CFD SIMULATION AND DESIGN OPTIMIZATION TO IMPROVE COOLING

PERFORMANCE OF OPEN COMPUTE ‘BIG SUR’ SERVER

(Reprinted with ASME Permission) [13]

by

MANGESH MOHAN DHADVE

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

August 2017

Copyright © by Mangesh Mohan Dhadve 2017

All Rights Reserved

Acknowledgements

I would like to take this opportunity to thank my thesis advisor, Dr. Dereje Agonafer for his continuous support and guidance throughout my research. I would also thank him for providing me the opportunity to work on latest technology, attend the conferences and present our work in front of industry experts.

I thank Dr. Abdolhossain Haji-Sheikh and Dr. Andrey Beyle for taking their valuable time and being part of my thesis defense panel. I would also like to thank John Stuewe, OCP who provided me with crucial information to create CAD model of this server. I thank William Tsu, NVIDIA to provide right direction for information of GPU cards. I would like to thank engineers at Future Facilities Ltd for their support during this research.

I would also like to thank EMNSPC team members and fellow lab mates Jimil Shah, Chinmay Kshirsagar for their continuous feedback and support during my research.

I would also like to express my gratitude to Ms. Sally Thompson, Ms. Debi Barton and Ms. Flora Pinegar for their continuous support and guidance throughout my journey at UT Arlington.

Finally, I must thank my parents, family and friends for providing me continuous support and encouragement throughout my journey till this day. This accomplishment would not have been possible without their help and guidance.

May 1, 2017

Abstract

CFD SIMULATION AND DESIGN OPTIMIZATION TO IMPROVE COOLING PERFORMANCE OF OPEN COMPUTE 'BIG SUR' SERVER,

Mangesh Mohan Dhadve, MS

The University of Texas at Arlington, 2017

Supervising Professor: Dereje Agonafer

In recent years, there have been a phenomenal increase in Artificial Intelligence and Machine Learning that require data collection, mining and using data sets to teach computers certain things to learn, analyze image and speech recognition. Machine Learning tasks require a lot of computing power to carry out numerous calculations. Therefore, most servers are powered by Graphics Processing Units (GPUs) instead of traditional CPUs. GPUs provide more computational throughput per dollar spent than traditional CPUs. Open Compute Servers forum has introduced the state-of-the-art machine learning servers "Big Sur" recently. Big Sur unit consists of 4OU (OpenU) chassis housing eight NVIDIA Tesla M40 GPUs and two CPUs along with SSD storage and hot-swappable fans at the rear. Management of the airflow is a critical requirement in the implementation of air cooling for rack mount servers to ensure that all components, especially critical devices such as CPUs and GPUs, receive adequate flow as per requirement. In addition, component locations within the chassis play a vital role in the passage of airflow and affect the overall system resistance. In this paper, sizeable improvement in chassis ducting is targeted to counteract effects of air diffusion at the rear of air flow duct in "Big Sur" Open Compute machine learning server wherein GPUs are located directly downstream from CPUs. A CFD simulation of the detailed server

model is performed with the objective of understanding the effect of air flow bypass on GPU die temperatures and fan power consumption. The cumulative effect was studied by simulations to see improvements in fan power consumption by the server.

Table of Contents

Acknowledgements iii

Abstract iv

List of Illustrations viiiii

List of Tables x

Chapter 1 Introduction........1

1.1. Open Compute Project........12

1.2. Motivation........13

1.3. Big Sur........15

1.1.1. Specifications........15

1.1.2. GP-GPU........16

1.1.3. Modularity........17

1.4. Types of cooling systems used in data centers........17

Chapter 2 3D Modelling........19

2.1. CAD Modelling........19

2.2. FAN Modelling........ 25

Chapter 3 CFD Simulation........28

3.1. Mesh Sensitivity Analysis........28

3.2. Baseline Simulation........28

3.2.1. Boundary Conditions........29

3.3. Simulation Results........30

3.3.1. Observations........34

Chapter 4 Modifications……………………………………………………………….….35

4.1. Air Duct Optimization……………………………………………………………35

4.2. Modifying middle partition……………………………………………………39

Chapter 5 Results……………………………………………………………………….41

Chapter 6 Conclusion and Future Work………………………………………………47

6.1. Conclusion…………………………………………………………………….47

6.2. Future Work…………………………………………………………………….47

Appendix A List of Abbreviations 48

References 49

Biographical Information 51

List of Illustrations

Figure 1-1 Facebook’s Prineville Data center 12

Figure 1-2 Energy Consumption Breakdown in data center 14

Figure 1-3 Big Sur Server 15

Figure 1-4 GPU acceleration 16

Figure 2-1 Creating drawing using Canvas draw feature 20

Figure 2-2 3D model of the ‘Big Sur’ 20

Figure 2-3 CAD model provided by OCP 21

Figure 2-4 3D model built with reference to OCP model 22

Figure 2-5 Front part of chassis 23

Figure 2-6 Intel Xeon E5 processor CAD model 24

Figure 2-7 Simplified model created in 6SigmaET 24

Figure 2-8 Final model used in CFD Simulation 28

Figure 2-9 Theoretical Fan curve for fans in series and for single fan 27

Figure 3-1 Mesh Sensitivity Analysis 29

Figure 3-2 Temperature distribution across components 30

Figure 3-3 Streamline plot from top view 32

Figure 3-4 Streamline plot from side view 32

Figure 4-1 Current Configuration 36

Figure 4-2 Proposed Configuration 36

Figure 4-3 Existing Air Duct 36

Figure 4-4 Existing Air duct 37

Figure 4-5 Modified air duct 37

Figure 4-6 Modified air duct 38

Figure 4-7 Middle Partition 39

Figure 4-8 Modified middle partition....................40
Figure 5-1 Streamline plot for modified air duct....................41
Figure 5-2 Streamline plot for modified middle partition....................42
Figure 5-3 CPU and GPU temperatures....................43
Figure 5-4 Fan flow rates....................44
Figure 5- 5 System Impedance curve....................44
Figure 5-6 Fan power consumption....................45

List of Tables

Table 3-1 Temperature across different components…………………………………………..….33

Table 3-2 Fan flow rates……………………………………………………………………..…..34

Table 5-1 Results………………………………………………………………………...…45

Chapter 1

Introduction

Today, our lives are made simple by online technologies such as email, navigation, banking, social networking. We rely greatly on these online services throughout our day. These services are possible only via data center which obtain and process information. Though data centers are part of an IT industry, it has many systems within itself. Cooling system is one of the major component of data center. To operate at best possible efficiency, data center atmosphere must be maintained at certain humidity level and temperature.

With rapid advancements in computing resources, size of the data centers and their processing power is increasing and it will not reduce in near future. International Data Corporation (IDC) predicted total number of deployed data centers will reach 8.6 million by the end of the 2017. Data centers are massive and consumes tremendous amount of energy. During their operation, IT equipment leaves significant amount of heat depending on their processing power. This heat needs to be taken away from the IT equipment for their efficient working. If this heat accumulated, it can build up to point where it can damage or destroy IT equipment as well as data center infrastructure. Therefore, to ensure safe and efficient working and to improve life of the data centers adequate cooling techniques are required.

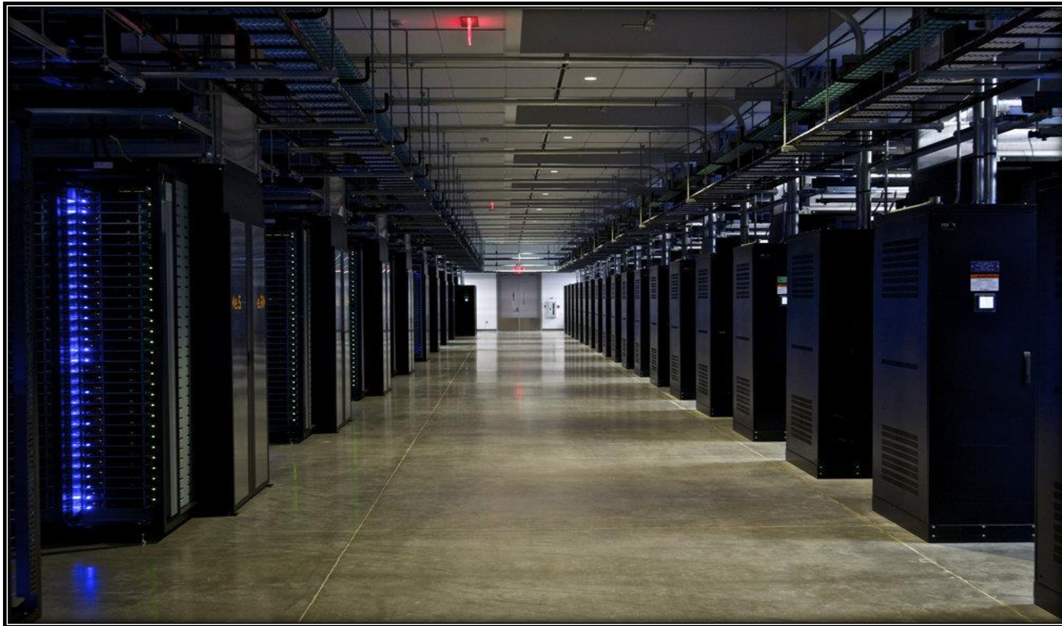

Figure 1-1 Facebook's Prineville Data center

1.1 Open Compute Project

Back in 2009, when Facebook started growing at immensely high rate and started to offer new features such as feed as per user's likes, photo sharing, they realized that their current hardware and facilities needs to be overhauled to support infrastructure and make data centers more efficient by controlling energy expenditure and other costs. In 2011, Facebook announced Open Compute Project initiative to share data center, server designs publicly. This initiative was later joined by various companies including Microsoft, Jupiter Networks, Goldman Sachs etc.

The redesigned data centers were 38% more energy efficient to build and 24% cost efficient when compared with previous generations of data centers. [1] In this project, members design and share efficient plans of data centers and its components. Being open source, anyone will be able to use these designs and modify these designs to improve efficiency. By openly sharing these ideas, it accelerates innovation with minimum complexity in data center components. This foundation enables its members to share their intellectual property. Aim of this foundation is to create efficient, flexible and modular hardware to reshape tech infrastructure.

These efforts helped Facebook to save $2 billion in infrastructure cost over the period of three years. [2] In 2014 alone, Facebook saved energy to power nearly 80,000 homes. [2]

Main projects of Open Compute Project are:

(i) Data Center- To maximize thermal and energy efficiency of data center by designing data centers in pair with Open Compute Servers. This covers five functional areas which are Power, Cooling, Layout and Design, Facility Monitoring and Control and Operations.

(ii) Storage- Being one of the key parts of any computing facility, this area offers numerous possibilities to make operations more efficient. OpenVault is the result of this foundation

(iii) Networking- This project is dedicated to make technologies modular and open, which helps in accelerating research. Scope of this project includes hardware and software development together.

(iv) Rack & Power- This project is dedicated to optimizing data center performance, infrastructure by utilizing OpenRack architecture

(v) Server- Key component of the data center. This project provides standard server designs for computing. Along with other OCP disciplines, server project makes sure wide adoption and optimization

(vi) High Performance Computing (HPC)- This project is aimed to develop open sourced heterogeneous computing and networking platform. Computing systems that are optimized for multi-node processors

1.2 Motivation

According to study published by datacenterknowledge.com [3], data centers in the United States alone consumed 70 billion kWh of electricity in 2014, which is roughly 2% of nation’s total energy consumption and Berkeley Lab estimates it to reach 73 billion kWh in 2020[4]. The data center industry survey by Uptime Institute shows global average of PUE is 1.7[5], that means data centers spent 0.7W extra energy per 1W of computing energy and most of this energy is used by

cooling systems. As per the Cisco's report [6], around 33% of energy is spent on cooling in data centers.

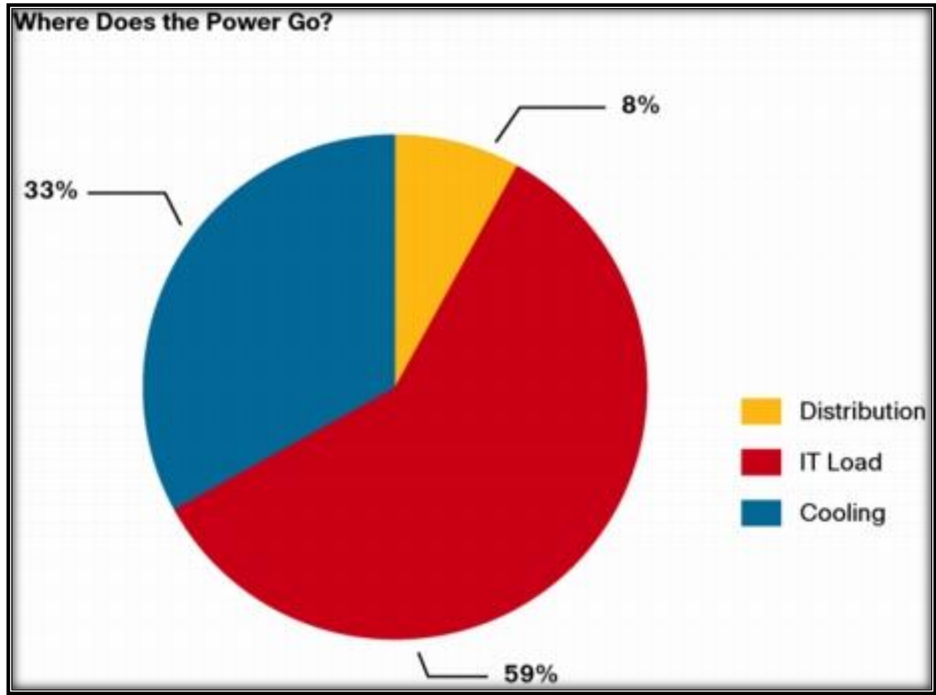

Figure 1-2 Energy Consumption Breakdown in data center

Due to rapid innovations, computing power of individual server is also on the rise. Previous generation Winterfell servers were design with maximum thermal design point (TDP) of 95 W then Yosemite server which was consuming total power of 400W. Above all these servers, there is Big Sur, although first of its GP-GPU OCP server with total designed power consumption of 2.5kW. Big Sur came in to existence as a need for better hardware to support advancements in Machine Learning and Artificial Intelligence. These tasks require tremendous computation horsepower. There is lot of innovation happening in the field of deep learning, therefore its application will require more advances in processing power as well [7]. As processing increasing, there will be a necessity of better heat dissipation. Now to remove heat from the electronic components, we can use number of different methods, depending on system requirement, atmosphere and amount of heat needs to dissipate.

Big Sur is the first OCP GPU server. This server is cooled by forced convection by four fans at the rear. Aim of this research was to improve cooling efficiency of this server without increasing system resistance significantly, therefore all the savings made will not go to mitigate increased system resistance.

1.3 Big Sur

Big Sur is Graphics Processing Unit (GPU) based server, used to run machine learning programs. Big Sur is designed to support eight full height gen3 PCI-e (Peripheral Component Interconnect Express) cards, two CPUs, 16 RDIMMs/LRDIMMs. Its form factor is 4OU.

Figure 1-3 Big Sur Server

1.3.1 *Specifications [8]*

- Processor- Intel® Xeon® E5-2600 v4 product (Broadwell Architecture) x 2
- Co-processor- PCI-e Gen3 full height cards x 8
- Memory- 2400 MHz DDR4 LRDIMM x 16
- Motherboard- Intel® C610
- Form Factor- 4OU
- Storage- 2.5" SSD/HDD
- Dimensions W*H*D(mm)- 533*189*800

- OpenRack v2 compati ble

1.3.2 GP-GPU

GP-GPU is acronym for General Purpose computing on Graphics Processing Unit. Big Sur relies on mostly on GPUs because machine learning tasks require more computing power. GPUs provide more computation power than CPUs. GPUs and CPUs work together to accelerate machine learning process. GPUs have more cores than traditional CPUs to perform operations parallel sequence. For example, traditional Intel Xeon E5 2694 v4 CPU have 14 cores while NVIDIA Tesla M40 have 3072 cores at its disposal. CPUs run most of sequential code while, GPUs handle compute-intensive part of the code. Figure 1-4 shows the exact working of GPU acceleration.

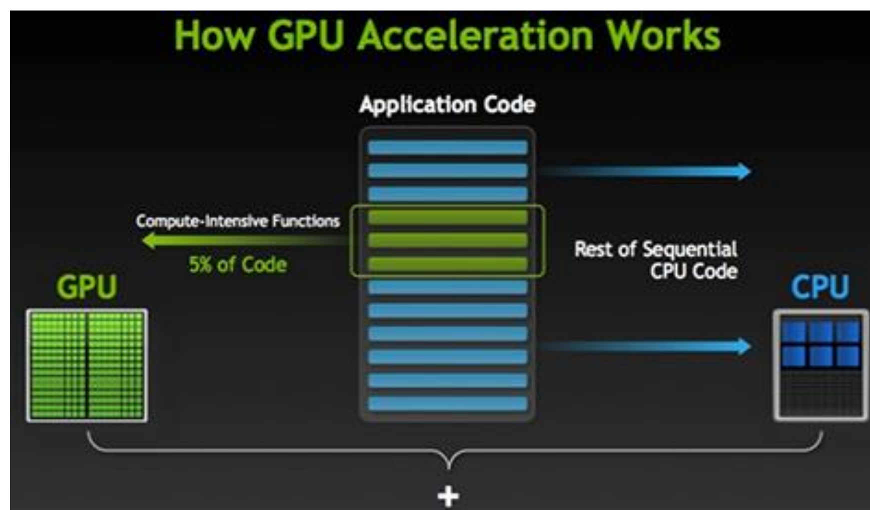

Figure 1-4 GPU acceleration

GPUs are now being used in various fields such as banking operations, self-driving vehicles, defense networks, drug discoveries and artificial intelligence. Big Sur can be modified to support these operations because of its high modularity.

1.3.3 *Modularity*

Various components of Big Sur such as Motherboard, linking board and storage pane are connected by connectors. Therefore, these components can be rewired and rearranged as per the user needs. Although it is designed around NVIDIA Tesla M40 cards, it can also support other full height Gen3 cards such as AMD FirePro, Intel Xeon Phi etc. This research focuses on server fitted with NVIDIA Tesla M40 cards. For this research, Intel Xeon E5 processor with TDP of 130 W is selected and NVIDIA Tesla cards with TDP of 250 W and passive cooling solution is used. This server is OpenRack V2 compatible. Topology of this server can be changed as per the use cases by rearranging CPUs and GPUs.

1.4 Types Cooling Systems used in Data Centers

ASHRAE (American Society of Heating, Refrigeration and Air-Conditioning Engineers) provides guidelines for operating temperature and humidity ranges for IT systems. For efficient operation and low energy consumption, IT equipment must be maintained within recommended limits. Various data centers use different cooling methods to cool their IT equipment. Choice of the cooling system depends on various factors such as geographic location, weather and load on the data center etc.

Various cooling systems used in data centers are:

1. Free Cooling
2. Air side economizers
3. Direct Evaporative Cooling
4. Indirect Evaporative Cooling

In its Prineville, OR data center, Facebook uses direct evaporative cooling system by utilizing outside air. Though servers can be cooled by liquid cooling methods Facebook utilizes air to cool server by forced convection.

In this study, individual server was considered with forced convection and atmospheric conditions were maintained as recommended.

Chapter 2

3D Modelling

Due to lack of hardware and the cost of the components, this study was limited to CFD simulation performed using commercially available CFD code. For this purpose, compact 3D model of the server was created using different CAD tools and then it was used to perform simulation.

2.1 CAD Modelling

When this research was commenced, I had little technical information on this server. We lacked geometrical data which was important for 3D modelling. Using the photographs of the server, drawing of this server was created through Autodesk Fusion 360. Known dimensions were used as benchmark to create datum then other dimensions were recorded by using Canvas draw feature. However, serious drawback of this feature is that it requires perfectly straight angled photograph to create accurate dimension and resolution of the photograph can affect the ability of the Canvas draw feature.

Figure 2-1 represent the Canvas draw feature and the photograph used to recreate the server drawing. As seen in the Figure 2-1, drawing was created using known dimensions as represented by blued lines. This drawing was then utilized to create 3D model of the server as shown in figure 2-2.

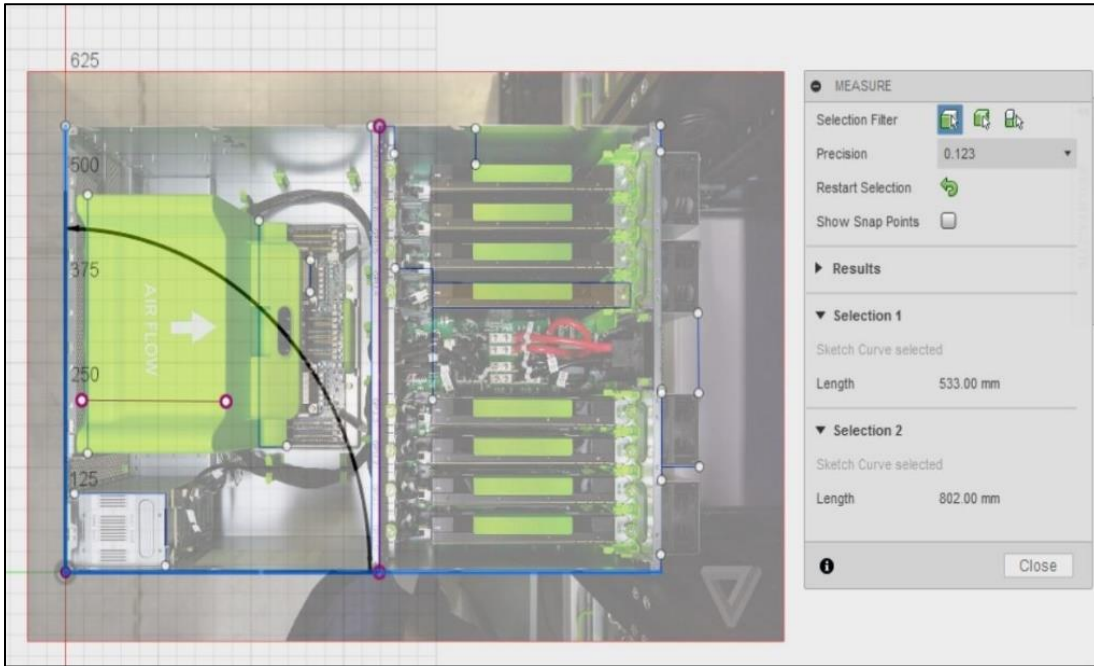

Figure 2-1 Creating drawing using Canvas draw feature

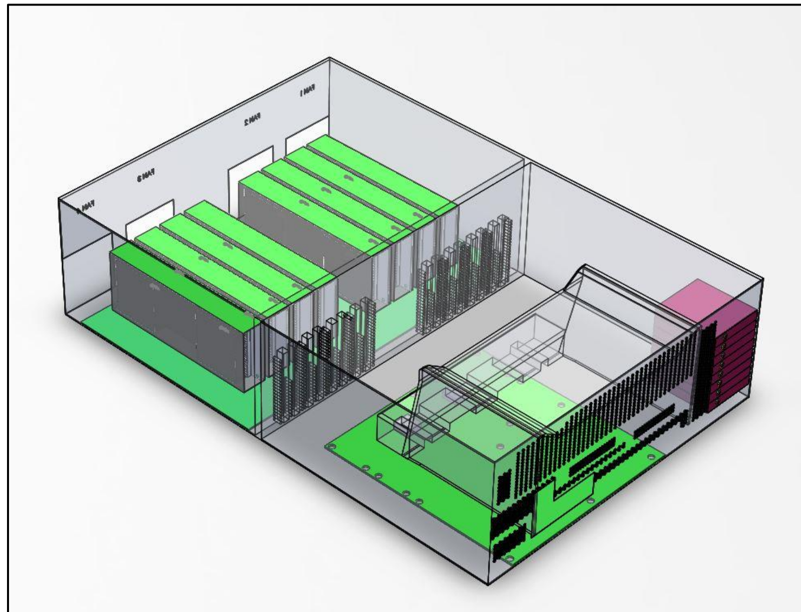

Figure 2-2 3D model of the ‘Big Sur’

After, this model was created. Open Compute Project also provided detailed 3D model of this server without 3rd party components. Upon comparison, it was detected that 3D model created was within ±2 mm accuracy.

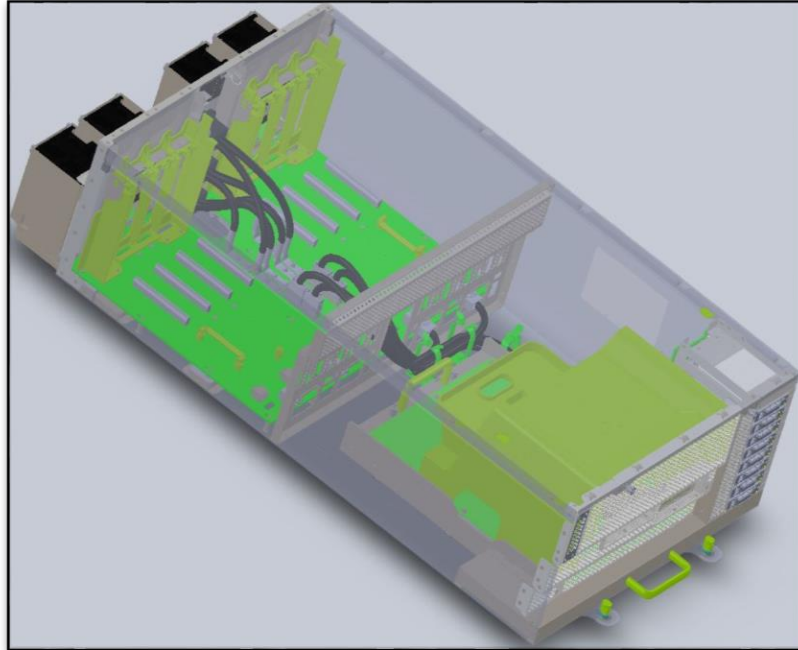

Figure 2-3 CAD model provided by OCP

Model provided by the OCP was detailed model representing each component. This increased model size and cannot be used as it is in the simulation. Using ANSYS SpaceClaim and SolidWorks, new 3D model was created using OCP model as reference. In this new model, motherboard, front part and middle partition was not included. Components which do not create significant flow resistance were also eliminated such as screws used to fix boards and panels, resistors, capacitors. This not only reduced the file size but also saved valuable computational time in the simulation process.

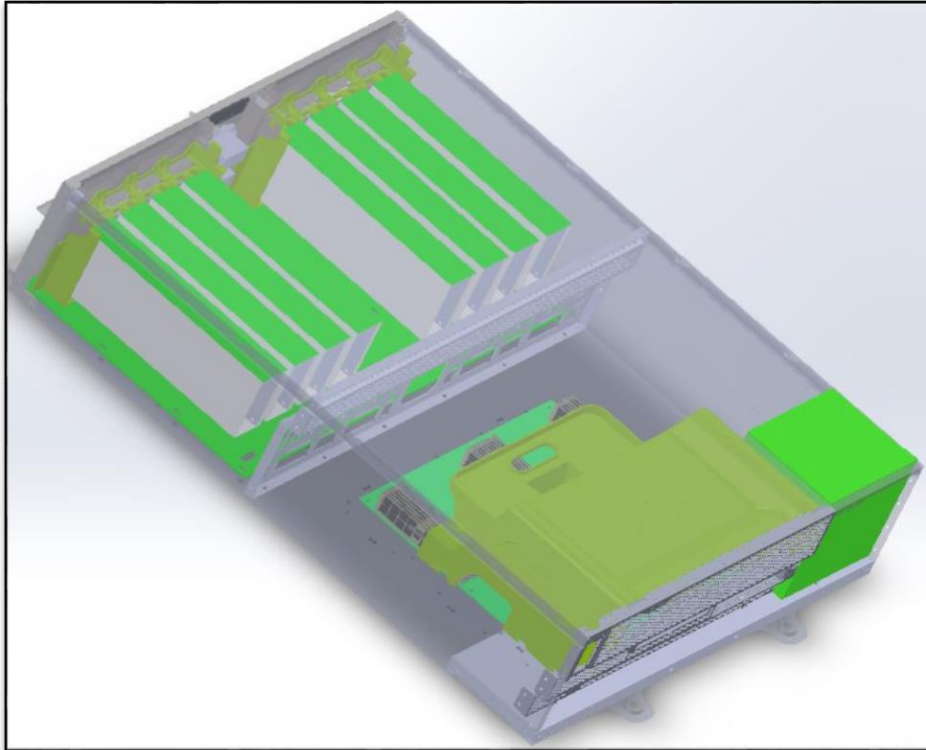

Figure 2-4 3D model built with reference to OCP model

Using CAD tools, simplified CAD model of NVIDIA Tesla card was created with reference to model obtained from GRABCAD. However, the processor and heatsink were modelled directly using 6SigmaET.

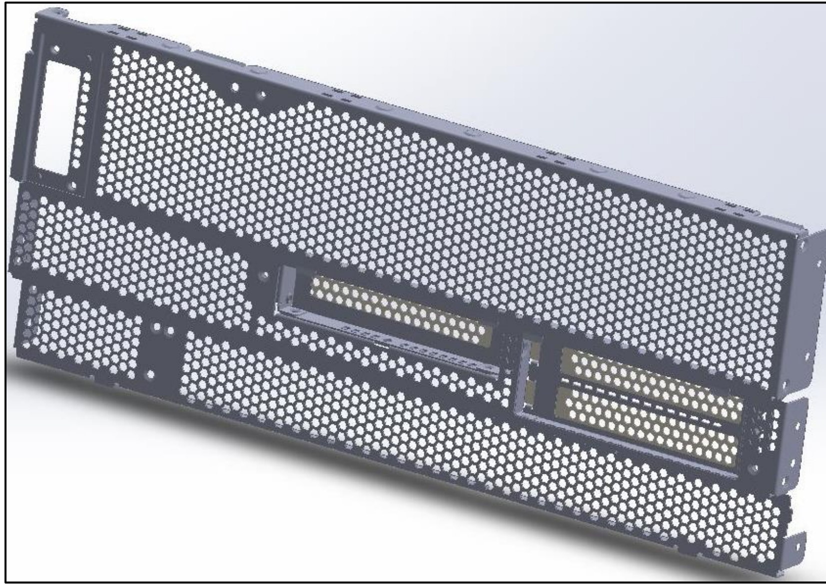

Figure 2-5 Front part of chassis

Figure 2-5 shows the actual front part of chassis. This model contained every detail which increased its file size. To simplify this part, its surface area was compared against the surface area of the solid plate of exact same dimensions created using CAD tool. Then in 6SigmaET front part modelled as perforated obstruction with 80% area opening and hole diameter size of 3mm. This simplified calculation.

Processor, processor heatsink, GPU chip and its heatsink and fans were directly modeled in 6SigmaET using the available reference models. Figure 2-6 and Figure 2-7 shows the comparison between processor models. Figure 2-8 shows the final model used in simulation.

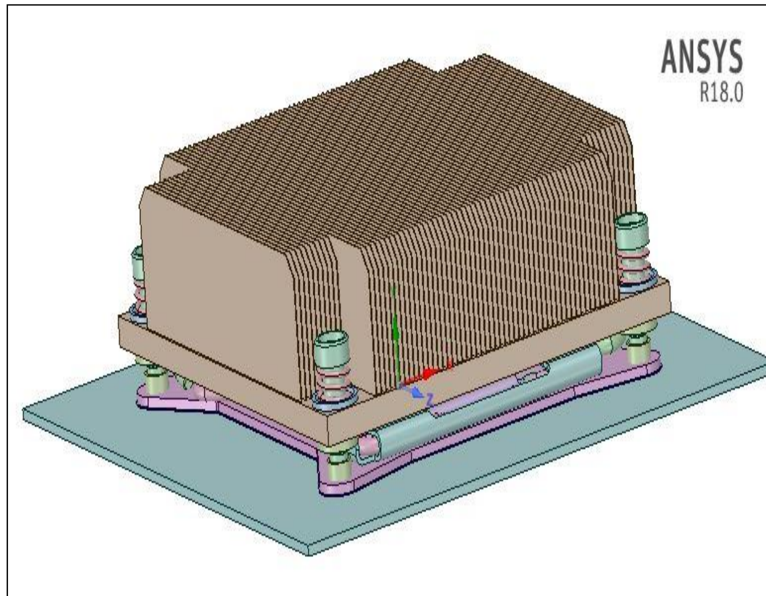

Figure 2-6 Intel Xeon E5 processor CAD model

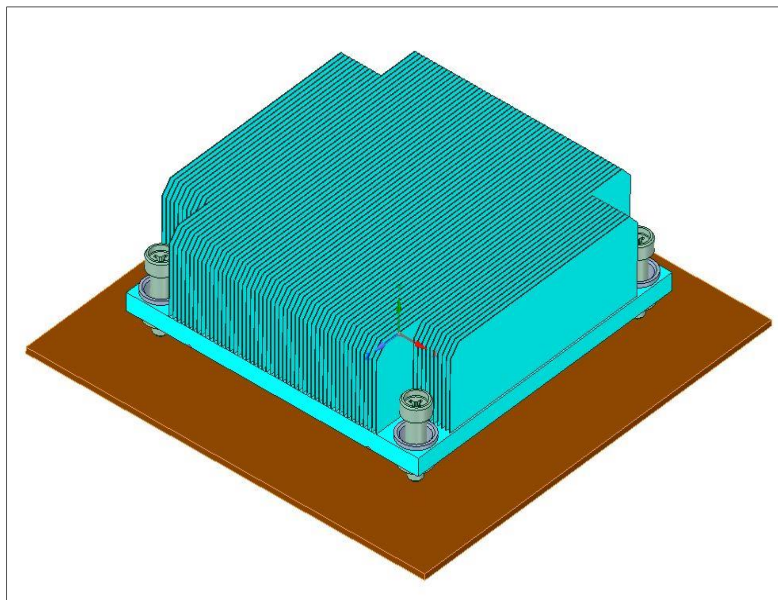

Figure 2-7 Simplified model created in 6SigmaET

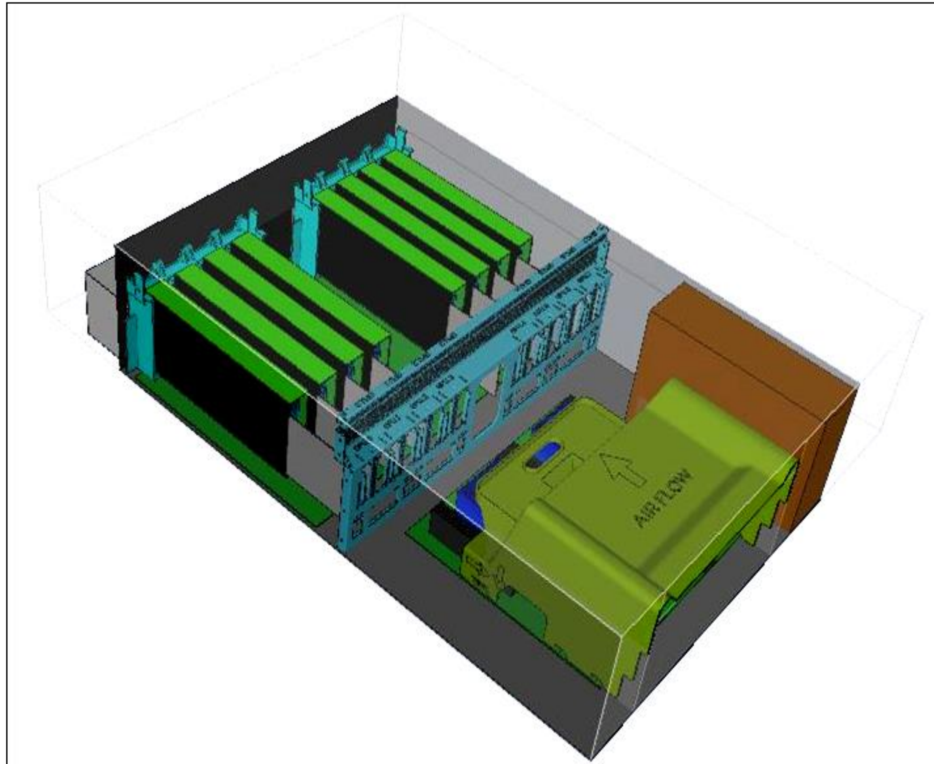

Figure 2-8 Final model used in CFD Simulation

2.2 FAN modelling

To model fans at the rear end of the server, enough data was not available. Therefore, the aiflow required was calculated using the empirical formula provided by Innovative Research Inc. [9]

Cooling Airflow rate in CFM= 154 × Heat load in kW × (Reference Pressure/Local Pressure)

In this study, local pressure was assumed (at sea level) as 1 atm same as the reference pressure.

Using this data Delta Electronics Fan AFB0912UHE-A was selected with flow rate of 160.22 CFM. This fan has following specifications [10]

- Model- Delta Electronics- AFB0912UHE-A (Single Rotor)
- Dimensions(mm)- 92*92*38

- 6000RPM± 10%
- Rated Voltage 12VDC
- Input Current (max)- 3A
- Input power (max)- 36 W
- Max air flow (at zero static pressure)- 4.537 m^3/min
- Max static pressure (at zero air flow)- 1.002 inch.H_2O

Big Sur uses four fan modules of dual fan rotor which essentially means two fans in series. When these fans are modeled separately giving fan curve, the solver struggles to find an operating point on fan curve for each fan. This happens because of constantly changing conditions at the cells between two fans. In this study, I tried increasing the distance between two fans, however the result was the same. Therefore, two fans in series were modeled as single. As two fans in series, their fan curves can be combined to model them as single fan. Theoretically, this should give us the double pressure difference, although in practice it is slightly less than double due to angular component of air flow at the exhaust of the fan. This component can be minimized by providing guide vane to redirect angular component back in the air stream. [11] Figure 2-9 shows the theoretical fan curve for two fans in series and for single fan.

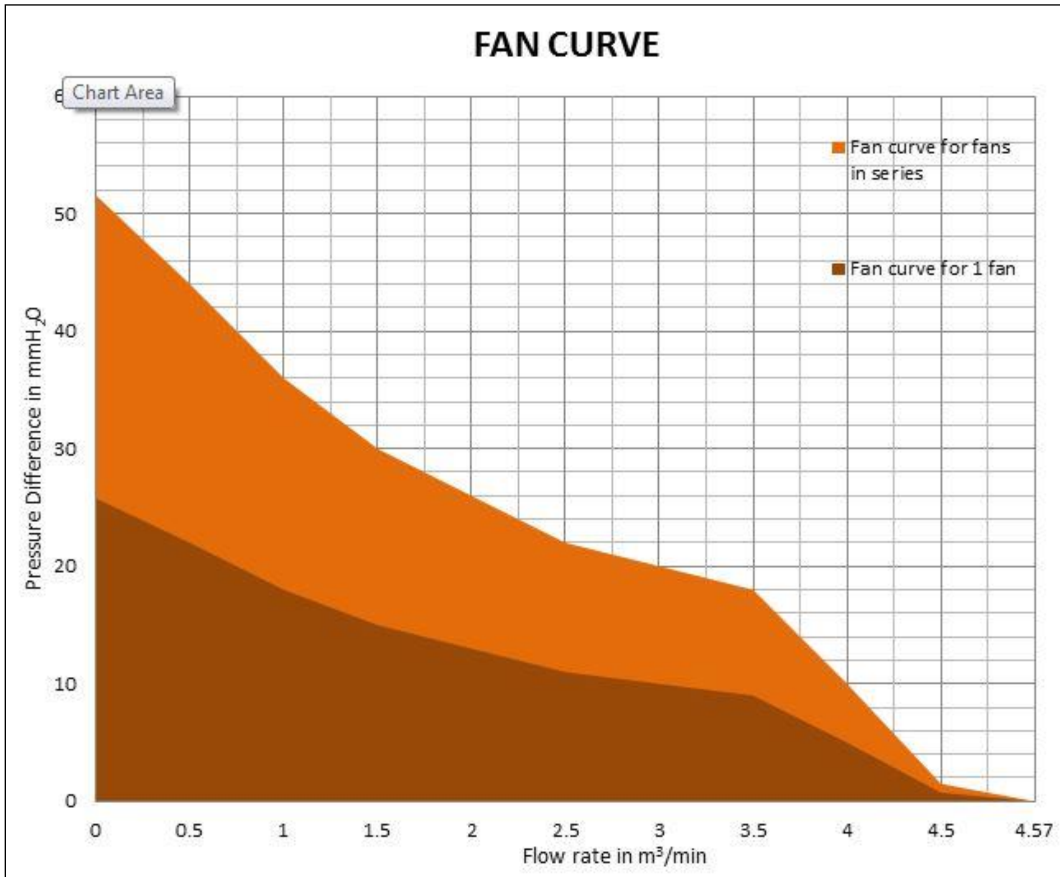

Figure 2-9 Theoretical Fan curve for fans in series and for single fan

Chapter 3

CFD Simulation

Due to lack of experimental data available on this server. A CFD model is created using commercially available CFD code, 6SigmaET. Simulation is performed on this model to establish baseline parameters such as die temperature, air velocity, pressure difference etc. In existing server, air is flowing over CPUs and DIMMs through air duct then this air carries heat from GPUs. Because of this some thermal shadowing occurs.

3.1 Mesh Sensitivity Analysis

To check the accuracy of the CFD model, mesh sensitivity analysis is performed. The temperature across all GPUs are considered for this analysis. The same simulation was run just by changing the number of cells (i.e. increasing mesh count) and temperature across all GPUs were noted. The rest of the boundary conditions are maintained same during all simulations. Figure 3-1 shows the GPU temperature versus cell count. Temperature of GPUs were constantly fluctuating even after crossing the usual cell count that we usually see. At about 82 million cell count, temperature across all the GPUs remain constant after increasing cell count.

3.2 Baseline Simulation

Baseline simulation is used establish benchmark due to lack of experimental data. This data was then used to compare against data obtained from modified server. Various material properties are assigned to different components as per existing server.

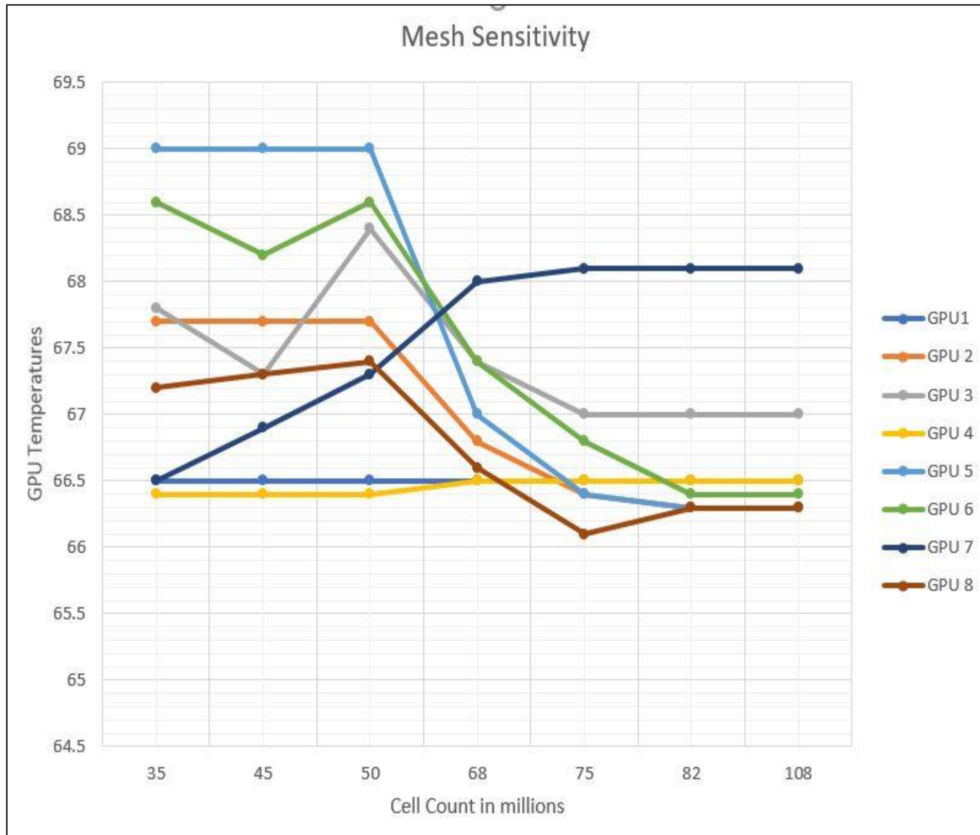

Figure 3-1 Mesh Sensitivity Analysis

3.2.1 Boundary Conditions

For this simulation, inlet air temperature was maintained at 20^{0}C as recommended by ASHRAE guidelines. Fan speed was fixed at 6000RPM due to lack of data in experimental fan controller.

Thermal Design Power (TDP) is assigned to each component. GPUs have TDP of 250W. CPUs are assigned TDP of 120W while each DIMM is assigned TDP of 3W. Each component is assumed to be working at its full capacity, this means they will utilize their assigned TDP value.

Standard k-ε turbulent model was used recommended by Future Facilities Inc. Standard k-ε model is modified version of exact k-ε model for practical approach. Standard k-ε model is used to determine to kinetic energy and dissipation. Equation for standard k-ε model is

Rate of change of k or ε + Transport of k or ε by convection = Transport of k or ε by diffusion + Rate of production of k or ε - Rate of destruction of k or ε.

To calculate kinetic energy, k

$$\frac{\partial(\rho k)}{\partial t} + \frac{\partial(\rho k u_i)}{\partial x_i} = \frac{\partial}{\partial x_j}\left[\frac{\mu_t}{\sigma_k}\frac{\partial k}{\partial x_j}\right] + 2\mu_t E_{ij} E_{ij} - \rho\epsilon$$

To calculate dissipation, ϵ

$$\frac{\partial(\rho\epsilon)}{\partial t} + \frac{\partial(\rho\epsilon u_i)}{\partial x_i} = \frac{\partial}{\partial x_j}\left[\frac{\mu_t}{\sigma_\epsilon}\frac{\partial \epsilon}{\partial x_j}\right] + C_{1\epsilon}\frac{\epsilon}{k} 2\mu_t E_{ij} E_{ij} - C_{2\epsilon}\rho\frac{\epsilon^2}{k}$$

Where,

u_i – velocity component in corresponding direction

E_{ij} –Component of rate of deformation

μ_t –Eddy Viscosity = $\rho C_\mu \frac{k^2}{\epsilon}$

$\sigma_k, \sigma_\epsilon, C_{1\epsilon} C_{2\epsilon}$ are constants

3.3 Simulation Results

Figure 3-2 shows temperature distribution across various components. In this view top cover of the server, cover of GPU cards are hidden while air duct is transparent.

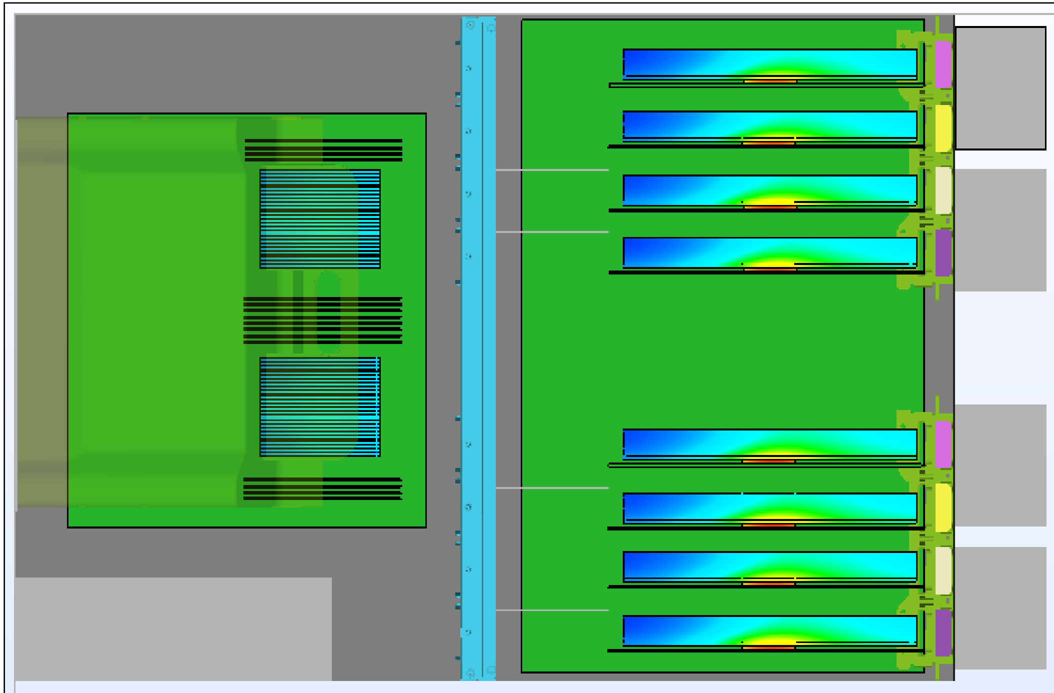

Figure 3-2 Temperature distribution across components

Figure 3-3 and figure 3-4 shows streamline plot depicting air flow across the server from top and side view. As seen in these two figures, middle partition creates significant resistance to the airflow. Also, much of the air flows through space between GPUs and above the GPU cards. In this simulation, maximum temperature of 68^0C was observed at GPU 7.

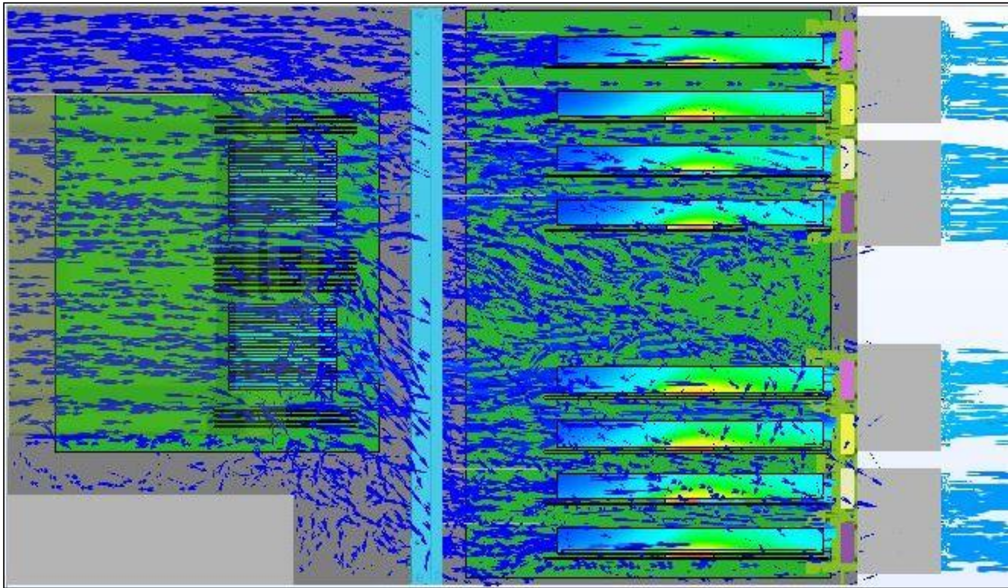

Figure 3-3 Streamline plot from top view

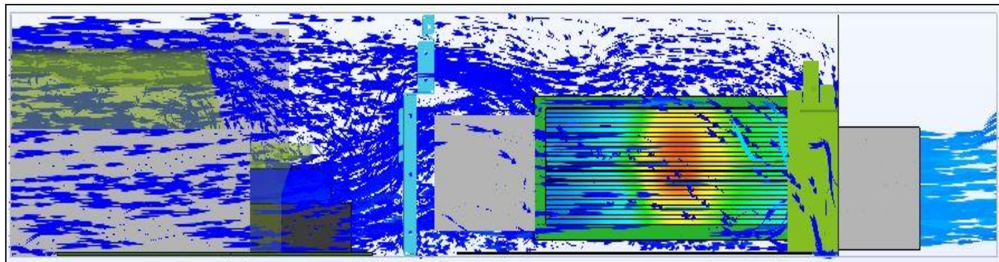

Figure 3-4 Streamline plot from side view

Table 3-1 and table 3-2 is used to discuss the results of baseline simulation.

Following are the results of baseline simulation:

Component	**Temperature (^{0}C)**
CPU_1	39.2
CPU_2	39.1
GPU_1	66.5
GPU_2	66.3
GPU_3	67.0
GPU_4	66.5
GPU_5	66.3
GPU_6	66.4
GPU_7	68.1
GPU_8	66.3

Table 3-1 Temperature across different components

FAN	FLOWRATE (m^3/s)
FAN_1	0.055
FAN_2	0.055
FAN_3	0.055
FAN_4	0.056

Table 3-2 Fan flow rates

3.3.1 Observations

As seen in figure 3-2 and 3-3, significant air is flowing through the space between in GPU cards and from top of the GPU cards. This air can be redirected to improve heat transfer. Furthermore, this air flows through the air duct which is then used to cool the GPUs. This creates thermal shadowing effect as heated air from the CPUs and DIMMs is used to cool GPUs. The goal of this study is to minimize the thermal shadowing as well as air flow resistance created by middle partition.

Chapter 4

Modifications

As seen in previous section, air duct and middle partition can be modified to optimize the air flow. Allowing adequate air flow to all critical components of the server is necessary and it needs to be kept in mind while making any changes in the system. Excessive bypass air will increase the temperature of CPUs and DIMMs which is undesirable. Optimum air flow must be maintained to the CPUs and DIMMs since they generate significant heat energy during their operation. Air flow resistance is also important factor that needs to be considered while making any changes. Since, increase in system will directly result in increase in fan power consumption and any savings made by those changes will be used partly to overcome increased system resistance. This will undermine the overall effect of changes made in the system.

4.1 Air duct optimization

Incoming air flows through the air duct, which is designed to give venturi effect to the air flow. Air flowing through the air duct is passed over the CPUs and DIMMs to carry heat. This air is then used to cool the GPUs. Now, as the air carries heat from CPU, it temperature is increased, when this air is used to cool the GPUs, it reduces its effective convective heat transfer due to reduced temperature difference between the air and the heatsink.

The air duct is optimized in a such way that it bypasses part of the air from CPUs. Part of the air is used to cool CPUs and GPUs is mixed with this cool air and then it is further used for the heat transfer from GPUs.

Figure 4-1 shows current configuration of the server where thermal shadowing effect occurs, figure 4-2 shows proposed configuration of the server with minimized thermal shadowing effect.

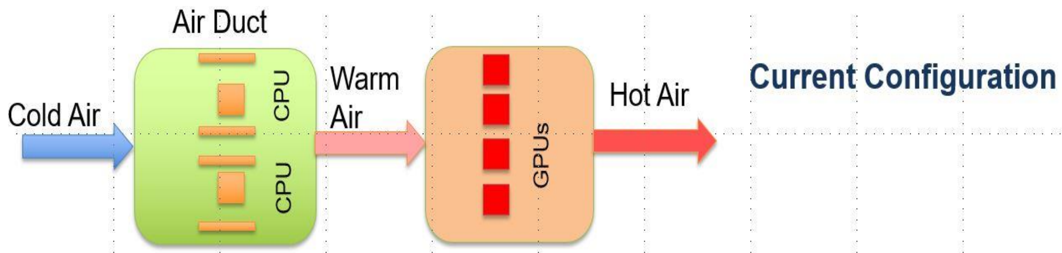

Figure 4-1 Current Configuration

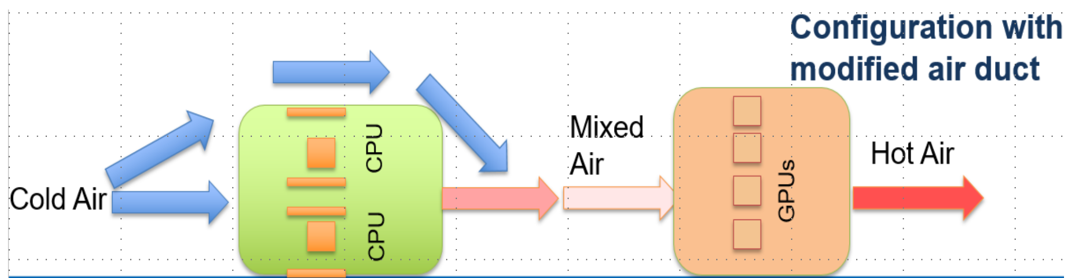

Figure 4-2 Proposed Configuration

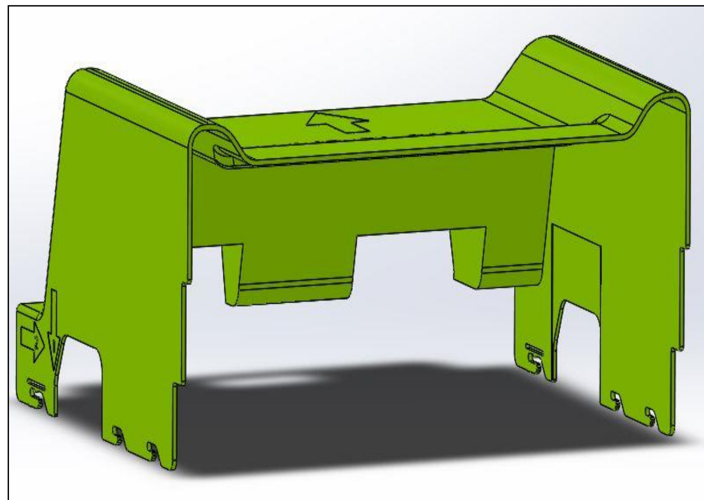

Figure 4-3 Existing Air Duct

Figure 4-3 and 4-4 shows the existing air duct with two clearly visible cavities to accommodate CPU heatsinks. This air duct has provision for connecting cable for motherboard. This airduct has reducing cross sectional area at the rear which creates venturi effect by acting

as restricted opening. This air duct is redesigned to allow bypass airflow. Figure 4-5 and 4-6 shows the modified air duct.

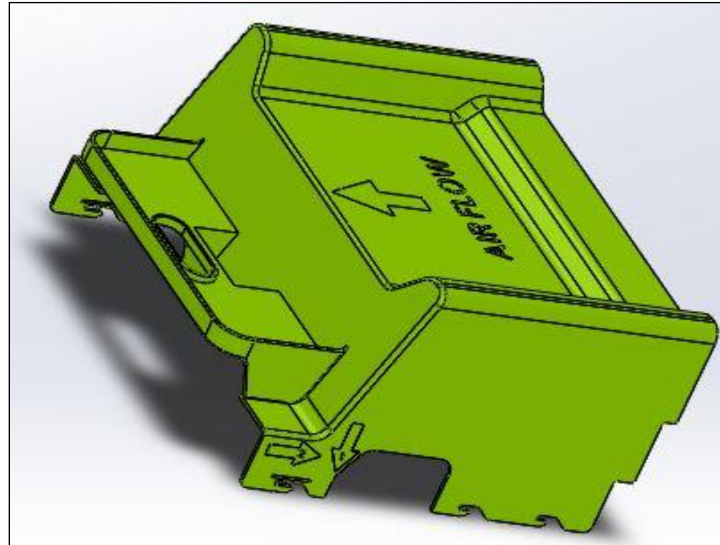

Figure 4-4 Existing Air duct

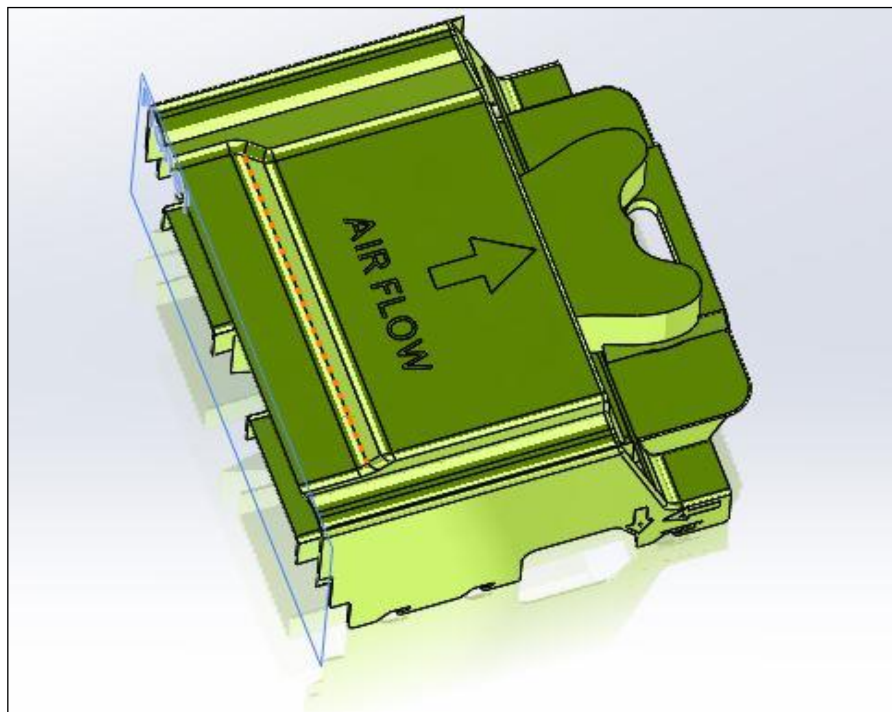

Figure 4-5 Modified air duct

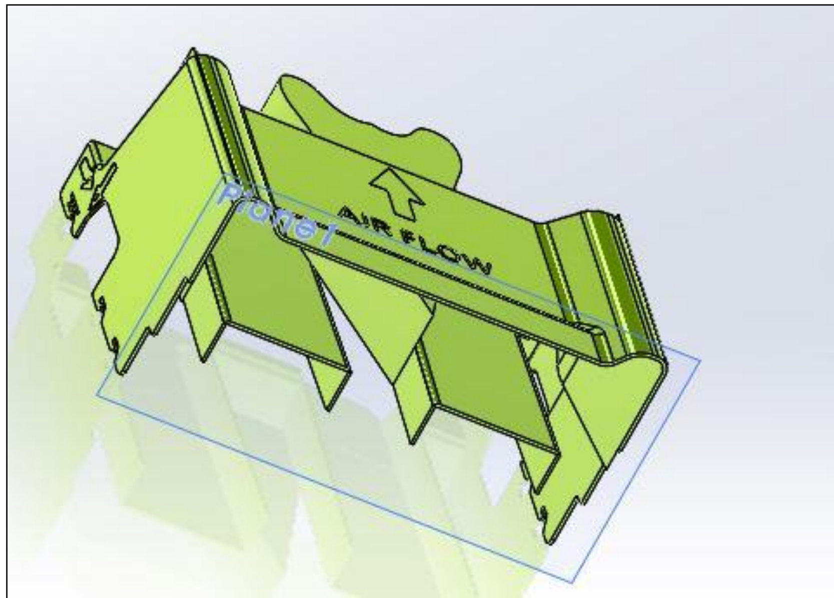

Figure 4-6 Modified air duct

As shown above, air duct is modified to bypass some CPU air. However, while doing that it is also required to maintain adequate airflow to the CPUs. Therefore, separate passage with wide opening is created for each CPU. This passage fits exactly on the CPU heatsink. Air duct walls are designed to have reducing cross sectional area to create venture effect. Furthermore, bypass air is directed towards the GPU cards by creating middle partition within air duct. This partition redirects air towards GPUs while acting as reduced area of opening. It is designed to minimize vortex formation at the trailing edge.

4.2 Modifying middle partition

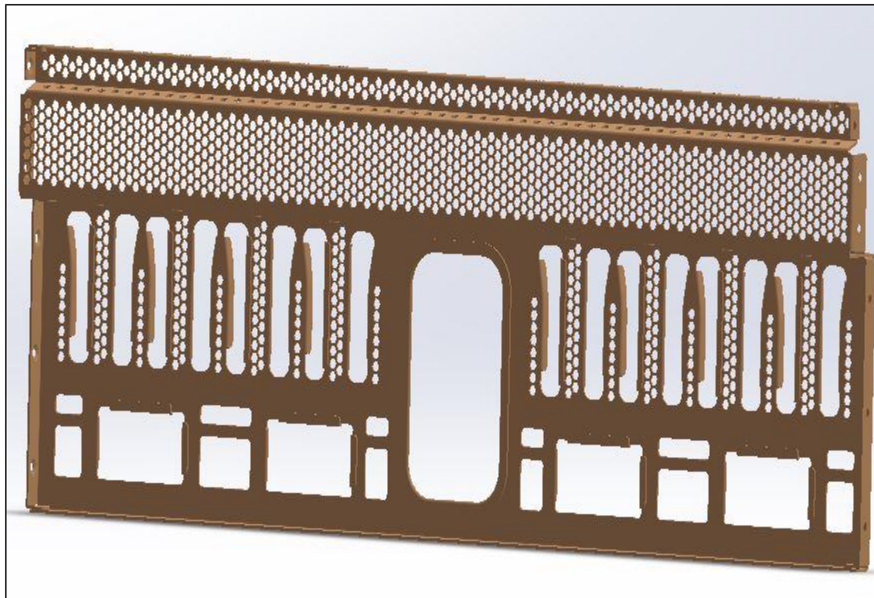

Figure 4-7 Middle Partition

Above figure shows middle partition used in this server. This partition has provision for connectors and GPU attachments. Due to its construction, this partition creates significant resistance for the air flow. This partition is modified to reduce the air flow resistance and directs air over the GPUs smoothly. Modified partition also have provision for connectors and to allow air flow over the FPGA board to cool smaller components. Figure 4-8 shows the modified middle partition. This partition eliminated the need for GPU attachments as GPU cards can directly rest over the slots provided.

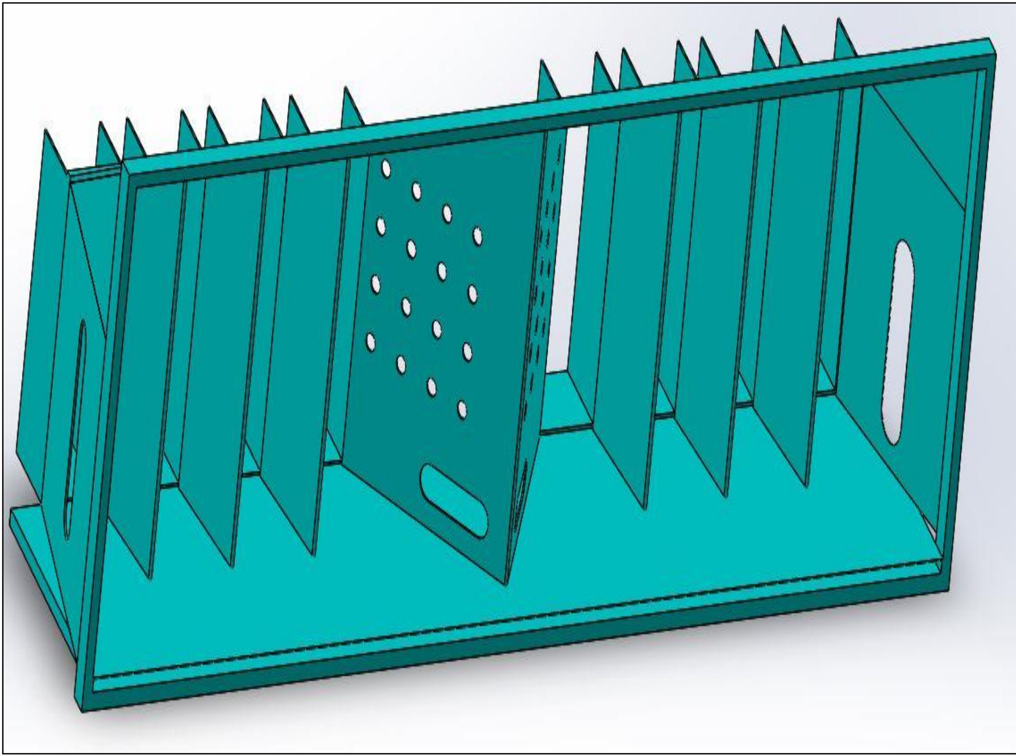

Figure 4-8 Modified middle partition

Chapter 5

Results

After careful study of air flow patterns, above mentioned changes were made to the server to reduce thermal shadowing effect and system resistance. CFD simulation was carried out on modified designs to observe the changes in performance of the server. The boundary conditions were maintained same as used in baseline simulation. Separate simulations were performed for modified air duct and modified middle partition cases. The reduction in temperatures of components were observed, although in case of modified air duct it was not significant.

Figure 5-1 shows the streamline plot of the case in which air duct is modified.

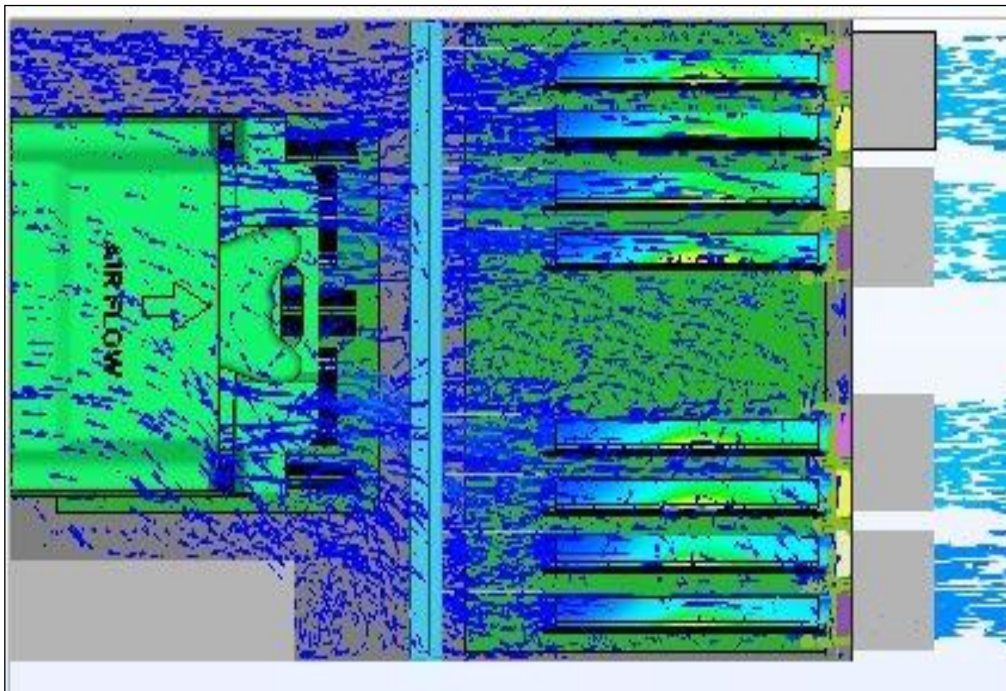

Figure 5-1 Streamline plot for modified air duct

It can be seen that air flowing through space between the GPUs is much less when compared with baseline simulation in figure 3-3. This is due to the partition created in air duct which gives directional momentum to the air.

Figure 5-2 shows the streamline plot for the case in which middle partition is modified.

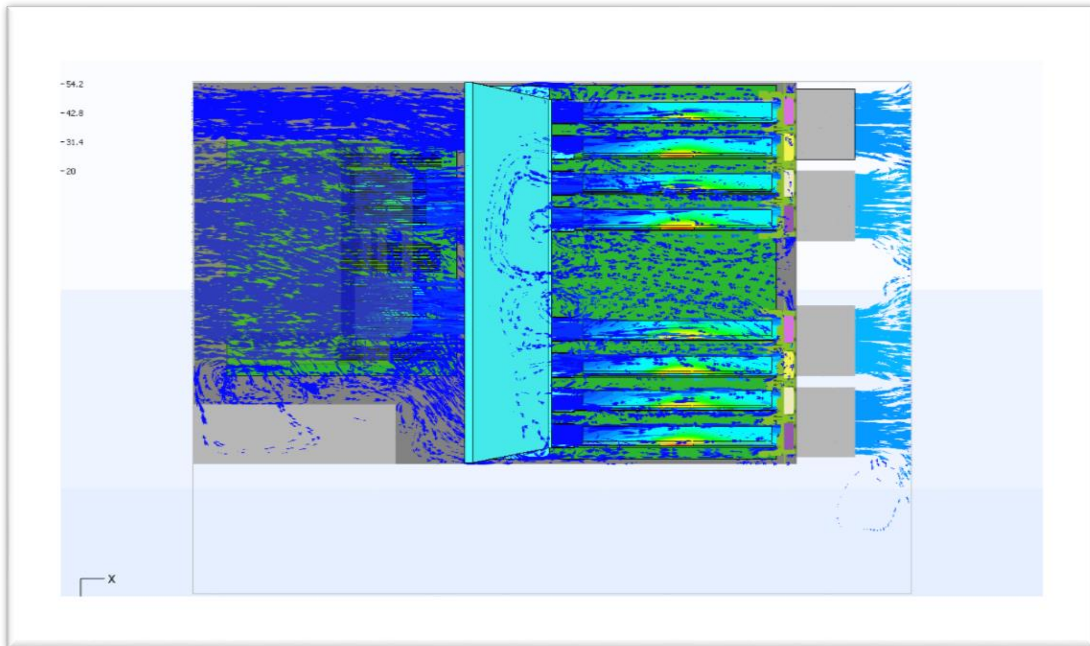

Figure 5-2 Streamline plot for modified middle partition

As seen in above figure, significant amount of air flows through the GPU cards. It is visible as GPU cover are hidden in this figure. Air flowing through space between the GPUs is much less compared to baseline and modified air duct case.

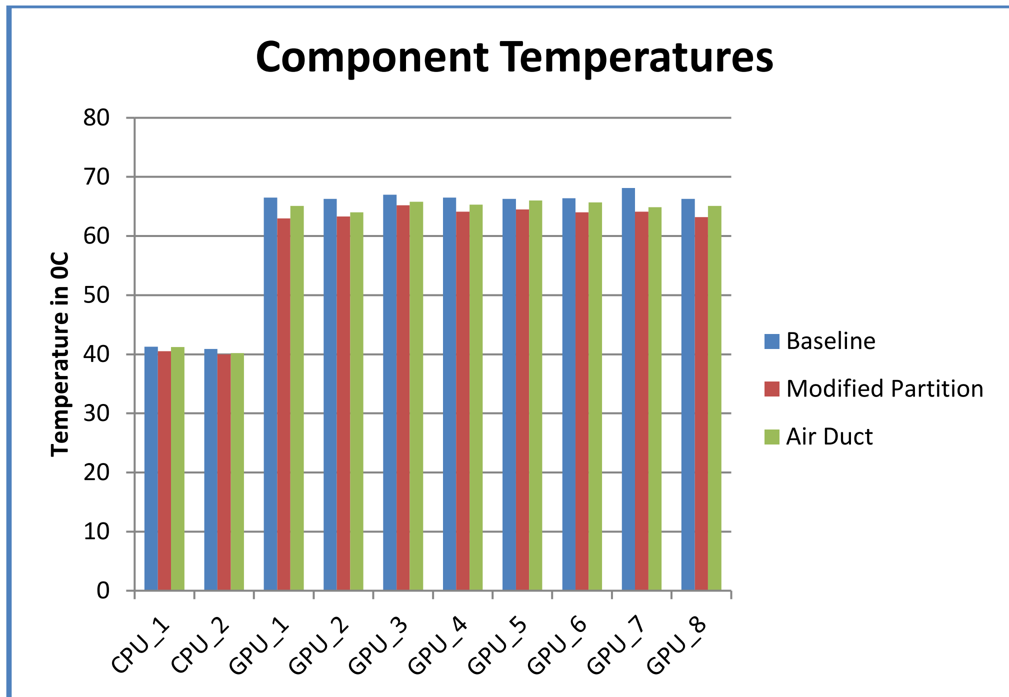

Figure 5-3 CPU and GPU temperatures

Figure 5-3 shows the comparison between the components temperatures in baseline simulation, modified partition and modified air duct cases. As mentioned earlier, temperature drop is observed in both cases. The reduction in temperature is not significant in case of modified air duct. Maximum temperature is 65.7^{0}C in case of modified air duct case against 68.1^{0}C in baseline simulation while it is 64.3^{0}C in modified partition case.

Changes in fan flow rates were also observed and noted in both cases. In baseline simulation, maximum fan flow rate was observed as 0.56 m^3/s while it increased in both case to 0.57 m^3/s for modified air duct and 0.59 m^3/s for modified middle partition. Figure 5-4 shows comparison between air flow rates of all fans for all the cases.

Fan power consumption is calculated by using formula: [12]

$$P = \Delta p \times q$$

Where, P= ideal fan power consumption, W

Δp= Pressure difference, N/m^2

q= fan flow rate, m^3/s

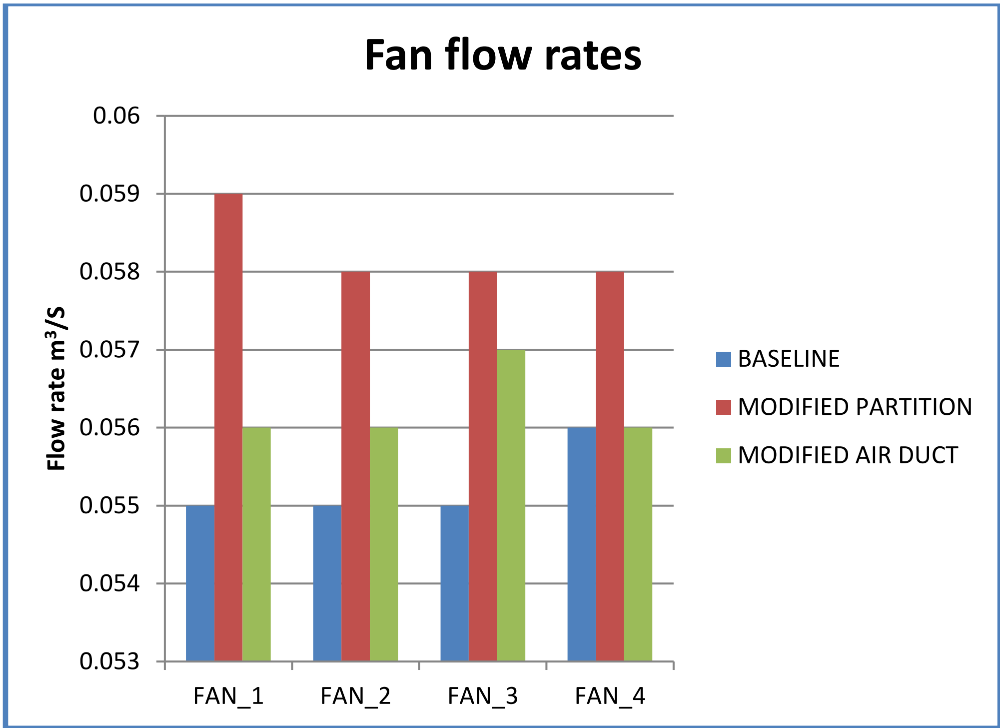

Figure 5-4 Fan flow rates

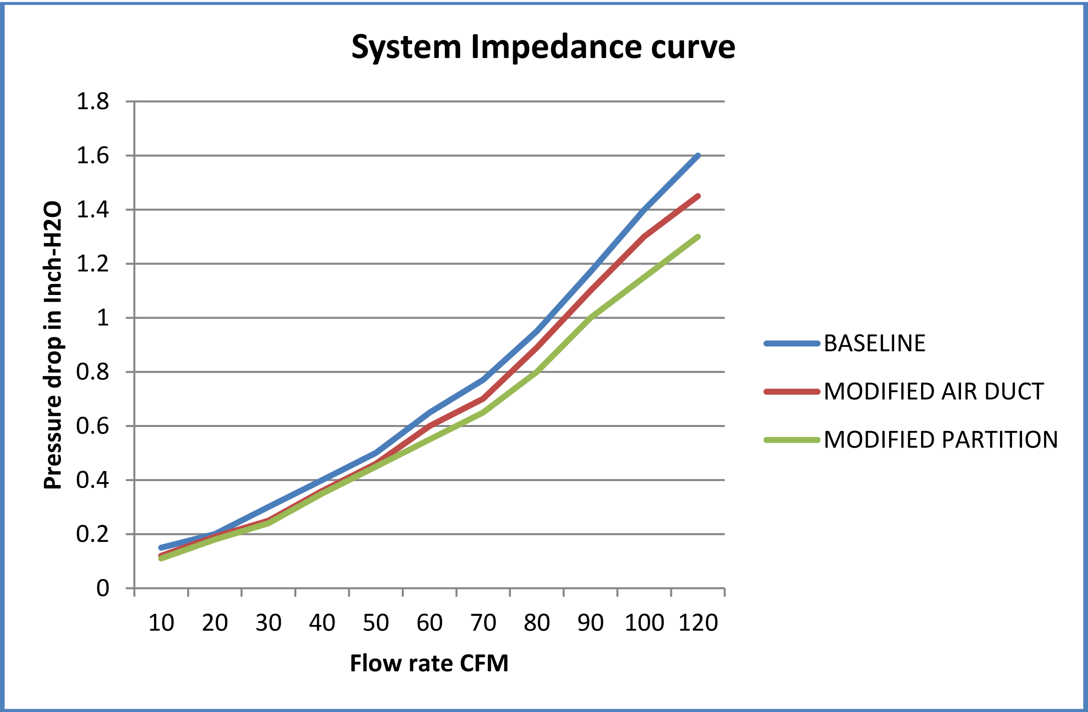

Figure 5- 5 System Impedance curve

Figure 5-5 shows the system impedance curve for three cases. Significant reduction in system resistance is observed when modified middle partition is used.

Using the fan power formula, power consumed by each fan is compared for all three cases. Figure 5-6 shows the comparison for fan power consumption in each case. The reduction in fan power consumption is due to the reduced system resistance as seen in figure 5-5.

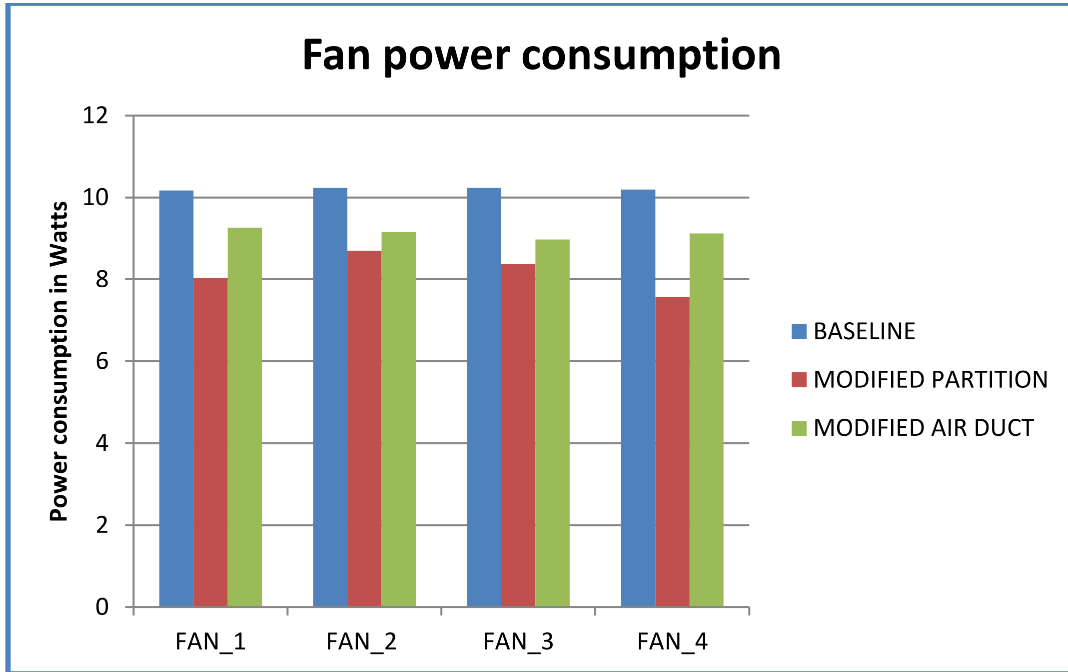

Figure 5-6 Fan power consumption

	Maximum Temperature (^{0}C)	**Maximum flow rate (m^3/S)**	**Minimum flow rate (m^3/S)**	**Total fan power consumption (Watts)**	**Fan Power Savings**
Baseline	68.1	0.056	0.055	40.82	
Modified Air Duct	65.7	0.057	0.056	36.5	10.98%
Modified Partition	64.3	0.059	0.058	30.05	25.14%

Table 5-1 Results

Table 5-1 compares the results of all three cases. The changes made in the system not only reduced the temperatures of the systems while it also reduced the fan power consumption by reducing system resistance. Total savings in fan power are 10.98% for modified air duct while 25.14% for modified middle partition.

Chapter 6

Conclusion and Future Work

In this study, cooling performance of 'Big Sur' server was evaluated by simulation in 6SigmaET. By carefully studying data obtained from baseline simulation, changes were made to the system to improve cooling performance of this server.

6.1 Conclusion

The fan power savings achieved through modifications are result of reduced system resistance. By changing design of middle partition significant savings can be made in fan power savings.

6.2 Future Work

1. Experimental testing- This study is limited to CFD simulation, in future experimental study can be performed to verify data obtained through CFD analysis
2. Study with different PCIe cards- Big Sur can be fitted with differenct PCIe cards. Study can be performed on PCIe cards with active cooling solution.
3. Liquid cooling- liquid cooling is one of the emerging technologies in this field. A CFD study can be performed using liquid cooling for cooling media instead of air. This will enable using more powerful CPUs and GPUs.

Appendix A

List of Abbreviations

GPU- Graphics Processing Unit

CPU- Central Processing Unit

OU- Open U

SSD- Solid State Drive

HDD- Hard Disk Drive

CFD- Computation Fluid Dynamics

OCP- Open Compute Project

IT- Information Technology

ASHRAE- American Society of Heating, Refrigeration and Air-conditioning Engineers

HPC- High Performance Computing

PUE- Power Usage Effectiveness

TDP- Thermal Design Power

GP-GPU- General Purpose Computation on Graphics Processing Unit

PCIe- Peripheral Component Interconnect express

RDIMM- Registered Dual Inline Memory Module

LRDIMM- Load Reduction Dual Inline Memory Module

DDR- Double Data Rate

CAD- Computer Aided Design

References

[1] About OCP, http://www.opencompute.org/about/

[2] Open Compute Project U.S. Summit 2015 – Facebook News Recap, https://code.facebook.com/posts/1538145769783718/open-compute-project-u-s-summit-2015-facebook-news-recap/

[3] Here's How Much Energy All US Data Centers Consume, http://www.datacenterknowledge.com/archives/2016/06/27/heres-how-much-energy-all-us-data-centers-consume/

[4] United States Data Center Energy Usage Report, https://eta.lbl.gov/publications/united-states-data-center-energy

[5] 2014 Data Center Industry Survey, https://journal.uptimeinstitute.com/2014-data-center-industry-survey/

[6] Power Management in the Cisco Unified Computing System: An Integrated Approach, http://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/unified-computing/white_paper_c11-627731.html

[7] Deep Learning with massive amounts of computational power, machines can now recognize objects and translate speech in real time. Artificial intelligence is finally getting smart, https://www.technologyreview.com/s/513696/deep-learning/

[8] Quanta "Big Sur" Contribution to Open Compute Project, http://files.opencompute.org/oc/public.php?service=files&t=c89644b25c2739107fee1f08a67ed5b3

[9] Calculation of Airflow Required for a Given Heat Load, http://tileflow.com/assets/files/tileflow/TF06-Heatload_cfm.pdf

[10] Delta Electronics Fan Specifications, http://www.delta-fan.com/Download/Spec/AFB0912UHE-A.pdf

[11] Using fans in series and parallel: performance guidelines, Gareth Jones, http://www.ebmpapst-ad.com/media/content/technical_articles/TA_Using-fans-in-series-and-parallel.pdf

[12] Fans - Efficiency and Power Consumption, http://www.engineeringtoolbox.com/fans-efficiency-power-consumption-d_197.html

[13] CFD Simulation and Optimization of the Cooling of Open Compute Machine Learning 'Big Sur' Server, Mangesh Dhadve, Jimil Shah, Dereje Agonafer, ASME InterPACK 2017 (IPACK2017-74331),

Biographical Information

Mangesh Mohan Dhadve is from Pune, India. He completed his Bachelors in Mechanical Engineering from Savitribai Phule Pune University (formerly known as University of Pune) in June 2014. Mangesh worked as project trainee at Nand Composites Pvt. Ltd during his undergraduate degree to work on his final year project from July 2013 to May 2014. He joined The University of Texas at Arlington in Fall 2015 for Master of Science in Mechanical Engineering. He joined EMNSPC team in Summer 2016 and started working on various projects. Mangesh graduated from The University of Texas at Arlington on May 01, 2017.