**Proceedings of the ASME 2018
International Mechanical Engineering Congress and Exposition
IMECE2018
November 9-15, 2018, Pittsburgh, PA, USA**

IMECE2018-88497

RAISING INLET AIR TEMPERATURE FOR A HYBRID-COOLED SERVER RETROFITTED WITH LIQUID COOLED COLD PLATES

Uschas Chowdhury
University of Texas at Arlington
Arlington, Texas, USA

Ashwin Siddarth
University of Texas at Arlington
Arlington, Texas, USA

Manasa Sahini
University of Texas at Arlington
Arlington, Texas, USA

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

ABSTRACT

In typical data centers, the servers and IT equipment are cooled by air and almost half of total IT power is dedicated to cooling. Hybrid cooling is a combined cooling technology with both air and water, where the main heat generating components are cooled by water or water-based coolants and rest of the components are cooled by air supplied by CRAC or CRAH. Retrofitting the air-cooled servers with cold plates and pumps has the advantage over thermal management of CPUs and other high heat generating components. In a typical 1U server, the CPUs were retrofitted with cold plates and the server tested with raised coolant inlet conditions. The study showed the server can operate with maximum utilization for CPUs, DIMMs, and PCH for inlet coolant temperature from 25-45 °C following the ASHRAE guidelines. The server was also tested for failure scenarios of the pumps and fans with reducing numbers of fans and pumps. To reduce cooling power consumption at the facility level and increase air-side economizer hours, the hybrid cooled server can be operated at raised inlet air temperatures. The trade-off in energy savings at the facility level due to raising the inlet air temperatures versus the possible increase in server fan power and component temperatures is investigated. A detailed CFD analysis with a minimum number of server fans can provide a way to find an operating range of inlet air temperature for a hybrid cooled server. Changes in the model are carried out in 6SigmaET for an individual server and compared to the experimental data to validate the model. The results from this study can be helpful in determining the room level operating set points for data centers housing hybrid cooled server racks.

INTRODUCTION

Data centers are infrastructures or space containing continuously operating servers, racks and IT equipment for compute, storage and network. Heat generated through the usage of IT hardware requires cooling for ensuring uninterrupted service, high performance and long-term reliability. According to the ASHRAE TC 9.9 2011, average power consumptions for IT equipment and cooling (power and related infrastructures) are around 50% and 45% [1]. Data centers account for about 3% percent of the global electricity consumption [2], and the sector exhibits a growth rate of 10.7% globally [3]. Power Usage Effectiveness (PUE), one of the metrics to measure to determine efficiency of data center, is the ratio of energy consumption of IT equipment compared to total energy consumption in a data center. In a recent report, the PUE is reported as 1.2 for hyperscale data center with improved cooling [4].

Traditionally IT equipment in data centers are cooled by air thorough CRAC or CRAH with or without containment. There have been many research works conducted on improving total cooling power consumption and improving cooling efficiency through various methods. To meet the limitations of air-cooling for high power density IT equipment and avoid critical issues, hybrid cooling can be employed with liquid cooling to the high heat generating components. Dividing the heat load thus helps to improve the cooling efficiency by reducing amount of heat load on air-cooling. This paper discusses about the options to improve cooling efficiency by providing increased inlet air temperature. Providing proper inlet air temperature can also reduce total cost for cooling. The conventional method of calculating the benefit of higher operational temperature is just to find the balance between the IT power consumption and IDC cooling energy savings [5]. The maximum and minimum recommended operating temperature are 18 C and 27 C for long term reliability [1]. For typical air-cooled data center, if the inlet air temperature rises the facility power goes down for reduction in power consumptions by fans and chiller. But the server power goes up as the fans ramp up based on the increase of critical component temperatures. As this study is about a hybrid cooled server, the heat load for fans has been reduced and there is a provision to extend the range of operating temperature for this server. So, with increase of inlet air temperature, optimum component temperature and reduction in power consumption must be observed to reduce cooling cost.

SERVER UNDER STUDY

The system used for experimentation and CFD simulation is Cisco UCS C220 M3 server. It's 1U rack mount server and the heat sinks on processors are retrofitted with cold-plates. This server is based on Intel motherboard used in this study [6]. And brief description of the components is represented in the table below. The specifications are obtained from the manufacturer's data sheets [6,7,8] which are available in various online resources.

Table 1. List of Components and Specification

Components	Spec Sheet Data
CPU (2 Processors)	Intel Xeon® E5-2600 v2 or E5-2600 (TDP-135W)
DIMMs (Total 16)	(8x8GB+8x4GB, 1.35V, 1600MHz-96 GB)
Fan (Total 5)	Delta fan, 40x40x56mm, (TDP-15.6W)
Pump (Total 2)	ASETEK Direct to Chip CPU cooler (Total 2, TDP- 4 W)

The critical heat generating components are 2 CPUs, 16 DIMMs and 1 chipset(PCH). For cooling, 5 Delta DC fans are located at the entrance to draws air through the server and the pumps with the cold plate removes the heat generated from CPUs.

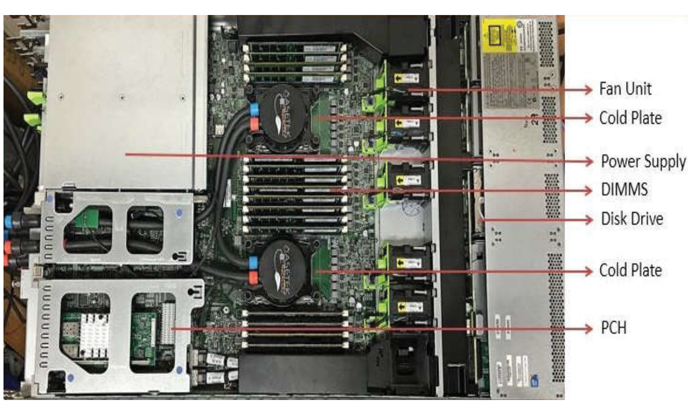

Figure 1: Cisco UCS C220 M3 server retrofitted with cold plates.

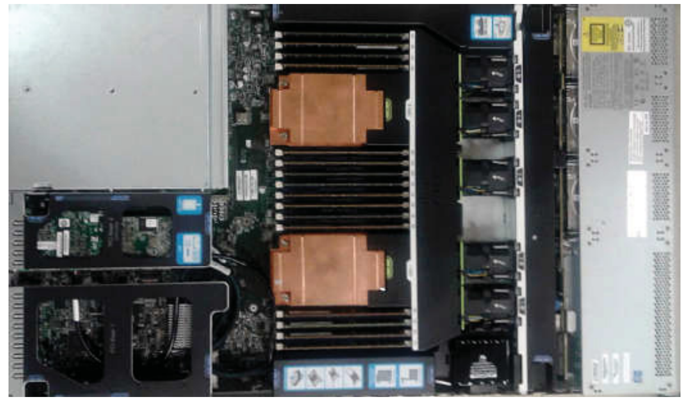

Figure 2: Cisco UCS C220 M3 server (air-cooled)

The setup for air-cooled server is different compared to hybrid-cooled server. The air flow from the fan to rest of the components is guided by a ducting placed before DIMMs for enhanced air-flow through the server. Both server is shown here in Figure 1 and 2.

EXPERIMENTAL STUDY

The air-cooled server is used in an experiment for raised inlet condition ranging from 15 C to 45 C in an environmental chamber to study the effect of leakage current and rise in component temperatures. The rise in temperature showed rise in both IT power and cooling power consumption. As CPUs are the maximum heat generating components in air-cooled server, the rise in inlet air temperature causes decrease the heat removal rate from the CPUs as the temperature difference shrinks. Moreover, the server is equipped with a duct to guide and accelerate the air flow through DIMMs. Even with precautions, the server can provide enough cooling as cooling efficiency drops after 35 ˚C [9].

In a separate study of the same server when equipped with hybrid cooling, the fan failure study showed that the air flow rate can be reduced without affecting the components (DIMMs and PCH). In other sense, there is provision for increasing the inlet air temperature providing same amount of air flow through the server by fixing a constant PWM for the fans. In the experimental study with five and three fans the flow rate was reduced by eliminating fans and varying PWM while maximum components temperatures were recorded for ambient condition (25˚C) [10]. This data can be used to validate a model to predict the component temperatures when inlet air temperatures are increased for a fixed flow rate. Moreover, the effect of raised inlet temperature with reduced number of fans can also be studied to determine cooling performance even at reduced flow rate.

MODELLING AND VALIDATION

K-Epsilon turbulence model is widely used model for turbulent flow modeling which is also commonly known as two-equation model. This model uses two variables; the kinetic energy of turbulence (k) and the dissipation rate of the kinetic energy of turbulence (ε) [11]. 6SigmaET is used for numerical analysis. The dimensions and specifications are measured and collected from the server and online resources as model data is not readily available. 6SigmaET's components like CPUs, DIMMs, HDD, fans, power supply, chipset and heat sinks are used from built-in object panel library. Properties like thermal, material, surface, etc. are defined as per requirements for the developing model [12]. Thermal dissipation power is applied to DIMMs, MOSFETs, chipset, PCI chipset and utilized same as in the experiments performed before. To simplify operations, and match behavior with actual components in a model, changes are made according to the specifications and experimental data. The model is validated with experimental data to define the flow rate and resistance through the server and opening at front, rear and top are adjusted to curve-fit the data. The model is also tested for different gridding to check discrepancy of data and the default grid setting is used while changing the target value for gridding. The final model contained about 8.5 million grids.

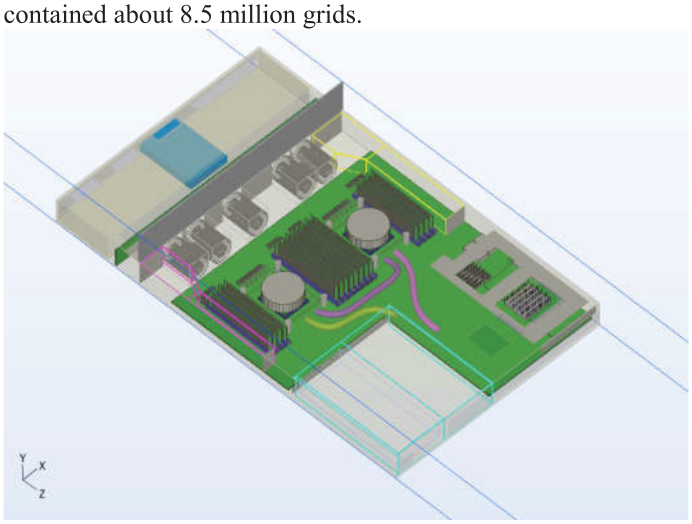

Figure 3: Iso-metric view of the model.

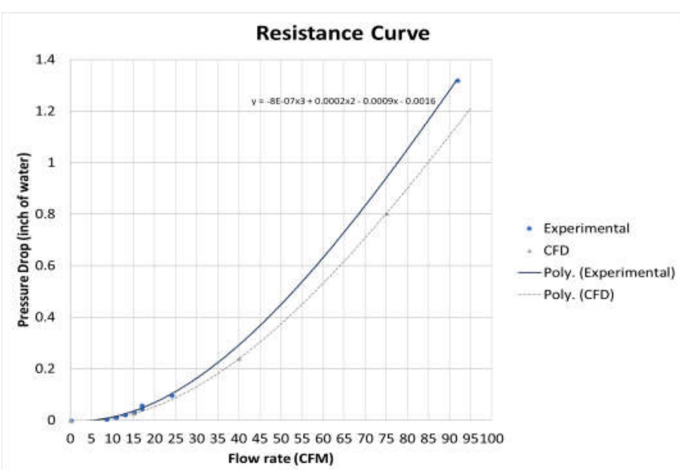

Figure 4: Resistance curve for the CFD Model.

As it is a simplified model, the details of the other components were not given priority and regarded as obstructions in model. The hybrid cooled server is modelled without the duct between fans and DIMMs as it is absent in hybrid cooled server during experimental studies. The pressure drop values across the server was found to vary up to 7%-8% of the measured values in lab. The pressure plots and velocity plots are shown for low flow rate when fans are operating at 3450 rpm.

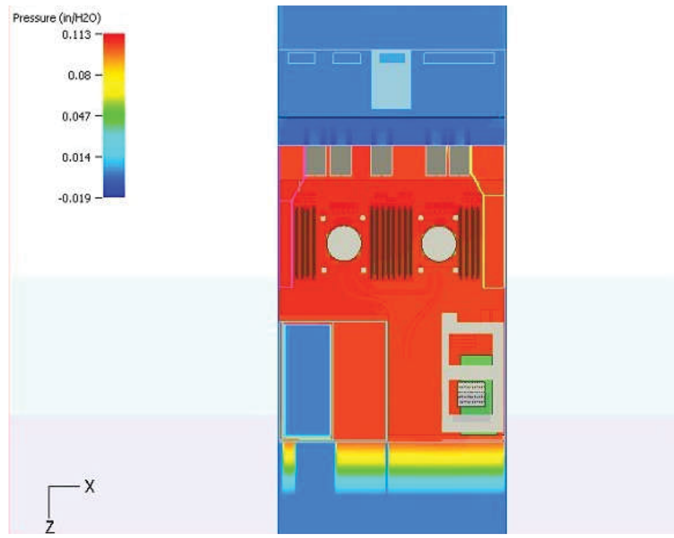

Figure 5: Pressure plot of the CFD model.

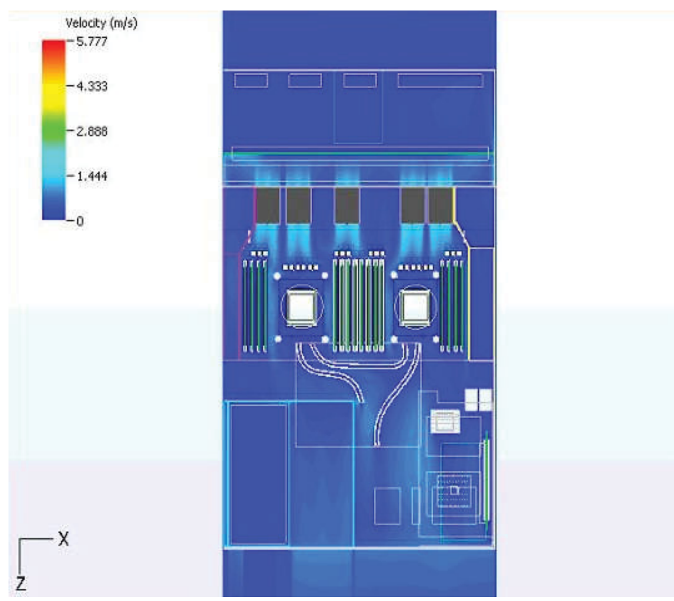

Figure 6: Velocity plot of the model.

CFD RESULTS

The model is updated with fan power curve from experimental results [9]. The inlet air temperature during the experiments was 25˚C ± 1˚C. But choosing the proper the rpm for the fans is a crucial criterion to match with experimental conditions. The fans in air cooled server ramped up for increase in inlet air temperature. The experiments with air cooled server can not be used to judge the situation in hybrid cooled server as there are changes inside the server for replacing the heat sinks with cold plates and removing duct in hybrid cooled server. But another study previous discussed above [10], showed the effect of flow rate with reduced fans. As the study included five fans with range of rpms for 25 ˚C room temperature, the idle state data is chosen to validate the CFD model data. According to the experimental data, the fans operate at an average rpm of approximately 3450 in a hybrid cooled server. Thus, the recorded temperature by IPMI tool is the baseline of comparison for CFD model. At 25 ˚C and 3450 rpm, the maximum DIMMs and PCH were found to be around 46˚C and 58˚C [10].

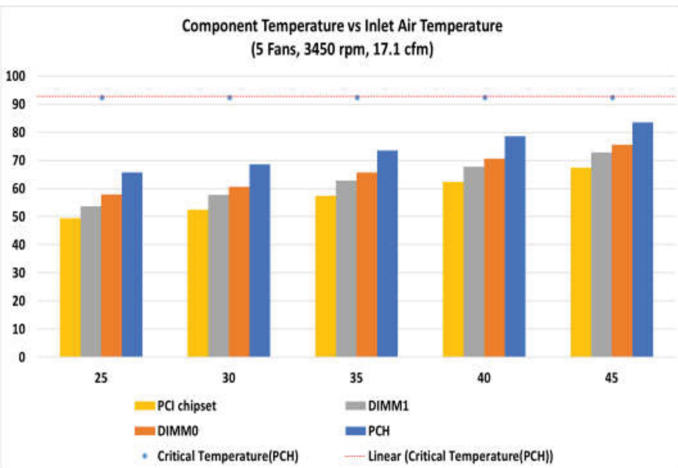

Figure 7: Component temperatures at different inlet temperatures when five fans are operating at 3450 rpm.

As memory was utilized up to 92.7%(maximum possible), the DIMMs, PCH, PCI chipset, MOSFETs are also stressed accordingly. The CFD results for five fans are plotted and shown here. The most critical component for the air-cooling is the PCH as it lies in a thermal showing zone where warmer air from the DIMMs cools the PCH. The critical value for PCH is shown in a dotted line in figure 4. Moreover, changing the heat sinks with cold plates changed the impedance to flow. The fans are operated at two different sets of rpms: 3450 and 7000 rpm. The 7000 rpm is chosen to compare performance and component temperature with the air-cooled server. The air-cooled server operated at 70% PWM with a rpm approximately 7000 rpm.

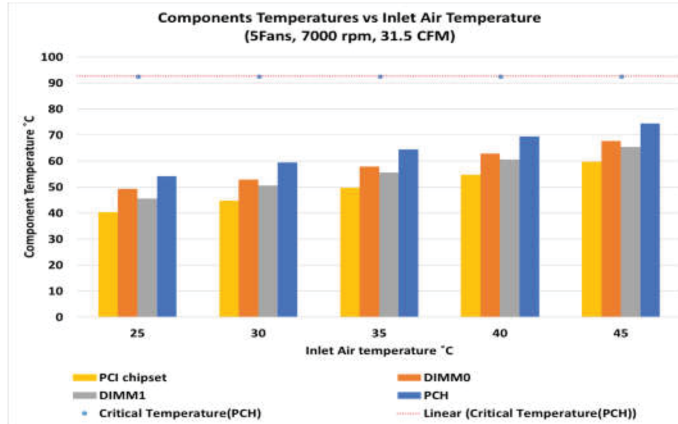

Figure 8: Component temperatures at different inlet temperatures when five fans are operating at 7000 rpm.

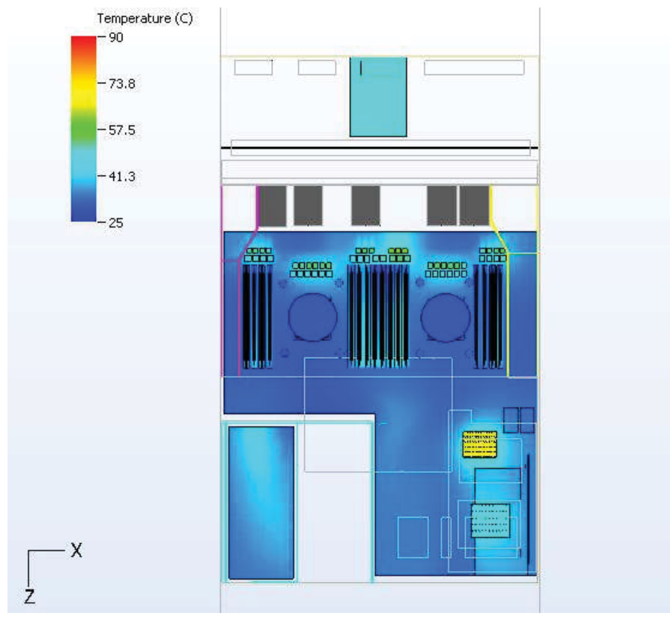

Figure 9: Contour Plot for surface temperature at 25˚C (3450 RPM)

As in this study we are increasing the inlet temperature, the overheating condition must be avoided, and the component must operate below the critical temperature at worst case scenario when the memory is stressed to the maximum. The contour plots are also provided for 25 ˚C and 45 ˚C for five fans at 3450 rpm.

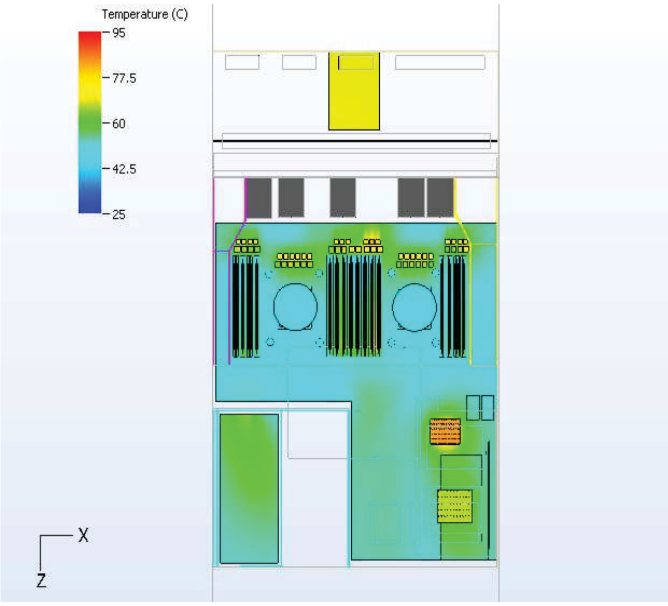

Figure 10: Contour Plot for surface temperature at 45C (3450 RPM)

With five fans it is expected that the component temperature should not exceed the critical temperature at stressed condition. Generally, in fans in server are internally controlled and ramps up according to the algorithm when component temperature exceeds the threshold values set in the coding. For minimum flow with five fans, the component temperature shows that PCH temperature is 83.51˚C for fans operating at 3450 rpm at 45˚C. For these PCB and DIMMs temperatures also increase. At 7000 rpm for 25 C case, DIMMs and PCH temperature are found to be around 49.2 ˚C and 54.15 ˚C.

EFFECT OF RAISED INLET AIR TEMPERATURE

The IT power consumption increases for raised inlet conditions to provide better heat transfer from the high generating components. On the other hand, increasing inlet air temperature raises the issue of leakage current and increase in power consumption for servers. The total data center power consumption as a function of coolant temperature, the computational state of chip, weather factors that employ free cooling has not shown improved energy efficiency even during maximized free cooling due to additional leakage power incurred by the associated higher coolant temperatures [13]. As the CPUs and other high heat generating components are cooled by liquid, temperatures of auxiliary components depend on inlet air temperature for a hybrid cooled server. This study is performed to understand the effect and determine range of inlet air temperature for a hybrid cooled server which will be eventually helpful to reduce cooling cost of a modular data center. Choice of operating temperature increases with a wider range of inlet air temperatures and data centers in a hot or warm climate can also utilize free cooling using the outside air. For raised inlet temperature the commonly faced challenges are performance, power and reliability. For optimum performance of IT equipment reasonable air inlet temperature and component temperatures should be maintained. Fan power is a function of the fan rpm and mass flow rate.

One of the way to reduce total cooling power ($Q_{air}=m_{air}C_p\Delta T$) is to increase inlet air temperature ($T_{ambient}$) while keeping the fan rpm and flow rate constant (i.e. m_{air} = constant). The reliability of components depends on junction and case temperature. In the previous experimental studies, it was found that component (DIMMs and PCH) temperatures are well below critical temperatures when this server is cooled by air. A study by El-Sayed et al. [14] indicate that the effect of temperature on hardware reliability is less than as it is assumed. In a separate study, it was found that DIMMs were operating reliably for elevated rack inlet temperature [15].

CONCLUSION

The CFD model is validated and compared with both air-cooled and hybrid cooled server. The inlet temperature for the model varied within 25-45 ˚C following ASHRAE guidelines (A1-A4). The limitations of inlet air temperature for air-cooled server can be overcome with hybrid cooled server and operating limits can be widened. So, this CFD study is conducted to determine the range of inlet air temperature for a typical 1U dense server. There is also a scope to increase operational hours for air-side economizers. True superiority quite often lies in the thermal management scheme and its capability for delivering precisely the performance needed while consuming the lowest power while meeting component specifications with the best acoustic signature [16]. From the study reliable operating temperatures can be determined for specific number of fan setup and respective rpms. The cooling power and acoustics depends on PWMs of fans and pumps. Operating at a minimum rpm without causing rapid rise in component temperature is the primary objective of this study. For future study, the data from this CFD study will help to experimentally test the same server to validate and make predictions for component reliability and safe operation in a data center.

REFERENCES

[1] ASHRAE TC 9.9 2011 "Thermal Guidelines for LiquidCooled Data Processing Environments", 2011 American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc.

[2] D. Keating, How green is your internet?, 2016, [online] Available: http://www.dw.com/en/how-green-is-your-internet/a-36691833

[3] Data Center Power Market by Solution Type (Power Distribution and Measurement Power Back-Up and Cabling Infrastructure) by Service Type by End User Type by Vertical and by Region - Global Forecast to 2020, 2015, [Online] Available:

http://www.researchandmarkets.com/research/bnfn4f/data_center_power.

[4] Shehabi Arman, Sarah Smith, A.Sartor Dale, E.Brown Richard, Magnus Herrlin, G.Koomey Jonathan, R.Masanet Eric, Nathaniel Horner, Inês Lima Azevedo, William Lintner, United States Data Center Energy Usage Report, Berkeley, CA:Lawrence Berkeley National Laboratory. LBNL-I005775, June 2016, [online] Available: http://eta.1bl.gov/publications/united-states-data-center-energy-usag.

[5] S. Zhang, N. Ahuja, Y. Han, H. Ren, Y. Chen and G. Guo, "Key considerations to implement high ambient data center," 2015 31st Thermal Measurement, Modeling & Management Symposium (SEMI-THERM), San Jose, CA, 2015, pp. 147-154.

[6] Cisco UCS C220M3 spec sheet. [Online]. Available: http://www.cisco.com/c/dam/en/us/products/collateral/servers-unified-computing/ucs-c-series-rack-servers/c220m3-sff-specsheet.pdf.

[7] Roy Zeighami, Winston Saunders, Henry Coles and Steve Branton, "Thermal performance modeling of hybrid liquid-air cooled servers," Thermal and Thermomechanical Phenomena in Electronic Systems (ITHERM), IEEE, pp. 583–587, 2014.

[8] Fan Manufacture Data. [Online]. Available: http://www.delta-fan.com/Download/Spec/GFB0412EHSDF00.pdf

[9] M. Sahini et al., "Comparative study of high ambient inlet temperature effects on the performance of air vs. liquid cooled IT equipment," 2017 16th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), Orlando, FL, 2017, pp. 544-550.

[10] Chowdhury U, Sahini M, Siddarth A, Agonafer D, Branton S. Characterization of an Isolated Hybrid Cooled Server with Failure Scenarios Using Warm Water Cooling. ASME. International Electronic Packaging Technical Conference and Exhibition, ASME 2017 International Technical Conference and Exhibition on Packaging and Integration of Electronic and Photonic Microsystems (): V001T02A002. doi:10.1115/IPACK2017-74028.

[11] S. A. Hall, G. V. Kopcsay, "Energy-Efficient Cooling of Liquid-Cooled Electronics Having Temperature-Dependent Leakage", *Journal of Thermal Science and Engineering Applications*, vol. 6, no. 1, pp. 12, 2014.

[12] M. M. Islam, U. Chowdhury, N. Inamdar and D. Agonafer, "Power consumption minimization in hybrid cooled server by fan reduction," 2017 16th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), Orlando, FL, 2017, pp. 850-856.

[13] S. A. Hall, G. V. Kopcsay, "Energy-Efficient Cooling of Liquid-Cooled Electronics Having Temperature-Dependent Leakage", Journal of Thermal Science and Engineering Applications, vol. 6, no. 1, pp. 12, 2014. [12] Prime95 Version 28.9 - CPU stress tool, [Online], Available: https://www.mersenne.org/download

[14] N. El-Sayed, I. A. Stefanovici, G. Amvrosiadis, A. A. Hwang and B. Schroeder, "Temperature management in data centers: why some (might) like it hot," in Proceedings of the 12th ACM Sigmetrics/ Performance joint international conference on Measurement and Modeling of Computer Systems, London, 2012.

[15] J. Fernandes, "Minimizing Power Consumption at Module, Server and Rack Levels Within a Data Center Through Design and Energy-Efficient Operation of Dynamic Cooling Solutions," University of Texas at Arlington, May 2015, pp. 65-69

[16] "IT Equipment Thermal Management and Controls" Whitepaper prepared by ASHRAE Technical Committee (TC) 9.9 Mission Critical Facilities, Technology Spaces, and Electronic Equipment. [Online], Available: https://tc0909.ashraetcs.org/documents/ASHRAE%202012%20IT%20Equipment%20Thermal%20Management%20and%20Controls_V1.0.pdf