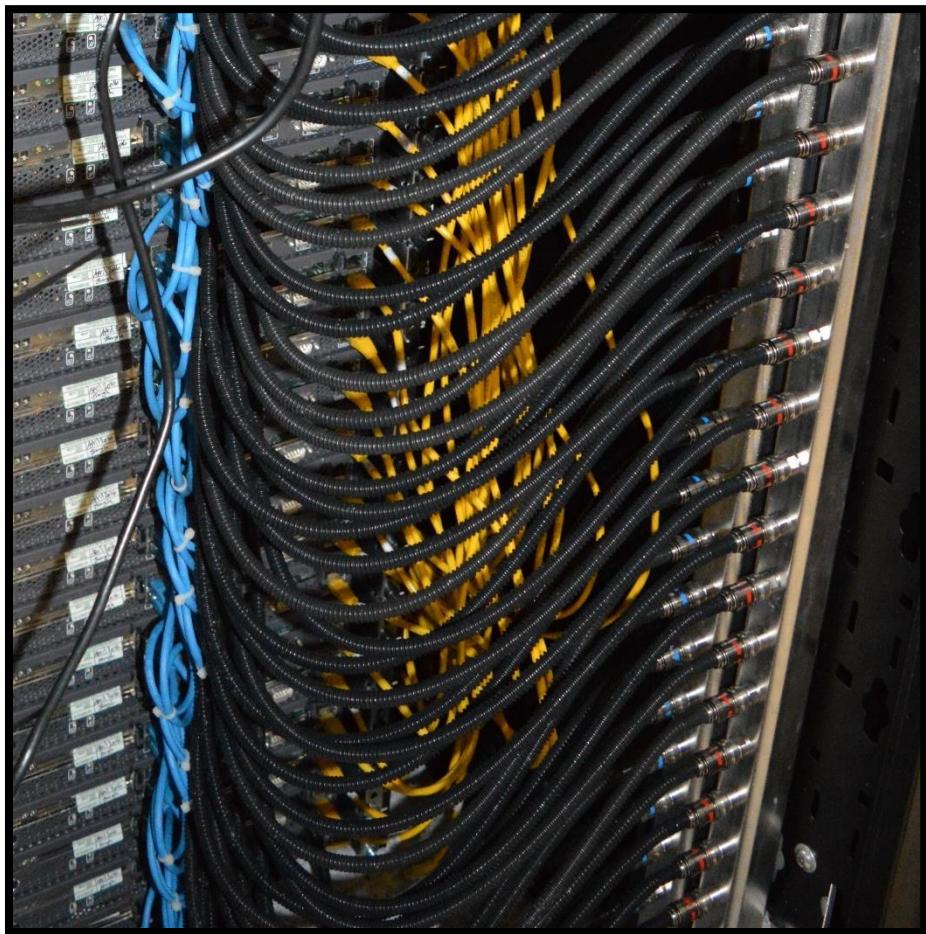


Figure 6.3 Back of the rack showing the networking cables and the coolant tubing, Lab 1.11, 2014.

Finally, as each server has slotted in its location all coolant tubing were connected to the manifold. So, the coolant is able to reject heat outside the

server. As shown in Figure 6.3, the manifold and tubes have been coloured in blue, for cold liquid flow to the servers, and red, for hot liquid flow out of the servers. The heat is rejected to the rack cooling systems as explained in the sext section.

6.3 The Rack Cooling System

The rack was designed to be cooled in either an air heat exchanger AHx or a coolant heat exchanger CHx. The AHx, which is provided by CoolIT and presented in Figure 6.4, represents a liquid-to-air heat exchanger. As the hot liquid collected via the manifold, it is directed to the AHx to reject its heat to the ambient air inside the lab. The coolant was circulated by means of two central in-line pump. The use of two pumps in series made the redundancy of the AHx being $n+1$. Therefore, the cooling at the server level can be either passive, i.e. rely on the central pumps, or active while the local pumps are on. The heat rejection causes the room temperature rises gradually which is not preferable for conducting the experiments. Therefore, an alternative solution was the CHx.



Figure 6.4 The air heat exchanger AHU, Lab 1.11, 2015.

As mentioned earlier, the liquid collects heat from the series of servers through the DCLC loops. This time the coolant loops from the manifold is directed to a liquid-to-liquid heat exchanger. The coolant heat exchanger CHx used in present study is supplied by CoolIT also, and the part model is CHx40 as shown in Figure 6.5.

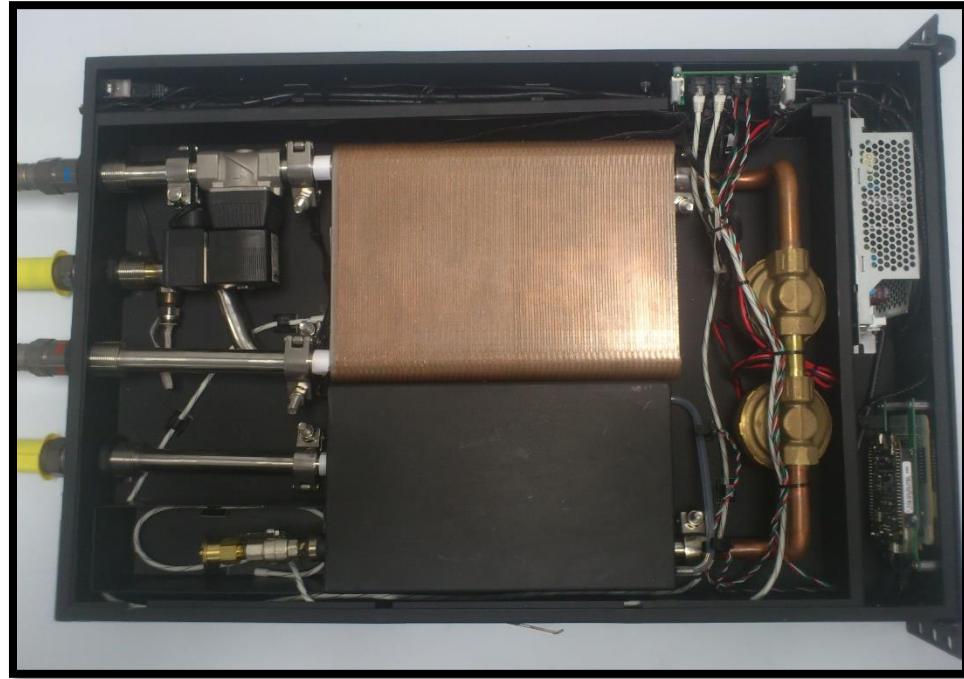


Figure 6.5 The coolant heat exchanger CHx, Lab 1.11, 2013.

The dimensions of the CHx40 are 481 mm wide, 785 mm deep and 89 mm high, which is 2U. The overall weight of the unit is 25.9 kg. CHx40 consists mainly of the plate-to-plate heat exchanger, two pumps in series for n+1 redundancy, primary and secondary piping systems, reservoir and electronic control system as shown in Figure 6.6 (CoolIT, 2015).

The term secondary piping refers to the loop which allows the hot coolant flow from the manifold to circulate. As the hot liquid pass through the CHx it cools down. The heat is rejected to the primary loop which is the facility water or cooling water. The cooling water is provided by the air handling unit AHU as will explain later. Both primary and secondary piping system are connected to the CHx via Stäubli quick connector of size 10 and 20, respectively.

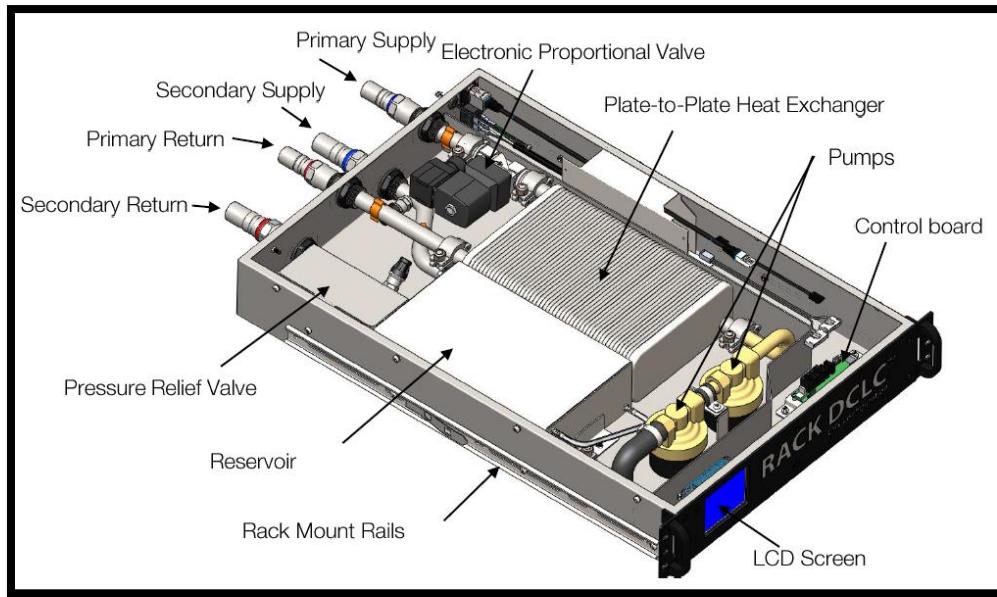


Figure 6.6 Illustration of CHx40 (CoolIT, 2015).

The CHx requires an AC 100 – 240 V with 50/60 Hz, which is available in the lab. The typical power consumption is 160 W (CoolIT, 2015). However, based on the experimental measurement at the current load, the rating power was 98 to 107 W.

This unit is designed for a 40-kW cooling capacity at facility flow rate and inlet temperature of 35 l/min and 30 °C. Figure 6.7 illustrates the effect of inlet water temperature on the maximum cooling load of the CHx.

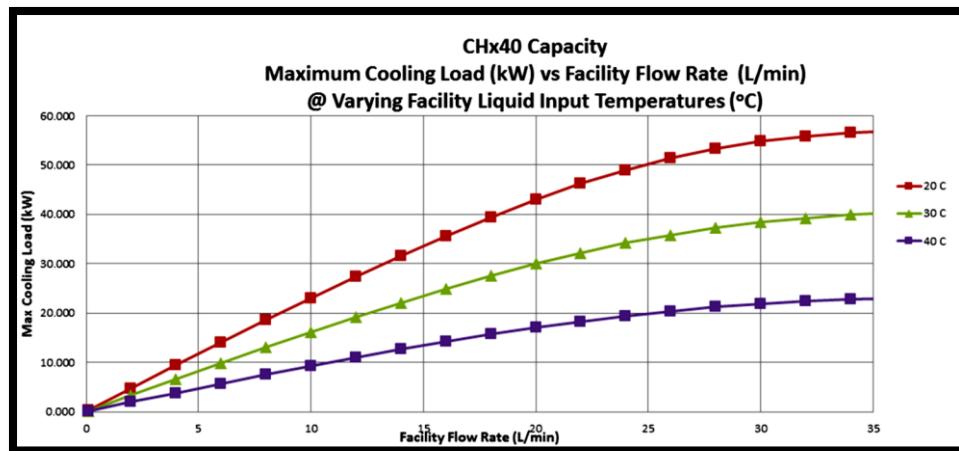


Figure 6.7 Effect of water inlet temperature on the cooling load (CoolIT, 2015).

The CHx is connected to the eng01 for data logging via standard simple network management protocol (SNMP). Whereas the management

information base (MIB) is used for the monitoring of hardware-related issues (CoolIT, 2015). As all required connection were conducted, the cloud computing rack is shown in Figure 6.8.



Figure 6.8 The cloud computing rack, Lab 1.11, 2016.

6.4 The Air Handling Unit (AHU)

The air handling unit (AHU) was designed according to the collaboration between KADHIM (2017) and Airedale Company. Airedale Company has provided the unit to investigate the evaporative cooling in datacentre. The AHU consists of liquid-to-air heat exchanger, water pump, fan, spray system, valves and piping system and control system. Therefore, there are three main working fluids within the control volume.

First, the hot coolant, which is provided by the primary loop of the CHx, enters the liquid-to-air heat exchanger. As the fan blowing ambient air the heat rejected to the flow passes across the heat exchanger. This represents the second working fluid. The fan speed is controllable to work either in manual mode or automatic mode. The manual mode enables a wide range of controlling the fan speed which is representing in the control system qualitatively by percentages. For example, the minimal fan working rate is at 25%, and this value increased until 100%. Whereas the automatic mode relies on the temperature set point, weather conditions and the heating load from IT side.

Third, the AHU has provided with ability to spray water to the air stream to increase the heat rejection efficiency through evaporative cooling. The evaporative cooling reduces the surface temperature due to removing the latent heat from the surface (Bejan and Kraus, 2003). Therefore, the vapour plays the working fluid for this case. For the automatic mode, the spray mode is activated when the fan speed exceeding 80%. So, the operating of the spray is very depending on the weather, temperature and relative humidity, and the heat load from the IT side.

6.5 Analysis of Energy-Efficient Rack-Level

The primary loop of the CHx is connected to the AHU as illustrated in the schematic diagram in Figure 6.9. The hot liquid of the primary loop rejects heat, via the CHx plate-to-plate heat exchanger, to the “return” pipe of the

AHU. The temperature set point of the controller plays the role of selecting the proper path for the liquid. For example, if the temperature value is set to 30 °C and the liquid temperature is below this threshold then the control valve will circulate the liquid to the lab again without entering the main heat exchanger of AHU. However, if the temperature is higher than the threshold, then the controller passes the flow towards the heat exchanger for heat rejection with fan only or fan and spray if required.

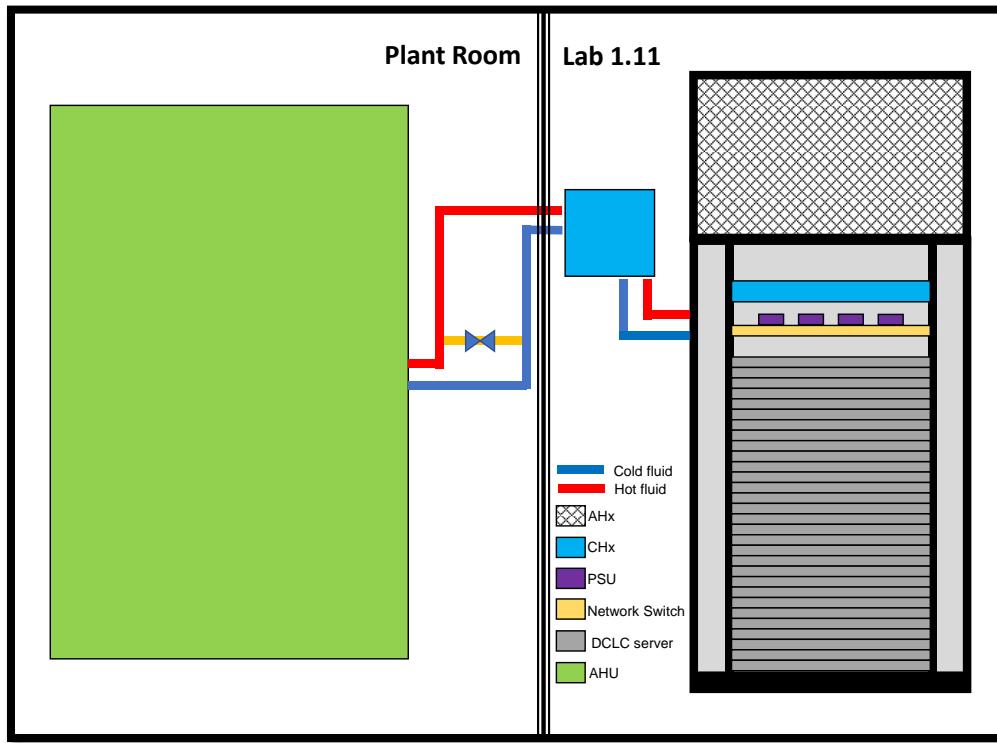


Figure 6.9 Schematic diagram of the connection between the rack and the AHU.

Many configurations allow monitoring and calculating different prospective in the liquid cooling of rack-level. The AHU operates in automatic mode, as mentioned earlier, and at various configurations in the manual mode. In addition, the differences in weather condition plays crucial role in evaluation the energy consumption of the evaporative cooling unit. Moreover, the temperature change annually affects the annual energy bill, as the evaporative cooling operates in summer at full load condition. For example, the average temperature in Leeds during July is higher than in December by about 15 °C as shown in Figure 6.10 (MetOffice, 2016).

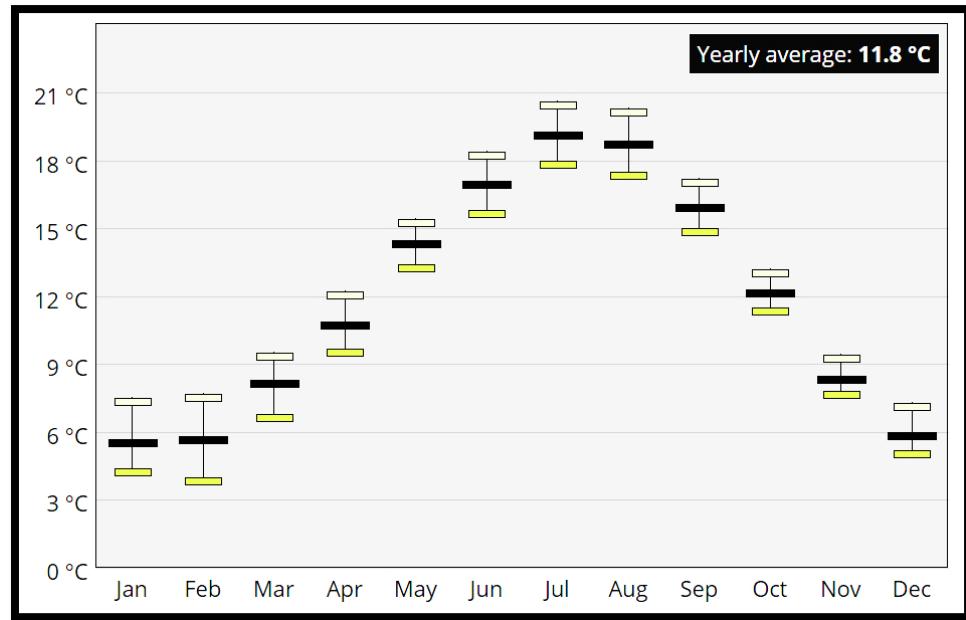


Figure 6.10 Temperature records for the city of Leeds, UK (MetOffice, 2016).

Experiments are conducted with different datacentre operational load scenarios and different AHU configurations. The response of the AHU to any difference in IT load perturbations is highlighted.

The analysis of the temperature balances at different level is conducted through running StressLinux script on the rack. The script produced a variety of IT load, as shown in Figure 6.11. This figure presents the measurements from the AHU side, therefore $T_{\text{sup(sec)}}$ refer to the outflow from the rack to the CHx secondary loop.

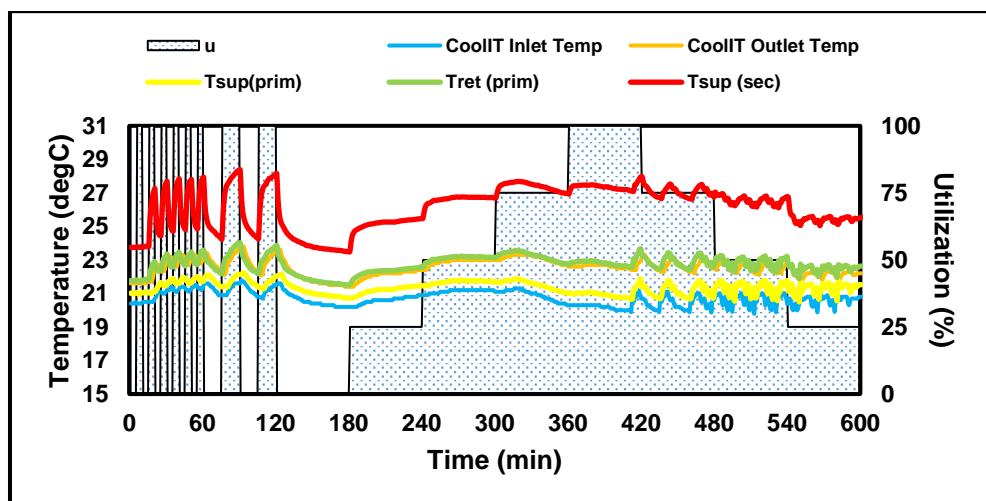


Figure 6.11 The monitoring of temperatures at different points for a virtual workload on the rack.

It is obvious that sharp increase of utilisation spikes the CPUs temperatures, which in turns increases the liquid leaving the rack towards the CHx. The figure show the mapping of behaviour between the proposed utilization and the nearest point of temperature measurements to the heat source. The sharp increase in the $T_{\text{sup(sec)}}$ activate the fan to speed up to dissipate the amount of heat. It is clear from the trends of the curves that there are shift in the time of reaching the peak of the temperature. This is explained as the time lag between the occurrence of the IT load and the reaching of the hot fluid volume to the AHU. The time lag was between 3 to 6 minutes depending on the utilization level and the environmental condition.

The stable load leads the temperature trends to follow the same behaviour but with a time lag due to the far distance between the rack and the AHU. Figure 6. 12 shows the monitoring of temperature at the main supply and return points with 100% stable utilization applied to the rack.

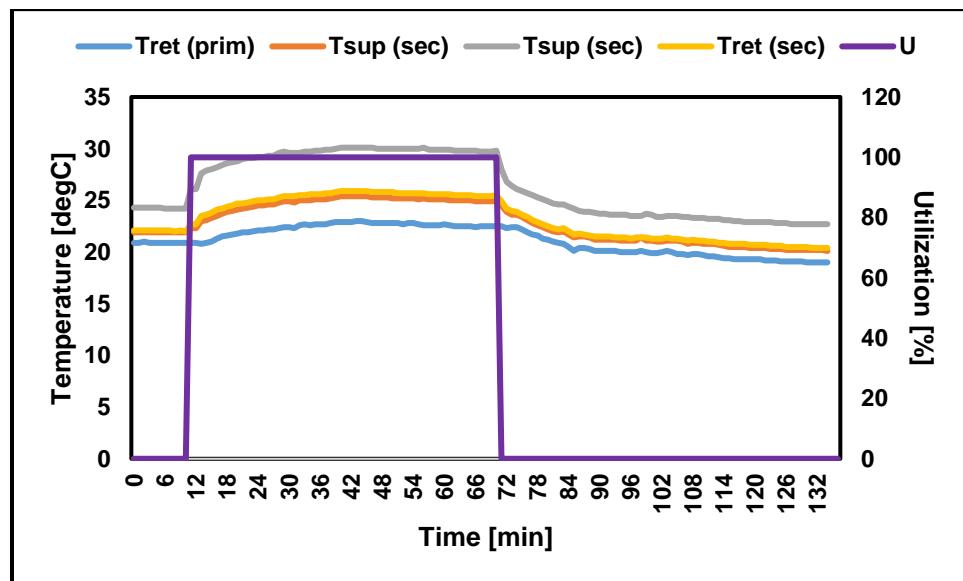


Figure 6. 12 The monitoring of temperatures at different points for a stable workload on the rack.

The key results of interest are the heat transfer calculations from the CPUs to the heat rejection system. Figure 6.13 shows a sample of the heat rejection in the secondary and primary loops of the CHx, the heat arriving to the heat exchanger of the AHU and the heat carried by the air to the environment. The IT produces a total heat Q_s which is supplied to the secondary loop of the CHx. The CHx heat exchanger carries the heat to the

primary loop Q_p , which is transferred to the AHU heat exchanger by water Q_w and finally carried away by the passing air Q_a to the external environment.

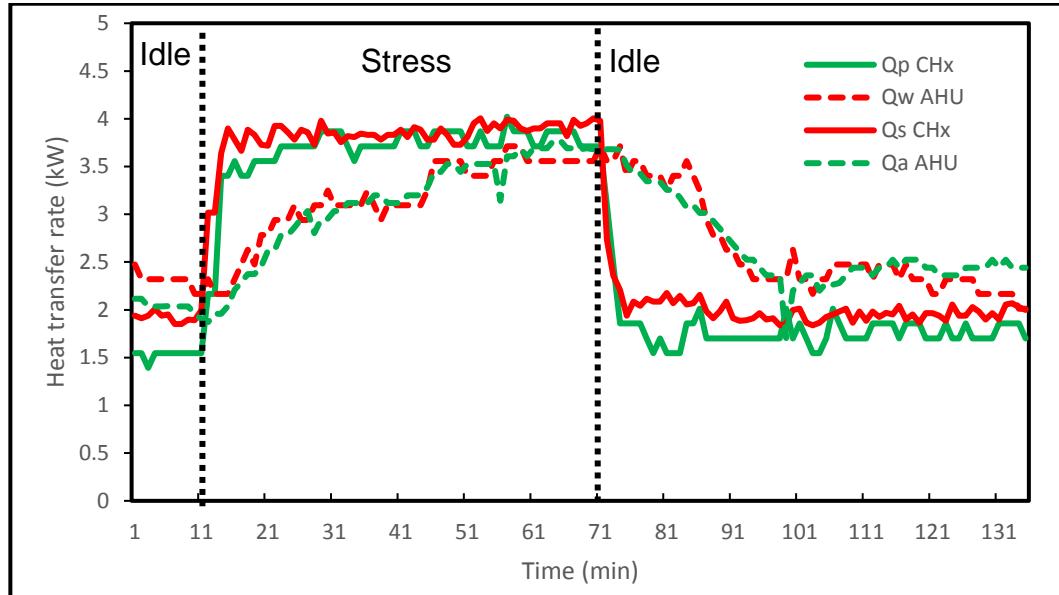


Figure 6.13 Heat transfer balance between the CHx and AHU.

The secondary loop of the CHx represents the closest coolant to the heat source, i.e. the CPUs. Therefore, as the IT load starts, the Q_s responds to the temperature rise rapidly. The second fast response occurs in the primary loop of the CHx. Since the AHU unit is installed in the plant room, which is in the third floor of the building. Therefore, there is a time lag, as mentioned earlier, in the response of the AHU towards the heat generated from IT side. The difference between the response of the AHU and CHx is obvious to recognise in Figure 6.13. In addition, the AHU operation relies on a combination between dry and wet cooling. For this reason, the time to dissipate the heat is longer than the IT load period. However, the overall heat rejected from the IT to the liquid is equivalent to the heat dissipated by the AHU when considering the operation time instead of IT load period.

Based on the heat balance, the effectiveness of the heat exchanger can be obtained by averaging the recorded value. The procedure of calculating the effectiveness is called the number of transfer unit (NTU) which define the ratio between the heat rejected (Q) to the maximum heat transfer (Q_{max}) in either side of the heat exchanger (Bejan and Kraus, 2003):

$$e = \frac{Q}{Q_{max}} \quad \text{where } 0 \leq e \leq 1 \quad (6.1)$$

Where, Q_{max} calculated from the following equation:

$$Q_{max} = C_{min}(T_{hottest} - T_{coldest}) \quad (6.2)$$

C_{min} is the minimum heat capacity when compare both fluids in the heat exchanger.

$$C_{min} = \min((\rho c_p)_{cold}, (\rho c_p)_{hot}) \quad (6.3)$$

Figure 6.14 and Figure 6.15 show the effectiveness monitoring and calculation for the CHx and AHU, respectively.

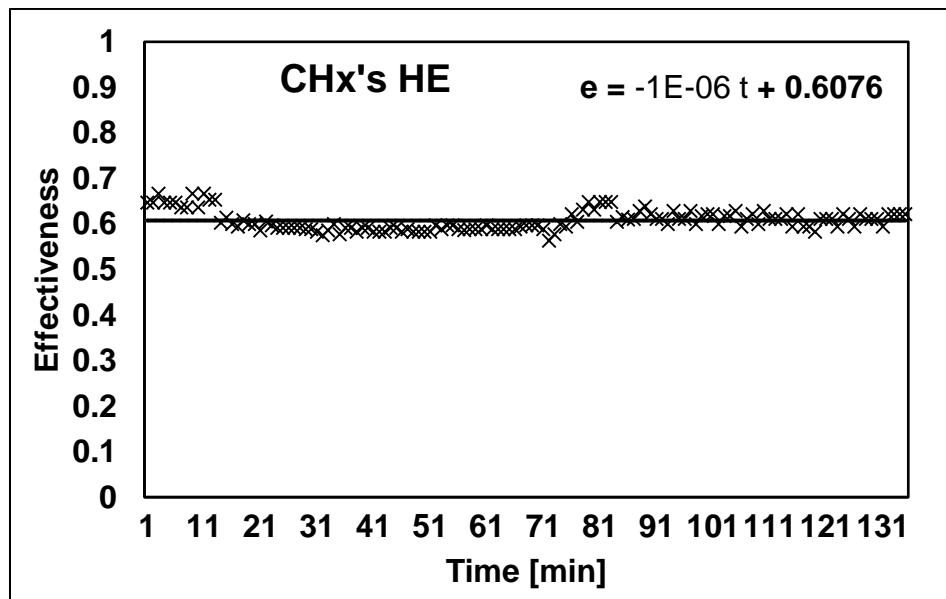


Figure 6.14 Effectiveness calculation of the CHx

Using the unified power model and the proposed PUE provide an evaluation of the whole system and compare the dry versus wet cooling. The results explore the variation of PUE for the system based on calculating the power consumption of the IT and the total power consumed for cooling. This cooling power consumption is divided into the three different heat transfer parts.

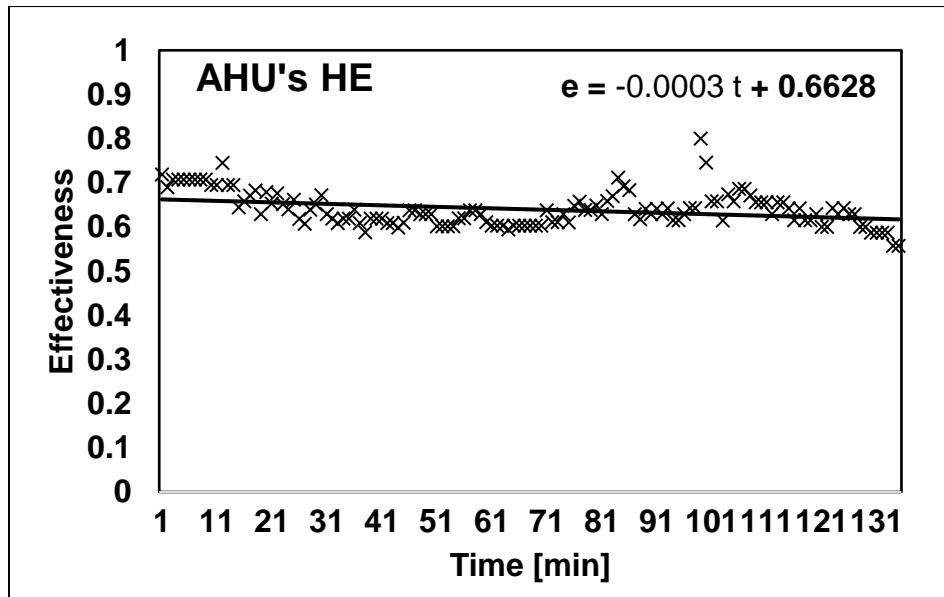


Figure 6.15 Effectiveness calculation of the AHU

Firstly, the server cooling level includes two main categories which are the CPU and the RAM. The current configuration of the servers only allows the CPUs to be cooled by the direct contact liquid. However, the RAMs are cooled by air which is pressurized by the server fans. The power consumption of the fans in each server is found to be between 3.44 to 6.66 W depending on the temperature of the RAM. This is calculated using the fan speed-power curves for each operational scenario to enable the power consumed by the fan for all the thirty servers to be measured.

Secondly, CHx power consumption includes the power consumed by two internal pumps on the secondary loop of the flow. These pumps have variable speeds that respond to the cooling requirements. As a result, the power consumption of the CHx is found to be in the range of 98 to 107 W.

Finally, the power consumed in the AHU is divided into three parts:

- Process water pump: The power consumed by the pump is almost constant over the experiments period as the flow rate is kept constant of 0.37 l/s.
- AHU fan: This is the major part of power consumption in the cooling system. The fan speed varies (from 250 to 1000rpm) depending on the conditions that the AHU is set to and the datacentre operation mode. The power consumption of the fan is from 256 to 1158 W.

- Spray pump: The spray pump power consumption varies from 0 to 197 W depends on the requirement and the operational modes of the AHU and datacentre in addition to the weather conditions, Figure 6.10.

The overall DCLC system is found to achieve a PUE of 1.25 and 1.17 for the dry and spray cooling respectively. This is due to the different power consumption of the AHU as shown in Figure 6.16. It is clear that changing from air cooling to evaporative cooling enhance the PUE by 6.4%.

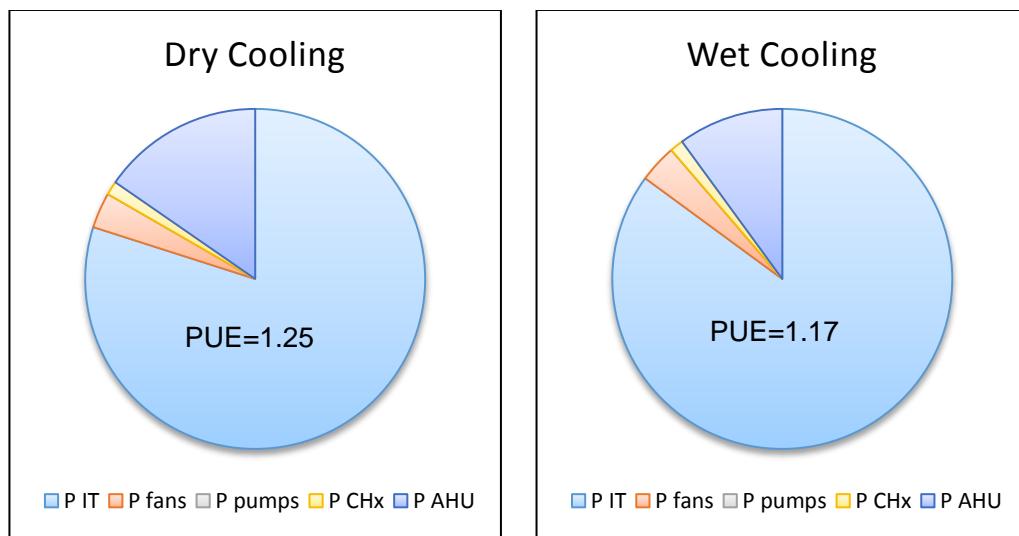


Figure 6.16 Power consumption breakdown and PUE

6.6 Harvesting the Rejected Heat

Connected the DCLC servers to the CHx means *harvesting* all amount of heat generated by the CPUs. This heat transferred through the heat exchanger outside the datacentre. It is very possible, with usage of liquid cooling, to benefit from more than 5.2 kW rejected from the current rack. This amount of heat can be utilized for different purposes to reduce part of power consumption of the infrastructure. Depending on the possibility of re-use of this heat, the PUE reduction is relies on the type of cooling technology and the IT performance level.

6.7 Summary

The experiments have shown that the power consumed by the AHU is higher than the other cooling units in the system even if it is designed to be a compressor-free unit. However, the high consumption is due to the energy requirement of the fan within the AHU. This can be decreased by spraying the AHU's heat exchanger. Although the operation of the spraying system adds more power consuming elements, however, the reduction in temperature reduces the overall PUE. Meanwhile, the spraying system requires certain amount of processed water to keep to maintain the desired supply temperature. Therefore, the overall evaluation of the AHU cooling performance should quantify the water usage of such a system. In addition, the current results are based on the weather of Leeds and during the hottest month only. Therefore, it is important to present solutions which being sufficient for wide range of weather and different locations.

Chapter 7 CONCLUSIONS AND RECOMMENDATIONS

7.1 Introduction

The overall power consumption of datacentres is increasing tremendously due to the high demand of digital services. Moreover, the cooling load stands alone for up to 50% of the power consumption (Joshi and Kumar, 2012). This high share of the cooling load is mainly because the production of a very high densities of newer versions of servers, as shown in Figure 3.1. However, there is an increase awareness in the operations of the sub-systems, i.e. IT workload, cooling load and power consumption. This awareness of the interactions between the sub-systems provides a better understand of maintaining the datacentre infrastructure being energy-efficient.

In the current work, multiple objectives were achieved and presented through the previous chapters. The focus of the current work was on the understanding, benchmarking and analysing the server-level sub-systems. In the first phase, a total air-cooled SunFire V20z server was benchmarked and tested to obtain a baseline for further comparisons with different technologies. The benchmarking was performed via SPECpower_ssj2008 workload.

Then similar procedure of experiments was conducted using a retrofitted direct contact liquid cooling DCLC technology to the same previous server. This retrofitted server has been benchmarked and tested following the

same previous procedure. Furthermore, a CFD simulation has been used to simulate both servers via 6SigmaET software.

Simultaneously, a novel approach was presented to model the fully immersed liquid cooled server. The server assumed to be encapsulated and modelled as a porous layer. The objectives were to investigate the thermal behaviour of a 1U server and optimize the geometrical design for better cooling criteria.

Next phase included data analysis of the tested servers and construct a holistic unified power model. The overall power consumption was divided into IT part and cooling part. The breakdown of the power results in quantifying the interactive effects between each sub-system. In addition, the holistic unified model has presented two approaches to calculate accurate partial power usage effectiveness.

Finally, a holistic evaluation of rack-level has been conducted for an IT rack and evaporative air handling unit.

In this chapter, the final concluding remarks are highlighted. The concluding remarks are categorised into four main groups. In addition, some recommendations for a future work are included.

7.2 Concluding Remarks

Due to various objectives conducted in this work, this concluding remarks are categorised depend on the objective. Firstly, the benefits of using the DCLC technology are listed below:

1. The power consumption of inside-server cooling reduces by 29% for the whole server scale. Whereas, the reduction in power consumption due to cooling of chipset, or CPU, scale is around 88%. Both figures are calculated when compared to the conventional air-cooled server.
2. The overall performance to power ratio increases by 10% over the air-cooled technology. This increase was accompanied by a

noticeable reduction in the die temperature of CPU0 and CPU1 by 6.5 °C and 10 °C, respectively.

3. Regarding the environment change inside the IT room, the noise has reduced by 88.1% comparing to the air-cooled CPUs. Moreover, the outlet temperature has reduced by 1.5 °C and saving the heat generated by CPUs for re-use purposes.

The main conclusions regarding the fully-immersed server are listed below:

1. The in-between CPUs distance has a greater effect on the thermal performance of a 1U server by 13.3% than the other server's dimensions.
2. Increasing the in-between CPUs distance increases the average Nusselt number of the upper CPU by more than 60% comparing to the lower CPU which does not affected. This is due to the passing of the cold fluid over the lower CPU first, then at a higher temperature the ability to gain more heat reduces.
3. The average Nusselt number of the lower CPU increases with increasing all dimensions of the fully immersed server, i.e. Lx, Ly1, Ly2 and S. In contrast, the average Nusselt number of the upper CPU increases for the narrowest server. The narrowest the server, the highest the velocity of flow occurs.

The most important conclusions from the holistic unified power models are:

1. The power consumption of an IT server depends on the utilization level, the die temperature and the supply temperature.
2. The hidden cooling load of air-cooled and DCLC servers were up to 10 and 13%, respectively.
3. Breaking down the power consumption into sub-systems provides better understand of the interactions between these sub-systems. In addition, a more accurate pPUE is presented based on available breakdown values of power consumption.
4. Lowering the cooling load by 7% and increasing the IT load by 5% results in reducing the pPUE by 1%.

The most important conclusions at the rack-level energy efficient are listed below:

1. The air handling unit AHU shows a successful chiller-less cooling system in the city of Leeds. Some experiments were conducted in the hottest weeks during July and August 2016. In addition, the heat generated from the fully utilized servers were dissipated in the capacity of the AHU unit.
2. It is important to consider the late response of the AHU since it is installed far away from the IT servers. This is avoided, or reduced, by selecting a proper threshold value for temperature. This threshold enables the fan to speed up and the spray to provide water droplet, within optimum levels, before IT reaches the peak load.
3. Although the wet cooling technique requires additional power, however, this balances the increasing of fan speed to overcome higher IT load. Thus, the PUE of the wet cooling is lower than the dry cooling by about 6.4%. This is due to the higher ability of liquid to dissipate more heat.
4. The current configuration provides some high opportunities to apply the heat harvesting for re-use within the datacentre.

7.3 Recommendation for Future Work

The recommendations for the future works are listed below:

1. Investigate the effects of direct contact liquid cooling technology on more microelectronics from various vendors.
2. Conduct experiments on real-time servers to evaluate their performance.
3. Extend the CFD simulation of the porous layer to a 3D model.
4. Extend the CFD simulation of the porous layer model to include non-Darcy effects.
5. Use the optimised design and apply it to a similar available server and validate the results.

6. Improve the accuracy of the holistic unified model using various statistical methods.
7. Apply the holistic unified power model on various servers to widen its range and improve its accuracy.
8. Compare the AHU operation and feasibility at different locations.
9. Evaluate the annual results of the AHU to introduce more realistic results for the energy-efficient datacentres.

References

- ABBASI, Z., VARSAMOPOULOS, G. & GUPTA, S. K. Thermal aware server provisioning and workload distribution for internet data centers. Proceedings of the 19th ACM International Symposium on High Performance Distributed Computing, 2010. ACM, 130-141.
- AL-ANII, Y. 2005. *Porous Medium Effect on Natural Convection Heat Transfer from Vertical Row of Cylinders*. M.Sc., University of Anbar.
- ALGER, D. 2009. *Grow a greener data center*, Pearson Education.
- ALMANEEA, A., THOMPSON, H., SUMMERS, J. & KAPUR, N. 2014. Cooling system analysis for a data center using liquid immersed servers. *International Journal of Thermal Technologies*, 4, 200-207.
- ALMOLI, A., THOMPSON, A., KAPUR, N., SUMMERS, J., THOMPSON, H. & HANNAH, G. 2012. Computational fluid dynamic investigation of liquid rack cooling in data centres. *Applied energy*, 89, 150-155.
- AMD 2006. AMD Opteron Processor Power and Thermal Data Sheet. 3.11 ed.
- ANANDAN, S. S. & RAMALINGAM, V. 2008. Thermal Management of Electronics: A Review of Literature. *Thermal Science*, 12, 5-26.
- ANDERSON, J. D. & WENDT, J. 1995. *Computational fluid dynamics*, Springer.
- ASHRAE 2005. *Datacom Equipment Power Trends and Cooling Applications*.
- ASHRAE 2012. *Thermal Guidelines for Data Processing Environments*.
- BAR-COHEN, A., ARIK, M. & OHADI, M. 2006. Direct liquid cooling of high flux micro and nano electronic components. *Proceedings of the IEEE*, 94, 1549-1570.
- BARTELS, A. 2011. *Data Center Evolution: 1960 to 2000* [Online]. Available: <https://blog.rackspace.com/datacenter-evolution-1960-to-2000>.
- BASH, C. E., PATEL, C. D. & SHARMA, R. K. 2003. Efficient thermal management of data centers—Immediate and long-term research needs. *HVAC&R Research*, 9, 137-152.
- BEITELMAL, A. & FABRIS, D. 2014. Servers and data centers energy performance metrics. *Energy and Buildings*, 80, 562-569.
- BEJAN, A. & KRAUS, A. D. 2003. *Heat transfer handbook*, John Wiley & Sons.
- BHOPTE, S., IYENGAR, M. K., SAMMAKIA, B., SCHMIDT, R. & AGONAFER, D. Numerical Modeling of Data Center Clusters: Impact of Model Complexity. ASME 2006 International Mechanical Engineering Congress and Exposition, 2006. American Society of Mechanical Engineers, 51-60.
- BLOCK, H., BECKETT, J., LANGE, K.-D., ARNOLD, J. A. & KOUNEV, S. Analysis of the Influences on Server Power Consumption and Energy Efficiency for CPU-Intensive Workloads. Proceedings of the 6th ACM/SPEC International Conference on Performance Engineering, 2015. ACM, 223-234.

- BOSTON. 2009. *Liquid Cooling For Servers* [Online]. Available: <https://www.boston.co.uk/technical/2009/12/liquid-cooling-for-servers.aspx>. [Accessed: 02/10/2016]
- BRADY, G. A., KAPUR, N., SUMMERS, J. L. & THOMPSON, H. M. 2013. A case study and critical assessment in calculating power usage effectiveness for a data centre. *Energy Conversion and Management*, 76, 155-161.
- BREEN, T. J., WALSH, E. J., PUNCH, J., SHIH, A. J. & BASH, C. E. 2010. From Chip to Cooling Tower Data Center Modeling: Part I Influence of Server Inlet Temperature and Temperature Rise across Cabinet. *Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm)*. Las Vegas, NV, USA.: IEEE.
- BREEN, T. J., WALSH, E. J., PUNCH, J., SHIH, A. J., KUMARI, N., BASH, C. E. & LYON, G. 2012. From Chip to Cooling Tower Data Center Modeling, Validation of a Multi-Scale Energy Management. *Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm)*. San Diego, CA, USA.
- BROWN, R. 2008. Report to congress on server and data center energy efficiency: Public law 109-431. *Lawrence Berkeley National Laboratory*.
- BRUNEAU, M. 2013. *Fundamentals of acoustics*, John Wiley & Sons.
- BURDETT, D. 2014. Energy Flows through Data Centres. University of Leeds.
- CHAN, B., IVEY, C. & BARRY, J. 1970. Natural convection in enclosed porous media with rectangular boundaries. *Journal of heat transfer*, 92, 21-27.
- CHEN, Y., GMACH, D., HYSER, C., WANG, Z., BASH, C., HOOVER, C. & SINGHAL, S. 2010. Integrated Management of Application Performance, Power and Cooling in Data Centers. *Network Operations and Management Symposium (NOMS)*. Osaka, Japan: IEEE.
- CHENG, P. 1979. Heat transfer in geothermal systems. *Advances in heat transfer*, 14, 1-105.
- CHESTER, D., HOPTON, P., BENT, J. & DEAKIN, K. 2013. Cooled electronic system. Google Patents.
- CHI, Y. Q. 2013. Typical Air Cooled Data Centre Compared to Iceotope's Liquid Cooled Solution. Iceotope.
- CHI, Y. Q., SUMMERS, J., HOPTON, P., DEAKIN, K., REAL, A., KAPUR, N. & THOMPSON, H. Case study of a data centre using enclosed, immersed, direct liquid-cooled servers. 2014 Semiconductor Thermal Measurement and Management Symposium (SEMI-THERM), 2014. IEEE, 164-173.
- CHO, J. & KIM, Y. 2016. Improving energy efficiency of dedicated cooling system and its contribution towards meeting an energy-optimized data center. *Applied energy*, 165, 967-982.
- CHO, J., LIM, T. & KIM, B. S. 2009. Measurements and predictions of the air distribution systems in high compute density (Internet) data centers. *Energy and Buildings*, 41, 1107-1115.

- CHO, J., YANG, J. & PARK, W. 2014. Evaluation of air distribution system's airflow performance for cooling energy savings in high-density data centers. *Energy and Buildings*, 68, 270-279.
- CHOI, J., KIM, Y., SIVASUBRAMANIAM, A., SREBRIC, J., WANG, Q. & LEE, J. 2008. A CFD-Based Tool for Studying Temperature in Rack-Mounted Servers. *IEEE Transactions on Computers*, 57, 1129-1142.
- CHU, R. C., SIMONS, R. E., ELLSWORTH, M. J., SCHMIDT, R. R. & COZZOLINO, V. 2004. Review of cooling technologies for computer products. *IEEE Transactions on Device and materials Reliability*, 4, 568-585.
- COOLIT 2011. Product Specification, Product: ECO III-120. CoolIT Systems Inc.
- COOLIT 2012. Product Specification, Product: ECO III. CoolIT Systems Inc.
- COOLIT. 2013. *Why Direct Contact?* [Online]. Cool IT Systems. Available: <http://www.coolitsystems.com/index.php/get-up-to-speed/why-direct-contact.html>. [Accessed: 15/06/2015]
- COOLIT 2015. Rack DCLC CHx40 Module - Operation & Maintenance Manual. A00 ed.
- CREMONESI, P., SANSOTTERA, A. & GUALANDI, S. Optimizing cooling and server power consumption. Intelligent Computer Communication and Processing (ICCP), 2011 IEEE International Conference on, 2011. IEEE, 455-462.
- CURTIS, P. M. 2011. *Maintaining mission critical systems in a 24/7 environment*, John Wiley & Sons.
- DAVID, M., IYENGAR, M., PARIDA, P., SIMONS, R., SCHULTZ, M., GAYNES, M., SCHMIDT, R. & CHAINER, T. 2012. Experimental Characterization of an Energy Efficient Chiller-less Data Center Test Facility with Warm Water Cooled Servers. *Semiconductor Thermal Measurement and Management Symposium*. San Jose, CA, USA: IEEE.
- DELTA 2015. DC Brushless Fans & Blower http://www.delta.com.tw/product/cp/dcfans/dcfans_main.asp. [Accessed: 10/02/2016]
- DRONIOU, J., EYMARD, R., GALLOUËT, T. & HERBIN, R. 2010. A unified approach to mimetic finite difference, hybrid finite volume and mixed finite volume methods. *Mathematical Models and Methods in Applied Sciences*, 20, 265-295.
- EC2. 2016. *The Cloud Market* [Online]. Available: http://thecloudmarket.com/stats#/by_platform_definition. [Accessed: 10/11/2016]
- EYMARD, R., GALLOUËT, T. & HERBIN, R. 2000. Finite volume methods. *Handbook of numerical analysis*, 7, 713-1018.
- FAN, X., WEBER, W.-D. & BARROSO, L. A. Power provisioning for a warehouse-sized computer. ACM SIGARCH Computer Architecture News, 2007. ACM, 13-23.
- FARKHANI, F. F. & MOHAMMADI, F. A. Temperature and power measurement of modern dual core processor by infrared thermography. Proceedings of 2010 IEEE International Symposium on Circuits and Systems, 2010. IEEE, 1603-1606.

- FLETCHER, C. A. J. 1988. *Computational Techniques for Fluid Dynamics - Specific Techniques for Different Flow Categories*, Springer-Verlag.
- FLETCHER, C. A. J. 1991. *Computational Techniques for Fluid Dynamics - Fundamental and General Techniques*, Springer-Verlag.
- FULPAGARE, Y. & BHARGAV, A. 2015. Advances in data center thermal management. *Renewable and Sustainable Energy Reviews*, 43, 981-996.
- FUTUREFACILITIES. 2015a. *6SigmaET* [Online]. Available: <http://www.futurefacilities.com/about/>. [Accessed: 15/12/2015]
- FUTUREFACILITIES 2015b. 6SigmaET Tutorial - Creating a Blade Server.
- FUTUREFACILITIES 2015c. 6SigmaET Tutorials - Multi-Fluid Tutorial.
- GARIMELLA, S. V., PERSOONS, T., WEIBEL, J. & YEH, L.-T. 2013. Technological drivers in data centers and telecom systems: Multiscale thermal, electrical, and energy management. *Applied energy*, 107, 66-80.
- GOUGH, C., STEINER, I. & SAUNDERS, W. 2015. *Energy Efficient Servers: Blueprints for Data Center Optimization*, Apress.
- HAM, S.-W., KIM, M.-H., CHOI, B.-N. & JEONG, J.-W. 2015. Simplified server model to simulate data center cooling energy consumption. *Energy and Buildings*, 86, 328-339.
- HAMANN, H. F., LACEY, J., WEGER, A. & WAKIL, J. Spatially-resolved imaging of microprocessor power (SIMP): hotspots in microprocessors. *Thermal and Thermomechanical Proceedings 10th Intersociety Conference on Phenomena in Electronics Systems, 2006. ITHERM 2006.*, 2006a. IEEE, 5 pp.-125.
- HAMANN, H. F., WEGER, A., LACEY, J., COHEN, E. & ATHERTON, C. Power Distribution Measurements of the Dual Core PowerPC/sup TM/970MP Microprocessor. 2006 IEEE International Solid State Circuits Conference-Digest of Technical Papers, 2006b.
- HASSAN, N., KHAN, M. M. K. & RASUL, M. 2013. Temperature monitoring and CFD analysis of data centre. *Procedia Engineering*, 56, 551-559.
- HEWLETT-PACKARD 2012. Applying 2001 ASHRAE Data Center Guidelines to HP ProLiant-based Facilities. *White Paper*.
- HOPTON, P. & SUMMERS, J. Enclosed liquid natural convection as a means of transferring heat from microelectronics to cold plates. *Semiconductor Thermal Measurement and Management Symposium (SEMI-THERM)*, 2013 29th Annual IEEE, 2013. IEEE, 60-64.
- ICEOTOPe 2013. Three Organizations Team to Win 2013 Green Enterprise IT Award for Energy Efficient Servers. In: LTD, I. R. D. (ed.).
- IPARAAILURY Digital Photo Tachometer - Operation Manual.
- ISMAEEL, M. E. 2011. Heat Transfer in a Square Porous Cavity With Partial Heating and Cooling for Opposite Vertical Walls. *Al-Rafadain Engineering Journal*, 19.
- IYENGAR, M., DAVID, M., PARIDA, P., KAMATH, V., KOCHUPARAMBIL, B., GRAYBILL, D., SCHULTZ, M., GAYNES, M., SIMONS, R., SCHMIDT, R. & CHAINER, T. 2012. Server Liquid Cooling with Chiller-less Data Center Design to Enable Significant Energy

- Savings. *Semiconductor Thermal Measurement and Management Symposium*. San Jose, CA, USA: IEEE.
- JIANG, G., DIAO, L. & KUANG, K. 2012. *Advanced thermal management materials*, Springer Science & Business Media.
- JOSHI, Y. & KUMAR, P. 2012. *Energy efficient thermal management of data centers*, Springer Science & Business Media.
- KADHIM, M., AL-ANII, Y., SUMMERS, J., KAPUR, N. & THOMPSON, H. Thermal Performance of a Mixed Mode Air Handling Unit for Direct Liquid Cooled Servers. SEMI-THERM 33, California, USA, 2017.
- KANOUN, K., CROUZET, Y., KALAKECH, A., RUGINA, A.-E. & RUMEAU, P. Benchmarking the dependability of Windows and Linux using PostMark/spl trade/workloads. 16th IEEE International Symposium on Software Reliability Engineering (ISSRE'05), 2005. IEEE, 10 pp.-20.
- KLADIAS, N. & PRASAD, V. 1989. Natural convection in horizontal porous layers: effects of Darcy and Prandtl numbers. *Journal of heat transfer*, 111, 926-935.
- KLADIAS, N. & PRASAD, V. 1990. Flow transitions in buoyancy-induced non-Darcy convection in a porous medium heated from below. *Journal of heat transfer*, 112, 675-684.
- KOOMEY, J. G. 2011. Growth in Data Center Electricity Use 2005 to 2010. Stanford University.
- LI, Z. & KANDLIKAR, S. G. 2015. Current status and future trends in data-center cooling technologies. *Heat Transfer Engineering*, 36, 523-538.
- MASSOLIN, B. 2013. *CoolIT Systems Partners with Staubli and AFCO Systems* [Online]. Available: <http://www.coolitsystems.com/index.php/news-events/press-releases/144-coolit-systems-partners-with-staubli-and-afco-systems.html>. [Accessed: 15/06/2015]
- METOFFICE. 2016. *Leeds Climate* [Online]. Available: <http://www.metoffice.gov.uk/public/weather/climate/gcwfhf1w0>. [Accessed: 10/11/2016]
- MINAS, L. & ELLISON, B. 2009. *Energy efficiency for information technology: How to reduce power consumption in servers and data centers*, Intel Press.
- MINYARD, C. 2004. IPMI—A Gentle Introduction with OpenIPMI.
- MÖBIUS, C., DARGIE, W. & SCHILL, A. 2014. Power consumption estimation models for processors, virtual machines, and servers. *IEEE Transactions on Parallel and Distributed Systems*, 25, 1600-1614.
- MOORE, G. 1965. Cramming More Components Onto Integrated Circuits, *Electronics*,(38) 8.
- MOORE, J., CHASE, J. S. & RANGANATHAN, P. 2006. Weatherman: Automated, Online and Predictive Thermal Mapping and Management for Data Centers. *ICAC'06*. IEEE
- MORENO, I. S., GARRAGHAN, P., TOWNEND, P. & XU, J. 2014. Analysis, modeling and simulation of workload patterns in a large-scale utility cloud. *IEEE Transactions on Cloud Computing*, 2, 208-221.

- MURSHED, S. M. S. 2016. *Introductory Chapter: Electronics Cooling — An Overview*.
- NADA, S. A., ELFEKY, K. E., ATTIA, A. M. A. & ALSHAER, W. G. 2015. Thermal Management of Electrooni Servers under Different Power Conditions *International Journal of Emerging Trends in Electrical and Electronics*, 11.
- NARAYANAN, A., TOROPOV, V., WOOD, A. & CAMPEAN, I. 2007. Simultaneous model building and validation with uniform designs of experiments. *Engineering Optimization*, 39, 497-512.
- NIEMANN, J., BROWN, K. & AVELAR, V. 2011. Impact of hot and cold aisle containment on data center temperature and efficiency. *Schneider Electric Data Center Science Center, White Paper*, 135, 1-14.
- NITHIARASU, P., SEETHARAMU, K. & SUNDARARAJAN, T. 1997. Natural convective heat transfer in a fluid saturated variable porosity medium. *International Journal of Heat and Mass Transfer*, 40, 3955-3967.
- OH, D., KIM, N. S., CHEN, C. C. P., DAVOODI, A. & HU, Y. H. Runtime temperature-based power estimation for optimizing throughput of thermal-constrained multi-core processors. Proceedings of the 2010 Asia and South Pacific Design Automation Conference, 2010. IEEE Press, 593-599.
- ORACLE. 2007. *Sun Fire V20z and Sun Fire V40z Servers - Server Management Guide* [Online]. Available: <https://docs.oracle.com/cd/E19121-01/sf.v20z/817-5249-17/index.html>.
- OUCHI, M., ABE, Y., FUKAGAYA, M., KITAGAWA, T., OHTA, H., SHINMOTO, Y., SATO, M. & IIMURA, K.-I. 2012. New thermal management systems for data centers. *Journal of Thermal Science and Engineering Applications*, 4, 031005.
- PANDIYAN, V. 2013. Development Of Detailed Computational Flow Model Of High End Server And Validation Using Experimental Methods.
- PATANKAR, S. 1980. *Numerical heat transfer and fluid flow*, CRC press.
- PEDRAM, M. 2012. Energy-Efficient Datacenters. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 31, 1465-1484.
- PELLEY, S., MEISNER, D., WENISCH, T. F. & VANGILDER, J. W. Understanding and abstracting total data center power. Workshop on Energy-Efficient Design, 2009.
- POULIKAKOS, D. & BEJAN, A. 1984. Natural convection in a porous layer heated and cooled along one vertical side. *International Journal of Heat and Mass Transfer*, 27, 1879-1891.
- PRASAD, V. 1991. Convective flow interaction and heat transfer between fluid and porous layers. *Convective heat and mass transfer in porous media*. Springer.
- REMSBURG, R. 2000. *Thermal design of electronic equipment*, CRC press.
- REMSBURG, R. 2011. *Advanced thermal design of electronic equipment*, Springer Science & Business Media.
- SAMPATH, S. 2013. Thermal Analysis Of High End Servers Based On Development Of Detail Model And Experiments.

- SHARMA, R., BASH, C. & PATEL, C. 2002. Dimensionless Parameters for Evaluation of Thermal Design and Performance of Large-Scale Data Centers. *Joint Thermophysics and Heat Transfer Conference*. St. Louis, Missouri, USA: ASME/AIAA.
- SHARMA, R., SHAH, A., BASH, C., CHRISTIAN, T. & PATEL, C. Water efficiency management in datacenters: Metrics and methodology. 2009 IEEE International Symposium on Sustainable Systems and Technology, 18-20 May 2009 2009. 1-6.
- SHEHABI, A., SMITH, S., HORNER, N., AZEVEDO, I., BROWN, R., KOOMEY, J., MASANET, E., SARTOR, D., HERRLIN, M. & LINTNER, W. 2016. United States Data Center Energy Usage Report. *Lawrence Berkeley National Laboratory, Berkeley, California. LBNL-1005775 Page*, 4.
- SINGH, K., BHADAURIA, M. & MCKEE, S. A. 2009. Real time power estimation and thread scheduling via performance counters. *ACM SIGARCH Computer Architecture News*, 37, 46-55.
- SPEC 2012a. Design Document: Control and Collect System SPECpower_2008. In: SPEC (ed.).
- SPEC 2012b. SPECpower_ssj2008 - User Guide. In: SPEC (ed.).
- SPEC. 2016. Standard Performance Evaluation Corporation - SPECpower_ssj2008 [Online]. Available: https://www.spec.org/power_ssj2008/. [Accessed: 10/04/2016]
- STAFFORD, J., GRIMES, R. & NEWPORT, D. 2012. Development of Compact Thermal–Fluid Models at the Electronic Equipment Level. *Journal of Thermal Science and Engineering Applications*, 4, 031007.
- STÄUBLI 2016. Fast Moving Technology.
- SUBRAMANIAM, B. & FENG, W. 2012. The Green Index: A Metric for Evaluating System-Wide Energy Efficiency in HPC Systems. *Parallel and Distributed Processing Symposium Workshops & PhD Forum (IPDPSW)*. Shanghai, China: IEEE.
- SUNMICROSYSTEMS 2008. Sun Fire™ V20z and Sun Fire V40z Servers: User Guide. CA, USA.
- SUNNYVALE, C. 2006. Sun Microsystems - Hardware.
- TARUTANI, Y., HASHIMOTO, K., HASEGAWA, G., NAKAMURA, Y., TAMURA, T., MATSUDA, K. & MATSUOKA, M. Temperature distribution prediction in data centers for decreasing power consumption by machine learning. 2015 IEEE 7th International Conference on Cloud Computing Technology and Science (CloudCom), 2015. IEEE, 635-642.
- TATCHELL-EVANS, M., KAPUR, N., SUMMERS, J., THOMPSON, H. & OLDHAM, D. 2016. An experimental and theoretical investigation of the extent of bypass air within data centres employing aisle containment, and its impact on power consumption. *Applied energy*.
- TONG, X. C. 2011. *Advanced materials for thermal management of electronic packaging*, Springer Science & Business Media.
- TONG, X. C. 2016. *Advanced materials and design for electromagnetic interference shielding*, CRC Press.
- TOP500. 2005. *TOP500 Supercomputer Sites* [Online]. Available: <http://top500.org>. [Accessed: 10/04/2016]

- TRIPPLITE 2012. Optimize Airflow Management in Your Data Center.
- VERSTEEG, H. K. & MALALASEKERA, W. 2007. *An introduction to computational fluid dynamics: the finite volume method*, Pearson Education.
- VOLTECH 2011. PM1000+ Power Analyzer - User Manual.
- WALSH, E. J., BREEN, T. J., PUNCH, J., SHIH, A. J. & BASH, C. E. 2010. From Chip to Cooling Tower Data Center Modeling: Part II Influence of Chip Temperature Control Philosophy. *Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm)*. Las Vegas, NV, USA.
- WANG, L., VON LASZEWSKI, G., DAYAL, J., HE, X., YOUNGE, A. J. & FURLANI, T. R. Towards thermal aware workload scheduling in a data center. 2009 10th International Symposium on Pervasive Systems, Algorithms, and Networks, 2009. IEEE, 116-122.
- WANG, Z., BASH, C., HOOVER, C., MCREYNOLDS, A., FELIX, C. & SHIH, R. 2010. Integrated management of cooling resources in air-cooled data centers. *Conference on Automation Science and Engineering*. Toronto, Canada: IEEE
- YEDAMALE, P. 2003. Brushless DC (BLDC) motor fundamentals. *Microchip Technology Inc*, 20, 3-15.
- ZHOU, R., BASH, Z. W. C. E., MCREYNOLDS, A., HOOVER, C., SHIH, R., KUMARI, N. & SHARMA, R. K. 2011. A Holistic and Optimal Approach for Data Center Cooling Management. *American Control Conference (ACC)*. San Francisco, CA, USA: IEEE.
- ZOTKIEWICZ, M., GUZEK, M., KLIAZOVICH, D. & BOUVRY, P. 2016. Minimum Dependencies Energy-Efficient Scheduling in Data Centers.

Appendix I: SPECpower_ssj2008 Reports

Air-cooled SUT Report

SPECpower_ssj2008
 Copyright (C) 2007-2016 Standard Performance Evaluation
 Corporation

SPEC System EM65
 SPECpower_ssj2008 = 198 overall ssj_ops/watt
 Test Sponsor: SPEC Company
 SPEC License #: 0
 Tested By: SPEC Company
 Test Location:
 Test Date: Apr 2, 2016
 Publication: Unpublished
 Hardware Availability: Jan-2008
 Software Availability: Jan-2008
 Test Method: Single Node
 System Source: Single Supplier
 System Designation: Server
 Power Provisioning: Line-powered

Benchmark Results Summary					
Target	Actual	ssj_ops	Average	Performance	
Load	Load		Active Power	to Power	
			(W)		Ratio
100%	98.7%	73,709	230	321	
90%	89.4%	66,715	223	299	
80%	80.3%	59,927	216	277	
70%	69.4%	51,793	207	251	
60%	60.6%	45,256	200	227	
50%	49.6%	37,052	190	195	
40%	39.7%	29,625	180	165	
30%	29.9%	22,314	171	130	
20%	19.8%	14,759	161	91.5	
10%	9.8%	7,330	152	48.1	
Active Idle	0.0%	0	135	0	
sum of ssj_ops / sum of power = 198					
System Under Test					
Set: 'sut'					
Set Identifier: sut					
Set Description: System Under Test					
f Identical Nodes: 1					
Comment: Here is a comment for this set.					
HARDWARE					
Hardware Vendor: SPEC					
Model: System EM65					
Form Factor: 0U					
CPU Name: Tacoma Spi					
PU Characteristics: Tri-Core, 1.3GHz, 37MB L3 Cache					
CPU Frequency (MHz): 1300					
CPU(s) Enabled: 6 cores, 2 chips, 3 cores/chip					
Hardware Threads: 12 (2 / core)					

CPU Description: Saturn ULP
 Memory amount (GB): 1.5

SOFTWARE

Operating System (OS): SPEC Open Doors 2006 F500
 JVM Vendor: IQ2 Corporation
 JVM Version: IQ2 Java VM v1.0
 CCS Version: 1.2.6

Measurement Devices

POWER ANALYZER PWR1
 Hardware Vendor: Energy Minder, Inc.
 Model: EM1000+ USB
 Serial Number: ser001122
 Connectivity: USB2
 Input Connection: Default
 Metrology Institute: NIST
 Accredited by: IQ2 Calibration Laboratory
 Calibration Label: N-32768
 Date of Calibration: 1-Jan-2007
 PTDaemon Host System: same as CCS
 PTDaemon Host OS: same as CCS
 PTDaemon Version: 1.6.3-a678ac2d-20140326
 Setup Description: Unknown

Notes

None

Aggregate Electrical and Environmental Data

Target Load	Active Power (W)
100%	230
90%	223
80%	216
70%	207
60%	200
50%	190
40%	180
30%	171
20%	161
10%	152
Active Idle	135

Line Standard | Elevation (m)

120V / 60 Hz / 1 phase /	132
2 wires	

Aggregate Performance Data

Target Load	Actual Load	ssj_ops Target	ssj_ops	Actual
Calibration 1				74,221
Calibration 2				74,339
Calibration 3				74,977
ssj_ops@calibrated				74,658
100%	98.7%	74,658		73,709
90%	89.4%	67,192		66,715
80%	80.3%	59,726		59,927
70%	69.4%	52,261		51,793
60%	60.6%	44,795		45,256

50%		49.6%		37,329		37,052
40%		39.7%		29,863		29,625
30%		29.9%		22,397		22,314
20%		19.8%		14,932		14,759
10%		9.8%		7,466		7,330
Active				0		0
Idle						

Copyright (C) 2007-2016 Standard Performance Evaluation Corporation
<http://www.spec.org> - info@spec.org
SPECpower_ssj2008 Reporter Version: [SSJ 1.2.10, May 9, 2012]

DCLC SUT Report

SPECpower_ssj2008
 Copyright (C) 2007-2016 Standard Performance Evaluation Corporation

SPEC System EM65
 SPECpower_ssj2008 = 218 overall ssj_ops/watt
 Test Sponsor: SPEC Company
 SPEC License #: 0
 Tested By: SPEC Company
 Test Location:
 Test Date: Apr 4, 2016
 Publication: Unpublished
 Hardware Availability: Jan-2008
 Software Availability: Jan-2008
 Test Method: Single Node
 System Source: Single Supplier
 System Designation: Server
 Power Provisioning: Line-powered

Benchmark Results Summary					
Target Load	Actual Load	ssj_ops	Average Active Power (W)	to Power Ratio	Performance
100%	99.8%	74,085	212	350	
90%	89.3%	66,315	205	324	
80%	80.2%	59,602	198	301	
70%	69.6%	51,679	189	274	
60%	60.1%	44,608	180	248	
50%	51.1%	37,923	173	219	
40%	39.1%	29,020	162	179	
30%	30.5%	22,665	154	147	
20%	20.0%	14,844	146	102	
10%	10.1%	7,471	134	55.7	
Active Idle	0.0%	0	118	0	
sum of ssj_ops / sum of power = 218					
System Under Test					
Set: 'sut'					
Set Identifier: sut					
Set Description: System Under Test					
f Identical Nodes: 1					
Comment: Here is a comment for this set.					
HARDWARE					
Hardware Vendor: SPEC					
Model: System EM65					
Form Factor: OU					
CPU Name: Tacoma Spi					
PU Characteristics: Tri-Core, 1.3GHz, 37MB L3 Cache					
CPU Frequency (MHz): 1300					
CPU(s) Enabled: 6 cores, 2 chips, 3 cores/chip					
Hardware Threads: 12 (2 / core)					
CPU(s) Orderable: 1,2 chips					
Primary Cache: 64 KB I + 64 KB D on chip per core					
Secondary Cache: 1 MB I+D on chip per chip					
Tertiary Cache: 37 MB I+D off chip per chip					

Other Cache: None
 Memory Amount (GB): 2.5
 # and size of DIMM: 5 x 512 MB
 Memory Details: 2GB 2Rx4 PC2-5300F ECC CL5; slots 1, 2, 3, 7, and 8 populated
 Power Supply Quantity and Rating 1 x 750 (W):
 Power Supply Details: PS-0815
 Disk Drive: 4 x 73GB 15K RPM SAS
 Disk Controller: Integrated SAS controller
 # and type of Network Interface 4 x SuperFast NIC Cards (NICs) Installed:
 NICs Enabled in Firmware / OS / 2/2/1
 Connected:
 Network Speed (Mbit): 100
 Keyboard: PS2
 Mouse: USB
 Monitor: Yes
 Optical Drives: Yes
 Other Hardware: ReallyFast Java Accelerator Card

SOFTWARE

Power Management: Enabled (see SUT Notes)
 Operating System (OS): SPEC Open Doors 2006 F500
 OS Version: 1.2.3.4 (64-bit)
 Filesystem: SPECFS
 JVM Vendor: SPEC
 JVM Version: SPEC Java VM 5.0 (build 1.2.3.4-tricore 20071111)
 JVM Command-line Options: -Xms3500m -Xmx3500m -XrunFast -XconsumeLessPower -Xmn3100m
 JVM Affinity: start /affinity [0F,F0,F00,F000]
 JVM Instances: 1
 JVM Initial Heap (MB): 3500
 JVM Maximum Heap (MB): 3500
 JVM Address Bits: 64
 Boot Firmware Version: 1.2.3.4
 Management Firmware Version: 1.2.3.4
 Workload Version: SSJ 1.2.10
 Director Location: Controller
 Other Software: None

Boot Firmware Settings

none

Management Firmware Settings

none

System Under Test Notes

- * Each JVM instance was affinitized to a socket.
- * Using the local security settings console, "lock pages in memory" was enabled for the user running the benchmark.
- * HT Technology disabled in BIOS.

Controller System

HARDWARE

Hardware Vendor: IQ2 Corporation
 Model: Meridian 38
 CPU Description: Saturn ULP
 Memory amount (GB): 1.5

SOFTWARE

Operating System (OS): SPEC Open Doors 2006 F500
 JVM Vendor: IQ2 Corporation
 JVM Version: IQ2 Java VM v1.0
 CCS Version: 1.2.6

Measurement Devices

POWER ANALYZER PWR1
 Hardware Vendor: Energy Minder, Inc.
 Model: EM1000+ USB
 Serial Number: ser001122
 Connectivity: USB2
 Input Connection: Default
 Metrology Institute: NIST
 Accredited by: IQ2 Calibration Laboratory
 Calibration Label: N-32768
 Date of Calibration: 1-Jan-2007
 PTDaemon Host System: same as CCS
 PTDaemon Host OS: same as CCS
 PTDaemon Version: 1.6.3-a678ac2d-20140326
 Setup Description: Unknown

Notes

None

Aggregate Electrical and Environmental Data

Target Load	Active Power (W)
100%	212
90%	205
80%	198
70%	189
60%	180
50%	173
40%	162
30%	154
20%	164
10%	134
Active Idle	118

Line Standard	Elevation (m)
120V / 60 Hz / 1 phase / 2 wires	132

Aggregate Performance Data

Target Load	Actual Load	ssj_ops Target	ssj_ops	Actual
Calibration 1				73,487
Calibration 2				73,935
Calibration 3				74,607
ssj_ops@calibrated				74,217
100%	99.8%	74,271		74,315
90%	89.3%	66,866		66,315
80%	80.2%	59,417		59,602
70%	69.6%	51,990		51,679
60%	60.1%	44,563		44,608

50%		51.1%		37,135		37,923
40%		39.1%		29,708		29,020
30%		30.5%		22,281		22,665
20%		20.0%		14,854		14,844
10%		10.1%		7,427		7,471
Active				0		0
Idle						

Copyright (C) 2007-2016 Standard Performance Evaluation Corporation
<http://www.spec.org> - info@spec.org
SPECpower_ssj2008 Reporter Version: [SSJ 1.2.10, May 9, 2012]

Appendix II: Building the 6SigmaET Models

Chamber Test

Creating the test chamber is done through the below steps:

1. Select **New Test Chamber** icon.
2. Define the required dimensions.
3. The rear side modifies to the following:
 - a. **Side Type** is Prescribed Flow.
 - b. **Flow Type** is Outflow.
 - c. **Specify Flow** Speed to 1 m/s.
4. The front side modifies to the following:
 - a. **Side Type** is Open.
 - b. **Environment** is Default Environment.
5. All other sides have **Side Type** of Symmetry.

The Server Chassis

First, it is required to either start new project or open existing project. Here, creating new server is explained. So, after select new project and save it as required, couple of steps need to be performed as bellow:

1. Expand the geometry node from the chassis and entre the dimensions of the SunFire V20z server as: **Height** of 42.93 mm, **Width** of 430.276 mm and Depth of 711.2 mm.
2. Select all sides and perform the following:
 - a. Set the **Thickness** to 1 mm.
 - b. Set the **Modelling Level** to Thick.

The Vents

The SunFire V20z has perforations at three sides; the front, the rear and the upper sides. The available options for the perforations is vent areas on the selected sides. The steps are as follow:

1. From the Chassis, select both, **Front Side**, **Rear Side** and **Top Side**.
2. At each side, create **New Vent Opening**.
3. The geometry of the vents is **Rectangular** with values correlate to actual rectangle perforated area.
4. Specify all required values of **Width**, **Length** and **Offset 1** and **Offset 2**.
5. Measure the diameter and number of perforations.
6. Find out the ratio of total perforation area to rectangular area from the prototype.
7. Change the **Open Area** in **Perforation Details** to calculated value.

The Printed Circuit Board PCB

The SunFire V20z has two possible cooling systems as mentioned earlier; the air cooled with extruded fins heatsink and the liquid cooled cold plate. Moreover, there are two CPUs and eight small daughter board for memory cards. The modeling steps are as follow:

1. From the **Chassis** node, expand **New** then **PCB**.
2. Use **Placement** tool to set the correct location of the **PCB**.
3. Select the overall rectangular geometry from the prototype server.
4. Use **Cut** and **Placement** tools to modify the final model of the **PCB**.
5. Specify the value of **Top Distributed Power** from **Power** depending on the thermal design power TDP.
6. The PCB is ready to next steps of attaching all chips.

The Chip Socket

The two CPUs of the studied server consist of chip socket, heatsink and thermal interface material.

1. From the **PCB**, expand **New Entity** to select **New Chip Socket**.
 - a. Expand **Geometry** to modify the **Width**, **Depth** and **Placement**.
 - b. Change the chip socket material as follow
 - i. From **Chip Socket**, expand the **Construction** node.
 - ii. Expand the **Library** to open **Choose Material Dialog**.
 - iii. Select the **Non-materials** on the **Materials** node
 - iv. Choose the **Polycarbonate**.
 - c. Add new **Component** from **New Entity**.
 - d. Expand **Geometry** to modify the **Width**, **Depth** and **Placement**.
 - e. Expand **Thermal Limit** to specify **Allowed Maximum Temperature** to the required value.
 - f. Expand **Power** node to enter the **Thermal Design Power** value.
 - g. Select the utilization level of the component up to **100%**.
2. From the **Component**, select **New Heatsink**.
 - a. Expand the **Geometry** and define **Base Width** and **Base Depth**.
 - b. Set **Fin Thickness**, **Fin Height** and **Number of Fins** to the required values.
 - c. For appearance and presenting purpose, the **Colour** attribute could be selected based of the material. For the current heatsink, which is made of Copper, the colour is **Orange** which has numeral value Red-Green-Blue RGB system of (255,170,0).
 - d. The geometry of the heatsink, which specified in (a) above, represent the whole volume of the heatsink. The actual heatsink has cutout, so it is important to modify the heatsink.

- e. Select the **Heatsink** and choose **New Heatsink Cutout**.
 - f. Expand the **Geometry** and define **Width**, **Depth** and **Offset** from **Placement** node.
 - g. If there are more than one cutout, it is possible to use **Duplicate** to generate copies of the defined cutouts. Then use **Placement** to move the duplicated cutout to its position.
3. There is thermal interface material TIM between the component and the heatsink. The TIM is related to the **Component** node. So, select the **Component** and add **NEW TIM**.
 4. Once the chip socket setting completed, use **Pattern** to duplicate the required number of units and allocate their positions by using **Placement**.

The Fans

The studied server has 3 sets of fans as mentioned earlier in Section 3.2.5. Both RAMs and CPUs six fans are attached to the PCB since they are powered when the server is running. However, the rest two fans are designed to be powered separately with the power supply unit PSU.

1. Select **New** from the **PCB** drop list.
2. Choose **Fan** and expand the **Identification** node.
3. Set the **Manufacturer** and **Model**.
4. Modify the characteristics of the fan from **Geometry** and **Placement**.
5. Expand the **Cooling** node and set the **Flow Direction** to **In**.
6. Set the **Flow Rate Option** to **Curve**.
7. Edit the **Curve** based on the given data in (DELTA).
8. Once the fan setting completed, use **Pattern** to duplicate the required number of fans and allocate their positions by using **Placement**.

The Memory Modules

The memory modules consist of socket and daughter board. The following steps are used to perform this part.

1. From the **PCB**, select **New**.
2. Choose **New Daughter Board Socket**.
3. Select proper geometry to accommodate the daughter board later.
So, the **Height**, **Width**, **Length** and **Socket Depth** are modified to the required values.
4. Use **Placement** to move the socket to its position.
5. Now, from the **Daughter Board Socket** select **New Daughter Board**.
6. Assign both **Top Distributed Power** and **Bottom Distributed Power** to the proper from TDP values.
7. Once the daughter board setting completed, use **Pattern** to duplicate the required number of units and allocate their positions by using **Placement**. The RAM bank need to duplicate per CPU, so there are two banks for the studied server.

The Disk Drive

The disk drives are added to the chassis level. The addition contains creating the drive bay first then adding the drives.

1. From the **Chassis** level, click **New Drive Bay**.
2. Select **2.5" Drive Bay** from the **Model** attribute.
3. Modify the geometry and placement to align the bay to the desired location.
4. Now, click **New Drive** from the **Drive Bay**.
5. Modify the dimensions of the **Drive**.
6. Ensure that the **Mounting Rails** in the **Construction** is **Yes**.
7. Define the **Thermal Design Power** for the drive.
8. Select the **Drive Bay** and choose **Duplicate** option.

9. Set the location of the duplicated drive bay.

The Power Supply

Servers have power supply unit PSU or multiple units. The PSU is a complex unit and consists of many components. However, 6SigmaET presents a simple model of PSU. Although the simplicity of the model, it reflects the correct characteristics of heat and fluid flow of the actual PSU. The only restriction is the prediction of hot spot inside PSU block, which is not one of interest of current study.

The procedure to add PSU is as listed below:

1. From the **Chassis**, select **New Power Supply**.
2. Expand the **Geometry** to modify the size; **Height**, **Width** and **Depth**.
3. In addition, define the correct location by setting the **Placement**.
4. From **Power** node, define both **Thermal Design Power** and **Utilization**.
5. To create the fans, select the **Power Supply** node and select **New** then click **Fan**.
6. Repeat steps 3 to 8 from **Fan** section above.
7. To create the vents, select the Power Supply node and select **New** then click **Vent Opening**.
8. Repeat steps 3 to 7 from section **Vent Opening** above.

Obstructions

Each server has different type of components which are not categorized above. So, those components are called obstructions. It could be wire extension, power joint or even mounting kits. The three generalized types of obstructions are:

1. Solid Obstruction: This type of components is a 3D solid body of any shape. These obstructions block the air flow totally.

2. Perforated Obstruction: It is either a flat plate or polyline plates of any shapes. These obstructions contain series of perforations which reduce the fluid flow.
3. Porous Obstruction: This obstruction is either 3D or plate of any shapes. In addition, the flow resistance in all direction can be adjusted independently.

The obstructions are part of the chassis or any other level in the model tree, and can be created using the following procedure:

1. Select the required level from the model tree, for example **Chassis**.
2. From **New**, select one type of the obstruction as explained earlier.
3. Modify the **Geometry** and **Placement**.

The Cold Plates

The cold-plate define under the obstruction blocks since there is no flow across. However, the design either performed within the geometrical facilities or imported from specific design, i.e. computer aided design CAD file.

1. Select **Obstructions** from the **Chassis** tree, and create Cold Plate.
2. Define the required dimensions from **Geometry** and align the block to its position using **Placement**.
3. From the **Cooling Node** define the **Material** properties and choose **Total Heat** from **Heat Option** drop list.
4. Inside the **Obstruction**, define a **Pump** and name it **Pump1**.
5. Under the **Pump** node, define the **Fluid** and specify the **Flow Rate**.
6. Use **Duplicate** option to create the second pump and name it **Pump2**.
7. Select the **Rear Side** and define the inlet pipe from **New** then **Pump Supply**.
8. Similarly, select the **Rear Side** and define the outlet pipe from **New** then **Pump Return**.
9. Adjust the **Pump Supply** and **Pump Return** size and location.

10. Click the **Pump Supply** node and add **New** then **Cooling Duct** to connect the **Pump Supply** point to the **Pump1**.
11. Modify the **Cooling Duct** to the required size and location via **Geometry and Placement**.
12. Similarly, click the **Pump Return** node and add **New** then **Cooling Duct** to connect the **Pump Return** point to the **Pump2**.
13. Create another **Cooling Duct** to connect the **Outlet of Pump1** to the **Inlet of Pump2**.