Proceedings of the ASME 2017 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems
InterPACK2017
August 29-September 1, 2017, San Francisco, California, USA

IPACK2017-74254

CFD OPTIMIZATION OF THE COOLING OF YOSEMITE OPEN COMPUTE SERVER

Aditya Gupta
University of Texas at Arlington
Arlington, Texas, USA

Ananthavijayan Sridhar
Future Facilities Inc.
San Jose, California, USA

Dereje Agonafer
University of Texas at Arlington
Arlington, Texas, USA

ABSTRACT

Over the past few years, there has been an ever increasing rise in energy consumption by IT equipment in Data Centers. Thus, the need to minimize the environmental impact of Data Centers by optimizing energy consumption and material use is increasing. In 2011, the Open Compute Project was started which was aimed at sharing specifications and best practices with the community for highly energy efficient and economical data centers. The first Open Compute Server was the 'Freedom' Server. It was a vanity free design and was completely custom designed using minimum number of components and was deployed in a data center in Prineville, Oregon. Within the first few months of operation, considerable amount of energy and cost savings were observed. Since then, progressive generations of Open Compute servers have been introduced. Initially, the servers used for compute purposes mainly had a 2 socket architecture. In 2015, the Yosemite Open Compute Server was introduced which was suited for higher compute capacity. Yosemite has a system on a chip architecture having four CPUs per sled providing a significant improvement in performance per watt over the previous generations. This study mainly focuses on air flow optimization in Yosemite platform to improve its overall cooling performance. Commercially available CFD tools have made it possible to do the thermal modeling of these servers and predict their efficiency. A detailed server model is generated using a CFD tool and its optimization has been done to improve the air flow characteristics in the server. Thermal model of the improved design is compared to the existing design to show the impact of air flow optimization on flow rates and flow speeds which in turn affects CPU die temperatures and cooling power consumption and thus, impacting the overall cooling performance of the Yosemite platform. Emphasis is given on effective utilization of fans in the server as compared to the original design and improving air flow characteristics inside the server via improved ducting.

INTRODUCTION

Over the years, the energy consumption by IT equipment in the Data Centers has increased a lot and thus, the need for finding better thermal solutions to reduce energy consumption in an efficient and cost effective way is growing. According to a report 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, data centers consumed about 61 billion kilowatt-hours (kWh), which was around 1.5% of total U.S. electricity consumption and equivalent to $4.5 billion in electricity costs [1]. According to a study done to evaluate growth in data center electricity use from 2005 to 2010, electricity consumption in global data centers in 2010 accounted for between 1.1% and 1.5% of overall electricity consumed and specifically for U.S., it had a range from 1.7% to 2.2% [2]. In 2013, the electricity consumption by U.S. Data Centers was around 91 billion kilo-watt hours which is nearly equivalent to annual output of 34 large coal fired power plants which would be sufficient to power homes in NYC twice over [3]. The growing energy demand for data centers has constantly made thermal engineers to think and redesign thermal solutions. In 2011, the Open Compute Project was introduced and started which was an initiative taken industrywide to share specifications and most efficient practices to build energy efficient and economical data centers [4]. The first Open Compute Server 'Freedom' was completely custom designed. Any component which was unwanted and didn't contribute to efficiency was removed, making the server a vanity free design and was first deployed in a data center in Prineville, Oregon in 2011. Design specifications for all hardware used in the data center which included components like motherboards, power supply, server chassis, battery cabinets, etc. were published online on Open Compute Project's website. Information like data center electrical and mechanical construction specifications was also shared on the website. Significant savings in energy consumption and costs were observed in the data center within the first few months of operation. Soon, a lot of big companies joined the Open Compute project. Since the introduction of the project, newer generations of Open Compute Servers have been introduced for applications like computing, storage, etc. The servers used for compute applications mainly had a 2 socket architecture. The Yosemite Open Compute server was introduced in 2015 and has a system-on-a-chip (SoC) architecture. It has a modular chassis that contains 4 high powered SoC processor cards providing immense improvement in performance per watt as compared to the previous generation servers and is better suited for heavy compute workloads [5].

In this study, Yosemite platform is taken into consideration and its design has been optimized using a commercially available CFD tool to improve its cooling performance. Yosemite is completely air cooled. An important factor in an air cooled server is proper utilization of fans which is directly affected by the system impedance. If there are a lot of obstructions in the server, fans have to do more work to get the required air flow to cool major components which increases the pumping power needed. Moreover, the air flow characteristics, flow rates and flow velocities play a major role in defining the cooling performance of a server by directly impacting the heat transfer rates. Researchers have extensive studied the effect of flow bypass on thermal performance of heat sinks [6 - 8]. Wirtz et al. studied the effect of flow bypass on the performance of longitudinal fin heat sinks [9]. Mani et al. studied the effect of improved ducting on cooling performance on a high end open compute web server specifically subjected to significant thermal shadowing. She showed significant improvement in cooling power consumption of the server but the system impedance of the improved model was higher than the original model and thus, some percentage of cooling power savings went into overcoming the increased system resistance [10]. The Yosemite Open Compute Server is a high density server and it is very important to make sure that the air flows efficiently inside such high density servers. It is also of critical importance that fans inside the server are used effectively as higher system impedance of some high density servers can significantly affect fan performance. This study mainly focuses on improving the cooling performance of Yosemite by improving air flow characteristics inside the server with the help of improved ducting coupled with effective fan utilization and other design changes to keep the system resistance of the improved design same so that no percentage of power savings achieved is used to overcome any increase in system impedance. This study only considers major heat dissipating components in Yosemite like CPUs, DIMMS and the Mezzanine chip.

KEYWORDS

Open Compute Server, Ducting, Cooling Performance, Yosemite, Air Flow Characteristics, CFD

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 1). For cooling, 2 - 80mm fans are present at the back and an air duct is present which runs sideways (Figure 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.

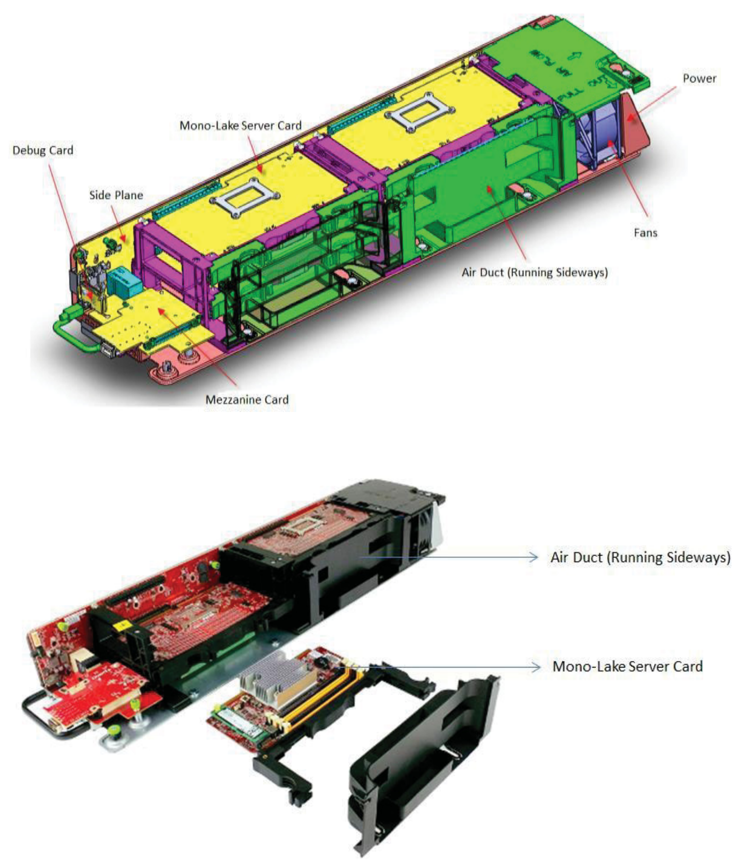

Figure 1 – Yosemite Platform

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) 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 – Monolake Micro Server Card

METHODOLOGY

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 [11]. For baseline testing, two fans RPMs and one air inlet temperature 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 better utilization of fans. The improved model is compared to the baseline model and savings in pumping power and flow rates are evaluated.

COMPUTATIONAL FLUID DYNAMICS

A commercially available CFD tool – 6Sigma ET is used to model the server and for solving the model, K-ε turbulence model is used. Navier-Stokes equation form the basis for a CFD problem. According to mass conservation and momentum conservation for a generalized case [12],

$$\frac{\partial \rho}{\partial x} + \nabla.(\rho u) = 0$$

$$\frac{\partial}{\partial t}(\rho u) + \nabla.(\rho u u) = \nabla.(\mu grad u) - \frac{\partial p}{\partial x} + B_x + V_X$$

And according to conservation of energy for a low velocity steady flow [12],

$$\nabla.(\rho u h) = \nabla.(k grad T) + S_h$$

K-ε turbulence model is the recommended turbulence model for this application. It is a 2 equation model and is widely used for modeling turbulent flows. The turbulent kinetic energy k and its dissipation rate ϵ are calculated from the 2 transport equations [13]:

$$\frac{\partial(\rho k)}{\partial t} + \frac{\partial \rho k u_i}{\partial x_i} = \frac{\partial}{\partial x_i}\left[\left(\mu + \frac{\mu_t}{\sigma_k}\right)\frac{\partial k}{\partial x_i}\right] + G_k + G_b - \rho\varepsilon$$

$$\frac{\partial(\rho\varepsilon)}{\partial t} + \frac{\partial \rho\varepsilon u_i}{\partial x_i} = \frac{\partial}{\partial x_i}\left[\left(\mu + \frac{\mu_t}{\sigma_k}\right)\frac{\partial \varepsilon}{\partial x_i}\right] + C_{1S}\frac{\varepsilon}{k}(G_k + C_{3S}G_b) - C_{2S}\rho\frac{\varepsilon^2}{k}$$

Figure 3 shows the server model created using CFD. The figure 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.

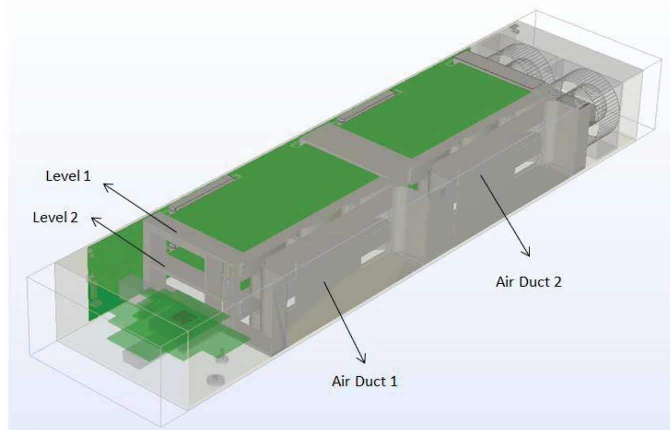

Figure 3 – CFD Model of Yosemite

Figure 4 shows side view of the server from left side showing the 4 CPUs and the fans at the back.

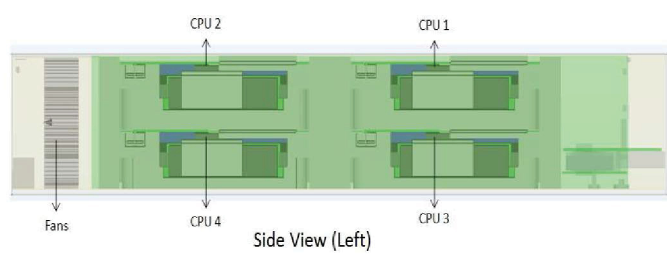

Figure 4 – Side view of Yosemite showing 4 CPUs

Table 1 shows the distribution of CPUs on level 1 and level 2.

Level	CPUs
Level 1	1, 2
Level 2	3, 4

Table 1

The main components taken into consideration on the micro-server card are CPU, DIMMs and 2280 M.2 form factor SSD as shown in figure 5. There is a heat sink mounted on the CPU.

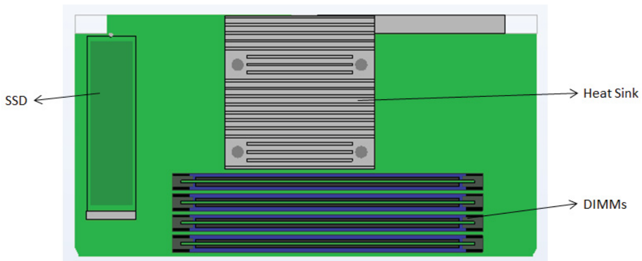

Figure 5 – CFD Model – Monolake Microserver Car

ORIGINAL SERVER DESIGN AND ITS OPTIMIZATION

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 6.

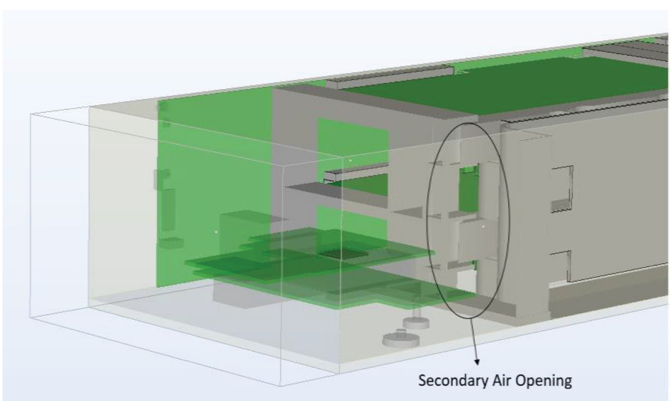

Figure 6 – Secondary Air Opening

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 7 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 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. Flow deflectors have also been added in the server for better guidance of air flow.

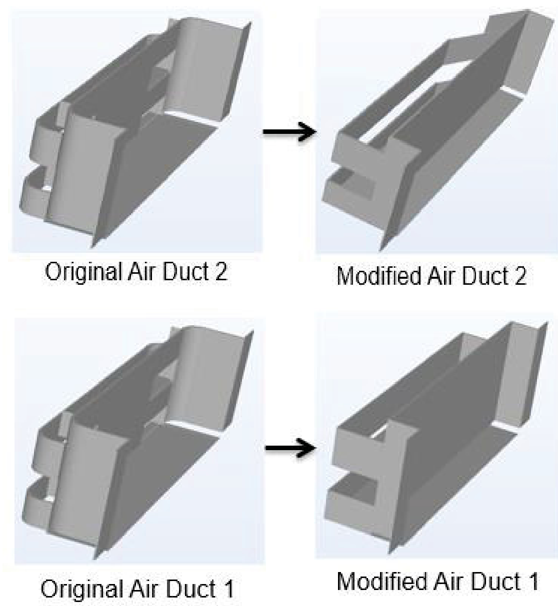

Figure 7
Left – Original Air Duct; Right – Modified Air Duct

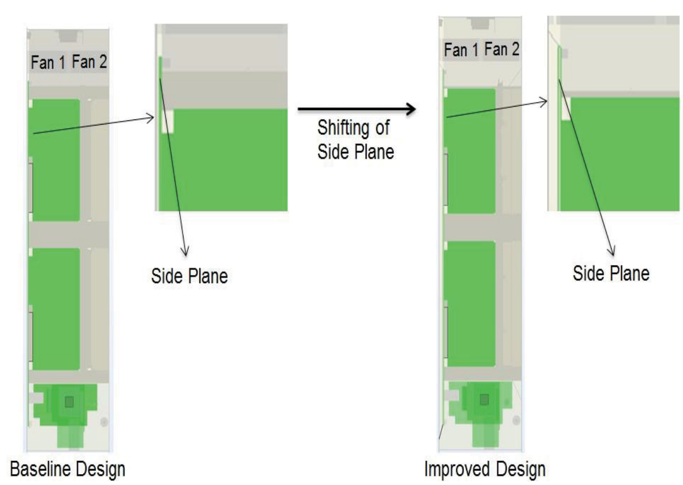

Figure 8
Left – Original Position of Side Plane;
Right – Side plane Shifted by 6.5 mm

There are 4 air passages present in Yosemite platform which guide the air as it flows through the server. The width of the fourth air passages has been reduced which makes the flow of air smooth in that region till it is pulled out by the fans. It was found that reducing the width of the fourth air passage helps in decreasing the system impedance. Figure 9 shows the original and modified air passages.

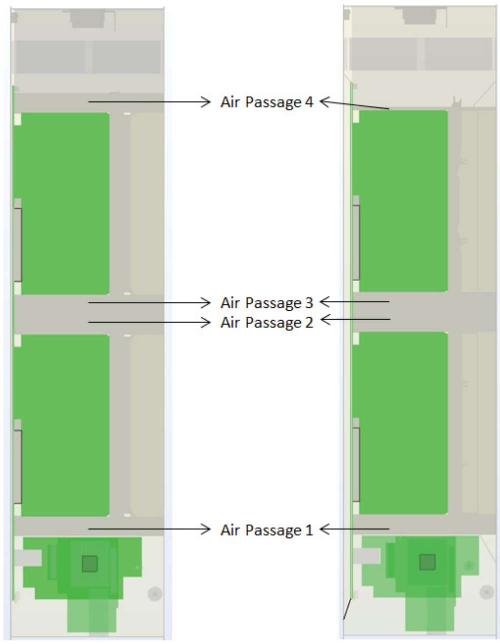

Figure 9
Left - Original Air Passages;
Right – Modified Air Passage

The mesh sensitivity analysis was performed for Yosemite platform to make sure that the results are independent of grid size and grid count. The grid count was varied from 4 million to 60 million and a CFM was prescribed at the server inlet. It was found out that the pressure drop remains constant from 40 million grid count onwards. The simulations were run at 40 million grid count for Yosemite platform.

RESULTS AND DISCUSSION

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.

Table 2 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 for the fans installed in the system and the system impedance of Yosemite platform. For the baseline case, the CPU die temperatures had a range between 68.4°C to 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 the same (Table 2). This is because the system impedance of both the systems were same as seen in Figure 10. Finally, the fan speeds were ramped down in the improved model till the CPU die temperatures matched the baseline case so that the CPUs meet the thermal specifications.

Test Case I

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
4000	20°C	68.4°C	77.3°C	73.6°C	82.1°C	0.63 W	51.9 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
4000	20°C	65.3°C	73.8°C	69.4°C	78.2°C	0.62W	53 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
3470	20°C	67.5°C	77.4°C	72°C	82°C	0.40W	45.7 CFM

Fan Speed Ramped Down

Table 2 - Results for Test Case 1

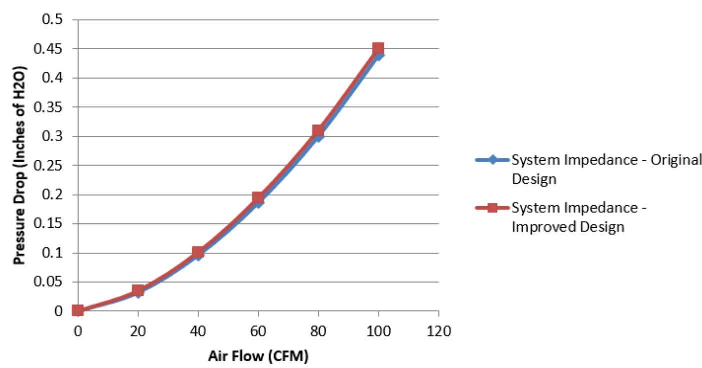

Figure 10 – System Impedance for Both Designs

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 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 figure 11. This coupled effect, along with increased air velocities aids in better cooling of the four CPUs in turn, bringing down the CPU die temperatures. Figure 11 shows the particle tracking for baseline and improved case for fan 1, fan 2 and both fans together. The figure clearly shows the change of air flow pattern especially in case of fan 2 in the improved case.

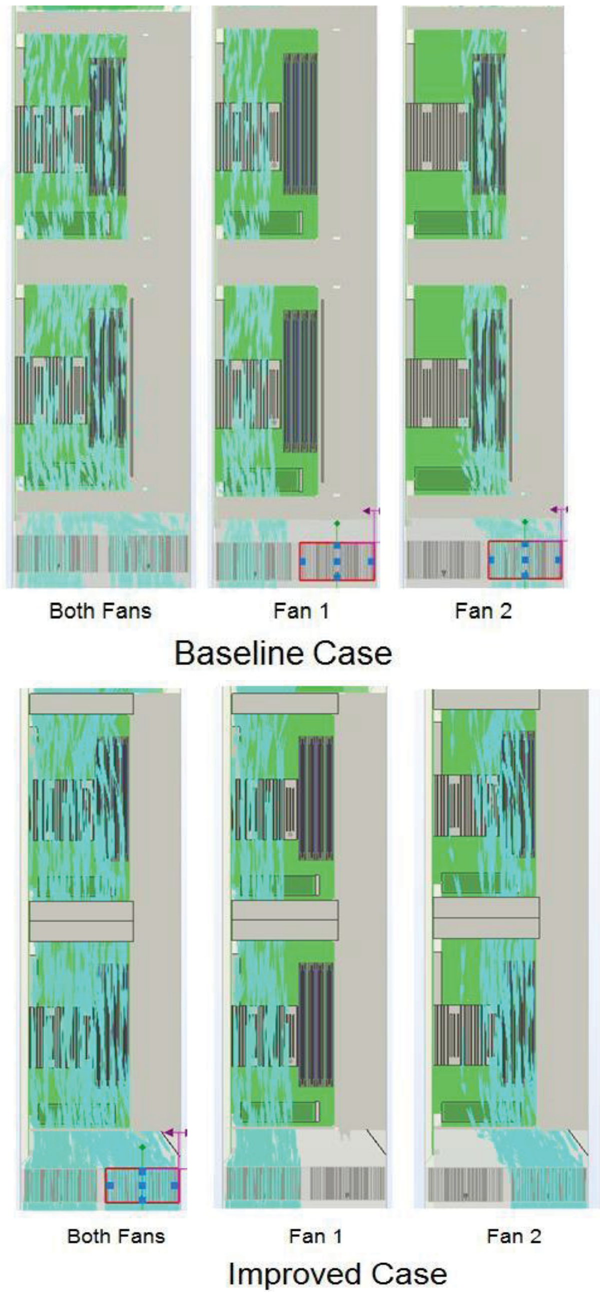

Figure 11 – Comparison of Air Flow Characteristics of Baseline and Improved Case

Table 3 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 12 shows a graphical representation of comparison of power savings in both test cases.

Test Case II

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 11 – Results for Test Case II

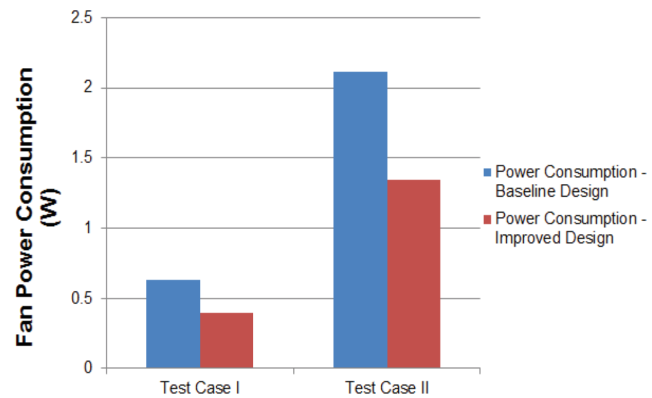

Figure 12 – Fan Power Consumption Comparison

CONCLUSION

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 .

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-on-server-and-data
2. Koomey, J. G., 2011, “Growth in data center electricity use 2005 to 2010”, A report by Analytical Press, completed at the request of The New York Times 9, from http://www.missioncriticalmagazine.com/ext/resources/MC/Home/Files/PDFs/Koomey_Data_Center.pdf
3. 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/data- center-efficiency-assessment-IP.pdf
4. Heiliger, J., 2011, “Building Efficient Data Centers with the Open Compute Project”, from https://www.facebook.com/notes/facebook-engineering/building-efficient-data-centers-with-the-open-compute-project/10150144039563920/
5. Li, H., 2016, “Introducing "Yosemite": the first open source modular chassis for high-powered microservers”, from https://code.facebook.com/posts/1616052405274961/introducing-yosemite-the-first-open-source-modular-chassis-for-high-powered-microservers-/
6. Radmehr, A., Kelkar, K. M., Kelly, P., Patankar, S. V., and Kang, S. S., 1999, “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, pp. 42–47.
7. Lee, R.S., Huang, H.C., and Chen, W.Y., 1990, "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 Measurement (SEMI-THERM VI), Phoenix, AZ, pp. 95- 102.
8. Leonard, W., Teertstra, P., Culham, J. R., and Zaghol, A., 2002, “Characterization of Heat Sink Flow Bypass in Plate Fin Heat Sinks,” ASME International Mechanical Engineering Congress and Exposition, American Society of Mechanical Engineers, Fairfield, NJ, Vol. 372, No. 5, pp. 189–196.
9. Wirtz, R.A., Chen, W., and Zhou, R., 1994, "Effect of Flow Bypass on the Performance of Longitudinal Fin Heat Sinks," ASME Journal of Electronic Packaging, Vol. 116, pp. 206-211.
10. Mani, D., Fernandes, J., Eiland, R., Agonafer, D., and Mulay, V., 2015, "Improving ducting to increase cooling performance of high-end web servers subjected to significant thermal shadowing – an experimental and computational study," 31st Thermal Measurement, Modeling & Management Symposium (SEMI-THERM), San Jose, CA, pp. 319-323.
11. Zhao, Y., Chu, R., Na, J., Dasari, S., 2016, “Facebook Multi---Node Server Platform: Yosemite Design Specification”, from http://files.opencompute.org/oc/public.php?service=files&t=7f5d8ba3c80744bc78a92d3cad46b7d2
12. Patankar, S. V., 1980, *Numerical heat transfer and fluid flow*, Hemisphere Publishing Corp., New York, Chap. 2.
13. Launder, B. E., and Spalding, D. B., 1974, “The numerical computation of turbulent flows” Computer methods in applied mechanics and engineering, 3(2), pp. 269-289.