CFD OPTIMIZATION OF THE COOLING OF THE

YOSEMITE OPEN COMPUTE SERVER

by

ADITYA GUPTA

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

May 2017

ABSTRACT

CFD OPTIMIZATION OF THE COOLING OF THE

YOSEMITE OPEN COMPUTE SERVER

Aditya Gupta, MS

The University of Texas at Arlington, 2017

Supervising Professor: Dereje Agonafer

Rising energy demands in data centers have constantly made thermal engineers to think and come up with innovative cooling solutions in data centers. It is of utmost importance to have control over the environmental impact of Data Centers. In 2011, the Open Compute Project was started which aimed at sharing energy efficient practices for data centers. The hardware and electrical specifications of the first open compute server – Freedom was shared on open compute project's website. It was a vanity free design and its components were custom designed. It was deployed in one of the data centers in Prineville, Oregon and within first few months of operation, considerable savings in every and cost were observed. Since then, many open compute servers have been introduced for applications like – compute, storage, etc. The open compute servers which were being introduced mainly had a 2 socket architecture. Yosemite Open compute server was introduced for serving heavy compute workloads. It provided significant improvement in performance per watt as compared to previous generations open compute servers.

Yosemite Open Compute Server has a system on a chip architecture and has 4 CPU's (1 CPU per sled). This study involves optimization of Yosemite Open Compute server to improve its cooling performance using. CFD tools are very useful for thermal

modeling of these servers and predict their efficiency. A commercially available CFD tool has been used to do the thermal modeling of the server and its optimization has been done to improve the cooling performance of the server. The model of the improved design has been compared to the existing design to show the impact of air flow optimization on the cooling performance of the server. The air flow characteristics and utilization of the fans have been significantly improved in the improved design.

Table of Contents

List of Illustrations

Chapter 1

Introduction

1.1 Data Centers: Introduction

A data center is a facility which stores and manages all the data of an organization. Businesses heavily rely on the large amounts of data in their data centers. Data Centers have IT equipment including servers, storage systems, networking switches, routers, firewalls, etc. They also have racks and cabling for proper organization and interconnection of IT equipment. Data Centers have infrastructure to support their daily functioning like power distribution, power subsystems, electrical cabling, backup generators, etc. Data Centers also require proper cooling infrastructure for adequate cooling of the IT Equipment present in the Data Centers.

In recent years, the energy consumption in data centers has risen and has in turn increased the need for better and energy practices for data centers. A report was submitted to congress in response to the request from Congress stated in Public Law 109-431 (H.R. 5646), "An Act to Study and Promote the Use of Energy Efficient Computer Servers in the United States.", in 2006, according to which data centers consumed about 61 billion kilowatt-hours (kWh). This was around 1.5% of total U.S. electricity consumption and equivalent to $4.5 billion in electricity cost [1]. The electricity consumption data centers in the U.S. was about 91 billion kilo-watt hours. This is equivalent to annual output of 34 large coal fired power plants which would be enough to power homes in New York City twice over [2]. Thermal engineers are constantly thinking and coming up with innovative thermal solutions for data centers to keep the energy consumption in data centers minimum so that energy can be utilized in an efficient and cost effective way with minimized material use and the environmental impact of data centers is lowered.

1.2 Cooling in Data Centers

Cooling in Data Centers ensures that environmental conditions are optimum for proper operation in a data center. The heat generated by IT equipment needs to be removed efficiently for continuous and reliable operation of the equipment in the data center. In an air cooling configuration in a data center, the cold air enters the servers from the cold aisle, picks up the heat from major heat dissipating components and exits the hot aisle. The heat transfer process in servers comprises of conduction, convection and radiation and the most important heat dissipating component in servers is the CPU. Heat sinks are mounted on the CPU which help is dissipating heat from the CPU to the ambient via Convection. Convective heat transfer plays an important in heat transfer in servers. The hot air from the hot aisle is sent to Computer Room Air Conditioning units (CRAC) and cold air from these units enters the cold aisle through perforated tiles. In air cooling, the management of air flow in the server is very important as it directly impacts the cooling efficiency of the server. Fans are present in the server to move the air for cooling. Fans are controlled by inbuilt fan control algorithms and get temperature inputs from the CPUs according to which the speed of the fans change. A typical cooling mechanism in a data center is shown in figure 1.1 [3].
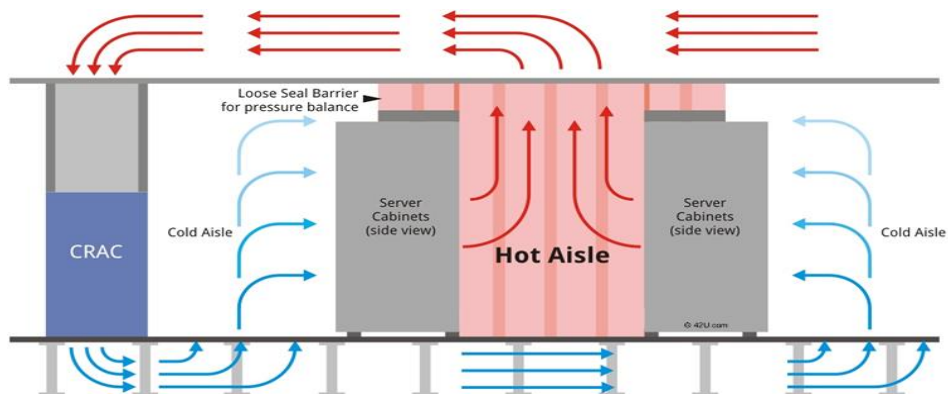

Figure 1.1 Cooling in Data Centers

1.3 The Open Compute Project

In 2011, in an effort to make energy efficient data centers and share best practices for data centers which are energy and cost efficient, the Open Compute Project was started in 2011 and the design specifications of the first Open Compute Server were made accessible for the community on the web [4]. All the components of the server were custom designed and the design of the first open compute server was vanity free – having minimum number of components. The open compute servers were deployed in one of the data centers in Prineville, Oregon and in first few months of operation, a lot of savings in energy consumption and cost were seen. Progressive generations of Open Compute Servers have been introduced since then which are used for purposes like computing, storage, network switches, etc.

1.4 Motivation of the work

Air flow characteristics play a critical role in defining how well the server is being cooled by directly affecting the heat transfer rates. Fans are present in the server to pull air and cool the major heat generating components. Fans overcome certain system impedance present in the server and cool the components which means that the system impedance directly affects the performance of fans. Fans have to do more work to pull the right amount of air if the number of obstructions are more, which increases the pumping power. A lot of researchers have studied flow bypass affects thermal performance of heat sinks [5-6]. Mani et al. modified the existing duct design and studied its on cooling performance on a high end open compute web server subjected to significant thermal shadowing. The improved design resulted in savings in cooling power but at the same time, increased the system impedance [7]. For high density servers like Yosemite, it is important that the air flow is very efficient and sufficient enough to cool the major heat generating

3

components. Moreover, proper utilization of fans is also important in high density servers. This research work focuses on improving the cooling performance of Yosemite Open Compute Server by improved ducting and proper utilization of fans. The optimization is done in such a way that the system impedance of the improved design remains same as that of the baseline design so that there is no power loss in overcoming the system impedance.

Chapter 2

Yosemite Open Compute Server

(Reprinted with permission © 2017 ASME) [10]

2.1 Yosemite Platform

Yosemite platform is a multi-node server which hosts 4 single socket micro-server cards. These 4 micro-server cards are mounted on a side plane. It has a form factor of 2 U. The overall dimensions of the micro-server card are 110mm*210mm. There are 2 levels in Yosemite and each level has 2 server cards (Figure 2.1). For cooling, 2 - 80mm fans are present at the back and an air duct is present which runs sideways (Figure 2.1). The server cards get single 12V power from the Yosemite platform and Yosemite gets 12V power from the rack via a bus bar. Yosemite's design is highly flexible and can have micro-server cards from different vendors. A mezzanine card is present right at the front end which acts as a connection interface from the server cards to the Top of Rack Switch (TOR) switch. Mezzanine card is hosted by an adapter card which serves as a carrier board. The side plane has a hot swap controller and a BMC (Baseboard Management Controller) which manages all server cards in the Yosemite.
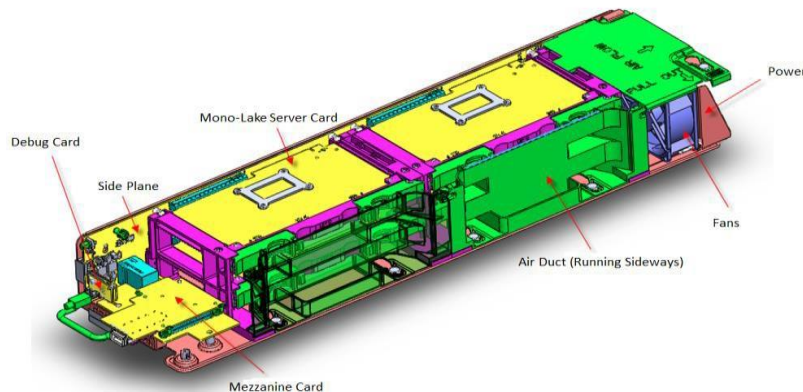
Figure 2.1 Yosemite Platform - Model

5

Figure 2.2 Yosemite Open Compute Server

A maximum of 15 W card power is recommended for the Mezzanine Card. Yosemite has 1 processor per card based on Intel Xeon-D processor series with a TDP of 65W. There are a total of 16 DIMM sticks in Yosemite (4 per micro-server card). The micro-server card is called Mono Lake server card (Figure 2.3) and can support up to two on card Solid State Drives (SSDs) with 2280 M.2 form factor. The recommended minimum capacity of the SSD is 128 GB.

Figure 2.3 Monolake Microserver Card

Chapter 3

Computational Fluid Dynamics

3.1 CFD: Introduction

CFD is a branch in fluid mechanics which deals with and characterizes fluid dynamics and heat transfer in a fluid flow problem. CFD IS capable of characterizing things like Pressure, Velocity, Temperatures, etc. The fundamental basis of problem in CFD is the Navier-Stokes equation. When we apply the laws of mechanics [8],

According to conservation of mass equation,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{V}) = 0$$

And according to conservation of momentum equation,

$$\rho \frac{\partial \vec{V}}{\partial t} + \rho(\vec{V} \cdot \nabla)\vec{V} = -\nabla p + \rho\vec{g} + \nabla \cdot \tau_{ij}$$

These 2 equations and the energy conservation equation are difficult to solve with analytical methods. But with the help of CFD, it is possible to obtain the solution of the governing equations.

CFD replaces a continuous problem domain with discrete domain via a grid [8]. The size and shape of grid has significant impact on the results and that is why it is important to do grid sensitivity analysis before a CFD problem is solved. In the case of continuous domain, every single flow variable is defined at every point in that region where as in the case of discrete domain, flow variables are defined at grid points. The flow variable are solved only at these grid points when a CFD problem is solved [8]. There are ways in which a CFD problem can be discretized like – Finite Volume Method, Finite Element

7

Method, etc. A commercially available CFD tool – 6Sigma ET was used to solve this CFD problem. The Finite Volume Method is adopted in 6Sigma ET for solving the problem wherein the measurement domain is discretized into control volumes.

3.2 Turbulence Modeling

Turbulence modeling is done in CFD to incorporate the effects of turbulence in a fluid flow problem. A turbulent flow field is characterized by fluctuations in flow velocity and pressure. It is different from laminar flow where in the fluid moved in parallel layers and the flow regime is smooth. There are various means by which turbulence can be modelled and the most common method is k-ε turbulence model. Figure 3.1 shows a comparison of Laminar and Turbulent Flow characteristics [9].
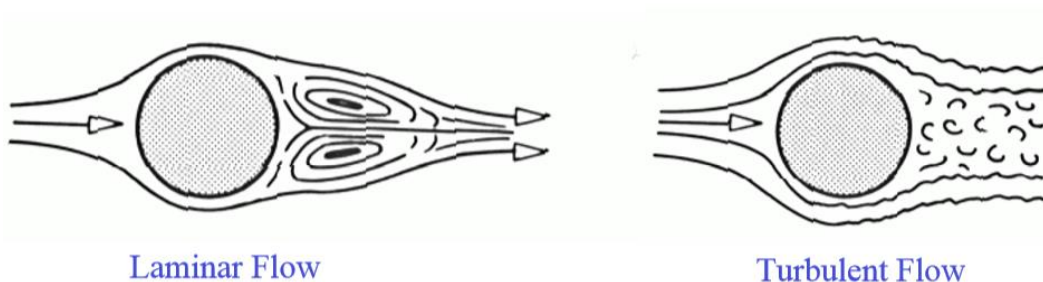

Figure 3.1 Laminar and Turbulent Flow

3.3 k-ε Turbulence Model

This model given a descriptive information about turbulent characteristics in a fluid flow problem. It has 2 additional transport equations. The first transport variable k gives the energy of turbulence and the second transport variable ε gives the scale of turbulence.

Chapter 4

Methodology

(Reprinted with permission © 2017 ASME) [10]

A commercially available CFD tool - 6Sigma ET has been used to create the baseline model of the Yosemite platform. The design specifications present on Open Compute Project's website were used to do detailed modeling of the server [4]. For baseline testing, two fans RPMs and one air inlet temperatures have been selected making 2 test cases. Parameters like CPU die temperatures, fan flow rates and pumping power are evaluated. The air flow characteristics are also studied for the baseline model. Then, optimization has been done in the baseline model which includes modified ducting and achieving symmetry in the design, resulting in better air flow characteristics and effective utilization of fans. The improved model is compared to the baseline model and savings in pumping power and flow rates are evaluated.

Chapter 5

Original Server Design and its Optimization

(Reprinted with permission © 2017 ASME) [10]

Figure 5.1 shows the air duct running sideways in the model comprising of 2 ducts – Air Duct 1 and Air duct 2. There are 2 levels in the server - Level 1 and Level 2, each having 2 microserver cards. Figure 5.2 shows side view of the server from left side showing the 4 CPUs and the fans at the back.
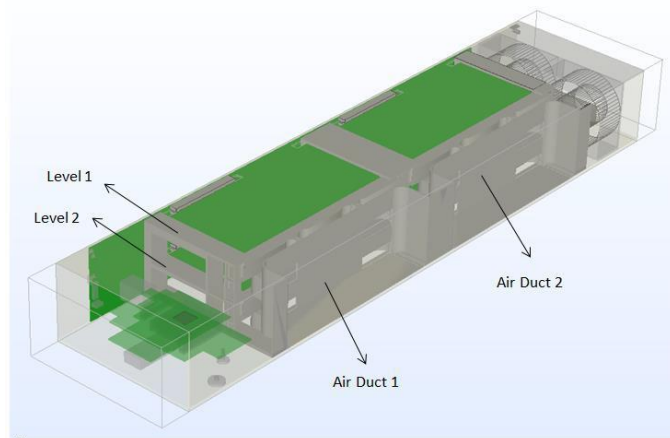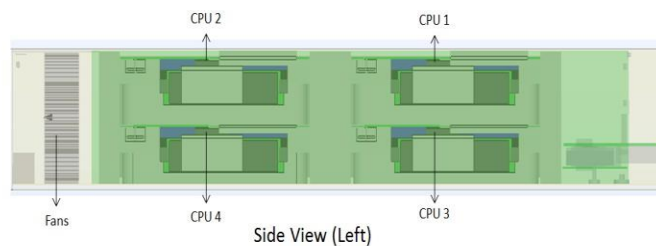

Figure 5.1 CFD Model of Yosemite

Figure 5.2 Side View of Yosemite showing 4 CPUs

The main components taken into consideration on the microserver card are CPU, DIMMs and 2280 M.2 form factor SSD as shown in figure 5. There are heat sinks mounted on the CPUs.

10
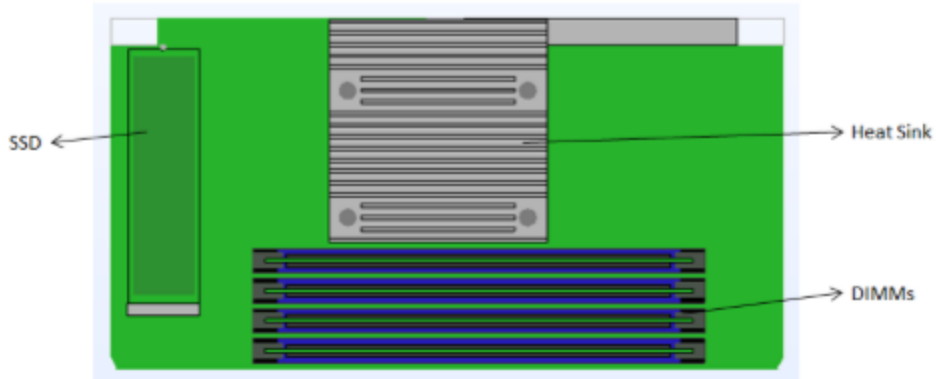
Figure 5.3 Yosemite Microserver Card

In this study, emphasis has been given on improving the air flow characteristics via ducting and how the fans can be used to cool the major components more effectively. The original ducting is done in such a way that it facilitates some air flow through the DIMMS via a secondary air opening apart from the primary air opening as seen in Figure 5.4.

Figure 5.4 Secondary Air Opening

But it is observed with the help of particle tracking that there is some air flow bypass and some air recirculation happening along the length of the duct. The improved ducting

completely blocks the secondary opening throughout the length of the duct and also eliminates all the air flow bypass through the duct. Therefore, with modified ducting, the air now flows via a comparatively smaller area because of presence of just one primary air inlet passage resulting 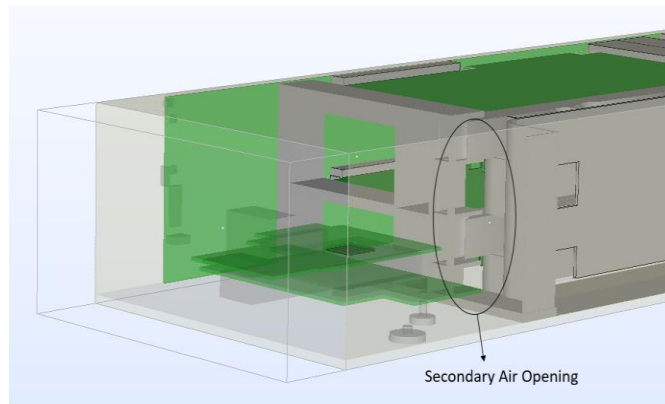in an increase in air velocity, resulting in better heat transfer rates and improved cooling. Figure 5.5 shows the improved ducting as compared to the original ducting. Moreover, the improved air duct 2 opens up diagonally towards the rear which facilitates smooth air flow exit towards the rear.

Another problem in Yosemite platform is symmetry of placement of CPUs and DIMMs with respect to the fans. It is of utmost importance that major heat dissipating components receive adequate amount of air so that they are maintained below their critical temperatures. With the help of particle tracking, it is seen that the air pulled by fan 2 mainly flows over the DIMMS and not over any of the heat sinks. In the improved design, the side plane has been shifted towards the right by 6.5 mm. This is shown in figure 5.6. This is done to achieve better symmetry of placement of major heat dissipating components with respect to fans. It facilitates better and effective utilization of fan 2.
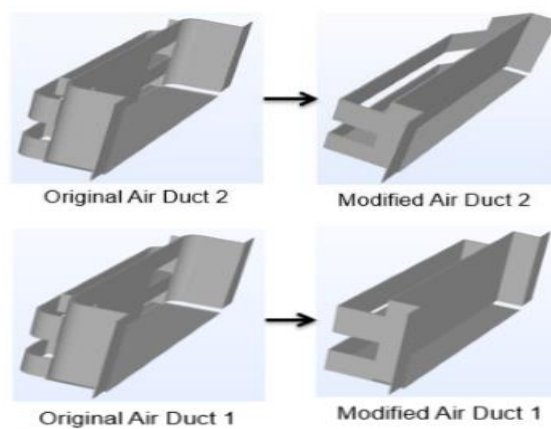

Original Air Duct 2   Modified Air Duct 2

Original Air Duct 1   Modified Air Duct 1
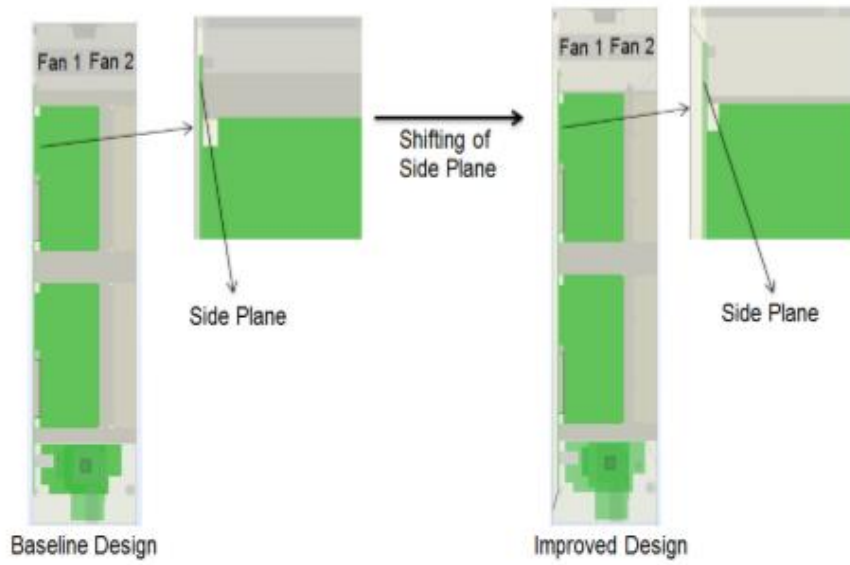
Figure 5.5 Original and Modified Ducting

Figure 5.6 Shifting of Side Plane

Chapter 6

Results and Discussion

(Reprinted with permission © 2017 ASME) [10]

The testing for the baseline case and improved case for Yosemite platform was done for one inlet temperature – 20° C and two fan RPMs – 4000 RPM and 6000 RPM giving two test cases. The CPU die temperatures for the baseline study for both test cases are recorded and compared with the CPU die temperatures in the improved design. The improved design gives lower CPU die temperatures as compared to the baseline study. Then, the fan speeds for the improved design were ramped down to achieve the original CPU die temperatures measured in the baseline case and finally, savings in fan power and flow rates were evaluated. Furthermore, the system impedance of the modified design was compared to the existing design.

Figure 6.1 shows the results for Test Case I. The fans were run at 4000 RPM and the CFM for both the fans was adjusted according to the fan curve and the system impedance. For baseline case, the CPU die temperatures had a range between 68.4°C and 82.1°C and for the improved case it had a range between 65.3°C and 78.2°C. The power consumption and flow rates for the baseline and improved case were nearly same (Figure 6.1) because the system impedance of both the systems were same as seen in Figure 6.2. Finally, the fan speeds were ramped down in the improved model till the CPU die temperatures matched the baseline case.

Fan RPM	Air Inlet Temp	CPU 1 Die Temp	CPU 2 Die Temp	CPU 3 Die Temp	CPU 4 Die Temp	Fan Power	Flow Rate
6000	20°C	62.3°C	68.3°C	66.4°C	72.2°C	2.11 W	78.7 CFM

Baseline Case

Fan RPM	Air Inlet Temp	CPU 1 Die Temp	CPU 2 Die Temp	CPU 3 Die Temp	CPU 4 Die Temp	Fan Power	Flow Rate
6000	20°C	59.7°C	65.6°C	62.9°C	69.2°C	2.09W	80.9 CFM

Improved Design Case

Fan RPM	Air Inlet Temp	CPU 1 Die Temp	CPU 2 Die Temp	CPU 3 Die Temp	CPU 4 Die Temp	Fan Power	Flow Rate
5200	20°C	61.5°C	68.2°C	65.0°C	72.1°C	1.34 W	69.7 CFM

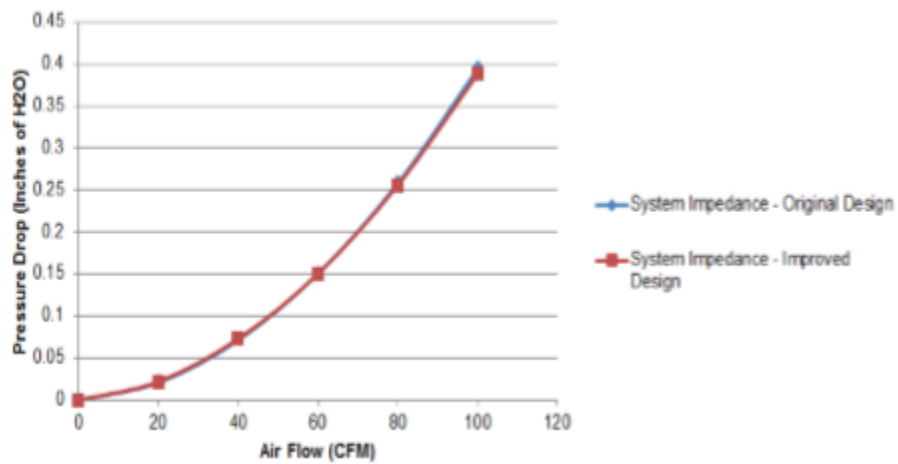Fan Speed Ramped Down

Figure 6.1 Test Case 1

Figure 6.2 System Impedance

The air flow characteristics in the baseline study and improved design are studied and analyzed with the help of particle tracking. There is a significant improvement in the air flow characteristics which is seen in the improved case. In the baseline case, the fan 2

only cools the DIMMS effectively because of the way the way ducting has been done in the baseline design as mentioned earlier. In the improved design, with a coupled effect of improved ducting and shifting of side plane, the CPU heat sinks receive some air from fan 2 as well, as compared to the baseline case in which the heat sinks only receive air mainly from fan 1 as seen in the particle tracking figure. This coupled effect, along with increased air velocities aids in better cooling of the four CPUs and in turn, bringing down the CPU die temperatures. Figure 6.3 shows the particle tracking for baseline and improved case for fan 1, fan 2 and both fans together.
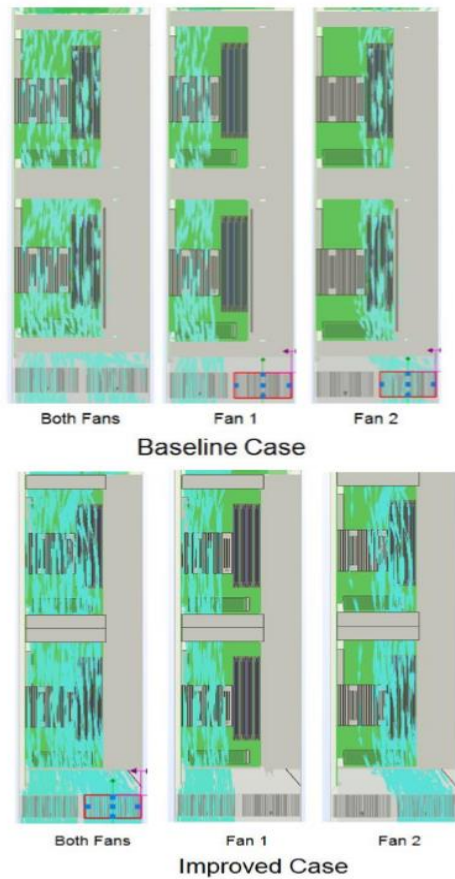

Figure 6.3 Air Flow – Particle Tracking

Figure 6.4 shows the results for Test Case II. The fans were run at 6000 RPM and the same procedure was followed as in Test Case I. The power savings for Test Case I was 36.5% and for Test Case II was 36.4%. Figure 6.5 shows a graphical representation of comparison of power savings in both test cases.

Mesh sensitivity analysis was done to make sure that the results are independent of cell count and it was found that at a cell count of 40 million, the results were independent of cell count.

Fan RPM	Air Inlet Temp	CPU 1 Die Temp	CPU 2 Die Temp	CPU 3 Die Temp	CPU 4 Die Temp	Fan Power	Flow Rate
6000	20°C	62.3°C	68.3°C	66.4°C	72.2°C	2.11 W	78.7 CFM

Baseline Case

Fan RPM	Air Inlet Temp	CPU 1 Die Temp	CPU 2 Die Temp	CPU 3 Die Temp	CPU 4 Die Temp	Fan Power	Flow Rate
6000	20°C	59.7°C	65.6°C	62.9°C	69.2°C	2.09W	80.9 CFM

Improved Design Case

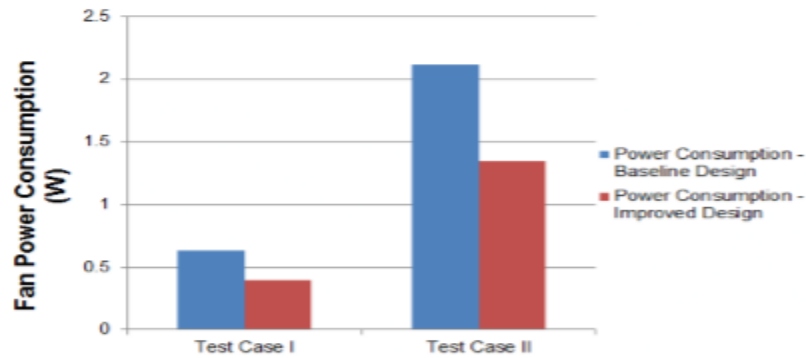| Fan RPM	Air Inlet Temp	CPU 1 Die Temp	CPU 2 Die Temp	CPU 3 Die Temp	CPU 4 Die Temp	Fan Power	Flow Rate
5200	20°C	61.5°C	68.2°C	65.0°C	72.1°C	1.34 W	69.7 CFM

Fan Speed Ramped Down

Figure 6.4 Test Case 2

Figure 6.5 Comparison of Fan Power Consumption

Chapter 7

Conclusion and Future Work

(Reprinted with permission © 2017 ASME) [10]

CFD modeling was done for Yosemite Open Compute Server using a commercially available CFD tool and the design was optimized with major changes in ducting and other design changes with an overall aim to improve the cooling performance of Yosemite platform. The air flow characteristics were studied with the help of particle tracking in the baseline design and improved by eliminating all bypass air flow through the duct and by improving the fan utilization via coupled effect of improved ducting and shifting of side plane. The fans were run at 2 different RPMs – 4000 RPM (test case I) and 6000 RPM (test case II). For test case I, 36.5% fan power savings were achieved and for test case II, 36.4% fan power savings were achieved in the improved design. The design was optimized in such a way that the system impedance of the improved design remains the same for Yosemite Platform, resulting in no loss of power to overcome any increase in system impedance.

In future work, experimental validation of the CFD results can be done and the server can be tested at higher inlet temperatures.

References

1. Brown, R. E., Masanet, E. R., Nordman, B., Tschudi, W. F., Shehabi, A., Stanley, J., Koomey, J. G., Sartor, D. A., and Chan, P. T., 2008, "Report to Congress on Server and Data Center Energy Efficiency: Public Law 109-431" LBNL-363E, from https://eta.lbl.gov/publications/report-to-congress-onserver-and-data

2. Whitney, J., and Delforge, P., 2014, "Scaling Up Energy Efficiency Across the Data Center Industry: Evaluating Key Drivers and Barriers", IP:14-08-a, from https://www.nrdc.org/sites/default/files/datacenter-efficiency-assessment-IP.pdf

3. http://www.42u.com/cooling/hot-aisle-containment.htm

4. www.ocp.org

5. Radmehr, A., Kelkar, K. M., Kelly, P., Patankar, S. V., and Kang, S. S., "Analysis of the Effect of Bypass on the Performance of Heat Sinks Using Flow Network Modeling (FNM)," 15th Annual IEEE Semiconductor Thermal Measurement and Management Symposium, IEEE, Piscataway, NJ, 1999, pp. 42–4

6. Lee, R.S., Huang, H.C., and Chen, W.Y., "A Thermal Characteristic Study of Extruded-Type Heat Sinks in Considering Air Flow Bypass Phenomena," Proceedings of the Sixth Annual IEEE Symposium on Semiconductor Thermal and Temperature

7. Mani, D., Fernandes, J., Eiland, R., Agonafer, D., and Mulay, V., "Improving ducting to increase cooling performance of high-end web servers subjected to significant thermal shadowing – an experimental and computational study," 2015 31st Thermal Measurement, Modeling & Management Symposium (SEMI-THERM), San Jose, CA, 2015, pp. 319-323.

8. http://dragonfly.tam.cornell.edu/teaching/mae5230-cfd-intro-notes.pdf

9. https://web.stanford.edu/class/me469b/handouts/turbulence.pdf

10. Aditya Gupta, Ananthavijayan Sridhar, Dereje Agonafer, "CFD Optimization of the Cooling of Yosemite Open Compute Server", IPACK2017-74254, ASME InterPACK 2017, San Francisco, California

Biographical Information

Aditya Gupta completed his Bachelor's in Mechanical Engineering from Amity University, Noida, India in 2013 and his Master's in Mechanical Engineering from University of Texas at Arlington in 2017. He specialized in the field of thermal fluid sciences