THERMAL ANALYSIS OF HIGH END SERVERS BASED ON DEVELOPMENT OF DETAIL MODEL AND EXPERIMENTS

by

SHREYAS SAMPATH

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE IN MECHANICAL ENGINEERING

THE UNIVERSITY OF TEXAS AT ARLINGTON

December 2012

To my father C. S. Sampath and mother S. Shashikala

who set the example and who made me who I am.

ACKNOWLEDGEMENTS

I would like to thank my supervising professor Dr. Dereje Agonafer for constantly motivating and encouraging me, and also for his invaluable advice during the course of my master studies and for the famous green mile walk around the campus. I wish to thank and really appreciate the support of my committee members Dr. Kent Lawerence and Dr. Haji Sheik for taking time to serve in my dissertation committee.

I am especially grateful to Dr.Veerendra Mulay and Dr. Saket Karajgikar for thier interest in my research and for the helpful discussions and invaluable comments. I wish also to thank John Ferenendes and Richard Eiland, PhD students for taking time to mentor and guide me throughout my research. I also could not have done this without the help of my colleagues and friends from the EMNSPC research group at UTA. Every single person in this group was always willing to offer insight or to help in any way they could. I would like to thank Sally Thompson for taking care of all official paper work and encouragement.

Finally, I would like to express my deepest gratitude to my father who have encouraged and inspired me and sponsored my education. I am extremely fortunate to be so blessed. I am also extremely grateful to my mother and sister for their sacrifice, encouragement and patience. I wish also to thank my roommates and several of my friends who have inspired, motivated and helped me throughout my career.

November 28, 2012

ABSTRACT

THERMAL ANALYSIS OF HIGH END SERVERS BASED ON
DEVELOPMENT OF DETAIL MODEL
AND EXPERIMENTS

SHREYAS SAMPATH, M.S.

The University of Texas at Arlington, 2012

Supervising Professor: Dereje Agonafer

Thermal management of an efficient data center can be primarily dictated by efficient servers. Continued increase in power density and power dissipation of high end processors, thermal analysis and management has made it strategically important in the challenge of advanced thermal solutions. Present day CFD software provides powerful tools to create very accurate fluid and heat transfer solutions. Validation of a detailed model of a server in CFD corresponds to accurate flow and temperature models and these detailed server models are the foundational blocks for accurate data center thermal management solutions.

In this research, an experimental study of the thermal behavior of a high end compute server is performed and discussed. Surface temperatures of key components are recorded and studied as a function of the processing utilization of the server and inlet air temperatures. Next, a detailed CFD model is developed, analyzed and validated for a subset of the experimental boundary conditions with the data, within an acceptable accuracy across all the considered design points. This validation work

serves as guidance in generating compact models or simplified models for rack and room level thermal management. Fixed design considerations for a system level or server level packaging which includes an understanding of effects of varying inlet temperatures on server operation, if the server equipped fans are capable to adequately cool the components and how the heat generated by the components affect other components can also be clearly visualized in this research. This research covers extensive validation at different operating points and not just for maximum Thermal Design Point (TDP).

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

Table Page

CHAPTER 1

Introduction

The main objective of this research is to develop detailed model of high end servers and analyze them from a thermal stand point and validate them through experiments. An enterprise developed CFD software tool is used as a means for demonstrating this detailed model. This chapter introduces the various definitions of data center, thermal management and the need for thermal management of electronic components, high end servers and CFD.

1.1 Data Center

A data center is a facility where servers and associated components, such as telecommunications and storage systems, redundant or backup power supplies and redundant data communications connections are co-located mainly for environmental requirement and also for physical security and ease of maintenance. A bird's eye view of a typical data center can be seen in the Figure 1.1 [1].

In order to protect the reliability parameter and the IT equipment manufacturer's limits the environmental conditions are essential. The ASHRAE TC 9.9 committee has classified data centers into four classes based on control of environmental parameters (dew point, temperature, and relative humidity) in data center [2].

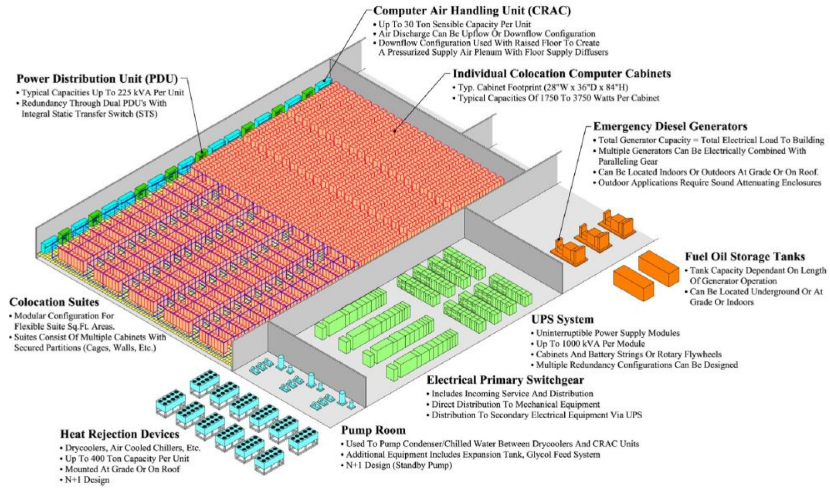

Figure 1.1. Typical Data Center Layout [1].

Data center Classes

- Class A1 data center: Tightly controlled environmental parameters and air conditioning.
- Class A2 data center: Some control of environmental parameters and air conditioning.
- Class A3 data center: No control of environmental parameters and air conditioning
- Class A4 data center: No control of environmental parameters and no air conditioning.

A psychrometric chart in Figure 1.2 would give a better detail of the data center classes.

1.2 Thermal Management

Heat transfer is a process that concerns the generation, use and exchange of heat between physical systems. Semiconductor devices result in significant internal

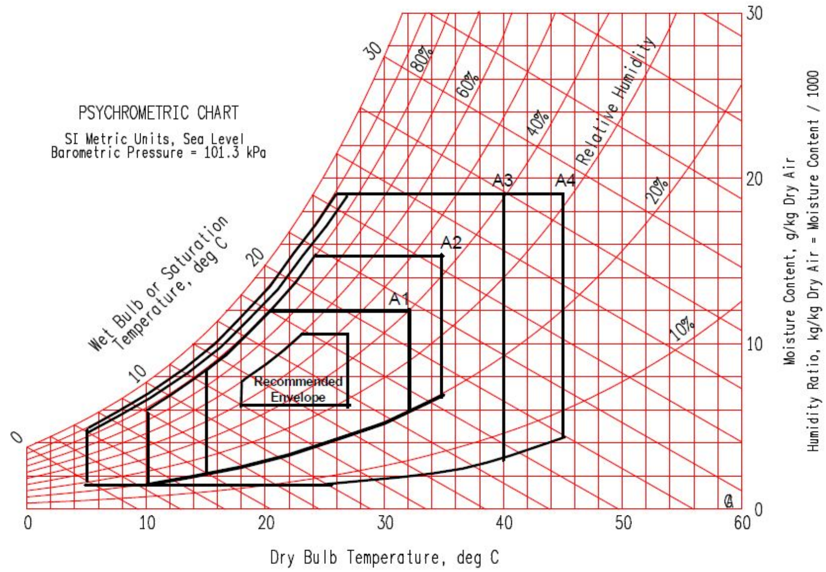

Figure 1.2. ASHRAE Environmental Classes for Data Centers [2].

heat generation by resistance to flow of electric current through transistors. The component in absence of cooling would rise to reach a value at which the device losses its physical integrity. Keeping the device in contact with a low temperature medium, the heat flows away from the component thereby temperature rise is moderated as it asymptotically approaches a steady-state value [3].

1.2.1 Need for thermal management

The need for thermal management of electronics is primarily driven by the need to prevent a catastrophic failure leading to immediate and complete loss of electronics function and package integrity. Catastrophic failure may lead to drastic deterioration in semiconductor behavior, and/or fracture, de-lamination, melting, vaporization and even combustion of the packaging materials. An insight into thermal control strategy can be established by understanding the catastrophic vulnerability

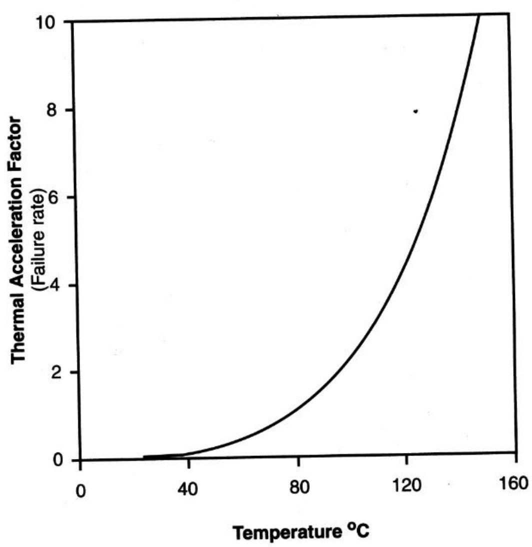

Figure 1.3. Effect of temperature on failure rate [3].

of specified components which makes it possible to select the appropriate fluid, heat transfer and inlet coolant temperature [3].

1.2.2 Failure rate increases with temperature

Reliability, the probability that the system will meet the required specification for a given period of time can be strongly affected by temperature. Electronic components often perform reliably with no moving parts at or near normal temperature. The IC's operate unfortunately at higher temperature and normally fail at prolonged exposure to elevated temperatures. Accelerated failure rate which result from things like mechanical creep in the bonding materials, parasitic chemical reactions and dopant diffusion.

Figure 1.3 reflects a near-exponential dependence of the thermal acceleration factor on component temperature. The failure rate increase exponentially with rise in ambient temperature from 75 to 125°C [3].

1.3 Packaging Level

Packaging can be commonly categorized into four different level

- Level 1 : Chip packages constitute level 1 of packaging.
- Level 2 : Printed Wiring Boards constitute level 2 of packaging.
- Level 3 : Motherboards constitute level 3 of packaging interconnecting the printed wiring boards.
- Level 4 : Racks or Cabinet constitute level 4 of packaging which comprises of the entire system

1.4 High End Servers

High end servers are high power density and high power dissipating computer hardware system dedicated to run multiple services or programs to serve requests of other users of the other computers on the network. Depending on the computer service that it offers it could be a different type of server. These high end servers contain on the order of one hundred million transistors with this number continuing to grow with Moore's Law. Moving towards a billion transistors, the growing power budgets of these chips must be addressed.

An ample increase in chip and module heat flux is witnessed with increase in both power dissipation and packaging density. The heat load trend in Figure 1.4 published by ASHRAE stands out to show the heat load for square feet of server foot print has increased extensively over the past 10 years. The heat load per compute servers alone*(server considered in this research)* have tripled for the period of 2002-2012.

Compute servers being the backbone of many IT industries, thermal management of these servers are of prime importance. Data centers of these companies are

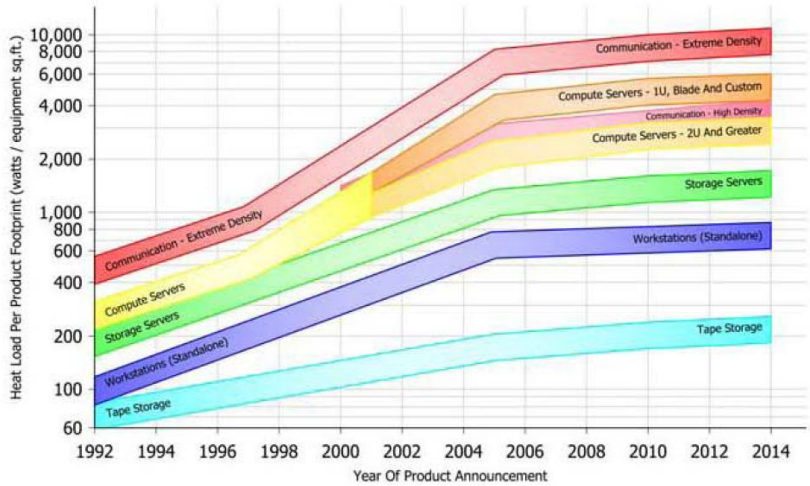

Figure 1.4. Heat Load Trends-ASHRAE Centers [4].

populated mainly with these servers. Most of the guidelines for a thermal management scheme are based on previous observation. In order to perk up on this practice, much effort of late has been placed on making predictions via CFD.

1.5 Computational Fluid Dynamics(CFD)

Computational Fluid Dynamics, CFD, is a technique used to model fluid flow using numerical analysis to help visualize vital parameters like temperature distribution, pressure, flow rate, system impedance and others. An iterative process predicts the flow of fluid based on how minute volumes interact with each other and with the surrounding surfaces to create a steady-state. The use of CFD analysis in data centers has a critical role in the design of data center. CRAC failure analysis, layout changes, contamination are some other parameters that can be analyzed. Growth in processor power has enabled the use of CFD modeling into many different applications and not just relegating it to high-end applications.

1.5.1 CFD Solution

CFD solution involves solving a set of partial differential equations that is used to describe the variation of dependent variable *(such as temperature or pressure or velocity)* with a number of independent variables *(such as time or distance)*. The variation of these dependent variables is continuous. This continuous nature of these non-linear variables can be approximated to linear relationships which are easily solved. This process is called "Discretation" [10].

The Discretation process starts with dividing the analysis space into a number of *non-overlapping* or *control volumes* and each control volume surrounds a grid point at which the dependent variables are evaluated.

1.5.2 Need for detail CFD model

The important objective of a CFD modeling is to predict the change in dependent variables with change in boundary conditions. The accuracy of which depends on the level of detail incorporated in modeling of the system to be analyzed.

However with greater detail, the control volumes and the grids surrounding them increase to a great extent thereby reaching an incomputable grid cell count. A data center model which contains hundreds of racks and thousands of servers, with each server modeled with great detail will generate a mesh with enough control volumes to make even the advanced computing platform short of computational resource. The general approach is to consider different assumptions in the design process to create a compact model and approximate the results, but at the cost of accuracy. Therefore, accurate predictive CFD models must be created either at the chip level, the rack level, or the room level in a data center environment but not at multiple levels simultaneously.

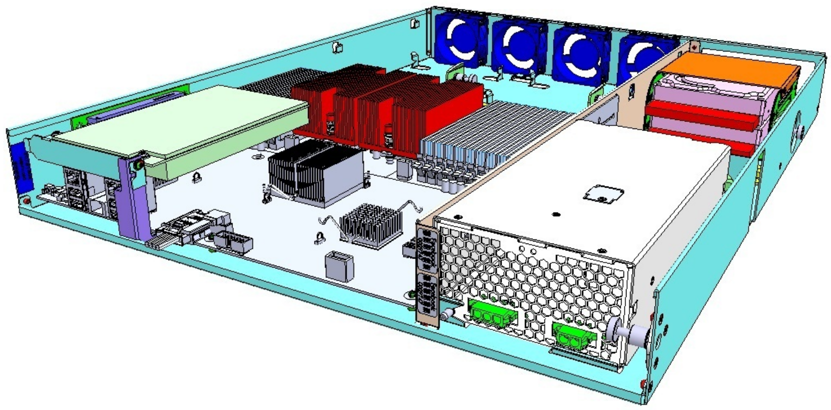

Figure 1.5. Open Compute Intel Server [5].

This research concentrates on modeling a server with great detail of all major heat generating and flow restricting components and experimentally validates its temperature parameter to an acceptable accuracy level. The now accurate model can be simplified to a black box with inflows, outflows and associated heat dissipation. Thereby providing the designer an accurate prediction of parameters associated with a data center facility leading the way to efficient data center management.

1.6 System under consideration

System under consideration is a Open Compute Intel based server*(Intel Motherboard v1.0)* from Facebook OPEN Compute Project [5]. It is a rack mount server of height 1.5 U. The conventional compute servers are usually 1 U but with 1.5 U height these servers provision a taller heat sink for better heat transfer as heat transfer is a function of area. The Open Compute Intel server which is considered can be seen in the Figure 1.5.

The server is equipped with 700W-SH AC/DC power converter, a single voltage 12.5VDC, closed frame; self-cooled power supply used in high efficiency IT applica-

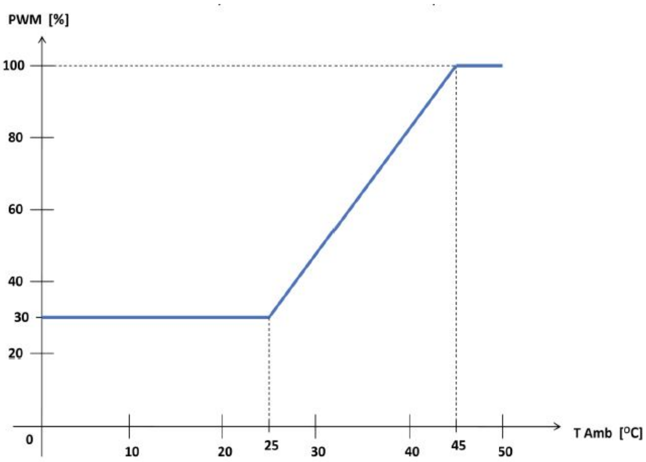

Figure 1.6. Fan Speed Control vs Ambient Temperature [5].

tions. The nominal AC input voltage is 277VAC RMS (200277 VAC). The supply is configurable to a 450W-SH power rating, as both models use the same PCB's, with just pin-to-pin component replacements. Based on the power rating and air flow direction four versions of the power supply can be used [11].

As mentioned, since the power supply is self-cooled it houses an internally powered efficient fan. The fan is a 12V component, size 60 x 60 x 25mm, 30 CFM (minimum), and is a 4-wire, double ball bearing type. A microprocessor controls the fan speed in relation to inlet air temperature. In spite of different power supply cooling conditions the minimum duty cycle of the PWM signal applied to the fan is at least 40% for the 700W-SH model and 30% for the 450W-SH model. Figure 1.6 clearly shows the behaviour of PWM signal due to fan speed control. This base speed is employed to avoid any hot spots inside the power supply [11].

The power supply efficiency on AC is seen to be high and is a function of CPU utilization.

- Efficiency > 90% at 20% load
- Efficiency > 94% at 50% load

Figure 1.7. Power Supply Unit.

- Efficiency > 91% at full load

Figure 1.7 shows the power supply unit equipped in the server.

In order to create an accurate CFD model of an electronic enclosure, all the major heat generating and flow restricting components where modeled. This is done not only in order to simplify the model development process, but also because the other small components contribution to the heat transfer results or flow results are negligible or minimum. The motherboard with all the components and the motherboard with major heat generating components and flow restricting components considered can be seen in the Figure 1.8 and Figure 1.9.

In order to account for small assumptions made in a design process and observe any discrepancy, the CFD results are validated with experimental tests. The CFD flow model validation is already been addressed previously so this thesis will present an effort to produce such an experimentally-validated CFD thermal model of this high end Open Compute server laying the foundation for a rack level solutions.

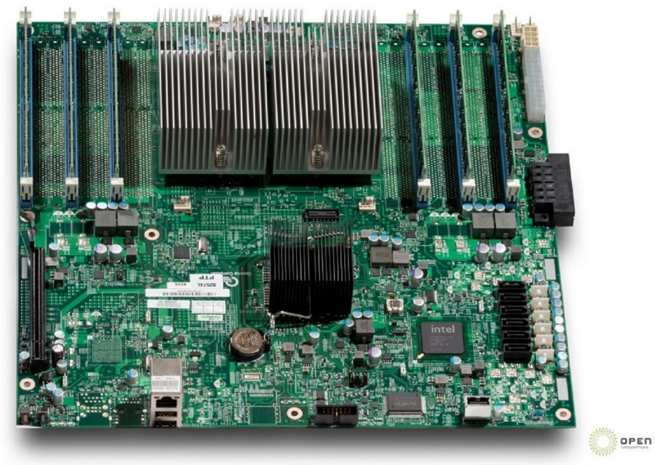

Figure 1.8. Motherboard with all the components [5].

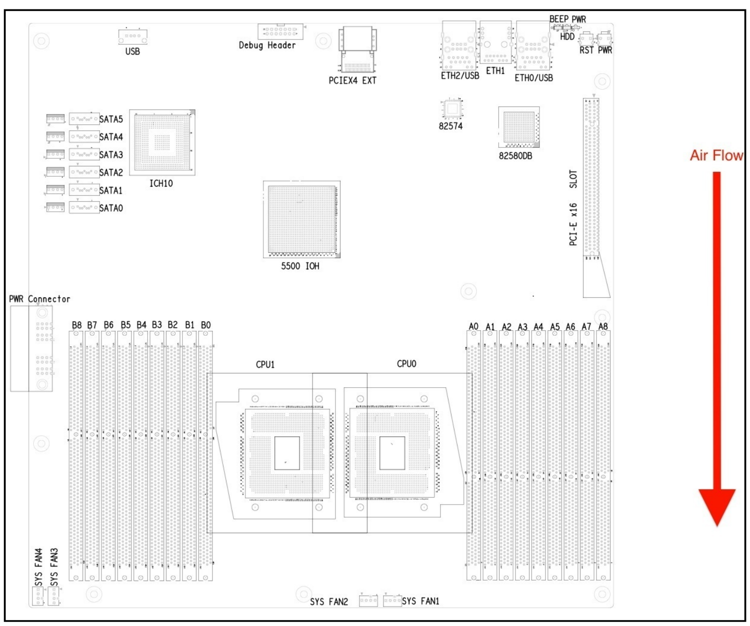

Figure 1.9. Motherboard with major heat generating components [5].

CHAPTER 2

Experimental Study

In previous studies, much of the validation for these CFD models has involved a comparison between the versions of detailed models and compact ones, but rarely experiments. This research work gives a true validation of the model with experimental results. The experiment procedure is divided into two stages in order to avoid a CFD validation of transient temperature problem, as the software cannot accurately simulate the fan control algorithm at this time.

- Step1 or Baseline Case- Fans internally powered, controlled and measured (using on board components). This sets as a base case of average fan speeds for step-2.
- Step2 or Final Case- Fans externally powered, controlled and measured (using external components). Average fan speed at their respective CPU utilization and ambient temperatures from step-1 is considered for step-2.

This chapter introduces the environmental chamber which is used to help test our server at different ambient temperatures. It also helps to maintain temperature and the relative humidity during the test procedure. The advantage of using an environmental chamber also helps us in creating a steady state temperature, so the CFD model can be simulated at steady state temperatures. Figure 2.1 shows the environmental chamber used in this test. Using an actual server in this test helps the model to be validated very close to its true physical conditions.

Figure 2.1. Thermotron Environemental Chamber.

2.1 Test Set-up

2.1.1 Test Set-up Overview

To accomplish the testing, several items must be set-up. This includes computers to assist in data recording and writing and executing test scripts, installing software's to assist in data acquisition, setting up supporting hardware for data acquisition, installing tools on servers for data monitoring and recording.

2.1.2 Test Case Selection

The number of test cases can vary with as many parameters in consideration. This research considers a few test cases where temperature and %CPU utilization are

Design Point	Target % CPU Utilization	Ambient Temperature (°C)		
		30°C	35°C	40°C
DP0	10			
DP1	20			
DP2	30			
DP3	40			
DP4	50			
DP5	60			
DP6	70			
DP7	80			
DP8	90			
DP9	98			

Table 2.1. Different test cases

the main parameters and humidity as a constant parameter. Table 2.1 indicates test cases considered in this research work.

2.1.3 Test Hardware and Software

Test Hardware

- Thermotron 7800 SE 600-Environmental Chamber
- Temperature Measurement

  Omega Thermocouples

  Omega USB Logger

  Agilent 34972A Data Loggers
- Power Measurement

  0-280V AC Staco Variac

  Yokogawa CW121 Power meter
- External Fan Operation

  Agilent E3633A DC Power Supply(20V,10A)

  Agilent 33210A function generator(10MHz)

```
[root@linux ~]# mpstat -P ALL
Linux 2.6.9-5.ELsmp (linux.itso.ral.ibm.com)    04/22/2005

03:19:21 PM  CPU   %user   %nice %system %iowait    %irq   %soft   %idle    intr/s
03:19:21 PM  all    0.03    0.00    0.34    0.06    0.02    0.08   99.47   1124.22
03:19:21 PM    0    0.03    0.00    0.33    0.03    0.04    0.15   99.43    612.12
03:19:21 PM    1    0.03    0.00    0.36    0.10    0.01    0.01   99.51    512.09
```

Figure 2.2. Output of mpstat command on multiprocessor system [6].

Aurduino MEGA R3 microcontroller

Test Software

- Linux Tools

  shell script

  mpstat (%CPU utilization)

  free (%MEM utilization)

  system health monitoring module (CPU die temperatures, fan speeds, etc.)

  lookbusy module

2.2 Data Acquisition

Data which are measured off the server are through utilizing LINUX tools in the server.

2.2.1 Linux Tools

2.2.1.1 mpstat command

It is native Linux tool used to measure or display processor utilization, report the activities of each of the available CPUs on a multiprocessor server. The mpstat utility enables you to display overall CPU statistics per system or per processor. An example output of mpstat command can be seen in Figure 2.2 [6]

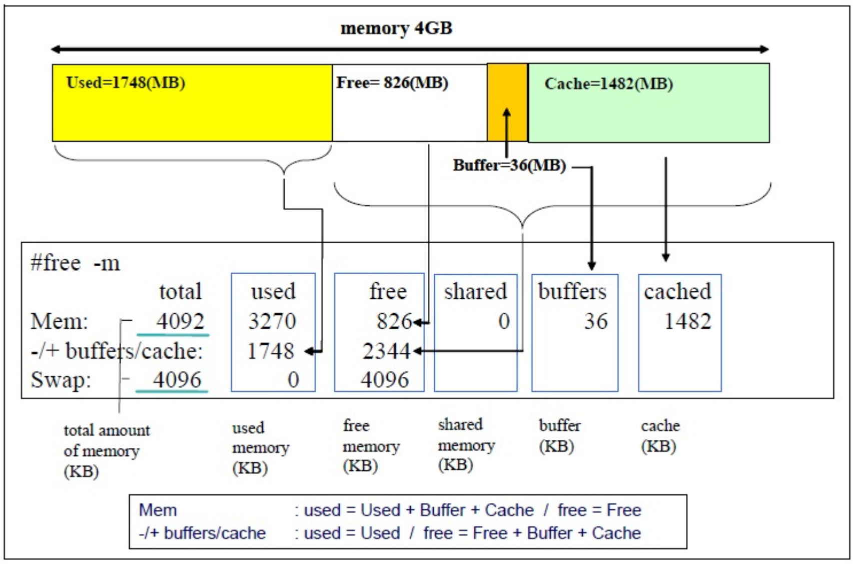

Figure 2.3. Output of free command [6].

2.2.1.2 free command

The command /bin/free displays information about the total amount of free and used memory (including swap) on the system. It also includes information about the buffers and cache used by the kernel. An example output of free command can be seen in Figure 2.3 [6]

2.2.1.3 system health monitoring module

The system health monitoring module is a Linux module used for controlling CPU die temperature under a certain threshold limit. The module algorithm key input is the CPU die temperature. If the CPU die temperature increases by a certain pre-set threshold value then the fan speed control algorithm is triggered and the motherboard sends a PWM signal to increase the duty cycle of the fans to increase their speed. The fans continue to run at this elevated fan speed till the CPU die temperature return to that pre-set threshold value. The motherboard sends a PWM

```
[root@localhost ~]# hir

********************************
*  Health Status Monitor V3.0  *
********************************
*Project: F01/F02              *
*Author : Derrick Huang        *
********************************
This is F02
TIME Record @ Tue Oct  9 11:32:19 2012
CPU0 Therm Margin: 27
CPU1 Therm Margin: 0
CPU0 Case Temp : 62
CPU1 Case Temp : 35
CPU0 Die Temp : 68
CPU1 Die Temp : 96
Right Outlet Temp: 33
Left Outlet Temp: 31
Inlet Outlet Temp: 26
CPU0 VR PROC HOT STATUS: currently inactive, count=0
CPU1 VR PROC HOT STATUS: currently inactive, count=0
CPUVcore1       : 1.208
CPUVcore2       : 1.208
AVCC-P3V3       : 3.360
3VCC-P3V3       : 3.360
VIN1-P1V5_DDRP1 : 1.504
VIN2-P1V5_DDRP2 : 1.504
3VSB-P3V3_STB   : 3.312
VBAT-P3V3_VBAT  : 3.072
SYSTEM FAN1: 1704
SYSTEM FAN2: 1739
SYSTEM FAN3: 1704
SYSTEM FAN4: 1638
===============================
[root@localhost ~]#
```

Figure 2.4. Output of hir command.

signal to decrease the duty cycle once the CPU die temperature fall below the preset threshold. An example of system health monitor module command line *" hir -f 1000000 -l 0"*. Figure 2.4 shows an example output of hir command.

2.2.1.4 lookbusy application

Lookbusy is a simple application for generating synthetic load on a Linux system. It is used to generate fixed, predictable loads on CPUs and keep specific amount of memory active. When generating CPU load in particular, lookbusy will attempt to keep the CPU(s) at the specified utilization level, adjusting its own consumption up or down to compensate for other loads on the system [12].

An example of lookbusy command line *"lookbusy c 60 -m 2000MB"* corresponds to 60% CPU utilization allocating 2000 MB of memory.

Data which are externally measured of different instruments.

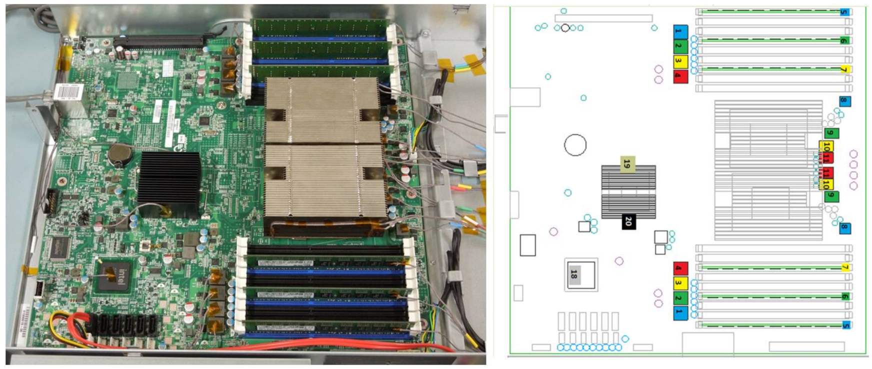

Figure 2.5. Server instrumented with thermocouples [7].

2.2.2 Temperature Measurement

2.2.2.1 Omega Thermocouples

Omega Type-T is a beaded wire thermocouple. Gauge of 0.127 mm (0.005 in.) were used to create a temperature map of the servers. Thermocouples produce a voltage output that can be correlated to the temperature that the thermocouple is measuring. Type-T can measure temperature of range 0 to 350°C with an error limit of ±0.5°C or 0.4%. Detailed temperature maps would require many data points, and an entire server could not be covered in thermocouple wires. Therefore, it was necessary to create a local system for thermal mapping, one that would be detailed enough to accurately map the device surface temperatures of critical components, but would also not be a significant flow obstruction, which would alter both the velocity and the temperature pattern. The server instrumented with thermocouples can be seen in the Figure 2.5.

2.2.2.2 Omega USB Loggers

The ambient temperature is controlled and maintained constant by the environmental chamber as discussed earlier. The inlet air temperature and outlet air

Figure 2.6. Omega USB loggers placed in front of the servers.

temperature just of the servers where measured using six Omega USB loggers. The USB loggers where positioned to the front and back end of the servers. The Figure 2.6 shows three Omega USB loggers placed at the front end of the servers. Three more USB loggers where placed in parallel at the rear end of the server.

2.2.2.3 Agilent Data Logger

Agilent Data Logger is used to log the surface temperature from the thermocouples instrumented on the server on major heat generating components. The thermocouples from the server are all connected to a multiplexer which is guided into the data logger. Figure 2.7 shows the Agilent Data Logger with multiplexer.

2.2.3 Power Measurement

2.2.3.1 Staco Variac

Staco Variac used here is a portable 120V input to 0-280V output variable transformer used to control voltage. It has an additional input voltage tap permitting

Figure 2.7. Agilent Data Logger with Multiplexer.

normal over voltage output, with half normal input voltage. The output current must be reduced when the output voltage exceeds 125% of the input voltage [13].

Figure 2.8 shows the Staco Variac used to step-up the voltage.

2.2.3.2 Yokogawa CW121 Power Meter

Power meter is used to measure the server power consumption. It is a 1-Phase, 2 -wire power meter. The voltage from the Variac is measured and the current before each power supply is measured using a current clamp. The power meter then calculates the power consumed by the server with all the phase angel calculations. It can communicate with a computer to help record power consumption of the server for specific utilization and for specific interval as small as 1 sec for a given time period.

Figure 2.8. Staco Variac.

2.2.4 External Fan Operation

In order to partially replicate the process of fan control algorithm by motherboard to have more control over the process the fans are externally powered and controlled. The server is equipped with four, 60 mm, 12V, 4-wire fans. The Figure shows the fan that is equipped in the server. Functions of 4 wires in the fans used are

- 1 is GND *(usually black)*
- 2 is +12VDC *(usually yellow or red)*
- 3 is Sense *(usually green)*
- 4 is Control *(usually blue or brown)*

A bread board is used to bring all of this hardware together. Figure 2.10 shows the complete setup of hardware connected.

Figure 2.11 gives a better understanding of the procedure.

Figure 2.9. Fan installed in Intel Open Compute server.

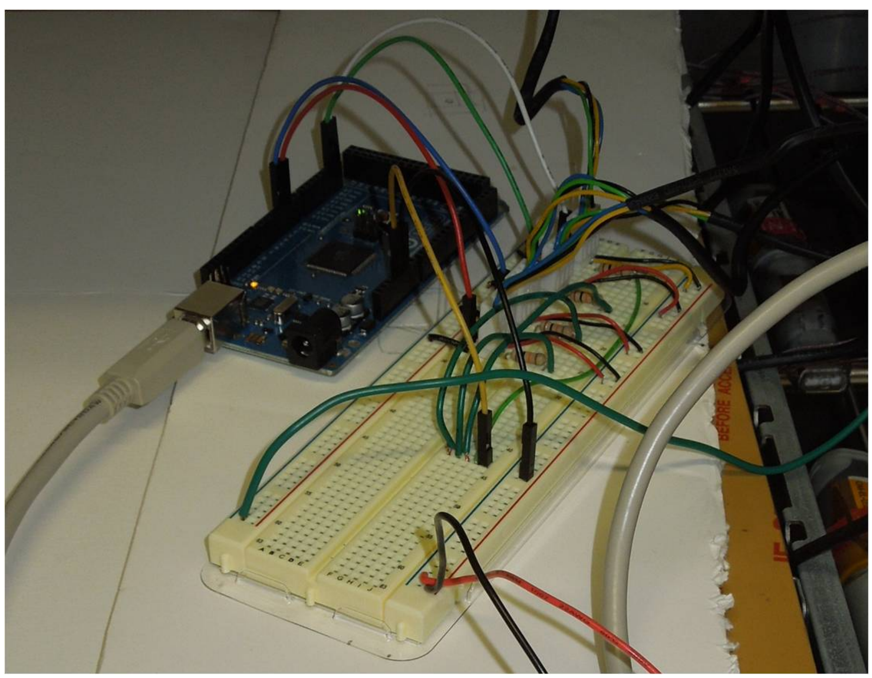

Figure 2.10. Externally Powered Fan Set-up.

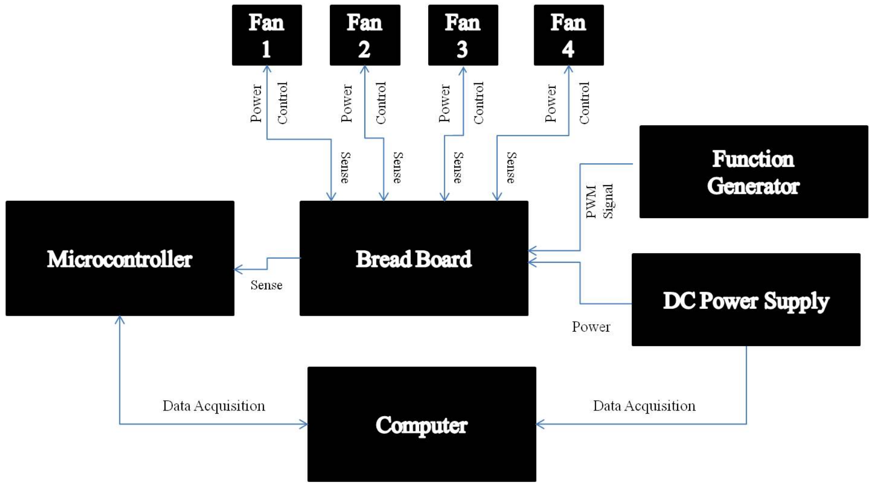

Figure 2.11. Externally Powered Fan Flowchart.

2.2.4.1 Agilent E3633A DC Power Supply

Agilent DC power supply used is 20V, 10A power supply. The power supply is connected to the bread board and hence powers the fans which are also connected to the board. The voltage and current supplied to the breadboard are measured using a computer which in turn gives the fan power $(P = V*I)$.

2.2.4.2 Agilent 33210A function generator

Agilent function generator is connected to the bread board. It replicates the Pulse Width Module (PWM) signal from the motherboard sent to control the amplitude, peak to peak voltage and percentage duty cycle of the fan to operate at the specified speeds.

Figure 2.12 shows the Agilent DC power supply with Agilent function generator

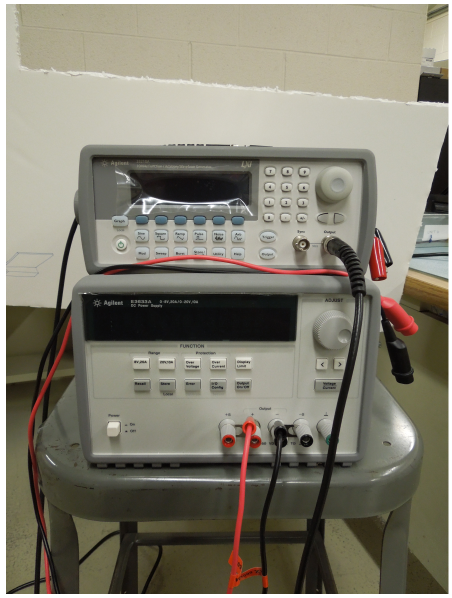

Figure 2.12. Function generator and DC power supply.

2.2.4.3 Aurduino MEGA R3 microcontroller

Aurdiono microcontroller can be powered via USB connection or with an external power supply. It works ideally for 7-12V. It has about 14 PWM output out of which 4 is connected to the bread board, through which it receives a tachometer signal from the fan to sense the fan speed. It can be communicated using a computer and can be used to record the fan speed measured.

Fan curves are generated for server equipped fans*(Delta QFR0612UH)* using these hardware which is used to measure the max flow rate provided by the fans and is also used in the CFD model during server validation. Figure 2.13 shows the plot for fan speed and PWM and Figure 2.14 shows the plot for power and fan speed.

2.3 Test Execution Procedure

Step 1 of the test procedure begins with placing the server in an environmental chamber to maintain the rack inlet temperature and relative humidity at less

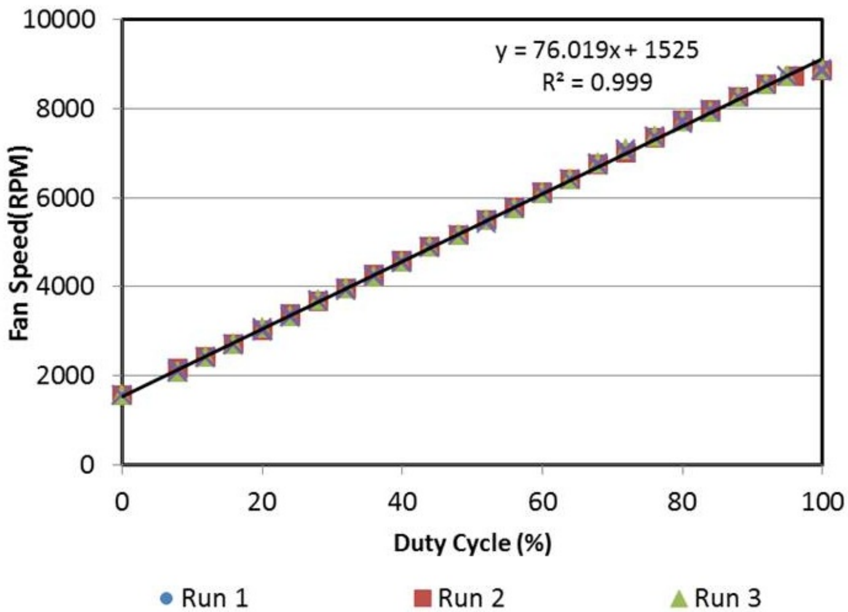

Figure 2.13. Fan Speed vs PWM.

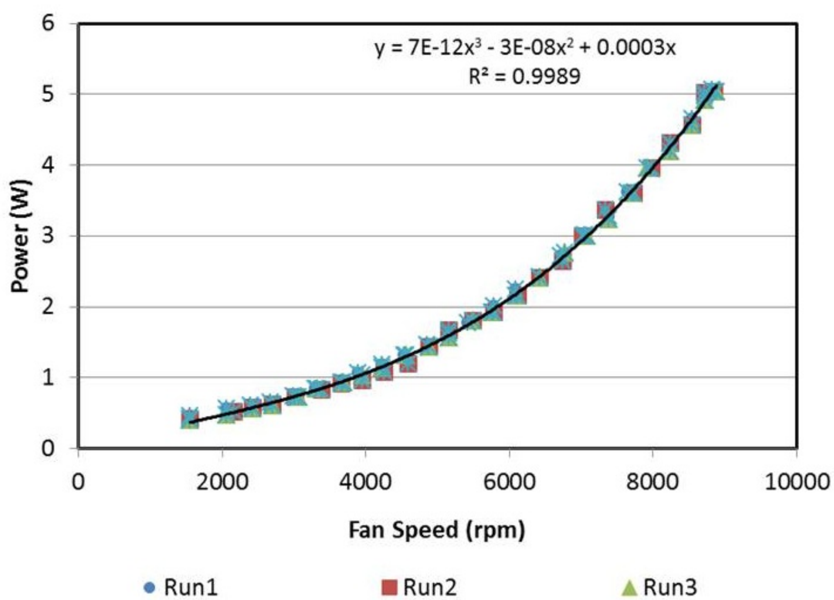

Figure 2.14. Power vs Fan Speed.

than 65%. The server is communicated using a desktop workstation through a LAN connection. A shell script deploys a pre-written code. The script applies a specific percentage of synthetic load on the CPU through the lookbusy application. The application also facilitates a fixed memory usage. This memory usage during the test is constantly measured using the free command. The 11 hour test for different CPU utilization from 10% to 98% is applied in increments of 10%. The CPU utilization is measured using the mpstat command. The free command and the mpstat command records data for a time interval of every 2 seconds for the entire time period of the test. A system health monitoring module measures the die temperature of both CPUs and server fan speeds and also controls the fan speed based on the fan control algorithm.

The server is instrumented with thermocouples to measure the surface temperature of major heat generating components. The server inlet and outlet air conditions (temperature, RH, dew point) are measured using USB data loggers.

The server is tested for a range of rack inlet temperatures; 30°C , 35°C and 40°C.

The entire experimental procedure is repeated at least three times for different percentage CPU utilizations for individual rack inlet temperature to check for repeatability and the average of all the three runs are calculated and plotted.

For step-2 of the experimental procedure, results from the base case are considered. For a specific CPU utilization, their respective average fan speeds are calculated. The server is now set to run at this average fan speed with fans powered and controlled externally for there respective CPU utilization. The experiment is repeated for different ambient temperature and different CPU utilization.

The server is tested for a range of rack inlet temperatures; 30°C , 35°C and 40°C.

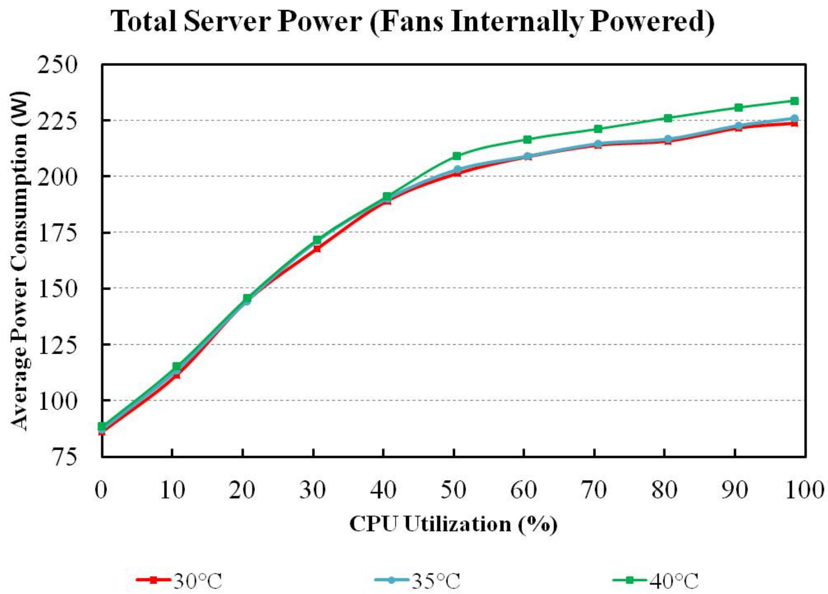

Figure 2.15. Average Server Power vs CPU Utilization.

2.4 Results

In this part we discuss the results of both cases of our experiments. Step-1 or baseline case results for are discussed first. This experiment gives a wide range of results. Important few results which are considered for step-2 of the experiments are server power consumption at different utilization and server fan speeds at different utilization. Figure 2.15 shows a plot of average of three runs of server power consumption over different utilization and different ambient temperature. Figure 2.16 shows a plot of average of three runs of fan speeds over different utilization and different temperature.

This average fan speed is kept constant by externally powering and controlling the fan at this speed.

Surface temperature of major heat generating components at different utilization is not considered with Step-1 as it's a function of fan speed control algorithm.

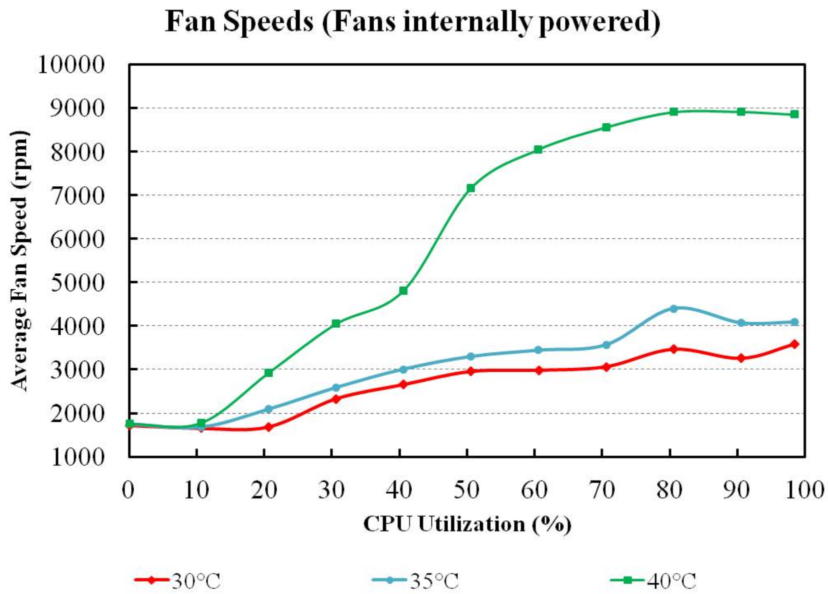

Figure 2.16. Average Fan Speed vs CPU Utilization.

Figure 2.15 clearly shows that the server power consumption is linearly increasing with increase in CPU utilization till about 50%. So the Step-2 of the experiment were sampled to only CPU utilization of 50%, 70% and 98%.

The fan speed is considered by taking an average fan speed of internally powered fan for each CPU utilization for their respective ambient temperature. Figure 2.17 shows the average server power consumption when the fans are running at fixed speed. Figure 2.18 shows average server fan speed at specific utilization as discussed before.

Figure 2.19 and Figure 2.20 shows the average server consumption and average fan speed at 30°C ambient temperature for previously discussed utilizations.

Figure 2.21 shows the surface temperature of major heat generating components at different utilization.

Figure 2.22 and Figure 2.23 shows the average server consumption and average fan speed at 35°C ambient temperature for previously discussed utilizations.

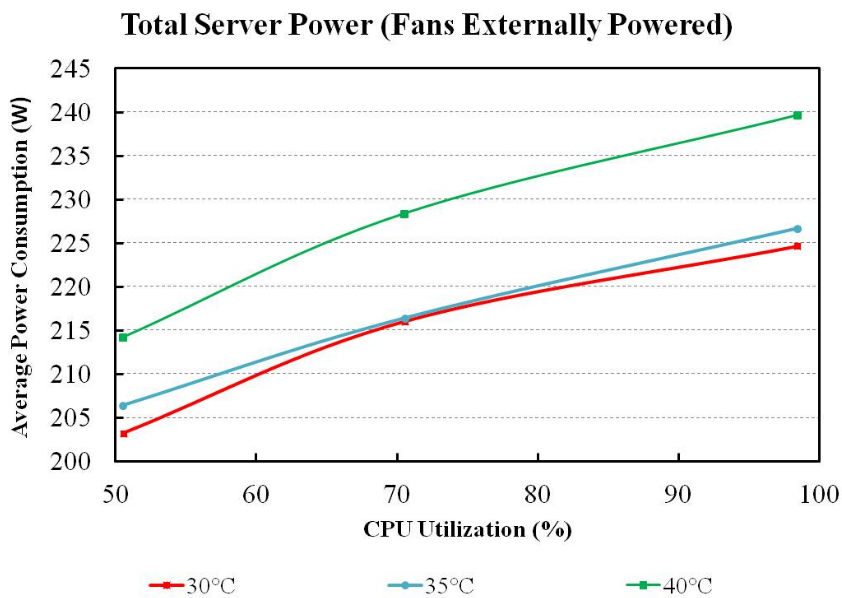

Figure 2.17. Average Server Power vs CPU Utilization.

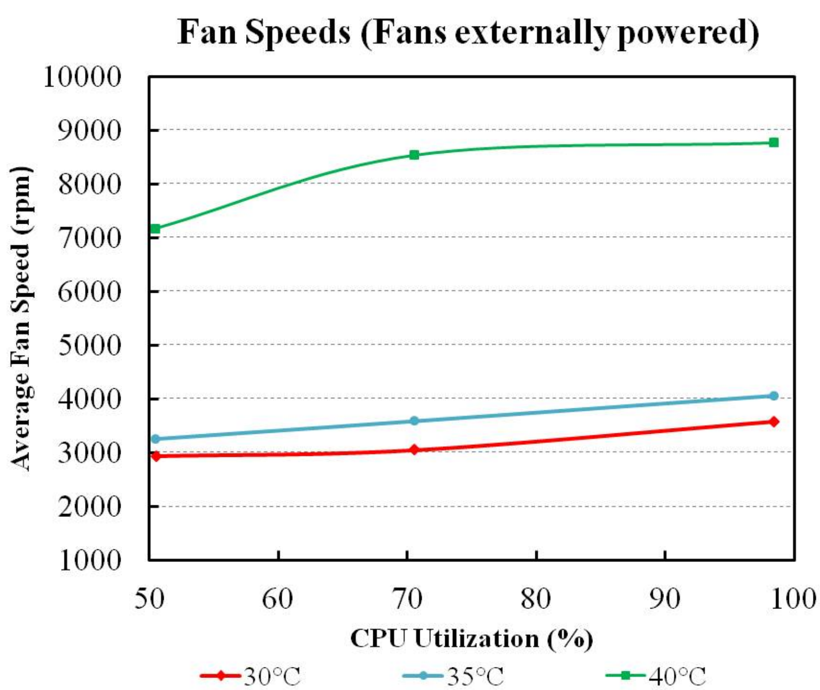

Figure 2.18. Average Fan Speed vs CPU Utilization.

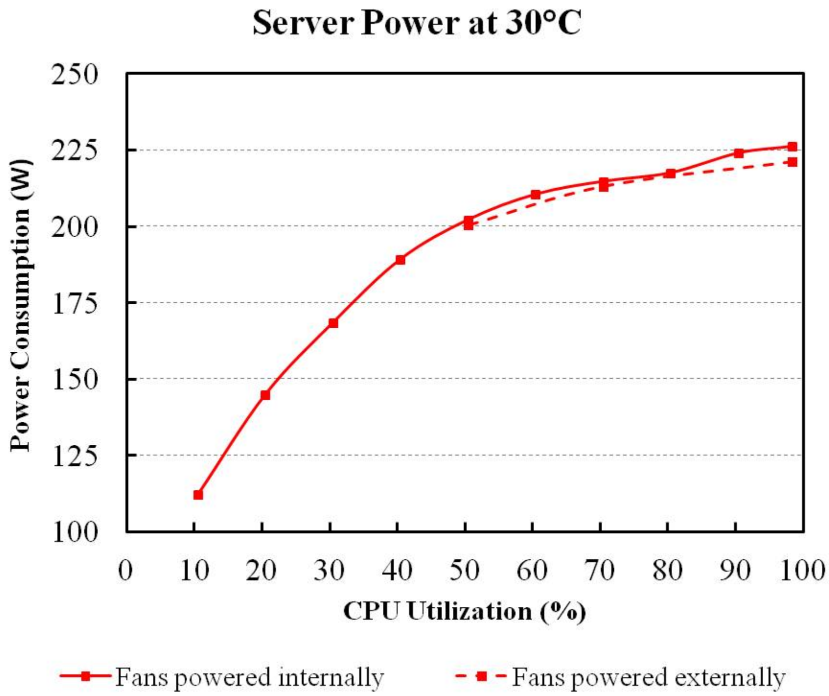

Figure 2.19. Average Server Power vs CPU Utilization.

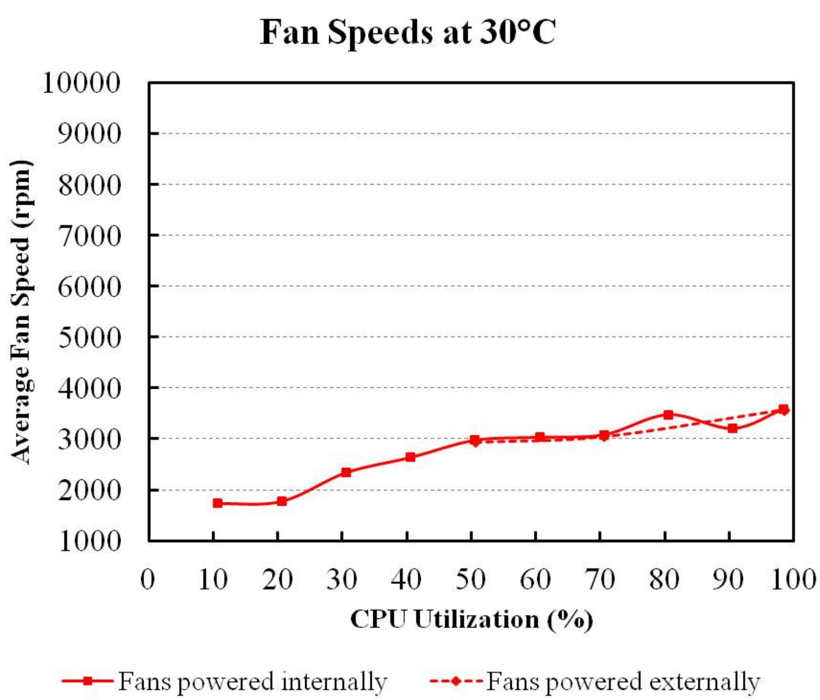

Figure 2.20. Average Fan Speed vs CPU Utilization.

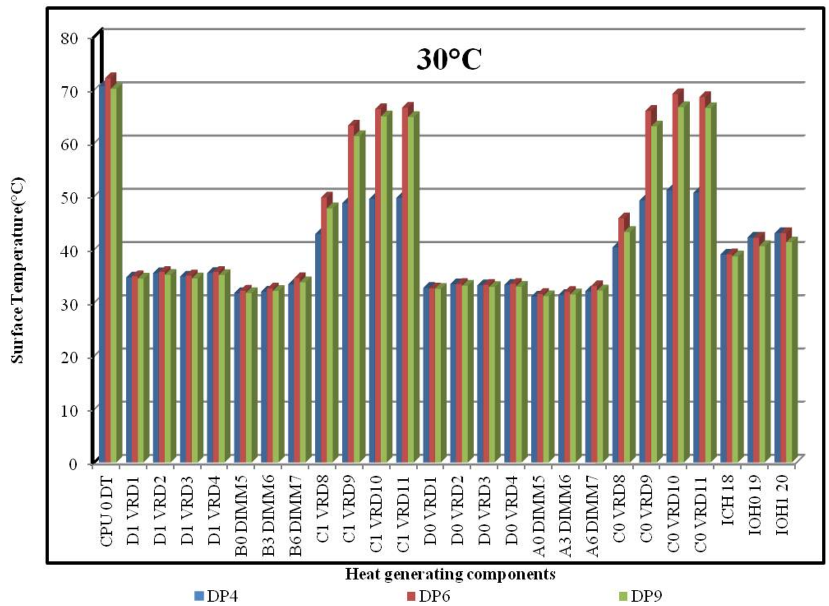

Figure 2.21. Surface Temperature of different components.

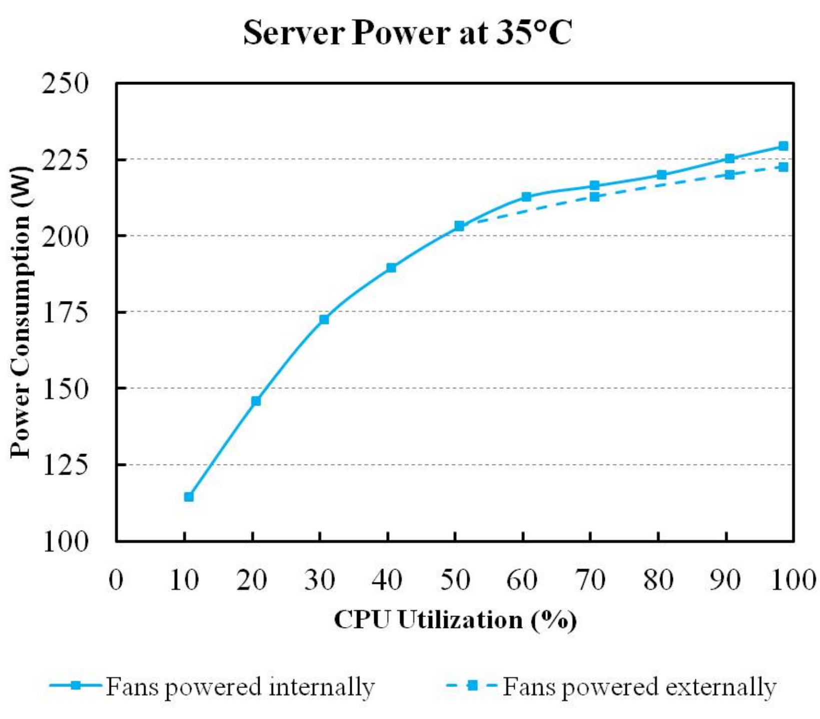

Figure 2.22. Average Server Power vs CPU Utilization.

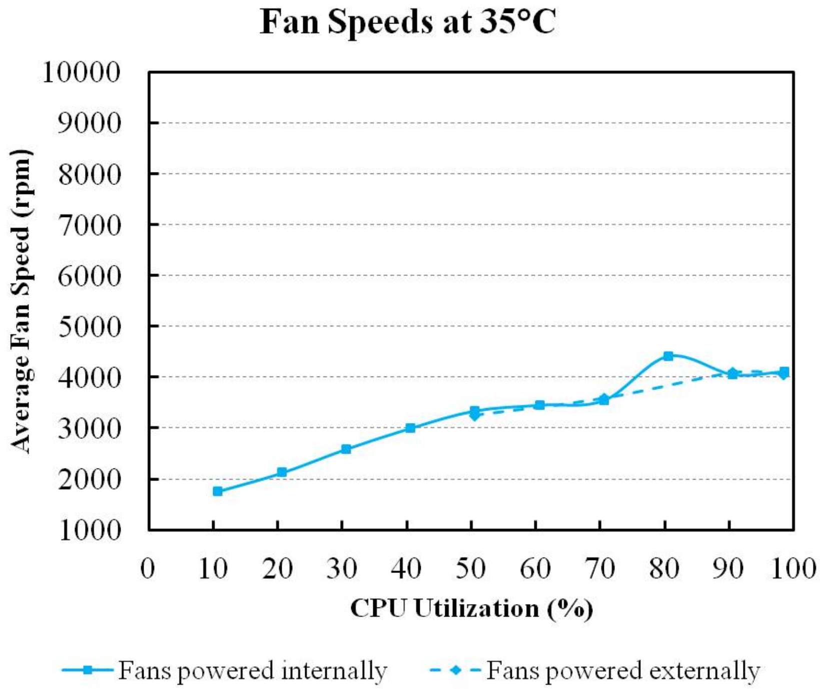

Figure 2.23. Average Fan Speed vs CPU Utilization.

Figure 2.24 shows the surface temperature of major heat generating components at different utilization.

Figure 2.25 and Figure 2.26 shows the average server consumption and average fan speed at 40°C ambient temperature for previously discussed utilizations.

Figure 2.27 shows the surface temperature of major heat generating components at different utilization.

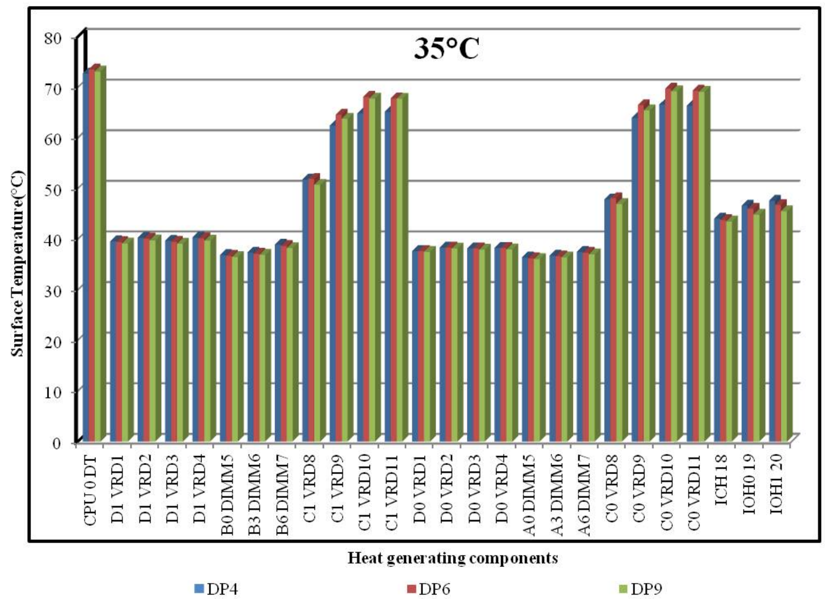

Figure 2.24. Surface Temperature of different components.

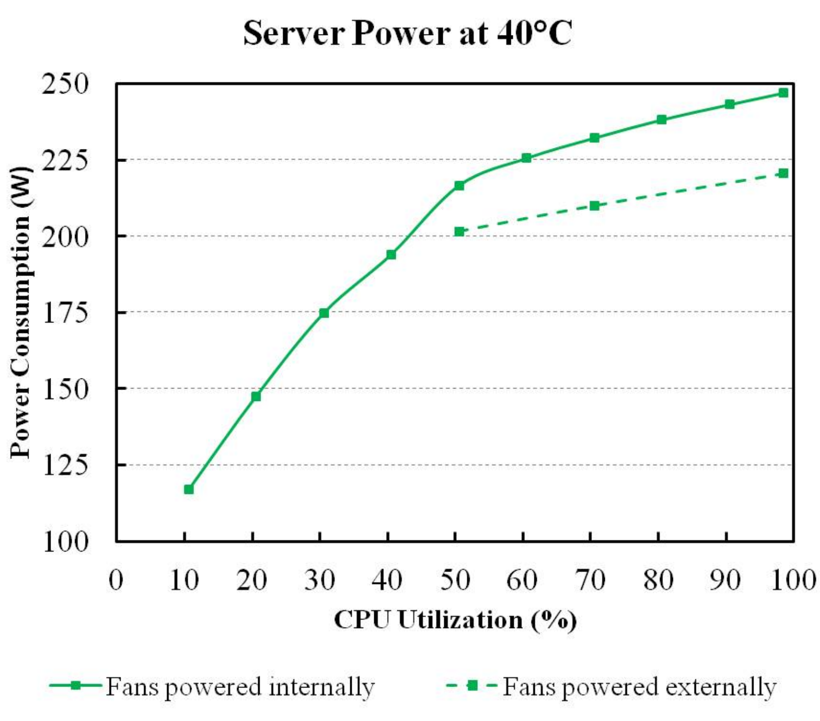

Figure 2.25. Average Server Power vs CPU Utilization.

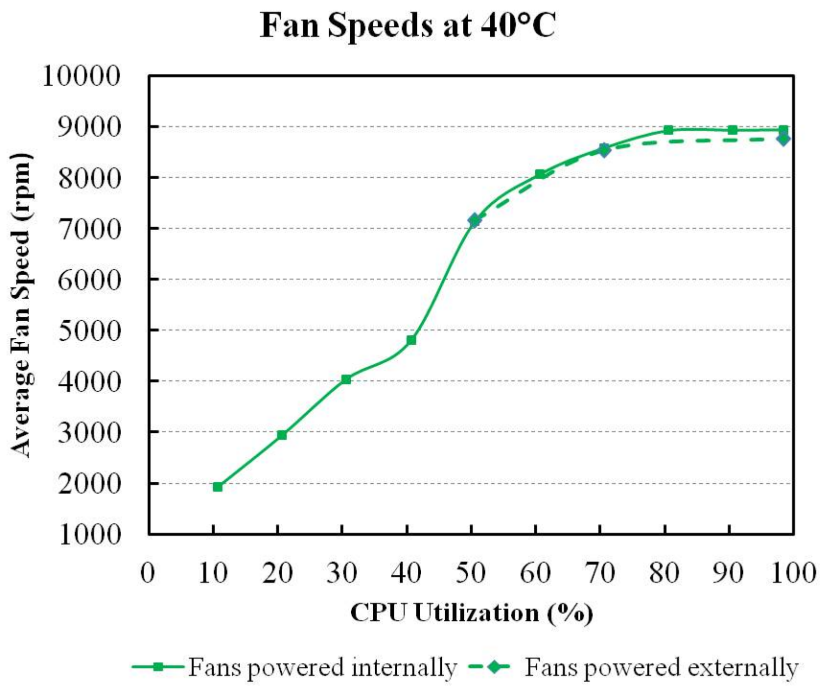

Figure 2.26. Average Fan Speed vs CPU Utilization.

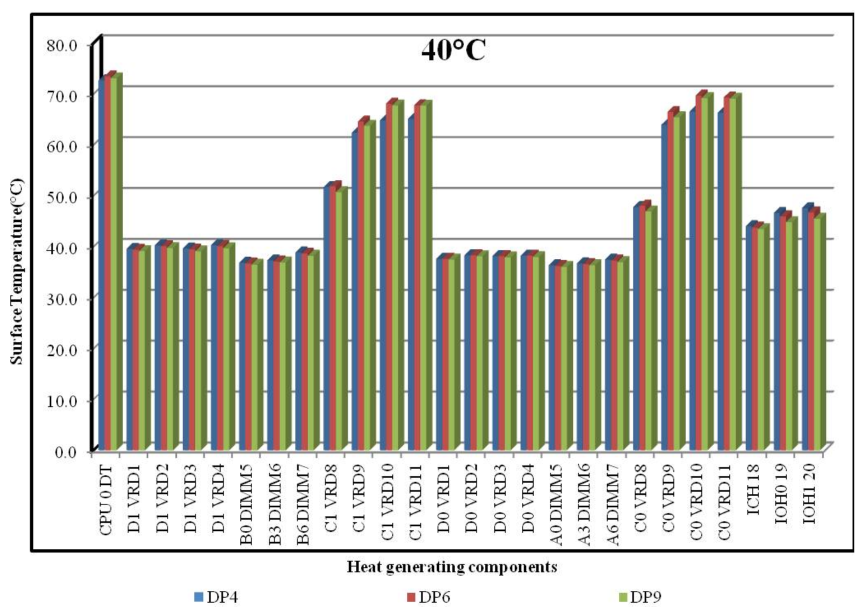

Figure 2.27. Surface Temperature of different components.

CHAPTER 3

CFD Modeling

This chapter explores the scope of modeling a server with reduced-order detail of all major heat generating and flow restricting components and experimentally validates its temperature parameters to an acceptable accuracy level. 6Sigma ET a commercial CFD software package was used to model the server. A description of modeling of each of the important components along with its material properties and dimensions is discussed. The model considered here is already validated for flow and focus is drawn only on thermal validation. The results of temperature profile are presented at the end of this discussion.

The concept of this modeling work would facilitate to observe impact of architectural change in the system level as well as to serve as a guidelines in rack and room level solutions. CFD model gives temperature results, in each point of a 3 dimensional space considered. Experimental comparison can be executed in different ways. In this research we concentrate on specific points method where specific point's or approximate hotspot surface points, of the major heat generating components are considered to validate. This would be a reasonable method as study is focused on specific heat generating components.

3.1 Server Model

The server under consideration is Intel Open Compute server which is a CPU dominant server as the CPU's consume the maximum of the total server power. It is a compute rack mount server of rack height 66mm or approximately of rack

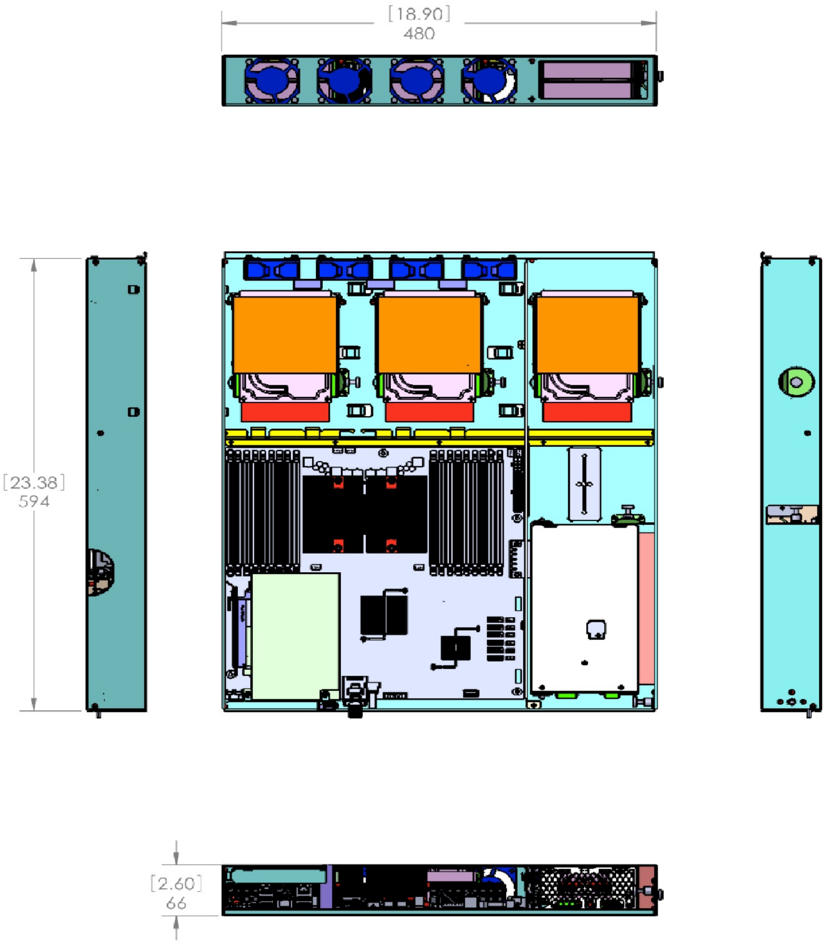

Figure 3.1. CAD representation of the server with Intel board,in[inches] and millimeters [8].

height 1.5U. "U" is the unit of measure used to describe the height of the rack equipment. Figure 3.1 shows the server considered for this research populated with all its components.

3.1.1 Chassis

The server is made of zinc pre-plated, corrosion resistant sheet metal of 1.2mm thick. The chassis are designed to accommodate both Intel motherboard and AMD

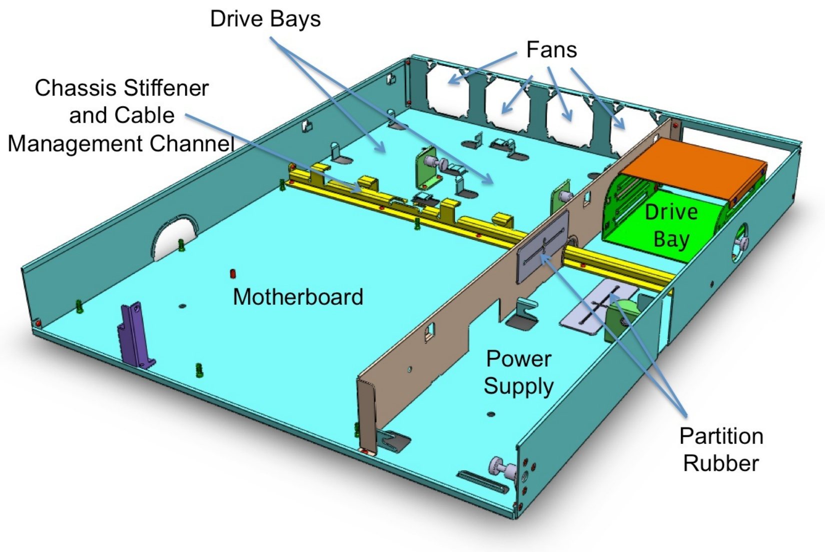

Figure 3.2. Chassis.

motherboard. The chassis is 66mm tall and has a depth of 594mm and width of 333mm.

It can also be clearly seen from the Figure 3.2 that the chassis is partitioned into two parts to divide the motherboard from the power supply unit. Partitions create a separate in-flow and out-flow for the motherboard part and power supply part and this is considered and the case is replicated while modeling the server.

The chassis modeled in 6Sigma is shown in Figure 3.3. Importance to detail in the model is given to all the chassis openings as to replicate the same condition during the simulation.

3.1.2 Motherboard

Intel motherboard is considered in this research work. The motherboard model is populated with only major heat generating and flow obstructing components. The power consumed by other small components which are not modeled are compensated by distributing the power to the complete motherboard. The motherboard modeled

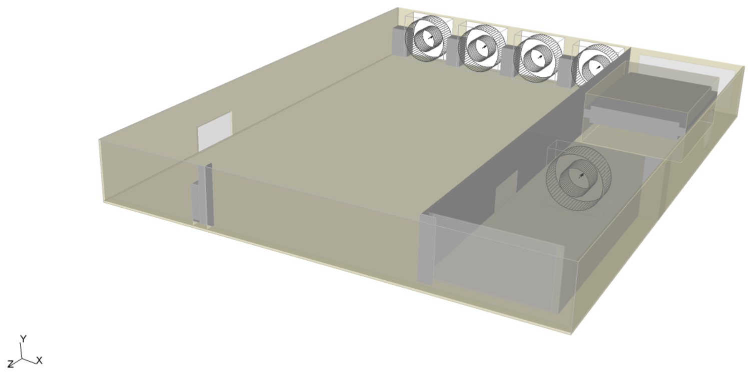

Figure 3.3. CFD Chassis Model.

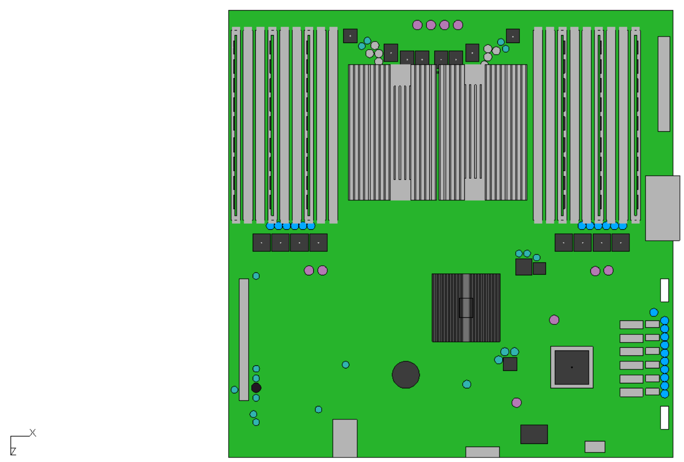

Figure 3.4. CFD Motherboard Model.

is as shown in Figure 3.4. The motherboard is 1.5mm thick with a width of 328mm and depth of 330mm. It has 12 layers in total of FR4 and copper.

3.1.3 Power Supply

Power supply unit is modeled with highest detail as it is partitioned in the chassis and has a separate inflow and outflow condition. Components for both AC supply and DC supply are considered in the modeling process. As discussed earlier the power supply efficiency is a function of CPU utilization the power distribution is

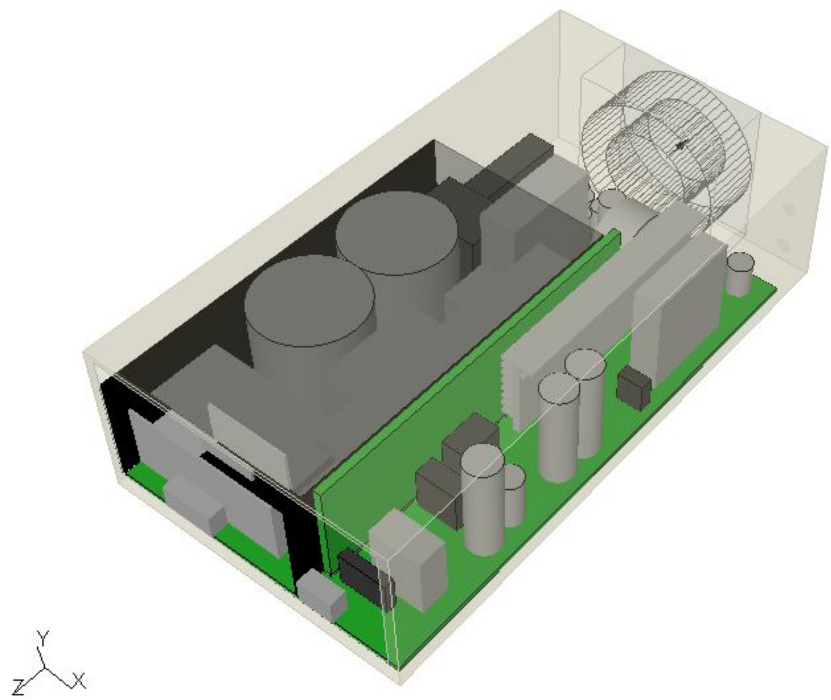

Figure 3.5. Power Supply Unit Model.

depended variable. For different test cases the power is distributed is a product of total server power consumption and efficiency at that case.Figure 3.5 shows a detailed power supply model.

3.1.4 Fans

The servers are equipped with four, 60mm x 60mm Delta QFR0612UH fans. The fans have a rated voltage of 12V and input current of 0.2 amps [9]. Figure 3.6 shows the fan specifications from the manufacturer which is implemented in the model. The fan curve which was experimentally calculated using an airflow bench to match the manufacturers data can be seen implemented in the model in Figure 3.7.

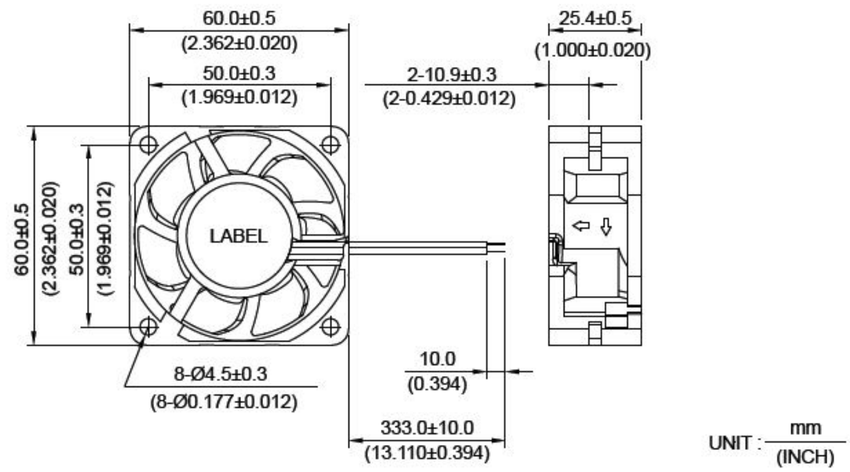

Figure 3.6. Fan Specifications [9].

3.1.5 Ducky

Ducky plays a very important role in the server operation. Forced convection is the major type of heat transfer in the server and the air flow over the components to be cooled dictates the heat transfer from those components. But air flow always takes the path of least resistance and this is avoided to a great extent by channelizing the air flow path to the components to be cooled. The Figure 3.8 shows the ducky, modeled with great detail. The ducky material is ABS-plastic.

3.1.6 Heat Sink

The heat sinks used are extruded aluminum heat sinks. The CPU heat sinks are 45mm tall and have a base width of 66mm and depth of 100mm. The extruded fins have a fin thickness of 1.8mm and have 18 fins. The modelled CPU heat sink is as shown in the Figure 3.9.

The IOH has a 45mm tall, 50x50mm anodized aluminum extruded heat sink with 22 fins of 1mm thick. The modelled ICH heat sink is as shown in the Figure 3.10.

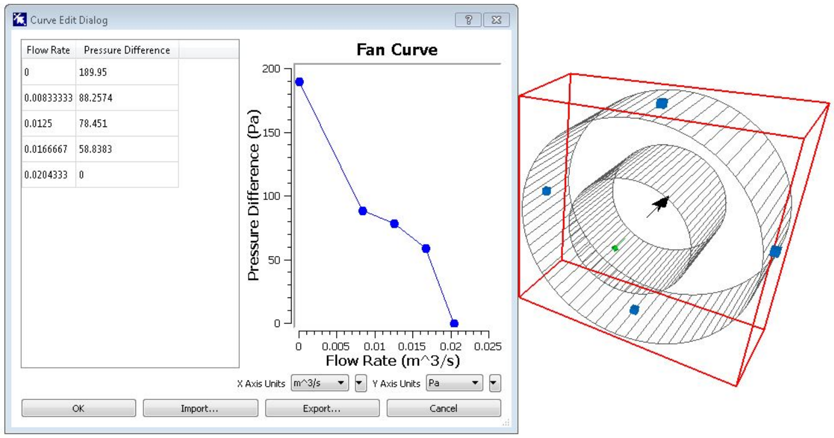

Figure 3.7. Fan Curve.

3.1.7 DIMM

DIMM acronym of Dual In-line Memory Module is a small circuit board the holds memory chips. The DIMMs are modeled with great detail. All the nine memory chips are modeled along with DIMM slots. Figure 3.11 shows the modeled DIMM.

The Complete CFD model of the server can be seen in the Figure 3.12.

3.2 Boundary Conditions

The initial conditions are not specified or considered as we not working on a time-dependent problem.

It is impossible and unnecessary to simulate the entire universe in a simulation and so a region of interest is chosen for the simulation. This chosen region has a boundary with the surrounding environment making it important to define the boundary condition.

One of the important boundary condition in this work is the ambient temperature. Ambient temperature plays a prime role in the server power consumption.

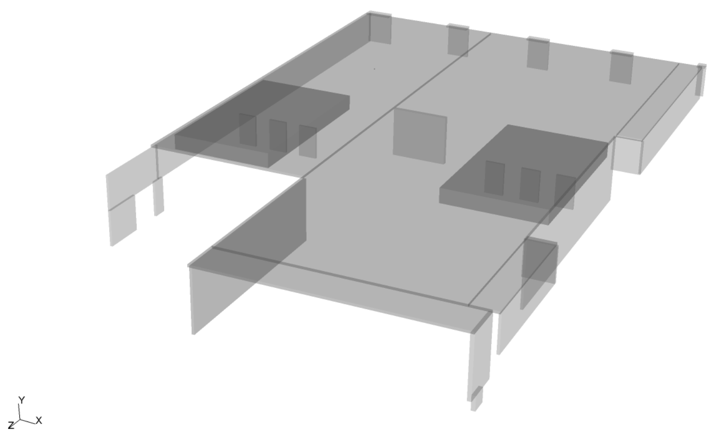

Figure 3.8. Ducky.

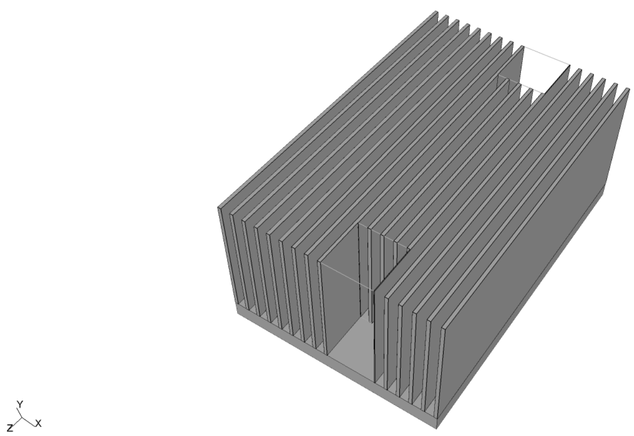

Figure 3.9. CPU Heatsinks.

Higher temperature would energize the molecules and hence the electrons leading to higher leakage current and hence higher power dissipation.

The other important boundary condition is the CPU utilization. The server considered is a CPU dominant and hence CPU utilization drives the server power consumption. Fan speed and the power consumptions of components supporting the CPU's like the voltage regulation modules are also function of CPU utilization. Fan speed is predetermined from step-1 of the experimental procedure.

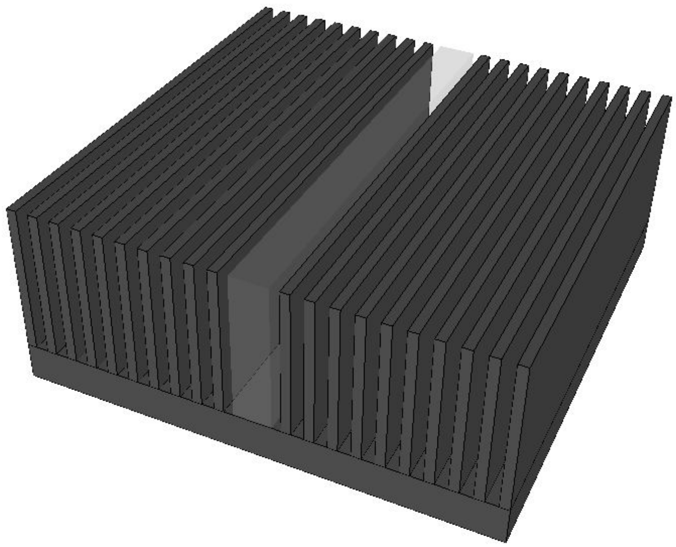

Figure 3.10. CPU Heatsinks.

3.3 Power Distribution

Power Distribution is of vital importance. The total server power was distributed to different components based on their Thermal Design Power (TDP). Intel defines the upper point of the thermal profile consists of the Thermal Design Power (TDP) and the associated T_{case} value. Thermal Design Power (TDP) should be used for processor thermal solution design targets. TDP is not the maximum power that the processor can dissipate [14].

Each component modeled in the server is assigned its maximum TDP. For each test case considered, a percentage utilization of this TDP is assigned for a specific ambient temperature, CPU utilization and their respective fan speeds.

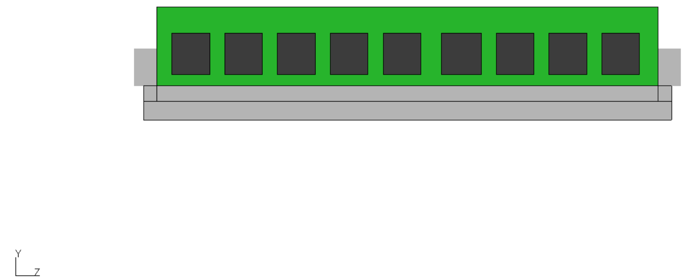

Figure 3.11. Dual In-line Memory Module Model.

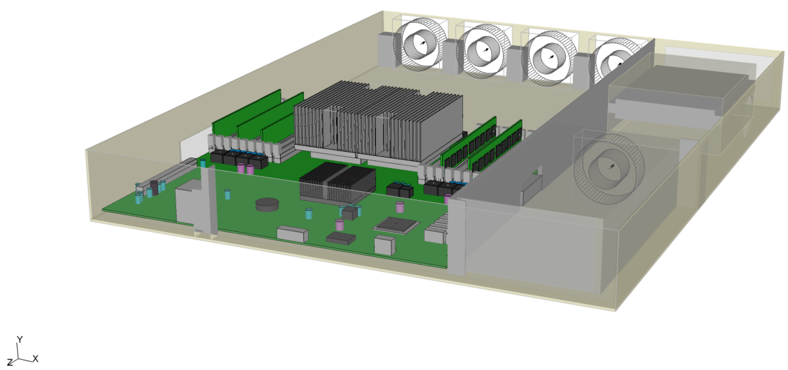

Figure 3.12. Complete CFD Model.

3.4 Results

In this part we discuss the simulation results. The server power consumptions and surface temperature of the modeled components for 50%, 70% and 98% utilization at ambient temperature of 30°C , 35°C and 40°C is observed and plotted.

Figure shows the server power consumption from CFD at three different ambient temperatures for previously discussed utilizations.

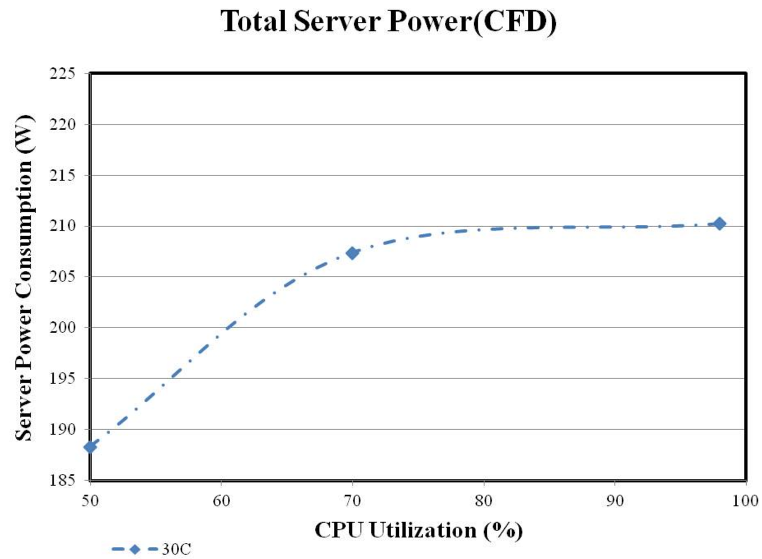

Figure 3.13. Server Power Consumption at 30°C.

Figure 3.13 and Figure 3.14 shows the server power consumption and surface temperature of modeled components respectively at 30°C ambient temperature for previously discussed utilizations

Figure 3.15 and Figure3.16 shows the server power consumption and surface temperature of modeled components respectively at 35°C ambient temperature.

Figure 3.17 and Figure 3.18 shows the server power consumption and surface temperature of modeled components respectively at 40°C ambient temperature.

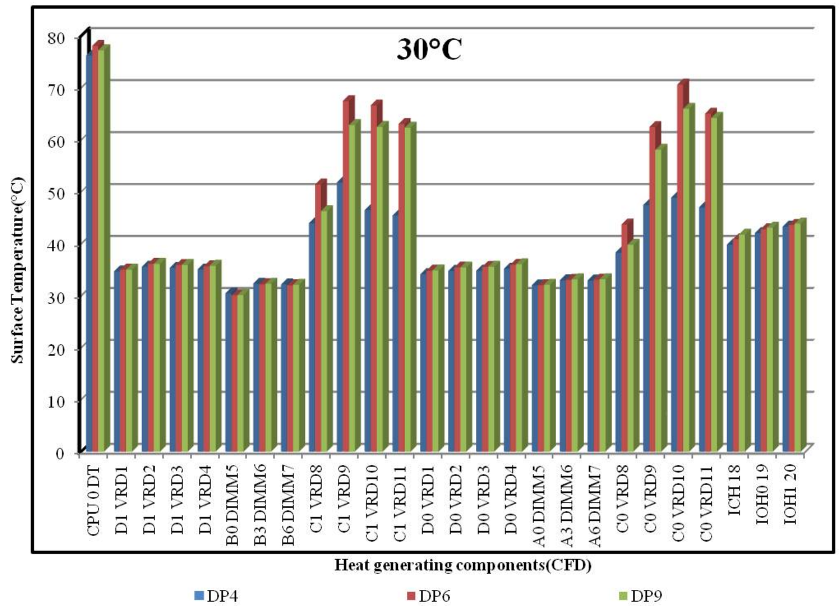

Figure 3.14. Surface Temperature of different components.

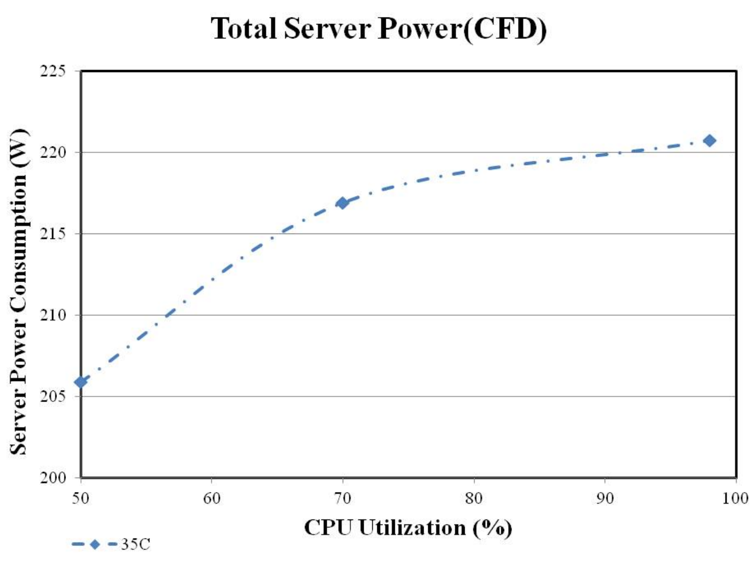

Figure 3.15. Server Power Consumption at 35°C.

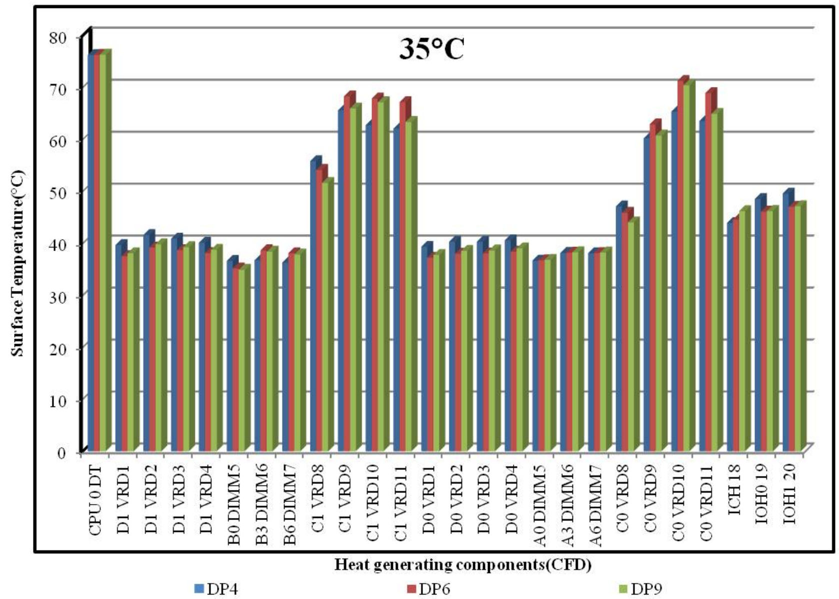

Figure 3.16. Surface Temperature of different components.

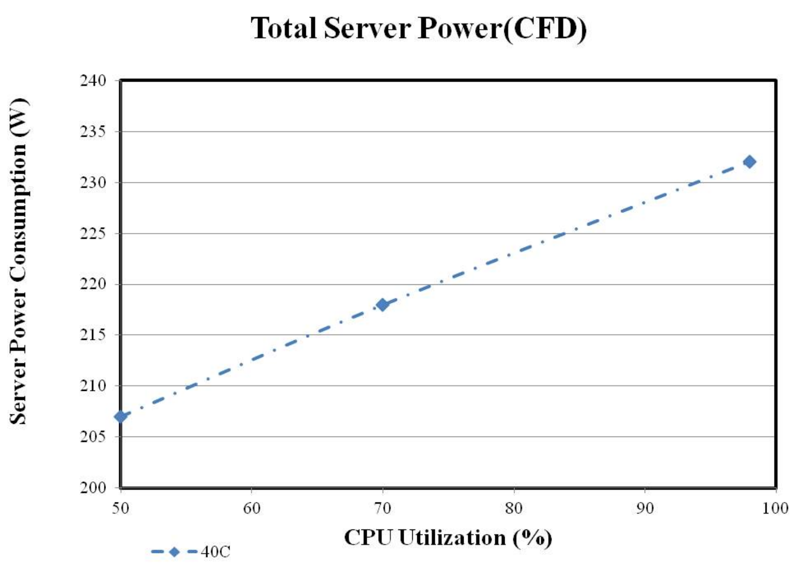

Figure 3.17. Server Power Consumption at 40°C.

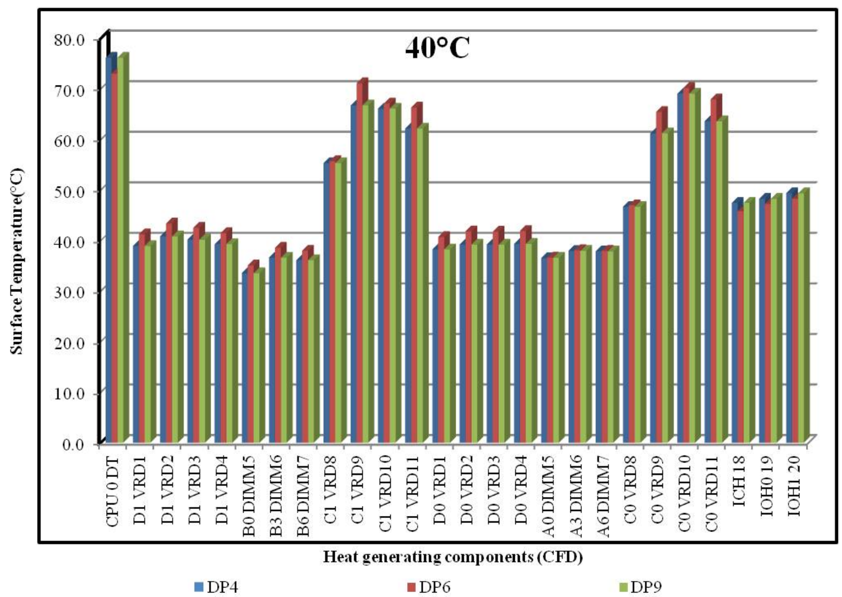

Figure 3.18. Surface Temperature of different components.

CHAPTER 4

Resutls and Conclusion

In this section we will discuss the results and compare it with experimental results and CFD simulation for the accuracy of our validation. In order to create an accurate CFD model, many approximations must be made in order to make the model development process simple and run simulations using current technologies. Discrepancies emerge in the design with those approximations between the CFD results and experimental observations; therefore, each assumption must be experimentally tested for validity.

Traditionally, compact models in all packaging levels are created and the thermal analysis and management is considered because detail model analysis of racks and data centers has not been possible due to limitation on computational resources. Result of this leads to not so accurate model.

This research serves as a foundation to perform accurate thermal analysis by creating a detail model at the system level which is later implemented as simplified (compact) model for the rack and room level analysis. A reduced-order detail model of single Intel Open Compute server was modeled using 6SigmaET. All the major heat generating and flow restricting components were considered for the analysis. The other heat generating components which were not modeled were considered as a source of heat in the motherboard and an average power is distributed to it.

4.1 Comparison of Experimental Study and CFD model

Ambient temperature and CPU utilization are the important boundary condition defined for each case considered. Fan curves were experimentally calculated using air flow bench and are fixed for all the test cases. CPU utilization with the influence from ambient temperature changes the power distribution over the motherboard for each design case. Server power consumptions and the surface temperatures of components are the two output parameters considered for validation. Because of some uncertainties in the measurement of empirical inputs along with geometric approximation some discrepancy between the simulated model and actual physical entity is expected. As per this expectation it is safe to assume a error percentage of 10 is within the limits of accuracy.

As experimental results are assumed to give an exact value the result comparison is expressed in percentage error between the simulated model and experimental results. Comparison which can also be expressed in percentage difference is just not considered in this research.

Serve power consumption and surface temperature at each pre discussed consideration of CPU utilizations for a specific ambient temperature is compared between experimental results and CFD simulation.

Figure 4.1 shows the comparison of experimental results of server power consumption when fans powered externally and CFD simulated server power at 30°C at all three different CPU utilization.

Figure 4.2, Figure 4.3 and Figure 4.4 shows the comparison of experimental results of surface temperature of major heat generating components when fans powered externally and CFD simulated surface temperature of modeled components at 30°C at all three different CPU utilization.

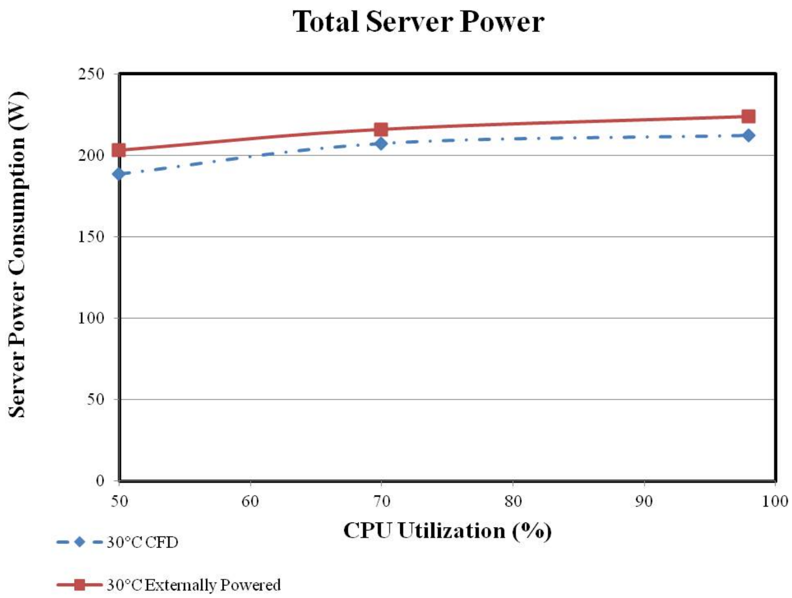

Figure 4.1. Server power consumption of experiment vs CFD.

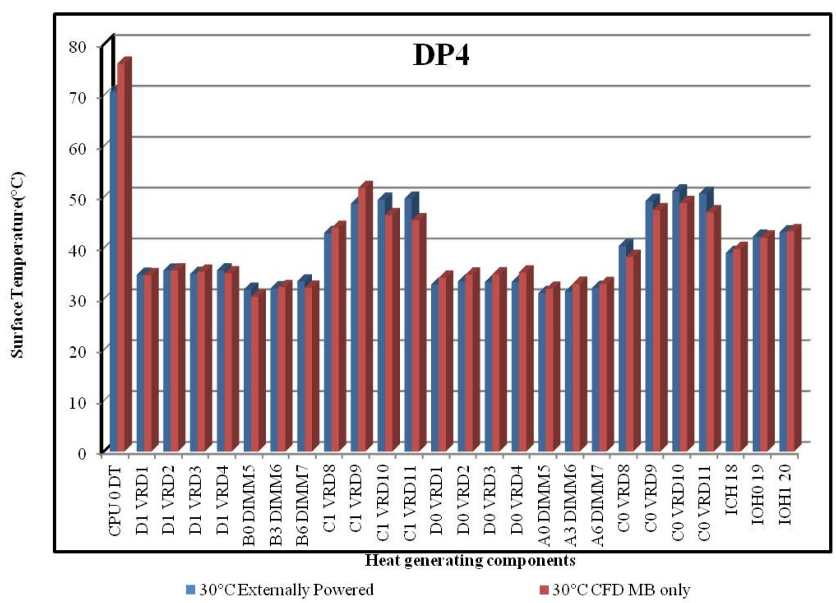

Figure 4.2. Surface temperature at 50% utilization.

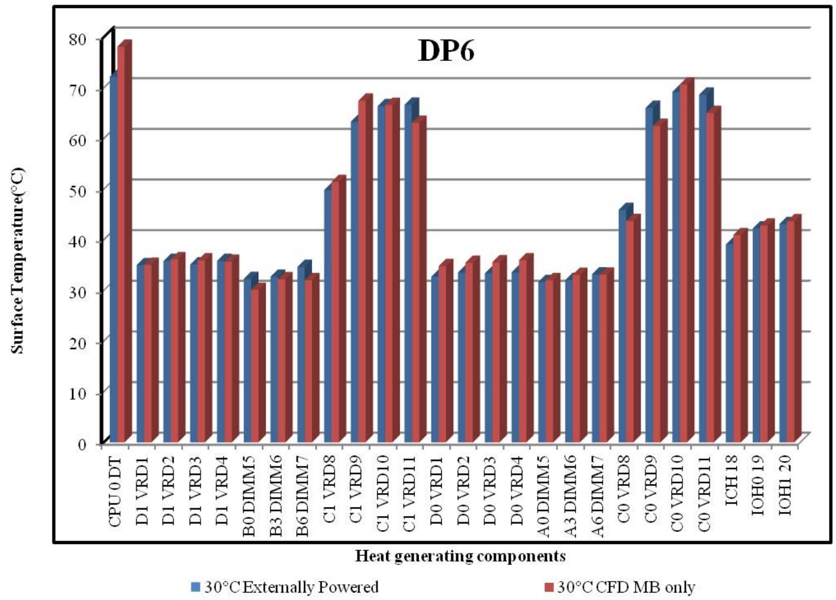

Figure 4.3. Surface temperature at 70% utilization.

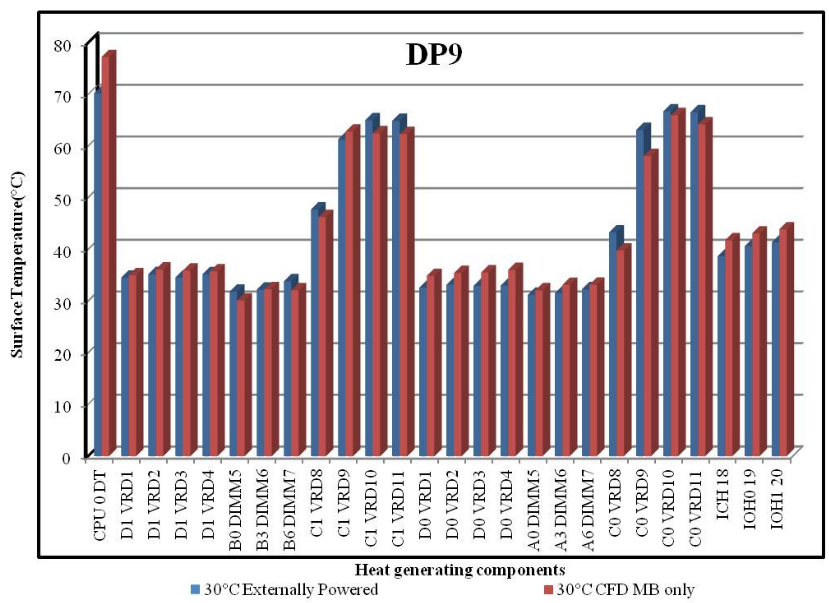

Figure 4.4. Surface temperature at 98% utilization.

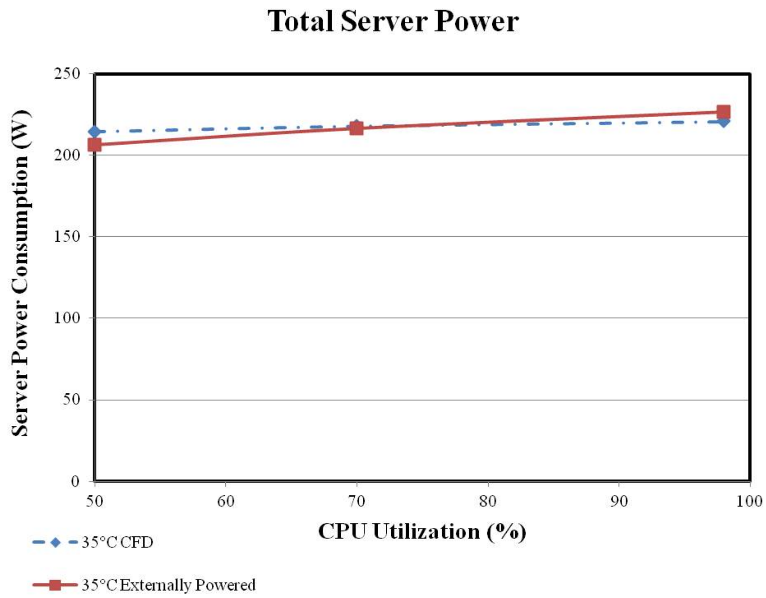

Figure 4.5. Server power consumption of experiment vs CFD.

Figure 4.5 shows the comparison of experimental results of server power consumption when fans powered externally and CFD simulated server power at 35°C at all three different CPU utilization.

Figure 4.6, Figure 4.7 and Figure 4.8 shows the comparison of experimental results of surface temperature of major heat generating components when fans powered externally and CFD simulated surface temperature of modeled components at 35°C at all three different CPU utilization.

Figure 4.9 shows the comparison of experimental results of server power consumption when fans powered externally and CFD simulated server power at 40°C at all three different CPU utilization.

Figure 4.10, Figure 4.11 and Figure 4.12 shows the comparison of experimental results of surface temperature of major heat generating components when fans powered externally and CFD simulated surface temperature of modeled components at 40°C at all three different CPU utilization.

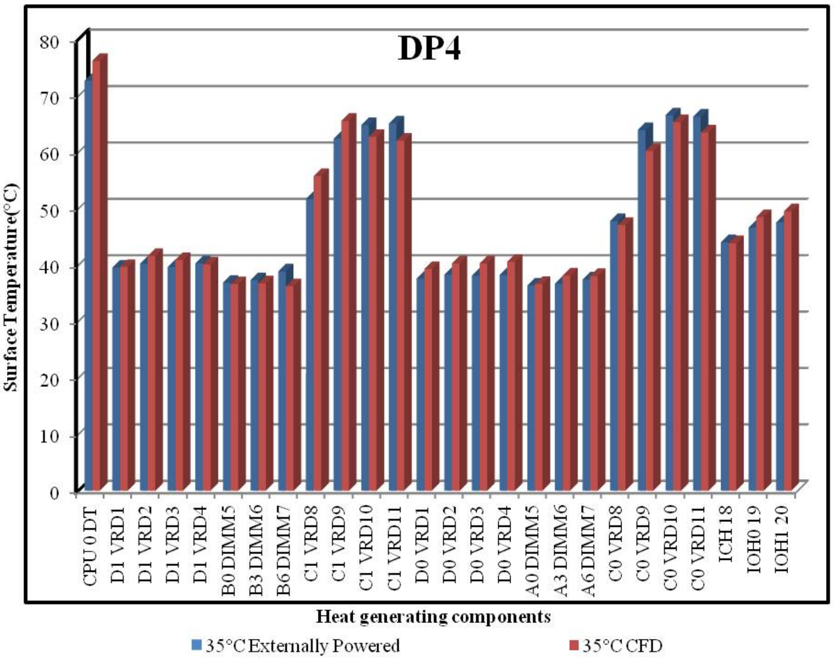

Figure 4.6. Surface temperature at 50% utilization.

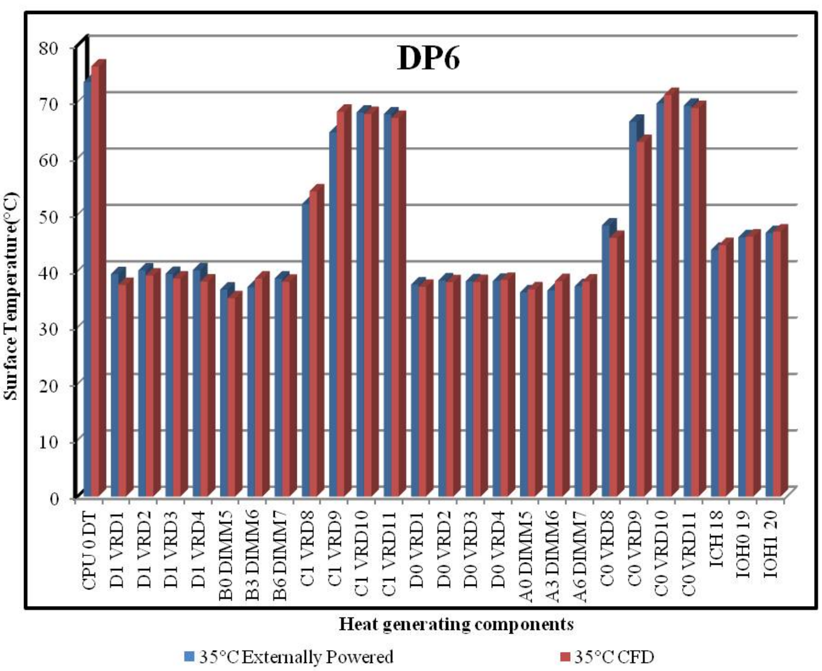

Figure 4.7. Surface temperature at 70% utilization.

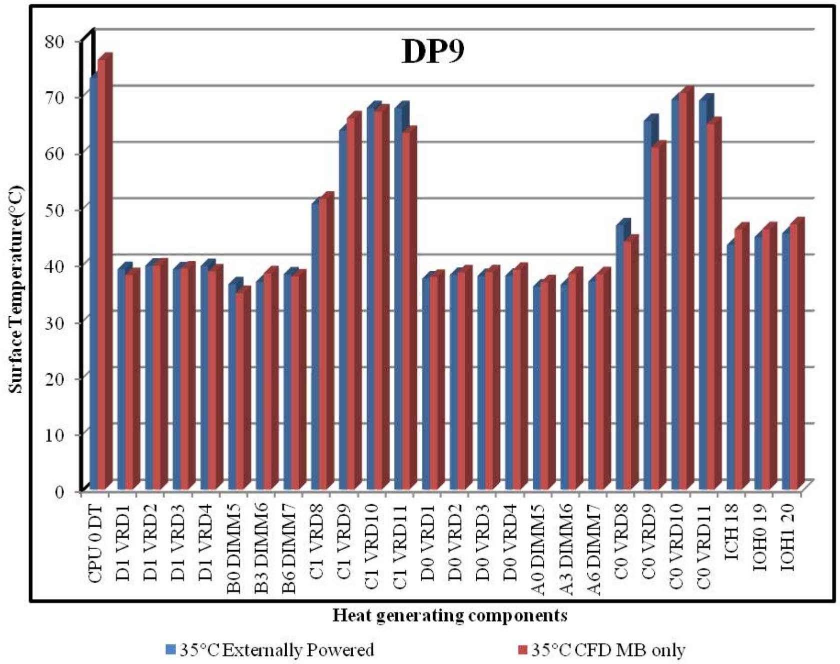

Figure 4.8. Surface temperature at 98% utilization.

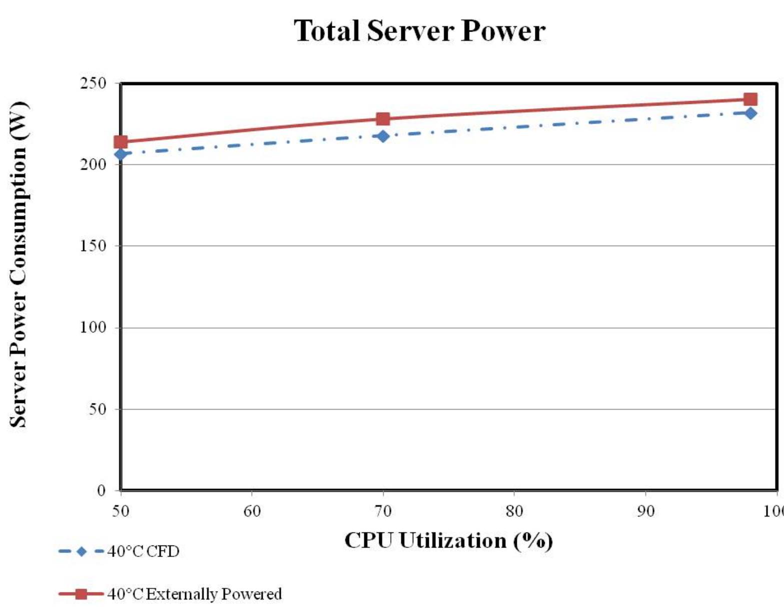

Figure 4.9. Server power consumption of experiment vs CFD.

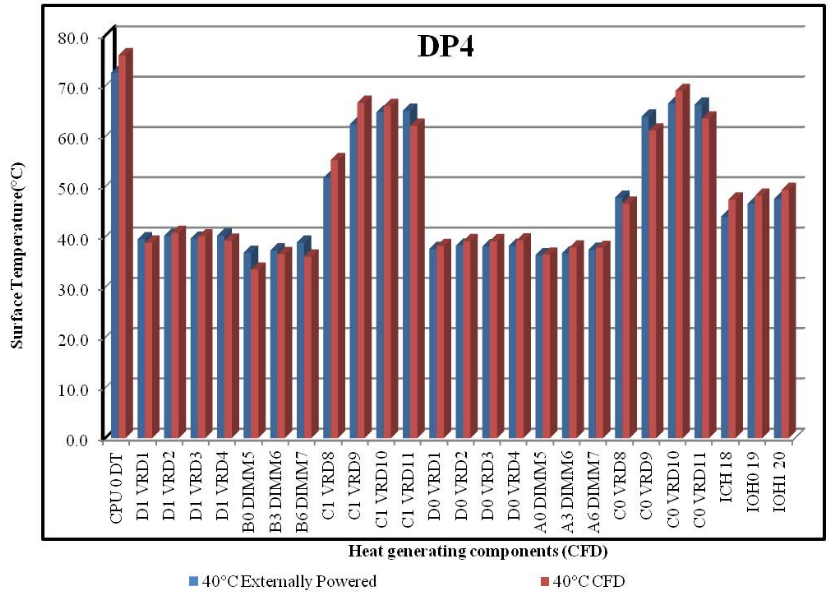

Figure 4.10. Surface temperature at 50% utilization.

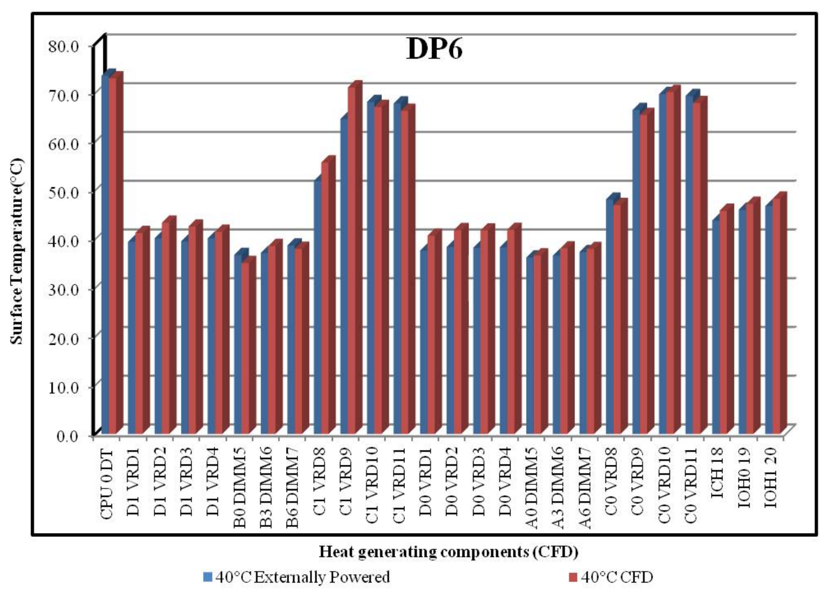

Figure 4.11. Surface temperature at 70% utilization.

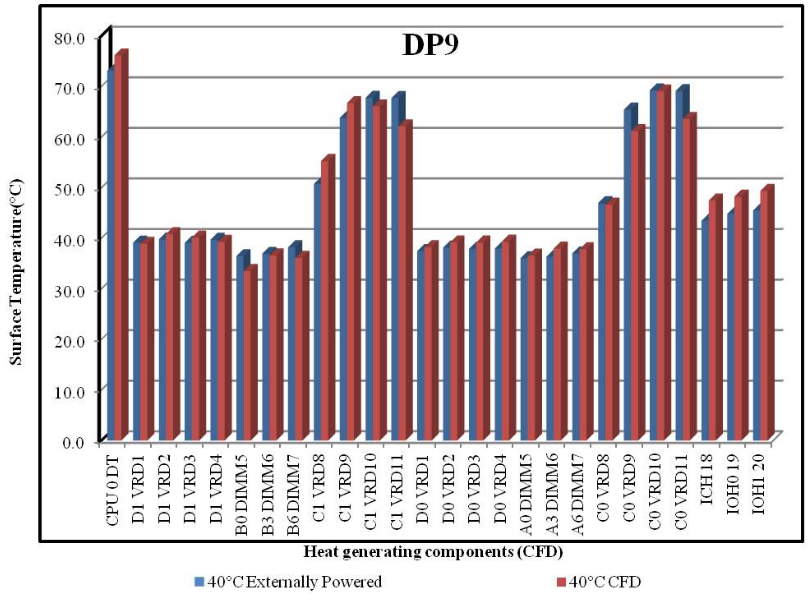

Figure 4.12. Surface temperature at 98% utilization.

4.2 Conclusion

CFD modeling plays a very vital role in analysis of data center operation. It is the only economical method to observe effects in a data center with parametric changes like air flow rate or other boundary conditions. Approach taken in this research will lead to use of CFD to observe hardware architecture change in server and their effect in a data center.

Data center CFD model involves modeling of racks in a room with CRAC units and due to limitations in computing resources the servers populated in the rack are usually modeled as cuboid blocks, which affects the accuracy of the analysis. If a reduced-order model is obtained for at least one of these level chip, rack and room then the models can be stitched together to obtain computationally efficient solutions.

The modeling technique used here is a detail component model of the system and not the detailed models of the component of a system. For this reason, experimental data is used as a base of comparison to validate the reduced detail model.An uncertainty in the measure of experimental data causes some discrepancies. To accommodate the discrepancies in thermocouple measurement $\pm 0.5°C$ and other geometrical approximation in modeling, it is safe to assume an error percentage of 10 to be within the limits of acceptable accuracy between experimental and CFD results. This research is concluded with the validation of CFD model to experimental results, based on server power consumption and surface temperature of modeled components.

The power distribution is based on server power consumption at different CPU utilizations. The power consumptions of the modeled components are based on their maximum thermal design power. Percentage of utilization is specified based on the CPU utilization and the expected surface temperature. Approximation of the power values is traditionally used with thermal design point as the constraint of approximation for designing thermal solutions. This approximation in this research for the CFD model is effective as the model is validated with the maximum of 10% error. It was seen the CPU1 part of the motherboard dissipated more heat. This is expected due to thermal shadowing from the IOH. This increases the ambient temperature around these components thereby increasing the electrons flow more aggressively to result in higher power dissipation due to leakage current.

6sigmaET software being exclusively designed for data center CFD modeling, it facilitates the modeling process. The library catalog feature contains smart parts that can be directly installed. The Intel server model was created by importing some of the smart parts directly and a lot more was built from scratch to exactly replicate the server model. Software enables us to create a test environment which helps to simulate the server at any ambient temperature. The percentage utilization feature facilitates to specify the percentage of power utilized for any given ambient temperature and CPU utilization.

The server was validated at three design points or CPU utilization at different ambient temperature, covering a total of 9 cases. Now the model can be converted or simplified to a cuboid box with inflows and outflows and can be used in rack level and room level solutions. The analysis of racks and rooms will be more accurate with incorporating the simplified detail model.

4.3 Future Work

The Intel Compute server validated model uses power distribution based on maximum thermal design point of each component. Procedure to calculate the actual power consumption of at least the chip and chipsets would facilitate to simulate the model to achieve higher accuracy at any design point and ambient temperature.

The validation procedure accommodated in this research was steady state. The server in reality works on fan control algorithm which regulates the CPU core temperature. When the CPU core temperature increases over a threshold limit the fans ramp up to provide more cooling. The software facilitates to incorporate this fan control algorithm which could be incorporated by designing a controller and replicate the actual server operation.

The model being validated for both flow and temperature now forms the foundation for rack and room level analysis. The software provisions to simplify this detail model to a cuboid black box with just inflows and outflow. This simplified model can be exported to the library which can be imported during rack and room level solution.

APPENDIX A

Governing Equations of Flow of Fluid and Heat Transfer

In this appendix, important principles of heat transfer and fluid flow like, the continuity, momentum, energy equations and also the Navier-Stokes equations are discussed.

A.1 Conservation of Mass Equation :

The conservation of mass equation basically states that the time rate of change of the mass in a region must be zero. In other words, the mass entering a region must equal the mass leaving a region plus any mass that is stored in the region.

The conservation of mass equation for single phase fluid may be represented as

$$\frac{\partial \rho}{\partial t} + \nabla \bullet (\rho \upsilon) = 0 \tag{A.1}$$

The conservation of mass equation for two-phase fluid may be represented as

$$\frac{\partial \rho}{\partial t} + \nabla \bullet (\rho \upsilon) = S_m \tag{A.2}$$

A.2 Continuity Equation :

The principle of continuity equation is derived from the fact that mass is always conserved in fluid systems regardless of the complexity or direction of flow. It is defined as "The mean velocities at all cross sections having equal areas are then equal and if the areas are not equal, the velocities are inversely proportional to the areas of the respective cross sections."

$$Q = A_1 V_1 = A_2 V_2 \tag{A.3}$$

A.3 Conservation of Momentum :

The conservation of momentum states that the time rate of change of the linear momentum of a region is equal to the sum of the forces on that region. There may be two types of sources acting on that region; body forces due to gravity that act on the entire bulk of fluid in the region and also surface forces, which act at the boundaries.

The conservation of momentum equation may be represented as

$$\frac{\partial \rho}{\partial t} \bullet (\rho\vec{v}) + \nabla \bullet (\rho\vec{v}\vec{v}) = -\nabla p + \nabla \bullet (\bar{\tau}) + \rho\vec{g} + \vec{F} \tag{A.4}$$

A.4 Energy Equation :

The energy equation can be defined as the rate of change of energy of a control volume is the result of amount of heat entering or leaving the system and the work done by the surroundings on the system [15].

$$dE = dQ + dW \tag{A.5}$$

(Rate of change of energy inside control volume)= (Rate at which enthalpy, kinetic energy and potential energy is entering the control volume) - (Rate at which enthalpy, kinetic energy and potential energy is leaving the control volume)+ (Shaft work entering the control volume per unit time) + (Shear work done at the control surface per unit time by surrounding fluid on fluid inside the control volume)

The continuity, momentum, energy equations combined together are often referred as Navier-Stokes equations which is extremely useful in determining the temperature and velocity in a region of interest.

REFERENCES

[1] L. Barroso and U. Hölzle, "The datacenter as a computer: An introduction to the design of warehouse-scale machines," *Synthesis Lectures on Computer Architecture*, vol. 4, no. 1, pp. 1–108, 2009.

[2] T. ASHRAE, "9.9.(2011). 2011 thermal guidelines for data processing environments–expanded data center classes and usage guidance." [Online]. Available: http://tc99. ashraetcs. org/documents/ASHRAE% 20Whitepaper

[3] R. Tummala, *Fundamentals of Microsystems Packaging.* McGraw-Hill, 2001. [Online]. Available: http://books.google.com/books?id=_7S-QwAACAAJ

[4] T. ASHRAE, "9.9.(2012). 2012 ashrae whitepaper-it equipment thermal mgmt."

[5] FACEBOOK, "Intel motherboard." [Online]. Available: http://opencompute.org/wp/wp-content/uploads/2011/07/Server-Intel-Motherboard-Specifications-v1.pdf

[6] E. Ciliendo and T. Kunimasa, *Linux performance and tuning guidelines.* IBM, International Technical Support Organization, 2007.

[7] "Type-t thermocouple error."

[8] FACEBOOK, "Intel motherboard." [Online]. Available: http://opencompute.org/wp/wp-content/uploads/2011/07/Server-Chassis-Specifications.pdf

[9] D. Electronics, "Fan curves." [Online]. Available: http://www.delta.com.tw/product/cp/dcfans/download/pdf/QFR/QFR60x60x25.4mm.pdf

[10] "How cfd works."

[11] FACEBOOK, "Power supply." [Online]. Available: http://opencompute.org/wp/wp-content/uploads/2012/05/Open-Compute-Project-700W-450W-Power-Supply-v1.0.pdf

[12] "lookbusy – a synthetic load generator."

[13] "Variable transformer general information."

[14] S. Huck, "April 2011,revision 1.1, measuring processor power." [Online]. Available: http://www.intel.com/content/www/us/en/benchmarks/resources-xeon-measuring-processor-power-paper.html

[15] S. Sukhatme, *Textbook Of Heat Transfer*. Universities Press, 2006.

BIOGRAPHICAL STATEMENT

Shreyas Sampath was born in Mysore, Karnataka, India in 1987. He received his B.S. degree from Visvesvaraya Technological University, Belgaum, in 2009 and his M.S. degree from The University of Texas at Arlington in 2012, all in Mechanical Engineering. He was a part of National Science Foundation (NSF), Energy Smart Efficient System (ES2) from January, 2012 and had been working for the Facebook research team and actively been involved in many data center related research areas. He is a member of ASHRAE society and also been a Treasurer for the ASHRAE Student Chapter in UTA.