energies

MDPI

Article

A Study on the Thermal Management Performance of Server-Oriented Memory Liquid Cooling Solutions

Yanling Chen [1], Zhongyun Tian [1,2,*], Mingzhi Kong [1], Lei Sun [1], Lizhi Zhou [1], Wujun Wang [1] and Mengyao Liu [1]

[1] IEIT Systems Co., Ltd., Jinan 250101, China
[2] School of Architecture and Design, Harbin Institute of Technology, Harbin 150001, China
* Correspondence: zhongyun_tian@163.com

Abstract

The rapid increase in memory power density has made memory thermal management a critical challenge in high-density servers, where extremely limited DIMM spacing significantly reduces the effectiveness of air cooling. Compared with CPUs and GPUs, memory-level liquid cooling has received less systematic study, particularly regarding the influence of cold plate structural design on thermal and hydraulic performance under realistic server conditions. In this paper, three engineering-feasible memory liquid cooling solutions (water-flowing cold plate, clamp-type cold plate and heat-pipe-based cold plate) are experimentally compared on a high-density server system. Experiments are conducted at coolant inlet temperatures of 37–50 °C with a fixed flow rate of 0.8–1.5 L/min. Memory, CPU, and voltage regulator temperatures, as well as system pressure drop, are measured. Results show that memory temperature increases with coolant inlet temperature for all configurations, while their relative performance remains unchanged. Memory temperatures range from 62.04 to 71.13 °C, 57.65 to 66.98 °C, and 66.22 to 76.07 °C, with corresponding pressure drops of 24.19–26.69 kPa, 32.73–35.98 kPa, and 27.00–29.96 kPa. These results provide insight into the role of coolant distribution and flow-path topology in memory thermal performance.

Keywords: liquid-cooled server; data center; high-density server; memory liquid cooling

1. Introduction

With the rapid growth of large-scale artificial intelligence model training, scientific computing, cloud workloads, and database applications [1], computing systems have continued to evolve toward higher bandwidth, higher capacity, and higher density [2].

Significant improvements in bandwidth and access latency have made DDR5 memory a key enabler of server performance scaling [3]. However, under high-frequency operation, the operating current and power consumption of DDR5 memory modules are substantially increased, resulting in a heat flux density that is approximately 30 to 60 percent higher than that of the DDR4 generation [4]. When multiple DDR5 modules are deployed in parallel, the power consumption of a single module can reach 12 to 18 W, while high-performance registered DIMMs can exceed 20 W [5]. In typical dual-socket or quad-socket servers, the combined thermal load generated by 16 to 32 memory modules operating simultaneously becomes non-negligible [6]. As a result, the memory subsystem has gradually emerged as a critical component in the thermal management of high-density servers [7].

Unlike CPUs or GPUs, memory modules exhibit inherent geometric disadvantages for heat dissipation. DDR5 memory chips are arranged in dense arrays on a single DIMM, and

Check for updates

Academic Editor: Mahmoud Bourouis

Received: 26 February 2026
Revised: 22 March 2026
Accepted: 24 March 2026
Published: 29 April 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

the spacing between adjacent DIMMs is typically less than 7.6 mm, which is much smaller than the characteristic dimensions of CPU cooling regions [8]. Such confined spacing significantly restricts convective pathways [9] and increases local flow resistance, making effective heat transfer difficult to achieve with air cooling under high-load conditions [10]. Previous studies have shown that when memory temperatures exceed 85 °C, error rates increase exponentially, particularly under high-frequency refresh modes, where device behavior becomes more sensitive [11]. In addition, temperature fluctuations accelerate thermomechanical fatigue, which shortens the service life of memory modules [12]. As server power density continues to increase, the limitations of conventional air cooling in managing memory thermal loads have become increasingly evident and are a key factor constraining system reliability and performance stability [13].

In parallel with the rapid increase in server power density, data centers have undergone a structural transition from traditional air-cooled architectures toward hybrid and fully liquid-cooled systems [14]. Direct-to-chip and cold plate-based liquid cooling have been increasingly adopted at the processor level to accommodate the rising thermal design power of CPUs and GPUs, while simultaneously reducing fan power consumption and improving overall energy efficiency [15]. At the system scale, liquid cooling enables higher rack-level power density, relaxed airflow constraints, and more flexible thermal zoning compared with conventional air-cooled designs [13]. As a result, liquid cooling is no longer limited to extremely high-performance computing platforms but is progressively being deployed in mainstream enterprise and cloud servers.

Despite this transition, the memory subsystem has largely remained air-cooled, even as its power density and thermal sensitivity continue to increase [16]. This mismatch between processor-level liquid cooling and memory-level air cooling leads to localized thermal bottlenecks and limits the effectiveness of system-level thermal optimization [17]. Consequently, extending liquid cooling to the memory subsystem is becoming a necessary step for achieving balanced thermal management in next-generation high-density servers.

Liquid cooling is regarded as one of the most promising thermal management technologies for high-heat-flux electronic systems due to its high heat transfer coefficient, large thermal capacity, and ability to remove substantial heat with minimal temperature differences [18]. In CPU, GPU, and system-level liquid cooling architectures, the effectiveness and energy-efficiency advantages have been well demonstrated, thereby reducing data center power usage effectiveness [19]. However, compared with processors, liquid-cooled memory remains in an early stage of research and industrial application. This is primarily attributed to several factors. First, the spacing between memory modules is minimal, which makes the layout of conventional cold plates difficult [20]. Second, memory modules are frequently replaced, which requires good assembly compatibility and modular design [21]. Third, liquid cooling pathways must be designed to avoid adverse impacts on the thermal management of voltage regulators, power-delivery components, and surrounding devices [22]. Fourth, cold plate structures must balance thermal performance, flow resistance, cost, and manufacturing complexity [23].

In recent years, only a limited number of studies have investigated the use of cold plates combined with flexible tubing for DIMM cooling [24]. However, most of these approaches rely on serial coolant flow configurations [25], which causes the coolant temperature to increase progressively along the flow path. As a result, the temperatures of downstream DIMMs remain consistently higher than those of upstream modules [14]. Some studies have proposed parallel flow strategies to improve temperature uniformity [26]. Nevertheless, challenges remain regarding flow distribution, structural arrangement, and manufacturing complexity [27]. In addition, other studies have explored the incorporation of phase change materials, heat pipes, and vapor chambers to enhance temperature uni-

formity [28]. However, the actual flow resistance, pressure requirements, and long-term reliability of these approaches in high-density server environments have not yet been systematically validated [29].

Given the challenges discussed above, this study uses a realistic high-density server system. It focuses on the thermal management of DDR5 memory and proposes three liquid cooling structures suitable for engineering deployment. Compared with conventional studies that primarily focus on microchannel parameters, flow path configurations, or the thermal performance of a single cold plate, this work makes several distinct contributions.

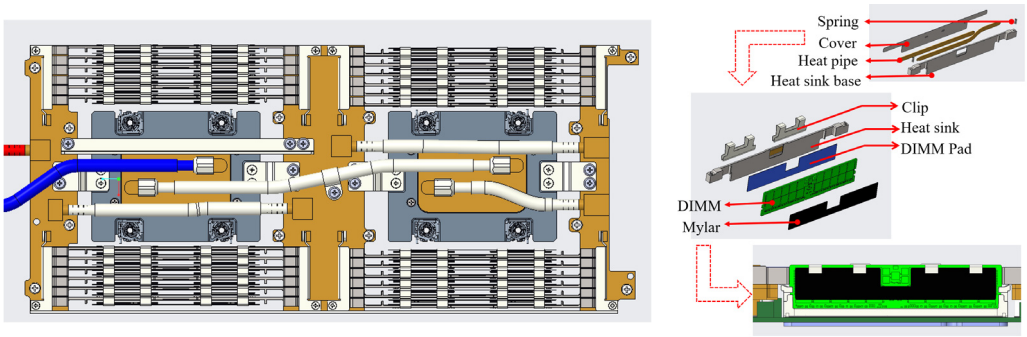
An experimental platform with controllable coolant temperature, flow rate, pressure drop, and airflow conditions is established, enabling the coupled thermal behavior of DDR5 memory, CPUs, and voltage regulator modules to be systematically measured. The three cooling structures are comprehensively compared with respect to thermal performance, temperature uniformity, hydraulic characteristics, cost and weight considerations, and system-level thermal coupling, thereby revealing the mechanisms by which structural design influences practical cooling performance. Finally, the scalability and deployment potential of liquid-cooled memory structures for future high-power DDR5 and DDR6 memory systems are discussed to guide modular liquid cooling design in data centers. The results of this study give systematic experimental evidence and clear structure–performance relationships for optimizing memory liquid cooling in high-density servers, thereby supporting the large-scale adoption of memory liquid cooling in practical data center applications.

2. Experimental System Design

2.1. Memory Liquid Cooling Structures

To compare the cooling effectiveness of different liquid cooling structures for DDR5 memory in high-density servers, three cold plate designs with practical engineering feasibility were designed and fabricated in this study. These include a water-cooled cold plate, a clamp-type cold plate, and a heat-pipe-based memory liquid cooling structure. The three designs differ substantially in flow channel configuration, heat transfer pathways, manufacturing processes, and assembly methods. The geometric configurations of the three memory liquid cooling structures are illustrated in Figure 1.

Figure 1 presents schematic diagrams of three memory liquid cooling configurations: the clamp-type memory liquid cooling structure (mode 1), the water-flowing memory liquid cooling structure (mode 2), and the heat-pipe-based memory liquid cooling structure (mode 3). These three designs differ significantly in structural layout, heat-transfer pathways, and coolant routing, corresponding to distinct thermal performance characteristics and engineering application scenarios.



(**a**)

Figure 1. *Cont.*

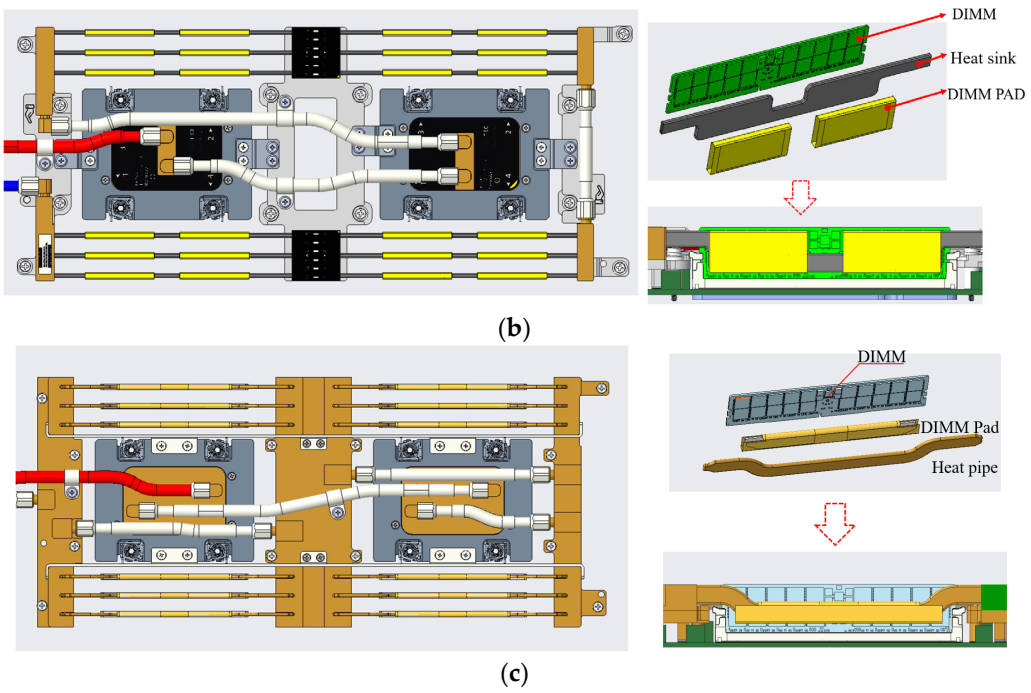


Figure 1. Schematic diagrams of the three memory liquid cooling configurations. (**a**) Clamp-type cold plate for memory liquid cooling (mode 1). (**b**) Water-flowing cold plate for memory liquid cooling (mode 2). (**c**) Heat-pipe-based cold plate for memory liquid cooling (mode 3).

The clamp-type memory liquid cooling structure (mode 1) comprises two symmetrically arranged high-thermal-conductivity metal clamping plates. Ultra-thin flattened heat pipes are embedded on the inner surfaces of the clamping plates, and the assembly is secured on both sides of the DIMM using an elastic clamping mechanism. This structure is compatible with memory modules from different manufacturers and specifications, providing good assembly adaptability. Thermal interface pads are applied between the memory chips and the clamping plates to reduce interfacial thermal resistance. The heat transfer path follows a multi-stage process involving a vapor chamber, heat pipes, and water-flowing cold plates. Heat generated by the memory chips is first spread laterally within the vapor chamber and then rapidly transferred through U-shaped heat pipes to water-cooling channels on both sides, where it is dissipated. Owing to the high effective thermal conductivity of the heat pipes, low flow resistance and good temperature uniformity are achieved on the coolant side. However, under high heat-flux conditions, the cooling capacity may be constrained by heat-transfer limits or by dry-out in the heat pipes.

The water-flowing, memory liquid cooling structure (mode 2) employs a composite design comprising stainless-steel coolant channels and an internal flow distributor. The cold plate body is typically manufactured from an aluminum alloy, and multiple coolant-flow channels are formed by precision computer numerical control machining. The coolant flows directly through regions corresponding to the memory chips. Thermal interface pads are placed between the heat sink and the DIMM to enhance thermal contact between the chips and the cold plate. A key feature of this structure is the direct or near-direct contact between the coolant channels and the memory chips, which significantly shortens the heat transfer path and reduces thermal resistance. Due to the relatively short flow paths, the coolant temperature rise along the channels is limited, thereby reducing temperature differences among chip locations and leading to a more uniform temperature distribution across the memory module. As a result, high cooling efficiency is achieved under high power-density and rapid thermal-response conditions.

The heat-pipe-based memory liquid cooling structure (mode 3) collects and transfers heat generated by the memory chips to a centralized water-flowing cooling region via flat heat pipes mounted on one or both sides of the DIMM. Effective thermal coupling between the memory chips and the evaporator sections of the heat pipes is achieved using thermal interface pads or metal contact structures. Heat is rapidly conducted through the heat pipes to the condenser sections and is ultimately dissipated via heat exchange with the coolant. In this configuration, the combination of heat pipes and coolant channels forms a compact installation structure that facilitates assembly and maintenance. Compared with direct water cooling, coolant routing is simplified, and flow resistance is reduced. However, the thermal performance of this structure is also limited by the heat-transfer capacity of the heat pipes, and performance bottlenecks may occur under extremely high heat-flux conditions.

To further illustrate the operation of the internal flow channels, Figure 2 shows the experimental setup for the three memory liquid cooling configurations installed on the server platform.

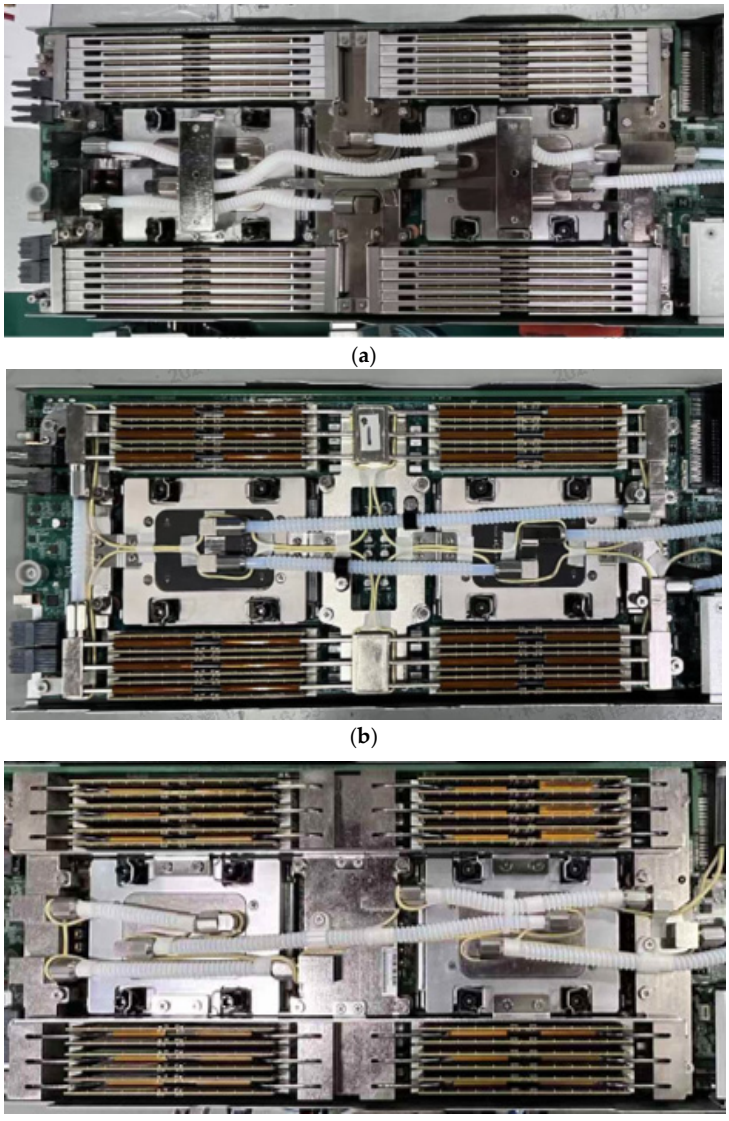


Figure 2. Experimental setup of the three memory liquid cooling configurations. (**a**) Clamp-type cold plate for memory liquid cooling (mode 1). (**b**) Water-flowing cold plate for memory liquid cooling (mode 2). (**c**) Heat-pipe-based cold plate for memory liquid cooling (mode 3).

2.2. Thermal Coupling Layout of the Memory and CPU Regions

The DDR5 DIMM chip arrangement is highly compact. The spacing between adjacent memory chips is approximately 7.5 mm. In comparison, the typical gap between two DIMMs is only about 4 mm. This constrained space not only severely limits air convection but also imposes strict constraints on cold plate design. The cold plate must cover the chip surfaces without interfering with DIMM insertion and removal, and it must avoid interference with voltage regulator modules, airflow paths, and CPU cold plates. To reproduce thermal coupling among multiple heat sources in a realistic server environment, a test platform is built on a server node. For all three memory liquid cooling configurations, a serial coolant-routing strategy is adopted, in which the coolant first flows through the memory region and then enters the CPU cold plate. This arrangement is selected because DDR5 modules are more temperature-sensitive under high-bandwidth operating conditions, and prioritizing memory cooling improves overall system stability. In addition, a dedicated cooling channel is designed for the voltage regulator region, where liquid cooling is applied to ensure stable operation of the VR modules.

2.3. Liquid Cooling Loop and Experimental Platform

A closed-loop liquid cooling system is used in the experiments. The system consists of a circulation pump, a plate heat exchanger, a temperature-controlled cooling source, a flow meter, differential pressure sensors, a temperature acquisition module, and insulated piping. The experimental system can supply coolant at temperatures accurately controlled between 25 °C and 55 °C. The maximum cooling capacity of a single node is 8000 W, and a vapor compression refrigeration cycle is used to ensure stable cooling for the server. The test system is connected to the server cold plates using quick-disconnect couplings. The schematic diagram and the experimental setup of the system are shown in Figures 3 and 4, respectively.

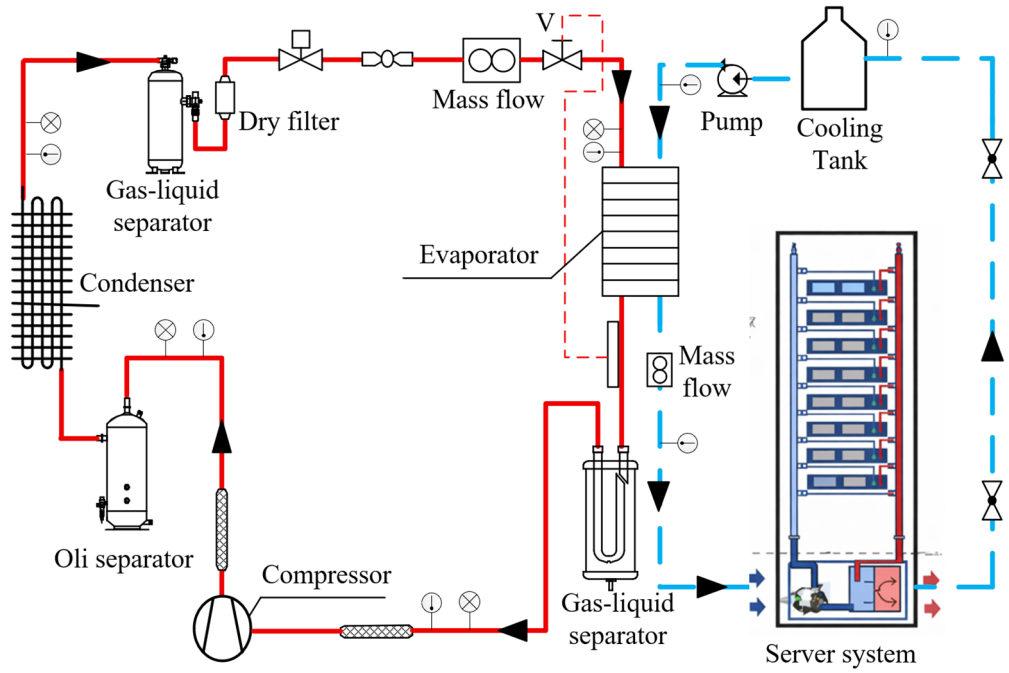


Figure 3. Schematic diagram of the server liquid cooling loop.

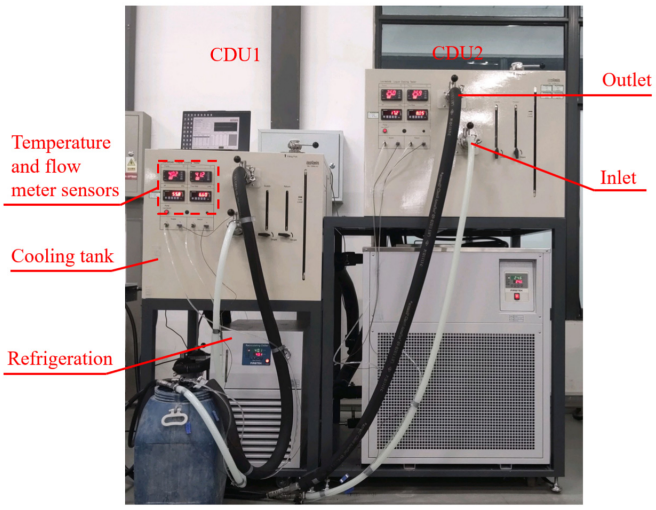


Figure 4. Experimental setup of the server liquid cooling test platform.

A 25% glycol–water solution provides adequate corrosion protection and good material compatibility for server liquid cooling loops. It also maintains stable viscosity and acceptable hydraulic performance under the tested temperature range. Accordingly, a 25% glycol–water solution is used as the coolant. The coolant inlet temperature is set to 37 °C, 40 °C, 45 °C, and 50 °C, while the flow rate is fixed at 1.0 L/min. All experiments are conducted in a temperature-controlled chamber maintained at 25 °C to minimize external disturbances.

2.4. Temperature and Flow Measurement System

To accurately determine the temperature distribution in DDR5 memory chips, this study employs two measurement methods. These include surface-mounted thermocouples and the internal temperature registers integrated within the DIMMs. All thermocouples are calibrated in a constant-temperature bath, yielding a nominal accuracy of ±0.2 °C. The thermocouples are attached to the top surfaces of the memory chips using high-temperature adhesive to minimize variations in contact thermal resistance. Coolant flow rate and pressure drop are measured using high-precision sensors. The measurement locations and sensor specifications are summarized in Table 1. The arrangement of temperature measurement points on the DDR5 chips is illustrated in Figure 5.

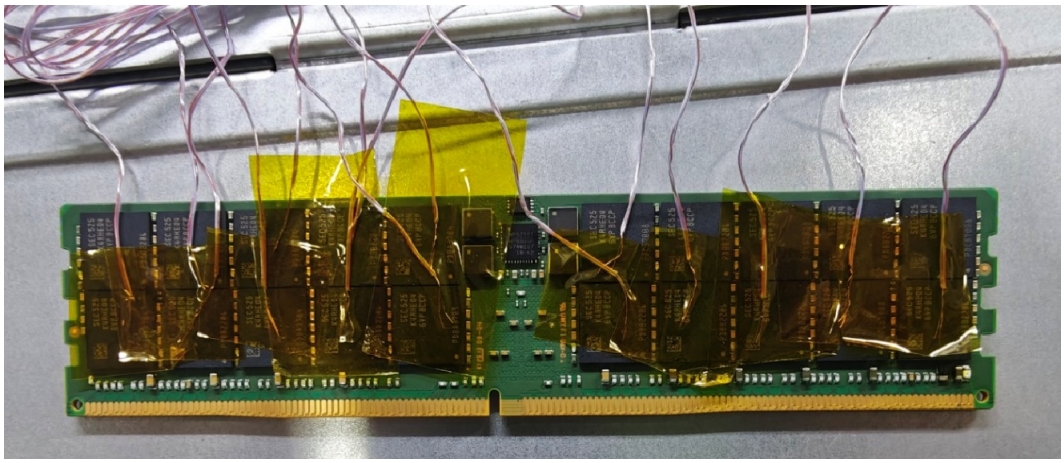


Figure 5. Layout of temperature measurement points on the DDR5 chips.

Table 1. Sensor locations and accuracy specifications for temperature, flow rate, and pressure drop measurements.

Measurement Location	Sensor Type	Accuracy	Range
DDR5 front die	T-type surface-mounted thermocouple	±0.2 °C	0.0~100.0 °C
DDR5 middle dies	T-type surface-mounted thermocouple	±0.2 °C	0.0~100.0 °C
DDR5 rear die	T-type surface-mounted thermocouple	±0.2 °C	0.0~100.0 °C
DIMM temperature register	Integrated module temperature sensor	±0.2 °C	0.0~100.0 °C
CPU hotspot	Motherboard digital temperature sensor	±0.2 °C	0.0~100.0 °C
Voltage regulator module	Motherboard digital temperature sensor	±0.2 °C	0.0~100.0 °C
Coolant inlet temperature	Pt100 resistance temperature probe	±0.1 °C	0.0~150.0 °C
Coolant outlet temperature	Pt100 resistance temperature probe	±0.1 °C	0.0~150.0 °C
Flow rate	Coriolis mass flow meter	±0.05 L/min	0.6~10.0 L/min
Pressure drop	Differential pressure sensor	FS ±0.1%	2.00~200.00 kPa

2.5. Experimental Conditions

A total of 12 operating conditions is designed, comprising three cold plate configurations and four coolant inlet temperatures. Each condition is tested three times. Steady-state operation is defined as a temperature variation of less than 0.2 °C within 30 min. After the steady state is reached, data is recorded for 20 min and averaged. The experimental condition matrix is summarized in Table 2. During all experiments, each DDR5 DIMM operated at a steady electrical power of 13.85 W, resulting in a total memory heat load of 332.4 W per node. The same memory workload and power level were maintained across all cooling configurations to ensure consistent thermal boundary conditions.

Table 2. Uncertainty assessment of measured parameters.

Measured Quantity	Main Sources of Uncertainty	Individual Uncertainty	Combined Uncertainty (95% Confidence Level)
DDR5 chip temperature	Thermocouple calibration error; sensor placement deviation; interface heat transfer variation	±0.2 °C; ±0.3 °C	±0.36 °C
SPD module temperature	Accuracy of the internal register sensor	±1 °C	±1 °C
CPU temperature	Motherboard digital sensor	±1 °C	±1 °C
VR temperature	Motherboard sensor error; local airflow fluctuation	±1 °C	±1.1 °C
Flow rate	Mass flow meter accuracy; pump speed fluctuation	±1% full scale; ±0.01 L/min	±0.02 L/min
Pressure drop	Differential pressure sensor error; zero drift	±0.5 kPa	±0.5 kPa
Coolant heat absorption	Temperature difference uncertainty; flow rate uncertainty	Temperature difference ±0.25 °C	±3%

2.6. Uncertainty Assessment

To verify the reliability of the measured data, the uncertainties associated with temperature, flow rate, pressure drop, and coolant heat absorption are evaluated using the error-propagation method. A heat balance between the electrical input power and the heat absorbed by the coolant is used for cross-validation. The typical deviation is within ±3 percent, which indicates that the overall experimental system is reliable. The combined uncertainty results are summarized in Table 2.

3. Results

This section presents a comprehensive analysis of the thermal performance of the three cold plate configurations under different operating conditions, including temperature control capability, temperature uniformity, hotspot distribution, coupled temperatures of the VR and CPU regions, pressure drop characteristics, and system-level heat balance consistency. Unless otherwise specified, all experiments are conducted at a flow rate of 1.0 L/min and an ambient temperature of 25 °C.

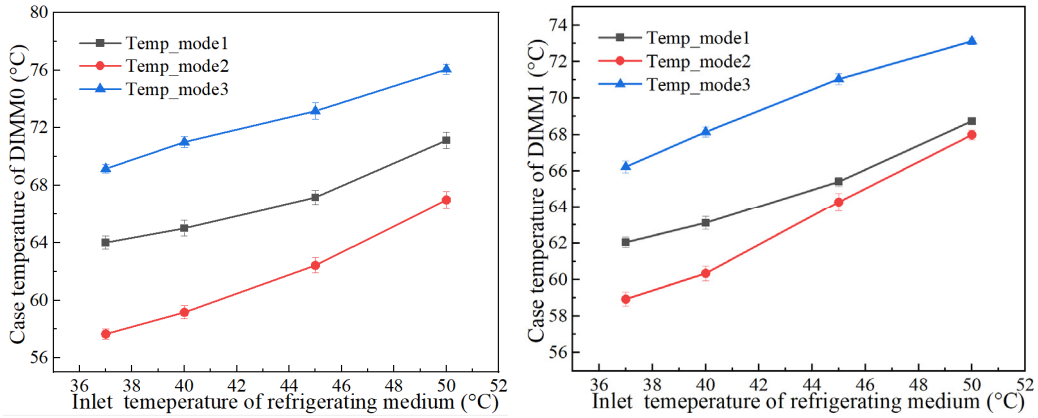
To facilitate heat-load-normalized comparison, an effective memory-to-coolant thermal resistance was introduced. The thermal resistance between memory and coolant ($R_{\text{mem-cool}}$) is defined as

$$R_{\text{mem}-\text{cool}} = \frac{T_{\text{mem,avg}} - T_{\text{cool,in}}}{Q_{\text{mem}}} \quad (1)$$

where $T_{\text{mem,avg}}$ denotes the average temperature of the monitored memory devices, $T_{\text{cool,in}}$ is the coolant inlet temperature, and Q_{mem} represents the applied memory heat load. This metric characterizes the overall heat transfer effectiveness from the memory devices to the coolant under the tested conditions.

3.1. Effect of Coolant Inlet Temperature on Memory Thermal Performance

Figure 6 shows the effect of the coolant inlet temperature on the memory temperatures of DIMM0 and DIMM1 across the three cooling modes. Across the investigated inlet temperature range, clear temperature separation among the three modes is maintained for both DIMMs, with mode 2 consistently exhibiting the lowest memory temperature, mode 3 the highest, and mode 1 falling in between.



(**a**) DIMM0 (**b**) DIMM1

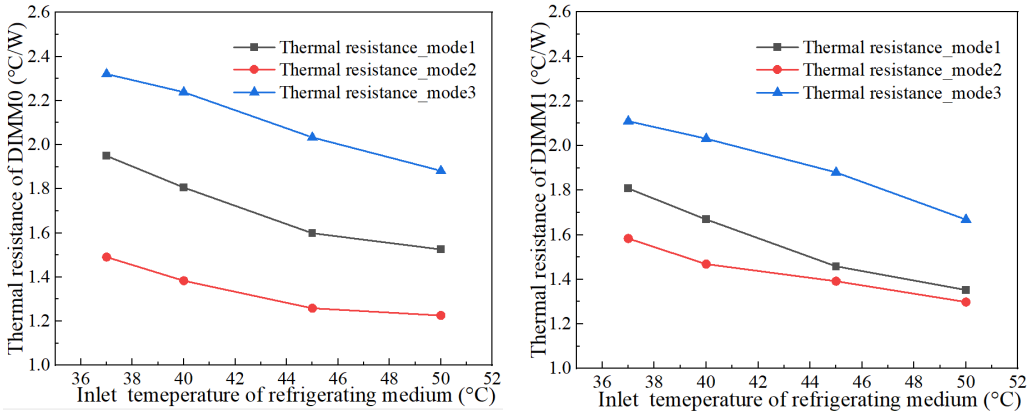
Figure 6. Effect of inlet coolant temperature on DIMM temperature.

Quantitatively, the total temperature increase of DIMM0 over the tested inlet temperature range is 7.12 °C, 9.33 °C, and 6.92 °C for modes 1, 2, and 3, respectively, while the corresponding increases for DIMM1 are 6.70 °C, 9.05 °C, and 6.89 °C. As indicated by the separation between the error bars, the observed temperature differences between the cooling modes exceed the combined measurement uncertainty for both DIMMs.

These results indicate that increasing the coolant inlet temperature primarily shifts the absolute memory temperature level. The differences among the cooling modes arise from their distinct memory-to-coolant thermal resistances, which lead to different temperature offsets as the inlet temperature varies.

As shown in Figure 7, the thermal resistance of both DIMM0 and DIMM1 decreases monotonically with increasing inlet coolant temperature for all three cooling configurations. This trend can be partially attributed to reduced coolant viscosity at elevated temperatures.

As the inlet temperature rises, the coolant's dynamic viscosity decreases, increasing the Reynolds number at the same volumetric flow rate. The enhanced turbulence intensity strengthens fluid mixing near the heat transfer surface, thereby improving the convective heat transfer coefficient and reducing the convective thermal resistance on the coolant side. In addition, the thermal resistance is also influenced by the changing temperature difference between the memory devices and the coolant. Although the absolute memory temperatures increase with higher inlet temperatures, the relative ranking among the three cooling modes remains unchanged, indicating that the dominant structural thermal resistance and contact conditions are not significantly affected by inlet temperature variation. Among the three configurations, the direct liquid-cooled design (mode 2) consistently exhibits the lowest thermal resistance, followed by the clamp-type cold plate (mode 1), while the heat-pipe-assisted solution (mode 3) shows the highest thermal resistance across the tested temperature range. This suggests that, under steady-state operating conditions, the overall cooling effectiveness is primarily governed by the module's structural thermal resistance, while the inlet coolant temperature mainly affects the convective component of the thermal resistance network.

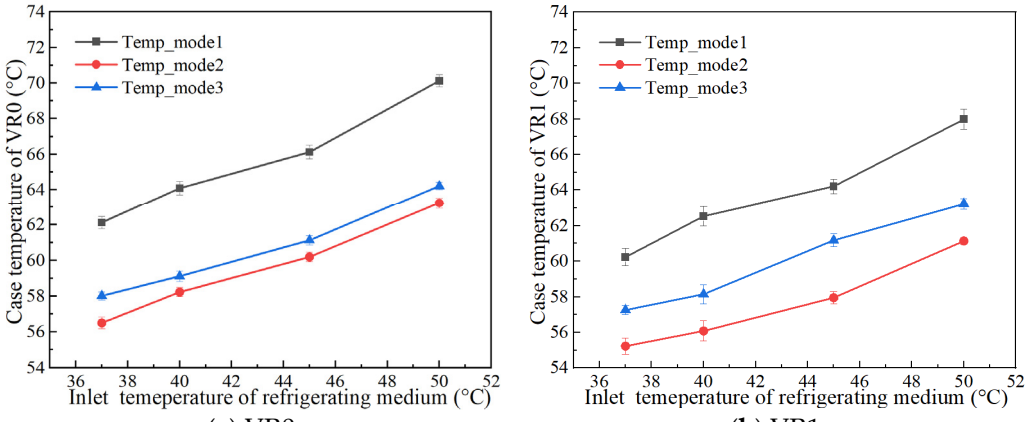


(**a**) DIMM0 (**b**) DIMM1

Figure 7. Effect of inlet coolant temperature on the thermal resistance of a DIMM.

3.2. Effect of Cold Plate Geometry on VR Temperature

Figure 8 compares the VR module temperatures under different coolant inlet temperatures for the three cooling modes. For both VR0 and VR1, the three modes exhibit distinct temperature levels across the entire inlet temperature range, with mode 2 operating at the lowest temperature and mode 1 at the highest.



(**a**) VR0 (**b**) VR1

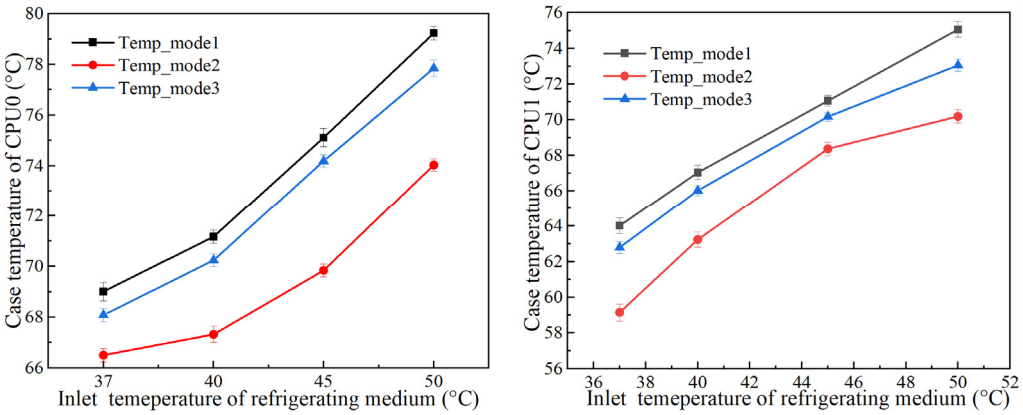
Figure 8. Effect of inlet coolant temperature on VR temperature.

The magnitude of temperature variation with inlet temperature differs among the cooling modes. For VR0, the total temperature increases from 37 °C to 50 °C and is 8.00 °C for mode 1, compared with 6.73 °C and 6.19 °C for modes 2 and 3, respectively. A similar trend is observed for VR1, where mode 1 shows a larger temperature increase (7.75 °C) than modes 2 (5.91 °C) and 3 (5.97 °C).

These differences can be attributed to variations in the local thermal environment around the VR modules induced by the different cold plate geometries, which affect heat spreading and convective heat removal in the vicinity of the VR components.

3.3. Coupled Effect of Cold Plates on CPU Temperature

Figure 9 shows the effect of coolant inlet temperature on the temperatures of CPU0 and CPU1 for the three cooling modes. For all modes, CPU temperature increases as the inlet temperature rises from 37 °C to 50 °C, while the relative temperature ranking among the three modes remains unchanged.



(**a**) CPU0 (**b**) CPU1

Figure 9. Effect of inlet coolant temperature on CPU temperature.

For both CPUs, mode 2 consistently exhibits the lowest temperature, whereas modes 1 and 3 show higher temperatures with only minor differences between them. Over the investigated inlet temperature range, the temperature rise of CPU0 is 10.24 °C, 7.53 °C, and 9.77 °C for modes 1, 2, and 3, respectively, while the corresponding values for CPU1 are 11.05 °C, 11.04 °C, and 10.26 °C. These results indicate that increasing coolant inlet temperature raises the absolute CPU temperature but does not alter the relative thermal performance of the three cooling modes.

3.4. Pressure Drop and Hydraulic Performance of the Cold Plates

Figure 10 presents the system frictional pressure drop (ΔP) of cooling system for the three cooling modes at different inlet coolant temperatures, with a fixed flow rate. Over the investigated temperature range, ΔP shows a moderate decreasing trend for all cooling modes, while the relative differences among the modes remain clearly distinguishable.

Across all inlet temperature conditions, mode 2 consistently exhibits the highest pressure drop, whereas mode 1 maintains the lowest, with mode 3 falling between the two. The variation in pressure drop with inlet temperature is comparatively slight in mode 1, indicating a relatively weak sensitivity of its hydraulic behavior to changes in coolant temperature.

These results suggest that differences in pressure drop among the cooling modes are primarily determined by their flow-path geometries rather than by the inlet coolant temperature. Consequently, at the fixed flow rate considered in this study, the hydraulic

penalty associated with each cooling mode is mainly determined by the cooling structure's inherent design, while variations in inlet temperature play a secondary role.

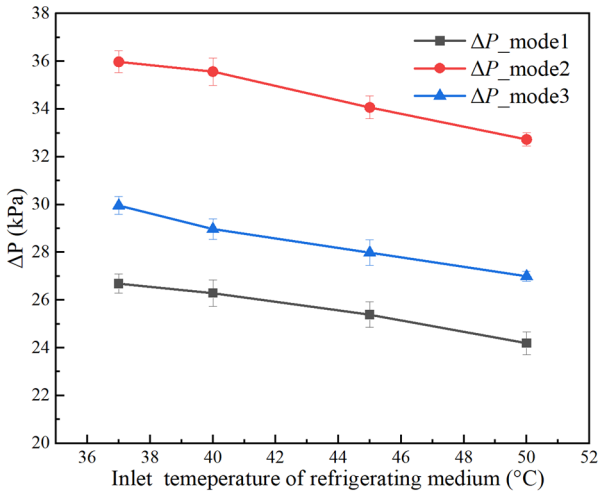
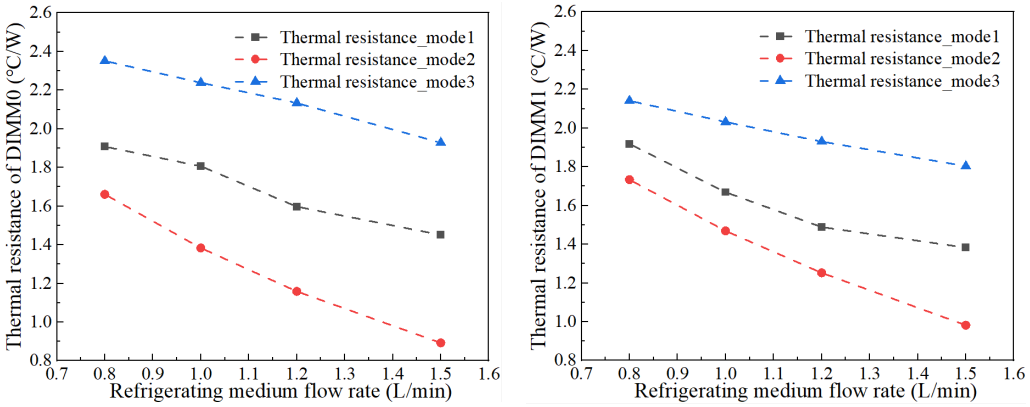


Figure 10. Effect of cooling configuration on frictional pressure drop (Δ*P*) in the flow channels.

3.5. Coupled Effect of Refrigerating Medium Flow Rate on Flow and Heat Transfer Characters of the Cold Plates

As illustrated in Figure 11, increasing the coolant flow rate results in a pronounced, nearly monotonic reduction in the thermal resistance of both DIMM0 and DIMM1 across all three cooling configurations. Among the three configurations, mode 2 consistently achieves the lowest thermal resistance, while mode 3 exhibits the highest values. The relative ranking among the three modes remains unchanged with increasing flow rate, suggesting that the intrinsic structural and contact thermal resistances of each cooling design govern their comparative performance.



(**a**) DIMM0 (**b**) DIMM1

Figure 11. Effect of coolant flow rate on DIMM resistance.

However, as shown in Figure 12, the improvement in thermal performance is accompanied by a continuous increase in frictional pressure drop (ΔP) with increasing flow rate. Mode 1 exhibits the lowest pressure drop, while mode 2 shows the highest hydraulic resistance across the tested range, with mode 3 lying between the two. This indicates that although mode 2 provides superior heat transfer performance, it imposes the largest pumping power requirement. The rate of thermal resistance reduction diminishes at higher flow rates, implying diminishing returns once the coolant-side resistance becomes less dominant within the overall thermal resistance network. Therefore, further increasing the flow rate yields limited thermal benefits while substantially increasing hydraulic penalties.

From a system-level perspective, an optimal design should balance thermal performance and pumping power consumption rather than pursuing maximum heat transfer alone.

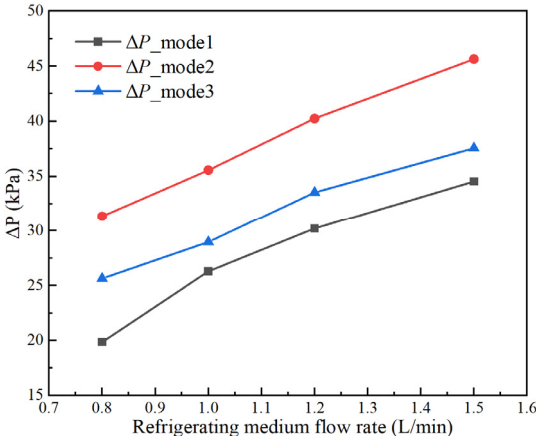


Figure 12. Effect of coolant flow rate on frictional pressure drop (Δ*P*).

4. Discussion

The performance differences among the three DIMM liquid cooling configurations are primarily due to variations in flow-channel topology and coolant distribution. Material selection and overall cold plate dimensions play a secondary role. Under identical operating conditions, DIMM thermal behavior is primarily governed by the ability to limit coolant temperature rise along the flow path in a constrained DIMM layout.

The water-flowing cold plate achieves the optimal DIMM thermal performance, which is enabled by its parallel short-channel design. This configuration effectively suppresses coolant temperature accumulation along the main flow direction and ensures sufficient low-temperature coolant is supplied to the memory region. Since DDR5 DIMM modules are linearly arranged with minimal spacing, even small coolant temperature gradients can lead to noticeable discrepancies in memory temperature. The results demonstrate that controlling along-path coolant temperature rise is more effective than introducing additional heat-spreading structures.

At the system level, VR temperature is influenced by cold plate geometry through airflow interaction in the DIMM region. Larger cold plate volume and height can partially obstruct airflow, thereby reducing local convective heat transfer around the VR modules. More compact structures introduce less airflow disturbance and consequently lead to lower VR temperatures. In the serial cooling configuration investigated in this study, the impact of memory liquid cooling on CPU temperature remains limited within the tested operating range. However, this coupling effect may become more pronounced as memory power density increases. Apparent differences are also observed in hydraulic performance. System pressure drop is mainly determined by flow path length and the degree of flow parallelization. Designs with higher flow parallelization achieve lower pressure drop. This characteristic is beneficial for system integration and for reducing pump power.

From a manufacturing and reliability perspective, the three liquid cooling configurations also exhibit distinct characteristics. The water-cooled plate uses a relatively simple channel topology and conventional metallic materials, which are compatible with mature machine or extrusion processes. Its structure contains no moving or phase-change components, resulting in predictable thermal behavior and favorable long-term reliability under continuous operation. The clamp-type cold plate improves serviceability by enabling repeated DIMM installation and replacement; however, its thermal performance and reliability are more sensitive to assembly tolerances, variations in contact pressure, and aging of

the thermal interface material. These factors may introduce performance dispersion across large-scale deployments.

The heat-pipe-based cold plate offers excellent temperature uniformity but requires more complex manufacturing processes, including vacuum sealing and charging the working fluid. Its long-term reliability is inherently linked to the integrity of the heat pipe and vapor chamber, as well as resistance to degradation under repeated thermal cycling. In high-density server environments with elevated and fluctuating heat flux, these reliability considerations become increasingly critical. Overall, the results indicate that manufacturing complexity, assembly robustness, and long-term operational stability must be considered alongside thermal and hydraulic performance when selecting liquid cooling solutions for practical DIMM deployment.

5. Conclusions

In this study, three memory liquid cooling configurations were compared under identical operating conditions: a water-flowing cold plate, a clamp-type cold plate, and a heat-pipe-based cold plate. The results show that coolant inlet temperature consistently affects memory, VR, and CPU temperatures across all configurations. As the coolant inlet temperature increases from 37 °C to 50 °C, memory temperature increases for all three cooling modes. Over the investigated range, memory temperature is maintained at approximately 57.65–66.98 °C for the water-flowing cold plate, 62.04–71.13 °C for the clamp-type cold plate, and 66.22–76.07 °C for the heat-pipe-based cold plate. The water-flowing cold plate exhibiting the lowest memory temperature and the heat-pipe-based cold plate exhibiting the highest.

A similar trend is observed for VR temperature. As the coolant inlet temperature increases, VR temperature rises to approximately 62.12–70.12 °C for the clamp-type cold plate, 57.26–64.21 °C for the heat-pipe-based cold plate, and 55.23–63.23 °C for the water-flowing cold plate. Despite differences in cold plate geometry, the relative temperature ranking remains stable across the investigated conditions. CPU temperature also increases with coolant inlet temperature for all three cooling modes. Across the investigated range, CPU temperature is maintained at approximately 59.15–70.19 °C for the water-flowing cold plate, 64.01–75.06 °C for the clamp-type cold plate, and 62.79–77.85 °C for the heat-pipe-based cold plate. The influence of upstream memory cooling on CPU temperature remains limited within the tested operating range.

Apparent differences in hydraulic performance are observed. As the coolant inlet temperature increases from 37 °C to 50 °C, the system pressure drop decreases for all three cooling modes. The pressure drop ranges from 32.73 to 35.98 kPa for the water-flowing cold plate, from 27.00 to 29.96 kPa for the heat-pipe-based cold plate, and from 24.19 to 26.69 kPa for the clamp-type cold plate.

Future work can be extended in three directions. First, multi-scale CFD methods can be applied to establish quantitative relationships between cold plate geometric parameters and coolant distribution, thereby supporting refined optimization of the flow channel topology. Second, the variation in interface thermal resistance under high heat-flux conditions should be investigated to develop low-resistance interface structures better suited to memory chip array layouts. Third, memory liquid cooling structures can be further studied in multi-heat-source coupled systems to develop liquid-loop design strategies for coordinated cooling of memory, CPUs, and accelerator devices, to meet the thermal requirements of DDR6 and future higher-power-density components.

Author Contributions: Conceptualization, Y.C. and Z.T.; methodology, M.K.; software, L.S. and M.K.; validation, Y.C. and W.W.; resources, Y.C.; data curation, Y.C.; writing—original draft preparation, Y.C. and L.Z.; writing—review and editing, Z.T. and M.L.; visualization, L.S.; supervision, M.K.; project

administration, Y.C.; funding acquisition, Y.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: All author were employed by the IEIT Systems Co., Ltd. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Nomenclature

$T_{cool,in}$	coolant inlet temperature (°C)
$T_{mem,avg}$	average temperature of the monitored memory devices (°C)
Q_{mem}	applied memory heat load (W)
$R_{mem\text{-}cool}$	thermal resistance between memory and coolant (°C/w)
Abbreviations	
DDR4	Double Data Rate 4
DDR5	Double Data Rate 5
DIMM	Dual In-line Memory Module
CPU	Central Processing Unit
GPU	Graphics Processing Unit
VR	Voltage Regulator

References

1. Yu, Z.; Wu, J.; Sun, Z. Development of a dynamic bottom-up model for city-scale residential building stock to estimate greenhouse gas emissions. *Environ. Dev. Sustain.* **2025**. [CrossRef]
2. Mutlu, O.; Olgun, A.; Yüksel, İ.E. Memory-Centric Computing: Solving Computing's Memory Problem. In Proceedings of the 2025 IEEE International Memory Workshop (IMW), Monterey, CA, USA, 18–21 May 2025; pp. 1–4. [CrossRef]
3. Guo, Z.; Yin, J.; Bai, Y.; Zhu, D.; Shi, K.; Wang, G.; Cao, K.; Zhao, W. Spintronics for Energy- Efficient Computing: An Overview and Outlook. *Proc. IEEE* **2021**, *109*, 1398–1417. [CrossRef]
4. Hassan, H.; Olgun, A.; Yağlıkçı, A.G.; Luo, H.; Mutlu, O.; Zurich, E. Self-Managing DRAM: A Low-Cost Frame-work for Enabling Autonomous and Efficient DRAM Maintenance Operations. In Proceedings of the 2024 57th IEEE/ACM International Symposium on Microarchitecture (MICRO), Austin, TX, USA, 2–6 November 2024; pp. 949–965. [CrossRef]
5. Kim, K.; Park, M. Present and Future, Challenges of High Bandwith Memory (HBM). In Proceedings of the 2024 IEEE International Memory Workshop (IMW), Seoul, Republic of Korea, 12–15 May 2024; pp. 1–4. [CrossRef]
6. Kumar, A.A. Semantic memory: A review of methods. models, and current challenges. *Psychon. Bull. Rev.* **2021**, *28*, 40–80. [CrossRef] [PubMed]
7. Olgun, A.; Osseiran, M.; Yağlıkçı, A.G.; Tuğrul, Y.C.; Luo, H.; Rhyner, S.; Salami, B.; Luna, J.G.; Mutlu, O. Read Dis-turbance in High Bandwidth Memory: A Detailed Experimental Study on HBM2 DRAM Chips. In Proceedings of the 2024 54th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), Brisbane, Australia, 24–27 June 2024; pp. 75–89. [CrossRef]
8. Modi, H.; Shahi, P.; Simon, V.S.; Chintaparthy, L.S.R.; Gupta, G.; Sivakumar, A.; Saini, S.; Bansode, P.; Ago-nafer, D. Experimental Study of Improved Chassis and Duct Redesign for Air-Cooled Server. In Proceedings of the 2023 22nd IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), Orlando, FL, USA, 30 May–2 June 2023; pp. 1–8. [CrossRef]
9. Choe, J. Memory Technology 2021: Trends & Challenges. In Proceedings of the 2021 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD), Dallas, TX, USA, 27–29 September 2021; pp. 111–115. [CrossRef]
10. Kandasamy, R.; Ho, J.Y.; Liu, P.; Wong, T.N.; Toh, K.C.; Chua, S.J. Two-phase spray cooling for high ambient tem-perature data centers: Evaluation of system performance. *Appl. Energy* **2022**, *305*, 117816. [CrossRef]
11. Kim, J.; Choi, H.; Lee, S.; Lee, H. Computational study of single-phase immersion cooling for high-energy density server rack for data centers. *Appl. Therm. Eng.* **2025**, *264*, 125476. [CrossRef]
12. Isazadeh, A.; Ziviani, D.; Claridge, D.E. Thermal management in legacy air-cooled data centers: An overview and perspectives. *Renew. Sustain. Energy Rev.* **2023**, *187*, 113707. [CrossRef]

13. Hnayno, M.; Chehade, A.; Klaba, H.; Bauduin, H.; Polidori, G.; Maalouf, C. Performance analysis of new liquid cooling topology and its impact on data centres. *Appl. Therm. Eng.* **2022**, *213*, 118733. [CrossRef]
14. Azarifar, M.; Arik, M.; Chang, J.-Y. Liquid cooling of data centers: A necessity facing challenges. *Appl. Therm. Eng.* **2024**, *247*, 123112. [CrossRef]
15. Ahmadi, V.E.; Barestrand, H.A.; Summers, J.; Yilmaz, C. Performance evaluation of two-phase direct-to-chip liq-uid cooling combined with air cooling for data centers. *Case Stud. Therm. Eng.* **2026**, *78*, 107650. [CrossRef]
16. Al Kez, D.; Foley, A.M.; Hasan Wong, F.W.B.M.; Dolfi, A.; Srinivasan, G. AI-driven cooling technologies for high-performance data centres: State-of-the-art review and future directions. *Sustain. Energy Technol. Assess.* **2025**, *82*, 104511. [CrossRef]
17. Orville, T.; Tajwar, M.; Bihani, R.; Saha, P.; Hannan, M.A. Enhancing Thermal Efficiency in Power Electronics: A Review of Advanced Materials and Cooling Methods. *Thermo* **2025**, *5*, 30. [CrossRef]
18. Chen, H.; Li, D. Current Status and Challenges for Liquid-Cooled Data Centers. *Front. Energy Res.* **2022**, *10*, 952680. [CrossRef]
19. Yuan, X.; Zhou, X.; Pan, Y.; Kosonen, R.; Cai, H.; Gao, Y.; Wang, Y. Phase change cooling in data centers: A review. *Energy Build.* **2021**, *236*, 110764. [CrossRef]
20. Liu, C.; Yu, H. Experimental Investigations on Heat Transfer Characteristics of Direct Contact Liquid Cooling for CPU. *Buildings* **2022**, *12*, 913. [CrossRef]
21. Ishimaru, K. Challenges of Flash Memory for Next Decade. In Proceedings of the 2021 IEEE International Reliability Physics Symposium (IRPS), Monterey, CA, USA, 21–25 March 2021; pp. 1–5. [CrossRef]
22. Dubey, P.; Pulugundla, G.; Srouji, A.K. Direct Comparison of Immersion and Cold-Plate Based Cooling for Au-tomotive Li-Ion Battery Modules. *Energies* **2021**, *14*, 1259. [CrossRef]
23. Gui, C.; Wang, Y.; Shen, C.; Wang, S.; Wang, R.; Lv, Z.; Lin, B.; Tian, W.; Li, Z.; Zhang, X.; et al. Advanced Cold Plate Liquid Cooling Solution for Hyper-scale Data Center Application. In Proceedings of the 2023 22nd IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), Orlando, FL, USA, 30 May–2 June 2023; pp. 1–6. [CrossRef]
24. Sun, Q.; Zhi, G.; Zhou, S.; Dong, X.; Shen, Q.; Tao, R.; Qi, J. Advanced Design and Manufacturing Approaches for Structures with Enhanced Thermal Management Performance: A Review. *Adv. Mater. Technol.* **2024**, *9*, 2400263. [CrossRef]
25. Xue, Z.H.; Ai, B.C.; Qu, W. Data center energy conservation study utilizing loop heat pipes as a chip-level cooling technique and its industrial application. *Appl. Therm. Eng.* **2023**, *220*, 119715. [CrossRef]
26. Asgari, S.; MirhoseiniNejad, S.; Moazamigoodarzi, H.; Gupta, R.; Zheng, R.; Puri, I.K. A gray-box model for re-al-time transient temperature predictions in data centers. *Appl. Therm. Eng.* **2021**, *185*, 116319. [CrossRef]
27. Regany, D.; Majós, F.; Barrau, J.; Rosell, J.; Ibáñez, M.; Fréchette, L.G.; Vilarrubí, M. Design and test of shape memory alloy fins for self-adaptive liquid cooling device. *Appl. Therm. Eng.* **2022**, *206*, 118010. [CrossRef]
28. Guo, R.; Li, L. Heat dissipation analysis and optimization of lithium-ion batteries with a novel parallel-spiral serpentine channel liquid cooling plate. *Int. J. Heat Mass Transf.* **2022**, *189*, 122706. [CrossRef]
29. Huang, X.; Yan, J.; Zhou, X.; Wu, Y.; Hu, S. Cooling Technologies for Internet Data Center in China: Principle, Energy Efficiency, and Applications. *Energies* **2023**, *16*, 7158. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.