

IPACK2017-74028

**CHARACTERIZATION OF AN ISOLATED HYBRID COOLED SERVER WITH
FAILURE SCENARIOS USING WARM WATER COOLING**

Uschas Chowdhury

University of Texas at Arlington
Arlington, Texas, USA

Manasa Sahini

University of Texas at Arlington
Arlington, Texas, USA

Ashwin Siddarth

University of Texas at Arlington
Arlington, Texas, USA

Dereje Agonafer

University of Texas at Arlington
Arlington, Texas, USA

Steve Branton

Data Center Solution at ASETEK
San Jose, California, USA

ABSTRACT

Modern day data centers are operated at high power for increased power density, maintenance, and cooling which covers almost 2 percent (70 billion kilowatt-hours) of the total energy consumption in the US. IT components and cooling system occupy the major portion of this energy consumption. Although data centers are designed to perform efficiently, cooling the high-density components is still a challenge. So, alternative methods to improve the cooling efficiency has become the drive to reduce the cooling cost. As liquid cooling is more efficient for high specific heat capacity, density, and thermal conductivity, hybrid cooling can offer the advantage of liquid cooling of high heat generating components in the traditional air-cooled servers. In this experiment, a 1U server is equipped with cold plate to cool the CPUs while the rest of the components are cooled by fans. In this study, predictive fan and pump failure analysis are performed which also helps to explore the options for redundancy and to reduce the cooling cost by improving cooling efficiency. Redundancy requires the knowledge of planned and unplanned system failures. As the main heat generating components are cooled by liquid, warm water cooling can be employed to observe the effects of raised inlet conditions in a hybrid cooled server with failure scenarios. The ASHRAE guidance class W4 for liquid cooling is chosen for our experiment to operate in a range from 25°C - 45°C. The experiments are conducted separately for the pump and fan failure scenarios. Computational load of idle, 10%, 30%, 50%, 70% and 98% are applied while powering only one pump and the miniature dry cooler fans are controlled externally to maintain constant inlet temperature of the coolant. As the rest of

components such as DIMMs & PCH are cooled by air, maximum utilization for memory is applied while reducing the number fans in each case for fan failure scenario. The components temperatures and power consumption are recorded in each case for performance analysis.

INTRODUCTION

A census conducted by Datacenter Dynamics in 2013 [1] indicated that worldwide data center power consumption grew up to 38.8 GW. The power density per square foot for data centers is increasing day by day with increased use of high density IT equipment. Thus, the cooling cost is increasing as cooling consumes approximately 25%-35% of the total energy of a production data center [2]. Improving the efficiency of data center cooling has become the motto to reduce the operating and maintenance costs. Typical costs for data centers can range from 10M to 20M per Megawatt. Of this cost, about 30% (within wide margins) is associated with cooling and heat removal [3]. For the increase in total power consumption, heat flux, and associated costs, data center managers are looking for an alternate to the traditional air cooling. A solution to reduce cooling costs that have been widely investigated in hybrid air-liquid cooling of data centers [4,6]. Hybrid cooling uses liquid cooling for high heat generating components while rest of the components are cooled by air. Hybrid cooling coupled with warm water cooling provides provisioning for reducing cooling power and costs, which is shown in a separate study [5]. Using coolant above ambient conditions can be a viable option for cooling IT equipment taking the advantage high specific heat and thermal conductivity of coolant. Characterization of hybrid

liquid-air cooled servers was done in a study showing key performance indicators including both heat recovery effectiveness and thermal loss parameters [6]. But operating at high temperatures also poses the risks of affecting the operation, serviceability or reliability of the IT equipment. Safety precautions and decisions for immediate response are necessary for smooth operation of the servers. Analysis of failure scenarios can provide a better perspective of the situations and prevent damage to any IT equipment. A study on failure scenarios of the dry cooler in chiller-less liquid cooled data center showed parametric analysis with a dynamic model [7]. In this experiment, a hybrid server is tested with both the pump and fan failure scenarios with warm water cooling to investigate options for reducing total cooling power.

TEST SETUP

The server for this experiment is Cisco UCS C220 M3 [8]. The server is 1U rack mount unit and retrofitted with ASETEK cold plates for processors. There are two CPUs with 10 cores in each, and have Thermal Design Power (TDP) of 135W. Additionally, there are 16 Dual in Line Memory (DIMMs), 8 DIMMs per CPU and installed memory is 96GB. The main heat generating components are CPUs, PCH and DIMMs. The server is also populated five 1U fans (40x40x56 mm) and 8 HDD. For our experiment, the fans and pumps are powered from outside and the Pulse Width Modulation (PWM) signal generator controls fan speeds. The coolant from the server is cooled by a miniature dry cooler (radiator). Two penetrable K-type thermocouples are used to measure the inlet and outlet temperature of the coolant through Data Acquisition (DAQ) unit. So, to perform the experiments with warm water cooling, the inlet temperature for coolant was controlled by changing the PWM of radiator fans via LabVIEW and Arduino UNO. The Schematic for the test setup is shown in Figure 1.

to the workstation. The power and RPMs for the pumps and fans are also measured by DAQ and Benchvue software. The power data is collected from time to time to monitor any unexpected change in power consumption of the server in response to the component temperatures.

TEST PROCEDURE

In an event of power failure, the most heat generating components needs cooling as UPS or generator is generally dedicated to providing power to IT equipment. The scope of the study is based on the effect of temperature rise of the server components to calculate the compensation in cooling in the event of failures. As the system is a hybrid cooled server, the main heat generating components are cooled by the inlet coolant and rest of the components such as DIMMs, MOSFETs, PCH and other auxiliary components are cooled by air drawn by the fans in the server. The server level study is performed in the sense to analyze temperature rise of individual components and find a baseline to predict safe service and operating condition for the failure scenarios on a server level. With respect to cooling, the failure scenarios were categorized as the failure of liquid cooling and air cooling.

The liquid cooling loop consist of two pumps connected in series and the coolant was cooled via an external heat exchanger (ASETEK miniature dry cooler). For failure scenarios, the cooling of CPUs is critical and pumps are driving force for the coolant in a loop. The pump failure scenario consisted of failing one of the pumps and observe the effect of temperature rise with different level server utilization. The inlet condition for coolant is also varied to characterize the system for high ambient conditions. The liquid cooling guideline for data centers provided by ASHRAE TC 9.9 [9] is followed for the experiments by varying inlet coolant temperature from 25°C to 45°C with an increment of 5°C. To evaluating the worst possible condition, the pump RPMs are varied as well for each coolant inlet temperature while running only one pump. Three RPMs are selected for the study and in each case, rpm is changed by setting a voltage (7V, 10V, 12V) in power supply unit. In each case, a computational workload of idle, 10%, 30%, 50%, 70%, 100% was applied for CPUs using a command line called “stress”, an internal tool for Ubuntu [10]. The effect of pump failure on the temperatures of CPUs and other components are monitored and recorded by Intelligent Platform Management Interface (IPMI) tool [11] in UBUNTU operating system of the server. The sensor data are collected every 3 seconds and server power data was also collected to observe the changes in power consumption. High temperatures are achieved for maximum server utilizations and the tests are repeated three times to ensure repeatability and reduce errors. Temperature rise is studied to decide safe operating conditions for an individual server in the event of single pump failure and to understand the control parameters.

There are five fans in the server and components other than CPUs are cooled by the airflow generated by the fans. Server fans are generally controlled by the built-in fan control

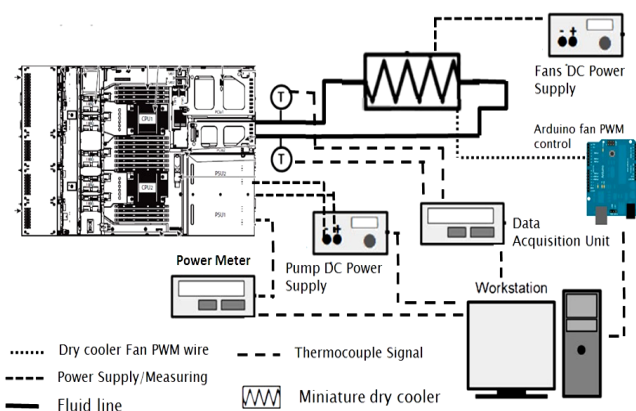


Figure 1: Schematic of test setup and data acquisition

The power consumed by the server is collected by an external power meter (Yokogawa CW121) which is connected

algorithm which depends mainly on the high heat generating components. As this server is retrofitted with cold plates, the high generating component, CPUs are cooled by the liquid coolant. So, the main heat generating components for air cooling are PCH and DIMMs. To understand the effect of temperature rise, fans are powered externally by a power supply and controlled by a PWM signal generating unit. For predictive fan failure scenarios, fans are disconnected one by one in each case. Fans in front of the pumps are turned off and later cases with fewer fans in front of the DIMMs are tried while keeping server utilization at maximum. In each case, the fan PWM is varied in five steps as 30%, 40%, 55%, 70%, 85% and 100% and maximum computational workload is applied for CPUs (100%) and memory is stressed up to maximum (92%) by a stressing tool “prime95” [12]. The component temperatures are recorded by IPMI tool and each stress test continued for 15 mins when steady state conditions are achieved. The tests are performed twice to check for repeatability and errors.

RESULTS

Pump Failure - The pumps are connected in series and provides coolant to the cold plates to cool the CPUs. Pumps are powered externally by a power supply. By changing the voltages, the pump rpm is varied and instantaneous rpms were measured through a DAQ. Three rpms are selected to run experiments ranging from 2500 to 3500 rpm. The test script for each case runs for 6 hours, with 30 minutes of stressing the CPUs at idle, 10%, 30%, 50%, 70%, 100% and 30 minutes of idling in between.

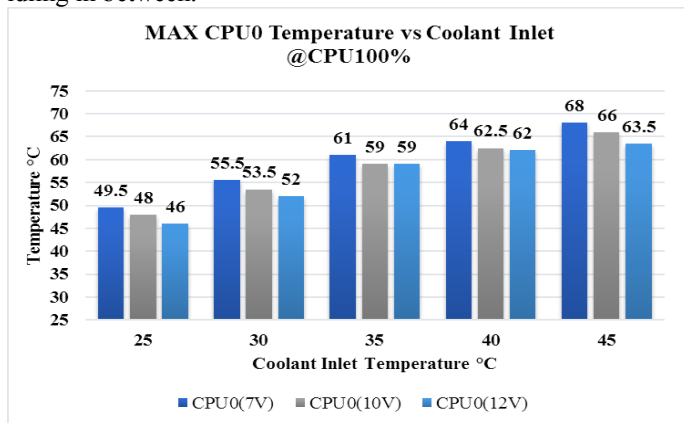


Figure 2: Maximum CPU0 temperature vs coolant inlet temperature with two pumps.

Each case is examined and CPU temperatures are plotted for different inlet temperature of the coolant to observe effect of warm water cooling with variable flow rate. The tests are repeated two times to ensure repeatability of data and changes in temperatures are recorded.

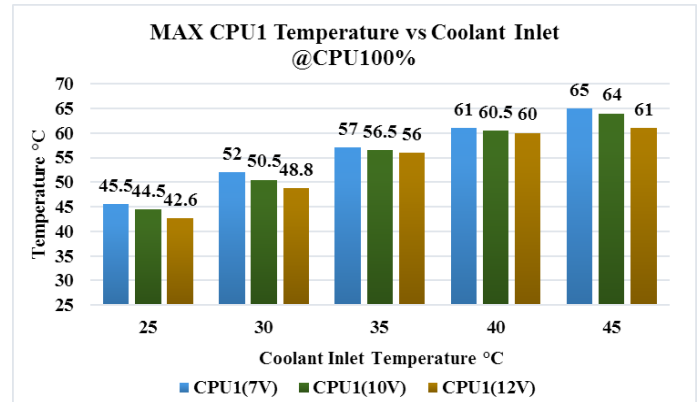


Figure 3: Maximum CPU1 temperature vs coolant inlet temperature with two pumps.

The temperature for CPU0 (downstream) was found to higher than CPU1(upstream). The maximum temperatures are observed for maximum CPU utilization. As the CPU0 exhibited higher temperature, the worst-case scenario will be the situation when pump for CPU0 fails. So, the downstream pump was turned off and the experimental procedure was followed as before.

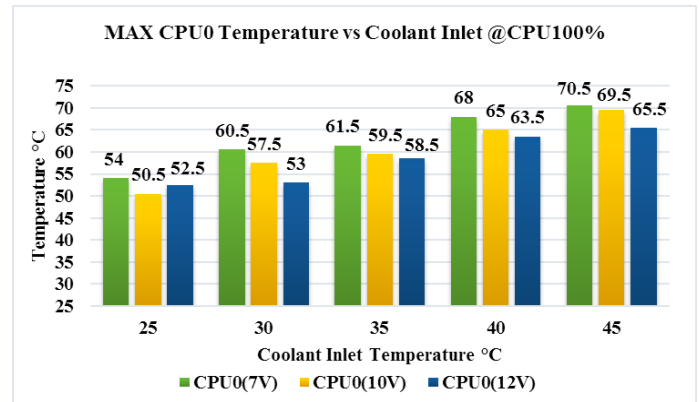


Figure 4: Maximum CPU0 temperature vs coolant inlet temperature with single pump in series.

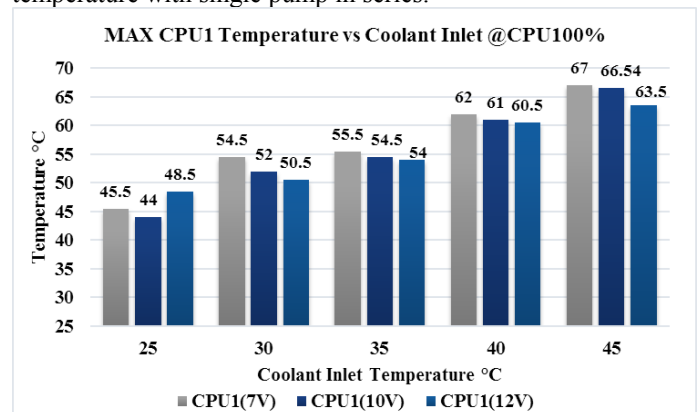


Figure 5: Maximum CPU1 temperature vs coolant inlet temperature with single pump in series.

By powering only one pump, same stress load has been applied and change in CPU core temperatures are monitored by “Psensor” tool [13]. Necessary precautions are followed while monitoring the CPU temperature. As the pumping power varies higher temperatures are observed for maximum CPU utilization. For failure of one pump in series, the maximum temperature for CPU0 is found to be 70.5°C at 45°C inlet and 7V supply voltage input (2500 rpm) which is for 7V power input (2500 rpm) and this case is considered to be the worst-case scenario. Due to pump failure, there is an average rise of 2.5°C in CPU temperatures. The temperature rise has been reducing from 25°C as we go to 45°C inlet coolant.

Fan Failure - The initial set of testing has shown that the fans are internally controlled based on the CPU temperatures. This setting is not applicable as the CPU are liquid cooled. Hence, for our analysis purposes, the fans are powered and controlled from outside. The study for fan failures is performed by turning off the fans and providing different PWM to the rest of the fans. The computational load is provided by “prime95” tool to both CPU and memory and sensor data for CPUs, DIMMs and PCH are collected by “IPMI” tool. The test is conducted for five (base line), four and three fans to predict the failure scenarios. In a previous study [14], it is found that a minimum flow rate of 10 cfm is required for this server. Fans in front of the CPUs were turned off one by one for the case study of five and four fans. Each case is experimented for different fan PWMs at 100%, 85%, 70%, 55%, 40%, 30%. Corresponding data is collected and plotted to observe the effect of reduced flow rate on the component temperatures as the fan number was decreased. Each test is continued for 15 minutes until steady state is achieved for the temperatures and test are performed twice to ensure repetition of the data

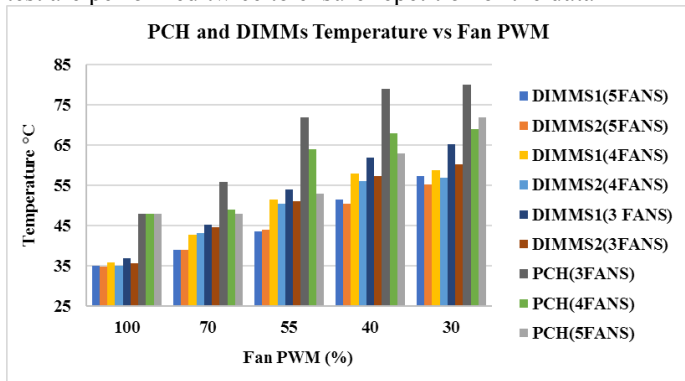


Figure 6: PCH and DIMMs temperature vs fan PWM for fans in front of cold plates

Location of fans is another factor for fans in a 1U server and fans ahead of DIMMs are of prime concern for air cooling. So, considering the worst-case scenarios, the fans in line with DIMMs are turned off and fan PWMs are changed ranging from 30% to 100%. Three cases (config 1,2,3) are studied by turning off fans for each side and in the middle, one at a time.

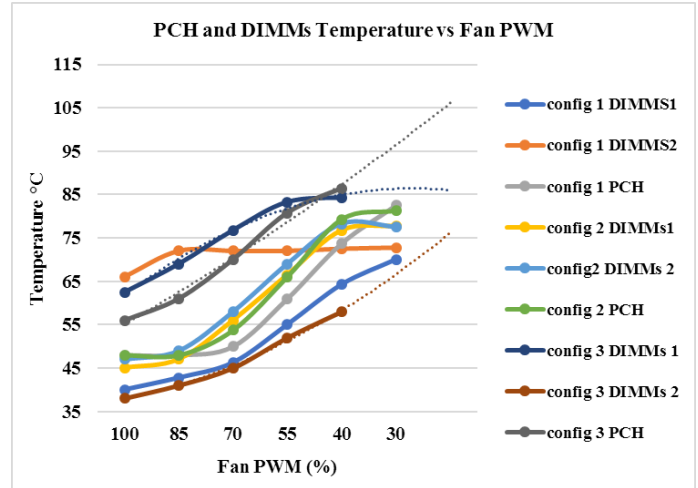


Figure 7: PCH and DIMMs temperature vs fan PWM for fans in front of DIMMs

Each test is continued for 15mins until steady state is achieved. For third case (config 3), 30% PWM may cause damage to PCH as reduced air flow has caused a rapid rise in temperature close to the critical temperature at 40% PWM. The data for the lowest PWM has been extrapolated. Sensor data is collected with IPMI tool. Maximum DIMMs temperature is calculated from the averaged values over the last five minutes of the test run.

TRANSIENT ANALYSIS

Full Pump Failure - The Pumps in this hybrid cooling generates flow rate for the coolant to pass through the cold plates and carry away the heat from both the CPUs. The CPU temperature depends mainly on utilization and rate of heat removal. The flow rate in conjunction with cold plate heat exchanger can remove heat from the surface of the processor very effectively [15]. Turning off one of the pumps showed the rise in temperatures for coolant as well as CPU junction temperature. But in the event of total cooling power failure, pumps will not be able to provide the amount cooling needed and the CPUs may be damaged due to overheating. In this experiment, the server is tested to approach the maximum CPU die temperature and calculate the time for temperature rise and monitor the pattern. Both the pumps are turned off and instantaneous sensor data was collected. The CPU was stress from idle to max with 25% increase in each step and results are plotted for idle, 50%, 100% CPU utilizations as shown in the graph. As the utilization increased, the die and junction increased and the lines for temperature rise for separate utilization can be noticed from the graph (right to left). Taking critical temperature into consideration, power supply to the pumps were controlled by monitoring the CPU core temperatures.

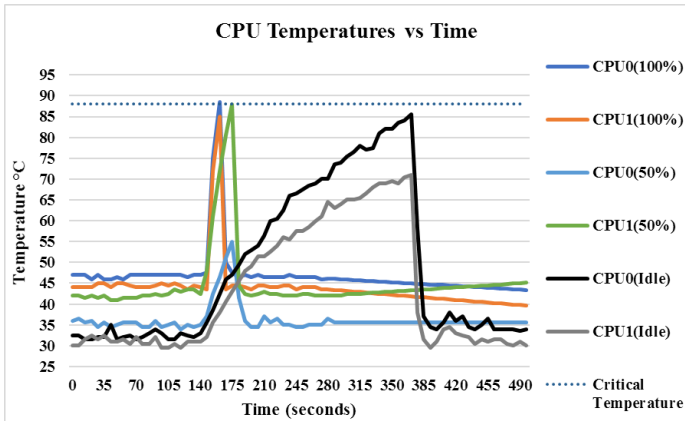


Figure 8: Transient analysis of the CPU temperatures with different CPU Utilisation for full pump failure scenario.

Full Fan Failure - Servers are provisioned for the failure scenarios of fans in air cooling and safety features for saving the components from such critical conditions can be found in reference [16]. But considering the worst case when all the fans fail, the temperature rise is expected but the risks are far less for hybrid cooled servers compared to air-cooled servers. To examine the condition experimentally and measure power consumption during the event, all the fans were powered off externally and the maximum computational load was given with “prime95” for both the CPU and memory. As the CPU is being cooled by the liquid coolant, the high temperature is not noticed for CPUs for liquid cooling. The fans are stopped around 147 seconds and powered again around 780 seconds. Temperature of other air-cooled components like PCH (blue) and DIMMs (labeled as in reference [8]) are plotted for transient analysis as shown in figure 9.

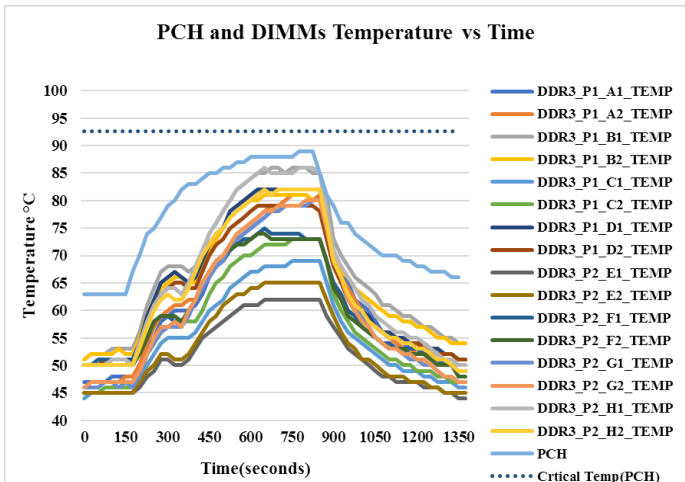


Figure 9: Transient analysis of the PCH and DIMMs temperatures with maximum CPU & memory Utilisation for full fan failure scenario.

POWER CONSUMPTION

Typical data centers are defined by the floor area of the IT facility and the power that the facility can support per unit area [5]. Calculating the power consumption from each of the server at various conditions provides a guideline for calculating the Power Usage Effectiveness (PUE) and characterize the data center cooling efficiency. In hybrid cooling most of the produced IT heat is transferred in liquid cooling loop, amount of air cooling needed for other components is reduced. But quantifying the amounts is beneficial for calculating total cooling power. As the experiment is conducted with an isolated server, the pPUE guideline [17] is followed to calculate total cooling power required. The total cooling power includes cooling power (fan power and pumping power) required within the IT equipment boundary. The powers consumption shows higher values when both CPUs and memory are utilized to the maximum. The graphs for pump and fan failure shows different but almost uniform power consumption for different cases. As the fans and pumps are externally connected the power for liquid and air cooling are measured separately and compared in the graphs.

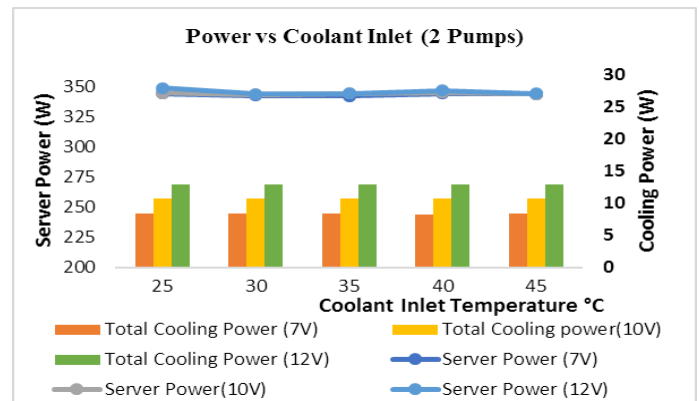


Figure 10: Server and total cooling power vs coolant inlet temperature for two pumps.

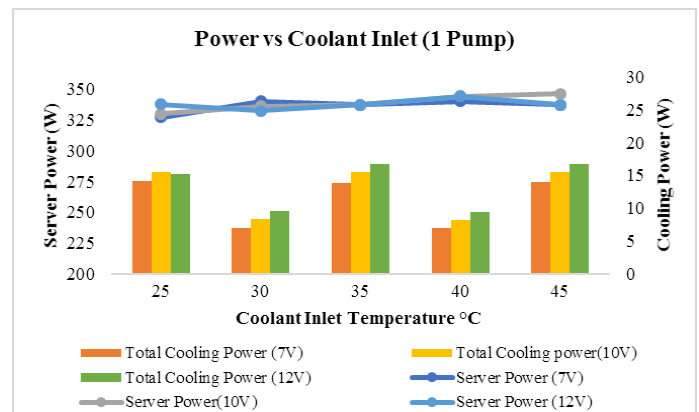


Figure 11: Server and total cooling power vs coolant inlet temperature for single pump.

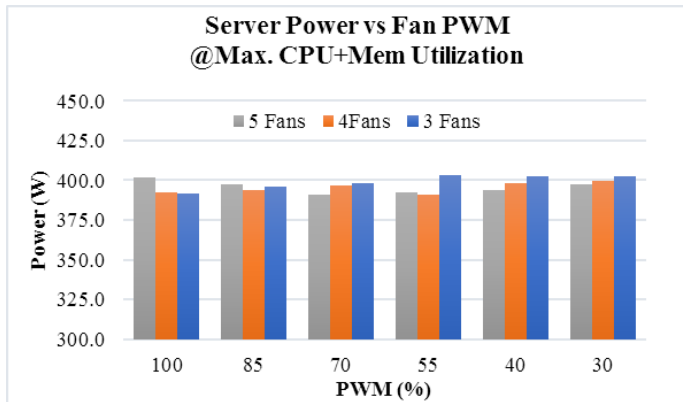


Figure 12: Server power consumption with different fan PWM for maximum CPU and memory utilization.

From the graphs, it is evident that the server consumes uniform power at different coolant inlet conditions both for the pump and fan failure scenarios. But the cooling power is varied in pump failure scenario (single pump) by changing the fan PWM to observe the changes in cooling power.

The power consumption for fans is calculated for both internally and externally controlled situation and consumed about 11.2W and 5.65 W respectively with maximum CPU utilization. The pumping power is calculated from the power supply unit. Two pumps require approximately 7.17W and single pump requires about 3.76 W. Again changing the rpms of pumps from 2500 rpm to 3300rpm and comparing single pump (7V) to two pumps (12V), maximum 81.32% reduction in pumping power is observed. With the use of only three fans and single pump, total cooling power reduces from 12.91W to 6.98W with a possible reduction of 45.93%.

SUMMARY

In this paper, different possible failure scenarios for an isolated hybrid cooled server has experimented and warm water cooling is employed to experimentally investigate the worst-case scenarios and opt for system redundancy. The system uses both liquid and air cooling and optimizing the cooling can provide a way to reduce cooling cost. Comparative analysis of two pumps to one pump shows the rise in CPU temperatures but the temperatures remain within safety zone well below the critical temperature even in the worst possible condition with warm water cooling. The transient analysis with pump failure shows less time than that for fan failure scenario. In other words, maintaining the pumps is more crucial for this hybrid cooled server. The time for response, repair and replacement can be calculated from transient analysis and safety measurements for hybrid cooled servers. Here, reducing the cooling power is another aspect to improve cooling efficiency. The power consumption analysis describes a gap of watts which shows a possible reduction in cooling power (45.93%) is possible on a server level without causing overheating and hampering IT equipment life. Pumps with low flow rate can be used for this cooling loop. But a rack level study can help to

determine the required flow rates for servers in a typical data center. So, future study can be based on the pressure drops and flow rate required for optimum cooling to evaluate the parametric performance of the servers running over a wide range of operational conditions.

ACKNOWLEDGMENTS

We would like to thank Steve Branton, Anderson Tsai, ASETEK Inc; Roy Zeighami, Microsoft, Cisco; Dr Robert Woods, UTA and IAB mentors, NSF ES2 I/UCRC for their input in preparing the document.

REFERENCES

- [1] "Is the industry getting better at using power?", *Datacenter Dynamics, Focus*, vol. 3, no. 33, pp. 16-17, 2014.
- [2] "Vision and Roadmap - Routing Telecom and Data Centers Toward Efficient Energy Use", May, 2009, www.eere.energy.gov/industry/datacenters/pdfs/vision_and_roadmap.pdf
- [3] Roy Zeighami, Winston A. Saunders, Henry Coles, "Thermal performance modeling of hybrid liquid-air cooled servers", *Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm)*, 2014.
- [4] Bruno Michel, "Energy efficiency and green technology: Direct use of waste heat to minimize carbon-dioxide emissions", www.zurich.ibm.com/st/energy_efficiency/zero_emission.html
- [5] Brandon A Rubenstein, Roy Zeighami, Robert Lankston, Eric Peterson, "Hybrid cooled data center using above ambient liquid cooling", 12th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), 2010.
- [6] 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.
- [7] Tianyi Gao, Milnes David, James Geer, "A dynamic model of failure scenarios of the dry cooler in a liquid cooled chiller-less data center", 31st Thermal Measurement, Modeling & Management Symposium (SEMI-THERM), San Jose, CA, 2015.
- [8] Cisco UCS C220M3 spec sheet, <http://www.cisco.com/c/dam/en/us/products/collateral/servers-unified-computing/ucs-c-series-rack-servers/c220m3-sff-specsheet.pdf>
- [9] Robin A. Steinbrecher, Roger Schmidt, "Data Center Environments ASHRAE's Evolving Thermal Guidelines", pp 48, December 2011.
- [10] Linux man Page, "Ubuntu Manuals", <http://manpages.ubuntu.com/manpages/zesty/man1/stress.1.html>

- [11] Intel, "Intelligent Platform Management Interface," Intel, <http://www.intel.com/content/www/us/en/servers/ipmi/ipmi-home.html>.
- [12] Prime95 Version 28.9 - CPU stress tool, <https://www.mersenne.org/download>
- [13] Psensor, A Graphical Temperature Monitoring Application, www.wpitchoune.net/psensor
- [14] M. M. Islam, "Experimental Characterization of Hybrid Cooled Cisco Servers Including the Effect of Warm Water Cooling", MS Thesis, University of Texas at Arlington, May 2016.
- [15] "How Liquid Cooling Works", www.asetek.com/desktop/technology/how-liquid-cooling-works
- [16] Cisco UCS C220 Server Installation and Service Guide, www.cisco.com/c/en/us/td/docs/unified_computing/ucs/c/hw/C220/install/C220/replace.html
- [17] Victor Avelar, Dan Azevedo, Alan French, "PUE™: A Comprehensive Examination of The Metric", The Green Grid, www.datacenters.lbl.gov/sites/all/files/WP49-PUE%20A%20Comprehensive%20Examination%20of%20the%20Metric_v6.pdf