# Universal Direct-to-Chip Cold Plates for Single- and Two-Phase Cooling

Qingyang Wang
*Accelsius*
Austin, USA
qwang@accelsius.com

Devdatta P. Kulkarni
*Intel*
Hillsboro, USA
devdatta.p.kulkarni@intel.com

Richard W. Bonner III
*Accelsius*
Austin, USA
rbonner@accelsius.com

John C. Gulick
*Intel*
Hillsboro, USA
john.c.gulick@intel.com

***Abstract*—For several decades, air cooling has been the dominant high power processor cooling method. Driven largely by the demand for higher power processors in artificial intelligence applications, liquid cooling has resurfaced as a cooling method for current and next generation high power processors due to its ability to more easily and efficiently remove high heat fluxes and higher socket thermal loads. Single-phase cooling using water and propylene glycol mixtures has taken the early lead due to its combination of adequate thermal performance and relative ease of installation compared to other emerging and higher performance alternatives. However, the insatiable demand for even higher power processors by the market over the next two to five years portends the eventual transition to higher performance cooling technologies, such as direct-to-chip two-phase cooling. To potentially hasten this transition, an investigation was performed to evaluate universal cold plates that could function in either single-phase or two-phase based cooling modes. The experimental results demonstrate that two-phase cooling outperforms single-phase cooling, even in a cold plate that was optimized for single-phase cooling. This result was consistent for R1233zd(E) (a lower pressure, low GWP working fluid) and R515B (a medium pressure, higher performance working fluid). The main benefit of a universal cold plate is that an end-user would not have to be as concerned with how and when to transition to two-phase cooling, but instead be inherently preparing for the transition over time. In the long run, cold plates optimized for two-phase cooling will be implemented to enable even higher power processors.**

***Keywords—liquid cooling, direct-to-chip two-phase cooling, high power processors, data center cooling, thermal management***

## I. Introduction

Data centers that house the higher power CPUs and GPUs performing state-of-the-art computations for artificial intelligence (AI) are critical for technological advancement in the modern age. The power and cooling requirements for AI workloads are far higher than the legacy power and cooling requirements of traditional data centers. Currently, data centers account for about 1% of the global electricity demand, with the expectation that AI workloads will double the demand over the next few years [1]. Air-cooling methods cannot keep up with the cooling requirements of AI optimized processors, servers, and racks. Liquid-based cooling methods can better handle these cooling requirements due to the inherently better heat transfer properties of liquids.

Direct-to-chip single-phase water emerged as one of the first liquid cooling strategies to replace air cooling at scale [2]. The scaling of water in the data center cooling market was likely driven by the extensive supply of water-based cooling technologies used in adjacent industries. The use of water in IT equipment comes with risks in the form of leaks, channel clogging, corrosion, bacterial growth, etc., although many of these issues have been sufficiently addressed by the industry [3]. Perhaps of greater concern, single-phase water direct-to-chip based solutions appear to be running out of headroom for thermal performance, as AI cooling demands continue to increase.

Pumped two-phase direct-to-chip addresses the performance limitations of single-phase cooling solutions by offering the capability to cool higher power sockets, higher heat fluxes, and higher rack level powers [4]. Cooling with refrigerants reduces risks associated with leaks (since refrigerants are dielectric) and eliminates both corrosion and bacterial growth issues. The improved thermal performance also leads to overall reduced energy consumption at the data center system level, mainly driven by the enabling of higher temperature facility coolant but also smaller electricity demands at the rack level due to lower pumping power requirements [5].

## II. Concept

Although the future is bright for direct-to-chip two-phase cooling in data centers, the transition to liquid cooling, mostly in the form of direct-to-chip single-phase water, has only just begun. Stakeholders making up the data center ecosystem (including processor and server manufacturers, integrators, data center operators, etc.) have a tough choice to make regarding liquid cooling, needing to solve the here and now but also desiring a liquid cooling technology that is future proof to the next generation of processors. One method of reducing the cooling choice burden on the data center ecosystem could be to have a universal cold plate capable of plugging into either single- or two-phase coolant distribution units.

In this experimental study, we attempt to demonstrate the feasibility of the universal cold plate concept by performance testing a cold plate optimized for single-phase water and subsequently performance testing the identical cold plate with several two-phase working fluids. The results show that all two-phase test results exceed the performance of single-phase water even though the cold plate has not been optimized for two-phase cooling. The study also includes some insights on some of the other requirements that must be met for a universal cold plate, including material and subcomponent compatibility. Finally, it should be emphasized that the universal cold plate is only intended to help with the transition of single- to two-phase cooling over the next two to five years and that cold plates optimized for two-phase cooling still need to be developed to meet future cooling demands.

## III. Experimental Setup

A thermal test vehicle (TTV) was designed to characterize the thermal performances of the universal cold plate under different working modes. The TTV is made of a copper block with a cartridge heater embedded inside to generate the desired heat load. A T-type thermocouple is attached to the center of the TTV surface to capture the case temperature. References [6-8] provide more details on the test setup and measurement techniques used for this study. In this work, PG25 was used for single-phase experiments and R1233zd(E) and R515B were used for two-phase experiments. The case-to-fluid thermal resistance is calculated as

$$R_{th} = \frac{T_{case} - T_f}{P} \quad (1)$$

where $T_{case}$ is the measured case temperature, $T_f$ is the fluid temperature, and $P$ is the heating power of the TTV and was set at 1,000 W during experiments. For single-phase mode $T_f$ is taken as the inlet fluid temperature, and for two-phase mode it is as the saturation temperature inside the cold plate.

## IV. Test Results and Discussion

Fig. 1 shows a plot comparing the performance of the cold plate tested with PG25 in single-phase mode, and R1233zd(E) and R515B in two-phase mode. The performance of PG25 improves with flow rate, but even at the higher flow rate tested it was not able to reach the performance of either two-phase test cases. Between both two-phase fluids tested, R515B tested the best. This was largely expected based on the thermal physical property differences between the two fluids, mainly vapor pressure. At 55°C, the saturation pressure of R515B is approximately 165 psi, while R1233zd(E) has a saturation pressure of 50 psi. Although R1233zd(E) demonstrates lower thermal performance, it should be noted that the fluid has some advantages over R515B, including lower global warming potential (1 vs 293) and lower pressure containment requirements [9].

Realizing a universal cold plate includes other considerations. The cold plate materials must be cross compatible. For instance, aluminum may be readily compatible with two-phase cooling systems, but may not be compatible will all variations of single-phase coolant. Elastomers used in gaskets, O-rings, and tubing must also be carefully considered. The pressure containment of all structures, most notably the flat evaporators that mate with the processors, must be able to hold the pressure of any of the intended fluids without deformation. Based on preliminary investigations, more than one solution set exists for the fluids discussed in this study.

## V. Conclusion

The possibility of having a universal cold plate appears feasible. Two-phase cooling performance was demonstrated to be better than single-phase at representative conditions, even when utilizing a cold plate optimized for single-phase cooling. The highest performance was demonstrated using R515B, a medium pressure, medium GWP refrigerant. However, R1233zd(E), a low pressure, low GWP alternative to R515B, also outperformed single-phase water. Hence, for two-phase cooling, customers can choose different working fluids prioritizing either cooling performance or environmental consideration. Preliminary indications conclude that developing a universal cold plate that is compatible with all three working fluids from a material compatibility and structural point of view is also achievable. The end result is that data center ecosystem decision makers could have the benefits of choosing a universal cold plate as the data center cooling market transitions from single-phase water to two-phase cooling over the next two to five years.

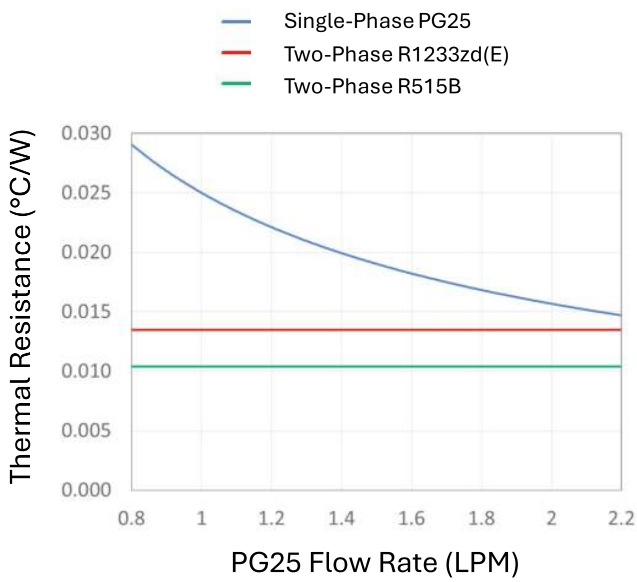


Fig. 1 The thermal performance of a cold plate optimized for single-phase cooling performance is plotted against flow rate. The two-phase thermal performance of the same cold plate tested with R1233zd(E) and R515B is also overlayed for comparison. Both two-phase cooling working fluids demonstrated improved performance over single-phase.


## Acknowledgment

The authors would like to acknowledge Aurelio Munoz for providing fabrication and testing support.


## References


[1] R. Bonner, et al., "High-Heat-Flux Rack Level Direct-to-Chip Two-Phase Cooling Using Sustainable Fluids", OCP Global Summit, 2023.

[2] A. Heydari, et al., "Experimental evaluation of direct-to-chip cold plate liquid cooling for high-heat-density data center", Applied Thermal Engineering 239 (2024): 122122.

[3] P. Shahi, et al., "Maximizing Cooling Potential: A Step-by-Step Guide to Commissioning Liquid-Cooled Data Centers", IEEE ITherm, 2024.

[4] S. Ozguc, Q. Wang, A. Narayanan, R. Bonner, "Investigation of Flow Restrictors for Rack Level Two-Phase Cooling under Nonuniform Heating", Semi-Therm, 2024.

[5] M. North, A. Kulkarni, D. Haley, "Effects of Datacenter Cooling Subsystems Performance on TUE: Air vs. Liquid vs. Hybrid Cooling", IEEE ITherm, 2024.

[6] Q. Wang, S. Ozguc, A. Narayanan, R. Bonner, "A Server-Level Test System for Direct-to-Chip Two-Phase Cooling of Data Centers Using a Low Global Warming Potential Fluid", IEEE ITherm, 2024.

[7] D. Kulkarni, "Thermal Perforamance of Common Cold Plate for Pumped Single- and Two-Phase Direct Liquid Cooling for Next Generation High Power Server Processors", ASME InterPACK, 2024.

[8] A. Narayanan, S. Ozguc, Q. Wang, R. Bonner, "Investigation of Server Level Direct-to-Chip Two-Phase Cooling Solution for High Power GPUs", ASME InterPACK, 2024.

[9] R. Bonner, B. Grieco, L. Cruz, "Understanding PFAS Concerns for Two-Phase Cooling of Data Centers", Data Center Frontier, 33035570.