

A Server-Level Test System for Direct-To-Chip Two-Phase Cooling of Data Centers Using a Low Global Warming Potential Fluid

Qingyang Wang^{1,*}, Serdar Ozguc¹, Akshith Narayanan¹, Richard W. Bonner III¹

¹Accelsius, Austin, TX, USA

*qwang@accelsius.com

Abstract— Cooling of data centers is becoming increasingly challenging with the surging power density, and direct-to-chip two-phase cooling is a promising solution with high cooling capacity and efficiency. Due to the difficulty in accurately modeling two-phase heat transfer under complex configurations, experimental investigations of direct-to-chip two-phase cooling under practical conditions are needed. In this work, we developed a server-level thermohydraulic experimental test system to characterize the performance of this novel and promising cooling solution. The system is a fluid circulation loop consisting of a test server sled, a refrigerant reservoir, a condenser, and a liquid pump. Refrigerant R1233zd(E) is used in this work, which has a low global warming potential of 1. The test server sled has two mounted thermal test vehicles (TTVs) with a maximum power rating of 1000 W each, which are used to simulate two Intel Sapphire Rapids CPUs. Two microchannel cold plates are attached onto the two TTVs to provide cooling by vaporizing the liquid refrigerant flowing inside. Case temperatures, fluid temperatures and pressures, pressure drops across multiple components in the loop, and liquid flow rate are measured using sensors implemented in the cooling loop. Preliminary tests demonstrated stable and efficient cooling of CPU power up to 1000 W with case-to-fluid thermal resistance below 0.02 K/W. At these power levels, we also showed that the flow rates between two parallel cold plates are balanced when the heat is applied nonuniformly. Our work establishes a server-level test system capable of characterizing various working parameters and system configurations, providing valuable insights into both the fundamental understanding and the practical realization of the two-phase direct-to-chip cooling solution for data centers.

Keywords— *direct-to-chip cooling, two-phase heat transfer, data center cooling, thermal test vehicle*

I. INTRODUCTION

With the surging power capacities of high-performance processors demanded by emerging technologies including artificial intelligence and machine learning, data centers are expected to face severe thermal management challenges resulted from increasing power density. The thermal power of the most advanced microchips is already over 300 W for central processing units (CPUs) and over 700 W for graphics processing units (GPUs) [1], bringing the power of a fully packed data

center rack to over 40 kW, which is beyond the capabilities of traditional air cooling. Moreover, with climate change causing frequent extreme weather conditions, especially extremely high outdoor temperatures in the summer, air cooled data centers could face significantly increased electricity cost, power grid instability, and water shortage, which also limits the ability of cooling the server racks using air and results in data center down time. Consequently, liquid cooling solutions for data centers are considered inevitable to complement or even replace air cooling and address the issues by offering higher cooling capacity and efficiency with minimal water consumption. [2].

Among the liquid cooling solutions, direct-to-chip two-phase cooling offers unique advantages [1, 3]. Compared with immersion cooling, direct-to-chip strategy allows easier retrofit of existing data centers. Compared with single-phase water cooling, by using dielectric fluids, two-phase cooling solutions could prevent the disastrous damage of IT equipment in the case of fluid leakage. In the academic community, microchannel-based two-phase heat sinks have been demonstrated to achieve high heat flux cooling in lab-scale systems [4-8], showing great potential for cooling high performance chips. Researchers have developed various models for microchannel flow boiling [9-12], analyzing the contributions from different heat transfer mechanisms and predicting the heat transfer performances. However, two-phase systems intrinsically have complicated interactions between thermal and fluid transport processes, making it extremely difficult to accurately model and predict the thermohydraulic behaviors in a complex system. Therefore, it is of great importance to conduct experimental investigations of the two-phase cooling solutions with configurations matching practical application conditions.

In this work, we developed a thermohydraulic test system at the server-level to study the performance of the two-phase direct-to-chip cooling solution for data centers. The system is a fluid circulation loop including a test server sled, a refrigerant reservoir, a liquid pump, and a condenser. Two thermal test vehicles (TTVs) are mounted on the server sled to simulate thermal power generation from high-performance CPUs, and two-phase microchannel cold plates are attached onto the TTVs

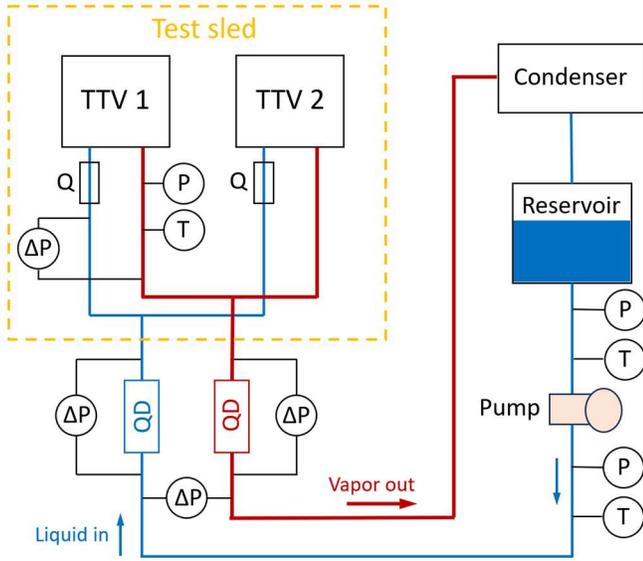


Fig. 1. Schematic of the test system

as heat sinks. Case temperatures, fluid temperatures, fluid pressures, pressure drops across multiple components in the loop, and volumetric flow rate of coolant in the loop are measured using sensors implemented in the cooling loop. A sustainable fluid with low global warming potential is used. Preliminary tests are conducted, and the results are analyzed and discussed. Stable and efficient cooling of CPU power up to 1000 W with case-to-fluid thermal resistance below 0.02 K/W is demonstrated, and the pressure drop of different components are reported. This work presents a server-level test system capable of characterizing various working parameters and system configurations, providing valuable insights into both the fundamental understanding and the practical realization of the two-phase direct-to-chip cooling solution for data centers.

II. EXPERIMENTAL SYSTEM

A. Fluid Circulation Loop

A fluid circulation loop is developed to characterize the thermo-hydraulic performance of a direct-to-chip cooling solution for a single server sled with two CPUs under a variety of conditions, including different cold plate geometries and layouts, different fluid types, different working parameters (pressure, temperature and fluid flow rate), different quick-disconnect (QD) fittings, and different heat loads. The schematic of the test loop is shown in Fig. 1, which contains a condenser, a refrigerant reservoir, a pump, a test server sled, and the fluid distribution tubings/hoses. The test server sled is connected to the loop by two QDs, which are used in practical server racks to allow easy installation and hot swapping of servers. The liquid pump drives liquid refrigerant from the reservoir to flow through the hoses and the inlet QD (termed liquid QD) and enter the test sled. Liquid vaporizes in the cold plates to dissipate the heat generated by the simulated CPUs in the test sled, and then passes through the outlet QD (termed vapor QD) and enters the condenser as vapor (or liquid-vapor mixture). Both QDs are placed horizontally, matching their position and orientation in a

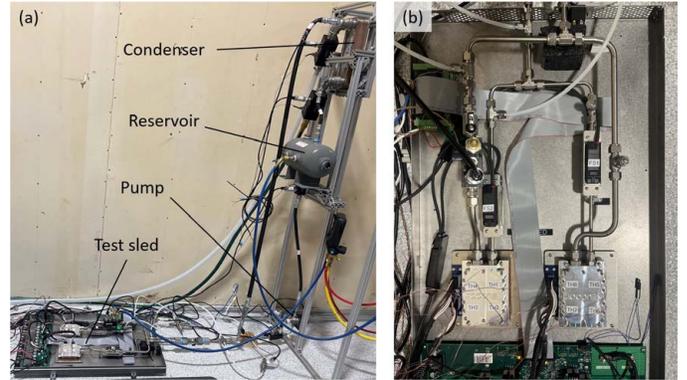


Fig. 2. Photographs of (a) the test system and (b) the test server sled

real server rack. The condenser used in this study is a plate heat exchanger, which condenses the refrigerant vapor back into liquid using cooling water. The condenser is oversized so that the heat exchanger capacity is not limiting the cooling performance at the server sled. The condensed liquid flows into the reservoir and completes the cycle. Fig. 2a shows the photo of the system, and Fig. 2b shows the test server sled including two mounted TTVs, which will be introduced below.

B. Thermal Test Vehicle

Each TTV used in this study contains a simulated CPU with performance matching the Intel Sapphire Rapids package. Four ceramic heaters are placed underneath a copper integral heat spreader to provide simulated CPU heat generation. Each ceramic heater can generate 250 W of heating power and has a surface area of 1 cm², delivering 250 W/cm² of heat flux and 1 kW of total power in each TTV. Four T-type thermocouples are attached onto the surface of the integral heat spreader of each TTV at positions directly on top of the ceramic heaters to measure the chip package case temperatures. The measured temperatures are then used to calculate the case-to-fluid thermal resistances of the cold plate. The microchannel cold plate is mated to the integral heat spreader with a layer of thermal interface material filled in between, with the cold plate substrate surface oriented horizontally.

C. Working Fluid

The working fluid selected in this study is R1233zd(E). It has an ultra-low global warming potential (GWP) of 1, making it an environmentally attractive replacement for R245fa with GWP value over 1000. It also has no ozone depletion potential. The fluid is non-flammable and non-conductive, which are desired properties for data center coolant to ensure safety of the chips and electronic equipment in rare cases of accidental coolant leakage. Table I lists the thermophysical properties of this fluid at 20 °C. Fig. 3 shows the saturation pressure-temperature curve of R1233zd(E), obtained from NIST REFPROP, showing a relatively low pressure (close to atmospheric pressure) under data center operating conditions (case surface temperature below 65 °C), which minimizes the possibility of leakage, lowers the pressure rating requirements for fittings and system components, and also enables the implementation of cold plates with small substrate thickness for

TABLE I. THERMOPHYSICAL PROPERTIES OF R1233zd(E) AT 20 °C

Property	Unit	Value
Liquid density	kg/m ³	1275
Vapor density	kg/m ³	6.1
Liquid viscosity	μPa·s	300.7
Vapor viscosity	μPa·s	10.1
Latent heat of vaporization	kJ/kg	193.7
Liquid specific heat	kJ/(kg·K)	1.2
Vapor specific heat	kJ/(kg·K)	0.81
Liquid thermal conductivity	W/(m·K)	0.084
Vapor thermal conductivity	W/(m·K)	0.010
Surface tension	mN/m	15.2

lower conductive thermal resistance. For the test results reported in this work, the fluid saturation temperature in the reservoir is close to 30 °C.

D. Measurement

In the current system (Fig. 1), the case temperatures of the two TTVs are measured using eight T-type thermocouples as described above. The temperature and pressure of the fluid at multiple positions in the loop are also measured using T-type thermocouple probes and Omega pressure transducers, respectively. As shown in Fig. 1, the measurement positions include the outlet of the fluid reservoir (inlet of the pump), the outlet of the pump, and the outlet of the cold plate. Four differential pressure transducers are placed in the loop to measure the pressure drop of different components, including the QDs for fluid entering and exiting the server, the cold plate itself, and the total pressure drop of the server sled. Two clamp-on ultrasonic flow rate meters are installed on the inlet tubes of the two cold plates to obtain the volumetric flow rate of coolant in each cold plate. The data from all thermocouples, pressure sensors and flow meters mentioned above are recorded using a Keysight data acquisition unit during the experiments. The measurement uncertainties of the quantities in this system are listed in Table II.

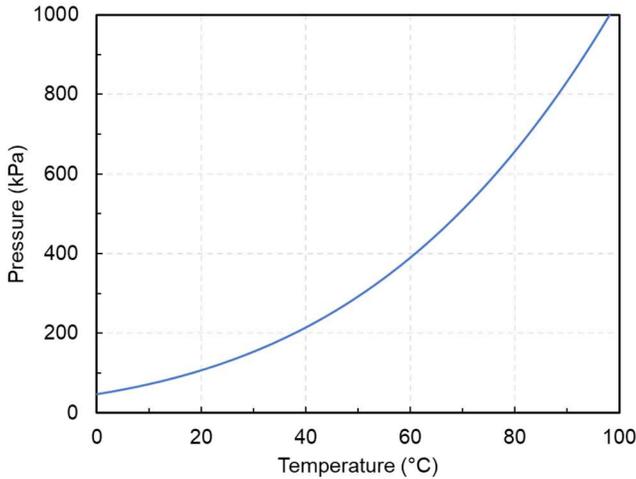


Fig. 3. Saturation curve of refrigerant R1233zd(E)

TABLE II. MEASUREMENT UNCERTAINTIES

Measurement	Sensor/Equipment	Uncertainty
Temperature	T-type thermocouples	±0.5 °C
Pressure	Omega PX309 Absolute	±1%
Pressure drop	Omega PX409 Wet/Wet Differential	0.08%
Flow rate	Keyence FD-X	0.3% (repeatability)
Heating power	Proprietary software	<10 W

With the case temperatures and fluid temperatures measured, the case-to-fluid thermal resistance for a specific TTV under a given condition is obtained as the steady-state temperature difference between the case temperature and the fluid saturation temperature divided by the heat load,

$$R_{th} = \frac{T_{case} - T_{sat}}{P_{CPU}} \quad (1)$$

where P_{CPU} is the CPU heat load (power). T_{case} here is calculated by averaging the four case temperatures measured using the four thermocouples embedded in one TTV. The fluid saturation temperature T_{sat} is measured by the thermocouple positioned at the outlet of the cold plate, assuming saturated two-phase mixture is exiting the cold plate and minimal pressure drop from the cold plate to the outlet (because the major pressure drop of the cold plate comes from the flow restrictor implemented at the inlet of the cold plate). It is worth noting that for single-phase convection-based heat sinks, the fluid inlet temperature is usually used to calculate the case-to-fluid thermal resistance. However, in our current work, the outlet fluid temperature (assumed to be the fluid saturation temperature) is used instead, because for our boiling-based microchannel cold plates, the heat transfer contribution from sensible heat is significantly smaller than latent heat and is nearly negligible. By neglecting sensible heat, the exit vapor quality x for a given heat load condition is calculated using the volumetric flow rate Q as:

$$x = \frac{P_{CPU}}{Q \rho_l h_{fg}} \quad (2)$$

where ρ_l and h_{fg} are the fluid density and latent heat, respectively.

III. RESULTS AND DISCUSSION

A. Case Temperature Profile

The steady state thermohydraulic performance of the cold plate under varying working conditions are characterized by fixing the fluid pumping power and increasing the CPU heat load in a stepwise manner with 100 W interval. The fixed pumping power provides a constant flow rate for each CPU power. Measurements are collected with a five-second interval. Each step of CPU power level is maintained for over 50 seconds to ensure a steady state has been reached. Fig. 4 shows the case temperature profile for a typical experiment with the refrigerant flow rate at 500~590 mL/min for one cold plate, with the top figure (a) showing the variation of CPU power over time, and the bottom figure (b) showing the case temperatures measured by two thermocouples embedded in the TTV. Here, only two

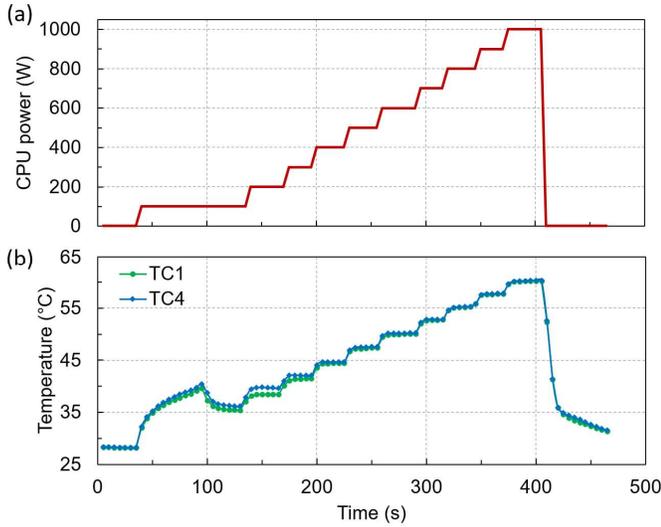


Fig. 4. Measured case temperature profile with CPU power variation

thermocouples out of the four are presented, named TC1 and TC4, which are positioned closest to the inlet and the outlet of the cold plate, respectively, as shown in Fig. 2b. Consequently, they represent the lowest and the highest temperatures among the four measurement positions on the TTV case surface, respectively. The temperature difference between TC1 and TC4 is minimal, indicating efficient heat spreading of the TTV.

Figure 4 shows that for most CPU power levels except 100 W, the system generally stabilizes within 10 seconds, represented by the stable case temperature profile. The case temperatures at the highest power of 1000 W are ~ 60 °C, showing efficient cooling of the CPU at such high power. When the CPU power is increased from zero to 100 W, the case temperatures rise to ~ 40 °C and then drops and stabilizes at $35\sim 37$ °C, showing a temperature overshoot, which is resulted from the boiling incipience delay. When the CPU power is increased from zero, the TTV is initially cooled by inefficient

single-phase convection and thus the case temperature keeps increasing. When the nucleation sites inside the cold plate is activated after the surface temperature reaches certain value, bubbles start to form and grow, and boiling occurs, which brings significantly enhanced two-phase heat transfer with smaller thermal resistance and consequently decreases the case temperature. Temperature overshoot during boiling incipience is observed quite commonly for two-phase heat transfer in the literature [13, 14], especially for highly wettable low surface tension fluids, which include the R1233zd(E) used in this work. Although an overshoot below 5 °C is not very significant, it can possibly cause CPUs to throttle performance under some conditions, which is not desirable. Future efforts to modify the cold plate geometry and surface condition are underway to improve the boiling incipience by promoting early bubble nucleation during startup operation.

B. Cooling Performance

For each power level, the data point is taken at the steady state segment. The data is then processed to calculate the thermal resistance. Fig. 5 represents the results corresponding to the working conditions of Fig. 4. The thermal resistance decreases with increasing CPU power, due to the increased nucleate boiling contribution with increasing power providing more efficient heat transfer. The case-to-fluid thermal resistance is below 0.03 K/W for power over 200 W and is below 0.02 K/W for power over 500 W, which is lower than most high-performance single-phase liquid cooling solutions. At 1000 W power, the thermal resistance is 0.017 K/W. The exit vapor quality is approximately 0.55, and no partial dry-out has been observed under this condition (Fig. 4), indicating the ability of the current cold plate to dissipate even higher power without cooling performance degradation. However, due to the limitation of the power capacity of the TTV, higher power levels were not tested. It is expected that the thermal resistance would increase when the exit vapor quality approaches 1 and partial dry-out occurs (especially near the outlet). This threshold vapor quality needs to be further investigated as future work and used

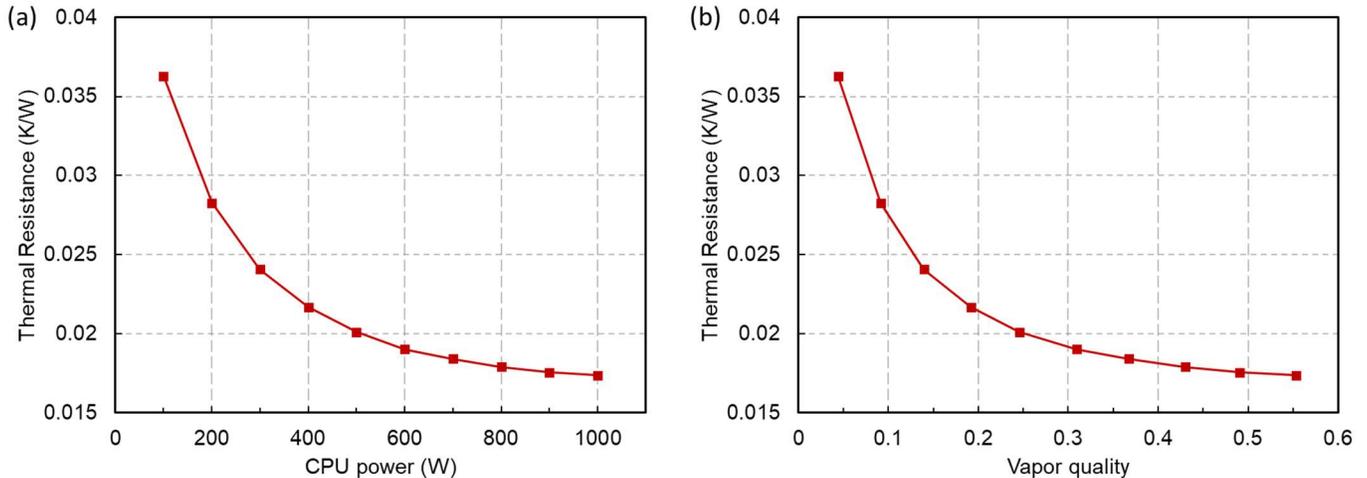


Fig. 5. Case-to-fluid thermal resistance of the TTV as a function of (a) CPU power and (b) exit vapor quality

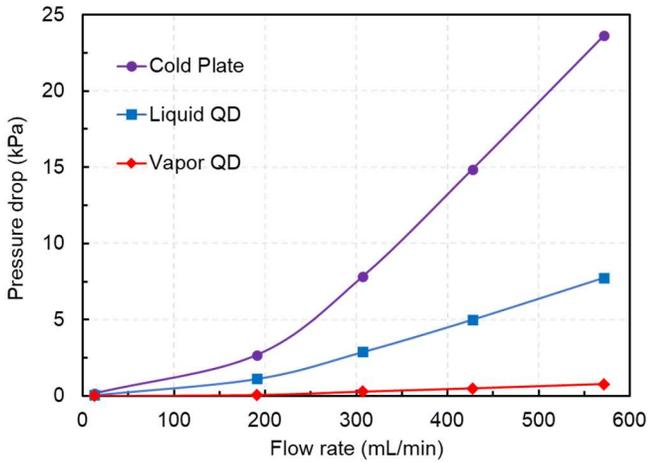


Fig. 6. Schematic of the test system

as a design guideline for the working conditions of the current cooling technology.

It is worth noting that in the TTV, the four thermocouples are placed directly on top of the four ceramic heaters, so that the measured case temperatures are likely to be the maximum values in the entire case surface, and their average is likely to be higher than the center point temperature of the case surface. Therefore, the actual case-to-fluid thermal resistance based on the center point temperature, as recommended by Intel, is likely to be even lower than the values reported here in Fig. 5.

C. Pressure Drop

The pressure drop of different components in a data center server rack determines the working conditions, since multiple server sleds are plumbed in parallel from the coolant pump and flow distribution and flow rate would be greatly affected by the pressure drop along the coolant pathway. Fig. 6 shows the pressure drop measurements corresponding to single-phase conditions (no heat load) with varying flow rates. As expected, the pressure drops of the cold plate and the two QDs increase with increasing coolant flow rate, and the trend is not linear due to the complex flow path inside the QDs. The cold plate has the highest pressure drop, which is because a flow restrictor is employed at the inlet of the cold plate to stabilize the flow distribution during imbalanced heat load conditions. The QDs upstream and downstream of the test server sled are termed liquid QD and vapor QD, respectively, since liquid refrigerant enters the sled through the liquid QD and two-phase mixture fluid exits the sled through the vapor QD. For single-phase conditions, the liquid QD has higher pressure drop than the vapor QD, due to the smaller size of the QD fitting and thinner tubing used for liquid line than for vapor line. The reason for its smaller size is that at any conditions, single-phase liquid passes through the liquid QD, which has a much smaller pressure drop compared with two-phase mixture (especially with high vapor quality) passing through the vapor QD. Therefore, smaller tubing size and QD component size are used to free up additional space in a 1U server without sacrificing too much pressure head. The pressure drop of these components under different vapor

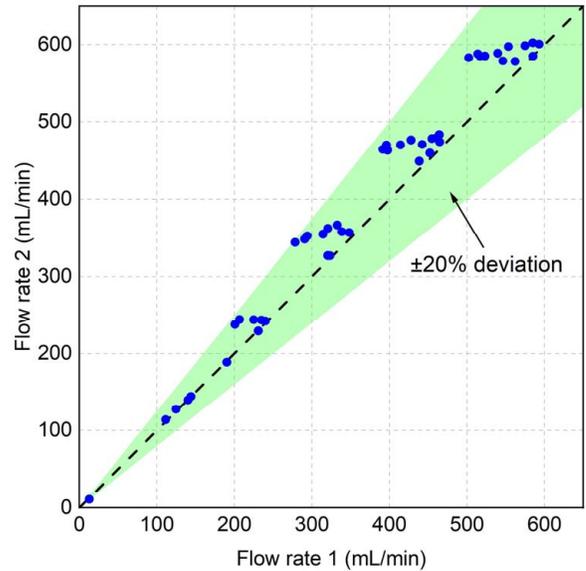


Fig. 7. Flow rate comparison of the two cold plates connected in parallel under nonuniform heating conditions

quality conditions are to be investigated as future work. Pressure drop from other fittings and bends that are not accounted for should also be taken into future consideration.

D. Flow Distribution

Each server sled has two CPUs, and the two cold plates cooling the two CPUs are connected in parallel. When the heat loads of the two CPUs are not balanced, the two cold plates could have different flow resistances, which results in flow maldistribution, i.e., the CPU with higher heat load gets less refrigerant flow. To characterize the flow distribution of the two parallel cold plates, imbalanced heat load conditions are tested, with one CPU having zero power and the other one having different power levels. Fig. 7 shows the flow distribution test results. With only one CPU having heat loads, the flow rates of the two cold plates do show certain differences, resulting in data points deviating from the diagonal of Fig. 7. For the CPU with heat load, higher vapor quality would result in more deviation due to larger flow resistance difference. Nonetheless, we have designed and implemented flow restrictors in the cold plates so that the pressure drop across the cold plates is dominated by the flow restrictors. Consequently, the flow distribution is mostly balanced, with less than 20% difference between the flow rates of the two cold plates even under extreme cases (one side showing partial dry-out and one side remaining purely single-phase liquid), ensuring that there is always sufficient flow to cool the heated CPU.

IV. SUMMARY AND CONCLUSIONS

In this work, we designed and established a test system to characterize the server-level thermohydraulic performance of the two-phase direct-to-chip cooling solution for data centers, using a sustainable fluid R1233zd(E). The system consists of a server sled, a refrigerant reservoir, a liquid pump, and a

condenser. The server sled contains mounted TTVs simulating Intel Sapphire Rapids CPUs, and microchannel cold plates are attached onto the TTVs (mated by a thermal interface material) to remove the generated heat by refrigerant vaporization. The following results are obtained through preliminary tests:

- The case temperatures of the TTVs are stable and below 60 °C under up to 1000 W of CPU power, demonstrating efficient cooling by the two-phase microchannel cold plates. Temperature overshoots are observed due to delayed boiling incipience, which needs to be minimized or eliminated.
- The case-to-fluid thermal resistance is below 0.02 K/W for heat loads in the range of 500~1000 W, which is expected to remain low with even higher power, and should be lower if the center point case temperature is used as the case temperature as recommended by Intel.
- The single-phase pressure drops of the QDs and the cold plate increase with refrigerant flow rate as expected. Under heat loads, the pressure drop would increase within the loop, especially with high-vapor-quality two-phase flow, which requires systematic and in-depth future study.
- Flow maldistribution is prevented due to the implementation of flow restrictors in the cold plates. The liquid flow rates of the two parallel cold plates are mostly balanced with below 20% difference even under extremely imbalanced heat load conditions.

With easy modifications, this system can be used to characterize the effects of different variables, including types of fluid, thermal interface materials, cold plate geometries, plumbing configurations (cold plates in parallel or in serial), pumping mechanisms (active/passive), piping sizes, and various components (QDs, valves, fittings, etc.). Further systematic study will allow this system to provide data and performance evaluations needed for the product design, parts selection, problem identification, and parameter optimization of this two-phase direct-to-chip cooling solution, as well as verification of various new ideas for cooling data centers. Hence, our work provides valuable insights into the development of a high-capacity, robust and efficient liquid cooling strategy for next-generation data centers with high performance processors.

ACKNOWLEDGMENT

The authors thank Aurelio Munoz, Tyler Richardson, Levi Jordan, and William Allai for their support in system setup and experimental testing.

REFERENCES

- [1] R. W. Bonner. "High-Heat-Flux Rack Level Direct-to-Chip Two-Phase Cooling Using Sustainable Fluids." Presentation in 2023 OCP Global Summit, San Jose, California, 2023.
- [2] A. Heydari, A. R. Gharaibeh, M. Tradat, Y. Manaserh, V. Radmard, B. Eslami, J. Rodriguez, and B. Sammakia. "Experimental evaluation of direct-to-chip cold plate liquid cooling for high-heat-density data centers." *Applied Thermal Engineering* 239 (2024): 122122.
- [3] A. Heydari, Y. Manaserh, A. Abubakar, C. Caceres, H. Miyamura, A. Ortega, and J. Rodriguez. "Direct-to-Chip Two-Phase Cooling for High Heat Flux Processors." In *International Electronic Packaging Technical Conference and Exhibition*, vol. 86557, p. V001T01A003. American Society of Mechanical Engineers, 2022.
- [4] K. P. Drummond, D. Back, M. D. Sinanis, D. B. Janes, D. Peroulis, J. A. Weibel, and S. V. Garimella. "A hierarchical manifold microchannel heat sink array for high-heat-flux two-phase cooling of electronics." *International Journal of Heat and Mass Transfer* 117 (2018): 319-330.
- [5] Y. Zhu, D. S. Antao, K.H. Chu, S. Chen, T. J. Hendricks, T. Zhang, and E. N. Wang. "Surface structure enhanced microchannel flow boiling." *Journal of Heat Transfer* 138, no. 9 (2016): 091501.
- [6] M. K. Sung, and I. Mudawar. "CHF determination for high-heat flux phase change cooling system incorporating both micro-channel flow and jet impingement." *International Journal of Heat and Mass Transfer* 52, no. 3-4 (2009): 610-619.
- [7] D. C. Moreira, V. S. Nascimento, G. Ribatski, and S. G. Kandlikar. "Combining liquid inertia and evaporation momentum forces to achieve flow boiling inversion and performance enhancement in asymmetric Dual V-groove microchannels." *International Journal of Heat and Mass Transfer* 194 (2022): 123009.
- [8] X. Yu, C. Woodcock, Y. Wang, J. Plawsky, and Y. Peles. "A comparative study of flow boiling in a microchannel with piranha pin fins." *Journal of Heat Transfer* 138, no. 11 (2016): 111502.
- [9] S. G. Kandlikar. "A general correlation for saturated two-phase flow boiling heat transfer inside horizontal and vertical tubes." *Journal of Heat Transfer* 112, no. 1 (1990): 219-228.
- [10] T. Harirchian, and S. V. Garimella. "Flow regime-based modeling of heat transfer and pressure drop in microchannel flow boiling." *International Journal of Heat and Mass Transfer* 55, no. 4 (2012): 1246-1260.
- [11] E. Costa-Patry, and J. R. Thome. "Flow pattern-based flow boiling heat transfer model for microchannels." *International Journal of Refrigeration* 36, no. 2 (2013): 414-420.
- [12] W. Qu, and I. Mudawar. "Flow boiling heat transfer in two-phase micro-channel heat sinks—II. Annular two-phase flow model." *International Journal of Heat and Mass Transfer* 46, no. 15 (2003): 2773-2784.
- [13] H. J. Cho, D. J. Preston, Y. Zhu, and E. N. Wang. "Nanoengineered materials for liquid–vapour phase-change heat transfer." *Nature Reviews Materials* 2, no. 2 (2016): 1-17.
- [14] W. Tong, A. Bar-Cohen, T. W. Simon, and S. M. You. "Contact angle effects on boiling incipience of highly-wetting liquids." *International Journal of Heat and Mass Transfer* 33, no. 1 (1990): 91-103.