Design and Experimental Validation of Partitioned Flow Strategies for Optimized Microchannel Cooling

Ron Zhang, Gill Fountain, KM Bang, Laura Mirkarimi, Suhail Sadiq and Arianna Avellan
Adeia, San Jose, CA
Ron.Zhang@adeia.com

***Abstract*— Rising power densities in advanced electronics make thermal management increasingly challenging. Because uniform power distributions are rare in processor chips, localized hot spots often govern system performance and reliability. Conventional microchannel liquid cooling with uniform flow is inefficient under non-uniform heat flux, as it does not address spatially concentrated thermal loads.**

***This work develops and experimentally validates a passive flow-partitioned silicon microchannel cold plate that maintains uniform channel geometry for manufacturability while redistributing coolant to match a strongly non-uniform XPU power map. The full three-dimensional cold plate geometry, including manifold and internal flow dividers, is modeled using a conjugate heat transfer framework and validated experimentally for both temperature and pressure drop.**

Under constant total volumetric flow rate and power dissipation exceeding 1 kW, the optimized partitioned-flow design reduces peak junction temperature by approximately 10–12 °C relative to uniform-flow operation while achieving approximately 70% lower pressure drop than temperature-minimizing split-flow configurations. The results demonstrate that constrained thermal–hydraulic optimization enables significant hotspot mitigation with substantially reduced pumping power, providing a practical and scalable cooling strategy for high-power multi-chip modules.

Keywords*— Microchannel cooling, hot spot mitigation, non-uniform flow distribution, flow partitioning, cold plate architecture, experimental validation.

I. Introduction

Continued scaling of semiconductor devices and increasing adoption of heterogeneous integration have resulted in steadily rising power densities. Liquid cooling using microchannel cold plates has emerged as an effective solution for high heat flux removal; however, thermal challenges persist due to highly non-uniform power dissipation in modern processors.

Uniform-flow microchannel cooling is inefficient under non-uniform heat flux, as coolant is oversupplied to low-power regions while high-power regions remain thermally constrained. Passive approaches that redistribute coolant flow in proportion to local thermal demand are therefore attractive. The influence of header and manifold design on flow distribution in parallel microchannel heat sinks has been widely reported, with non-uniform flow leading to spatial temperature gradients and degraded thermal performance [1], [2]. More recent studies have explored tailored flow maldistribution strategies to improve hotspot mitigation under non-uniform heat sources while balancing hydraulic cost [3]. Multi-objective optimization of inlet and outlet placement further demonstrates that thermal and pumping-power performance are strongly coupled through flow architecture decisions [4].

This work builds upon the integrated cooling system (ICS) platform previously introduced by the authors [14] and extends it through the development of a passive, geometry-compatible flow partitioning strategy for mitigating hot spots in high-power multi-chip modules. We design and experimentally validate a silicon microchannel cold plate that maintains uniform channel geometry for manufacturability while redistributing coolant to match a strongly non-uniform XPU power map.

At constant total coolant flow rate and power dissipation exceeding 1 kW, the proposed approach reduces peak junction temperature by approximately 10–12 °C relative to uniform-flow operation, while achieving approximately 70% lower pressure drop compared to temperature-minimizing split-flow configurations. Experimental measurements across multiple embedded sensor locations agree with conjugate heat transfer simulations within 6%, validating the modeling methodology.

The primary contributions of this work are:

1. A manufacturable partitioned-flow silicon cold plate concept that preserves uniform microchannel geometry.
2. A systematic methodology for tuning regional hydraulic resistance to match spatial power dissipation.
3. Quantification of thermal–hydraulic tradeoffs that identify a practical operating “sweet spot” under system-level pumping constraints.

4. Experimental validation on a molded multi-chip module test vehicle under realistic boundary conditions.

II. Background and Related Work

Manifold microchannel (MMC) configurations have been proposed as a means of shortening effective flow paths and reducing pressure drop while maintaining high heat transfer coefficients [5], [6]. By redistributing coolant through alternating inlet and outlet manifold channels, such architectures mitigate hydraulic penalties associated with long parallel flow paths. Hotspot-focused routing strategies incorporating split-flow or localized inlet concepts have also demonstrated significant reductions in peak temperature, often accompanied by increased hydraulic resistance [7].

The thermal impact of non-uniform heat flux has also been examined in the context of flow maldistribution mitigation, where geometric modification or hydraulic resistance variation is used to reduce spatial temperature variation [8]. At the system level, recent reviews of cold plate liquid cooling technologies emphasize the growing importance of manifold-aware design strategies for high-power computing applications [9].

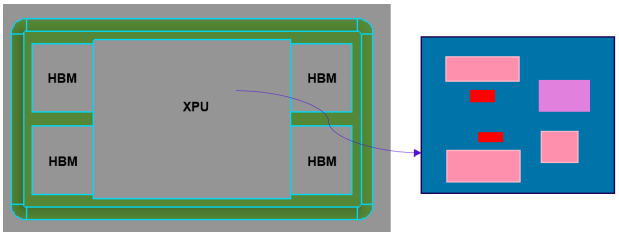
The foundation of microchannel cooling originates from the seminal work of Tuckerman and Pease [10], who demonstrated high heat flux removal using etched silicon microchannels. Subsequent research established scaling relationships for high-flux heat removal and identified key challenges in pressure drop and flow stability [11]. The influence of non-uniform heating on microchannel performance has also been examined, highlighting the importance of matching flow distribution to spatial heat generation [12], [13].

Directly bonded silicon cold plates eliminate TIM-related issues and enable flexible channel design. However, uniform-flow implementations remain inefficient for non-uniform power maps. From a manufacturing perspective, varying microchannel density within a silicon cold plate introduces fabrication complexity. Thus, there is strong motivation to achieve adaptive thermal performance through flow optimization rather than channel geometry variation.

III. Problem Description and Partitioning Strategy

A. System Description and Boundary Conditions

The MCM is implemented as a molded package, representative of advanced high-volume packaging approaches used in high-performance computing applications. Heat dissipation is assumed to occur only through the cold plate side of the molded package, which represents a conservative boundary condition. This assumption neglects heat removal through the interposer or package substrate; however, it is justified because an effective backside liquid cooling solution is expected to dominate overall heat removal. As a result, the predicted junction temperatures represent an upper bound for the thermal performance of the cooling architecture.



(a) (b)

Figure 1. (a) Multi-chip module (MCM) layout showing XPU and HBM placement. (b) Spatial power density map of the XPU under representative workload conditions. The maximum local power density exceeds 3 W/mm², with significant intra-die non-uniformity driving hotspot formation.

TABLE I. MCM Chip Dimensions & Power Dissipation

Component	Material	Size (mm × mm)	Power (W)
XPU	Silicon	26 × 32	1000
HBMs	Silicon	10 × 11	15 (each)
Interposer	Silicon	30 × 56	0

The power map shown in Fig. 1(b) is representative of workload-driven activity patterns commonly observed in high-performance XPUs. The spatially varying power distribution applies to the XPU only, as shown in Fig. 1(b). For the HBM devices, a uniform power distribution is assumed based on their comparatively low and evenly distributed power density. This distinction allows the study to isolate hotspot-driven flow redistribution effects primarily within the XPU region. Localized regions of elevated power density result in spatially concentrated hot spots that dominate peak junction temperature, even when average power density is moderate. Under such conditions, uniform-flow cooling is inherently inefficient, as it delivers equal coolant flow to regions with vastly different thermal demands. This motivates the need for spatially adaptive cooling strategies capable of delivering coolant preferentially to high-power regions.

B. Thermal Modeling Approach

A conjugate heat transfer (CHT) framework is employed to model the coupled heat conduction within the silicon structures and forced convection within the coolant. Simulations are performed using a commercial finite-volume solver (ANSYS Icepak), which is widely used for thermal analysis of electronic systems. The use of a commercial solver ensures numerical robustness and facilitates direct comparison between simulation and experimental results.

Heat conduction is modeled through the silicon devices, molded package materials, and the silicon cold plate, while convection is resolved within the microchannels using water as the working fluid. The numerical model simultaneously solves the solid and fluid domains, enabling conjugate heat transfer effects to be captured without reliance on empirical

thermal resistances. The full three-dimensional geometry of the cold plate, including all internal flow dividers and manifold features, is explicitly represented in the computational domain. No geometric simplifications are introduced for the partitioning structures, ensuring that the predicted pressure drop and flow redistribution accurately reflect the implemented design. Fig. 2 shows the cross section of the cold plate where the trapezoidal channel details are provided proportionally.

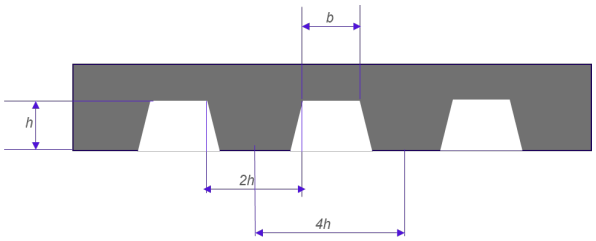


Figure 2. Dimensioned cross section of the silicon microchannel cold plate showing trapezoidal channel geometry, channel width, height, wall thickness, and pitch.

Thermal contact resistance between the molded package and the cold plate is neglected due to the direct bonding configuration. The silicon cold plate is directly bonded to the silicon device with a thermally grown SiO_2 layer less than 1μm thick at the interface. Comparative simulations performed with and without an interfacial thermal resistance indicate negligible impact on peak junction temperature. Experimental measurements further confirm that the interfacial resistance is below the sensitivity of the measurement setup. The interface is therefore modeled as ideal contact.

The simulation is conducted under steady-state conditions to reflect the experimental operating point after thermal equilibrium is reached. In the fluid domain, laminar flow is assumed based on microchannel dimensions and operating flow rates, and heat transfer is resolved using the coupled energy equation. In the solid domain, three-dimensional heat conduction is solved throughout the device stack and cold plate, allowing lateral heat spreading effects that influence local hot spot temperature to be captured.

Boundary conditions are defined to closely reflect the experimental setup. A prescribed heat flux corresponding to the spatial power map is applied at the device level, while a constant inlet coolant temperature and total volumetric flow rate are specified at the cold plate inlet. Pressure outlet conditions are applied at the manifold exit. Grid refinement studies are performed to ensure numerical convergence of peak junction temperature ($T_{j,max}$) and cold-plate pressure drop (ΔP).

To assess model robustness, sensitivity checks are performed on boundary conditions that influence peak junction temperature, including inlet coolant temperature, total flow rate, and applied power. These checks confirm that the reported thermal trends—particularly the relative performance of the uniform-flow baseline, alternative layouts, and the optimized partitioned design—are insensitive to reasonable variation in operating conditions. The primary metrics extracted from simulation include peak junction temperature, temperature non-uniformity across the MCM, and pressure drop across the cold plate.

C. *Modeling Assumptions and Limitations*

The numerical model employs several assumptions that are appropriate for the present study but should be noted when interpreting the results. Flow within the microchannels is assumed to be laminar and fully developed, which is consistent with the channel dimensions and operating Reynolds numbers considered. The Reynolds number within the microchannels is less than 2,000 across all evaluated configurations, remaining within the laminar regime. Turbulence effects and entrance-region development are therefore neglected.

In addition, material properties are treated as temperature-dependent where available; however, potential variations due to manufacturing tolerances and interfacial imperfections are not explicitly modeled. Radiation heat transfer is neglected due to the dominance of conduction and convection under the operating conditions studied. While these assumptions may influence absolute temperature predictions, the comparative trends between baseline, alternative layouts, and optimized designs are expected to remain valid.

D. *Manifold Design for Flow Partitioning*

Flow partitioning is achieved through manifold-enabled flow distribution while maintaining uniform microchannel geometry. Inlet and outlet placement is optimized to reduce pressure drop, but detailed manifold optimization is treated as an enabling consideration.

E. *Design Constraints*

Designs are constrained by fabrication limits, footprint, and allowable pressure drop. Uniform microchannel geometry is maintained to facilitate silicon fabrication.

IV. Flow Partitioning and Optimization Methodology

To evaluate the effectiveness of flow partitioning under non-uniform power dissipation, a sequence of cold plate layouts and flow distribution strategies is examined. The objective is to reduce peak junction temperature and improve temperature non-uniformity while maintaining an acceptable

thermal–hydraulic operating point under a fixed total coolant flow rate.

As a starting point, a baseline uniform-flow configuration is defined. In this study, the baseline refers to the uniform-flow cold plate layout shown in Fig. 4(a), in which separate cold plate regions are used for the XPU and HBMs and coolant flow is distributed evenly without regard to local power density. Under this configuration, each device region receives an equal fraction of the total coolant flow, simplifying flow delivery but neglecting spatial variation in thermal demand.

Simulation results for the baseline configuration exhibit peak-to-minimum temperature differences of up to 30 °C, indicating inefficient cooling under strongly non-uniform power dissipation. These large temperature gradients arise from a mismatch between coolant delivery and local heat generation: regions containing hot spots experience insufficient convective heat removal, while low-power regions are overcooled. Increasing total flow rate under these conditions provides limited benefit, as the additional flow is not preferentially directed toward the dominant thermal bottlenecks.

To explore the impact of coolant footprint and flow routing, multiple cold plate layout options are evaluated, as illustrated in Fig. 2(b)–(d). These layouts vary inlet and outlet placement as well as liquid coverage over the MCM, enabling assessment of their influence on both thermal performance and hydraulic cost. The configuration that minimizes peak junction temperature while maintaining reasonable pressure drop is selected as the reference design for subsequent optimization.

Following layout selection, the cold plate is conceptually partitioned into multiple flow regions corresponding to distinct thermal zones on the MCM. Regional flow rates are then adjusted to reduce hot spot temperatures and minimize temperature non-uniformity (ΔT), subject to a constant total flow rate. Importantly, microchannel geometry and density are kept uniform across the cold plate to preserve manufacturability.

Flow partitioning is implemented through the combined use of manifold design and internal flow dividers. Similar approaches based on controlled hydraulic resistance and manifold shaping have been explored for microchannel heat sinks subjected to non-uniform heat loads [3], [6]. However, many prior implementations rely on channel geometry variation or topology optimization that may complicate manufacturability. In contrast, the present approach preserves uniform microchannel geometry while achieving redistribution solely through passive manifold and divider design. These features introduce controlled hydraulic resistance between regions, enabling targeted flow redistribution without active control. The optimization process is performed iteratively by adjusting regional hydraulic resistance to redistribute flow among the partitioned regions while maintaining constant total flow rate.

Candidate designs are evaluated using two primary metrics: peak junction temperature and pressure drop. Designs that reduce peak temperature at the expense of excessive pressure drop are rejected in favor of configurations that provide balanced thermal–hydraulic performance. Rather than pursuing a single global optimum, the optimization seeks a practical operating window that minimizes temperature non-uniformity while remaining compatible with system-level pumping power constraints. This approach reflects realistic design tradeoffs encountered in high-power cooling systems.

V. EXPERIMENTAL SETUP

A dedicated experimental setup is developed to validate the numerical predictions of the proposed flow-partitioned cooling strategies. The test vehicle utilizes the same molded package architecture described previously, ensuring consistency between simulation and experiment. A silicon microchannel cold plate is bonded to the backside of the molded MCM, and coolant is supplied through a controlled flow loop as shown in Fig. 3.

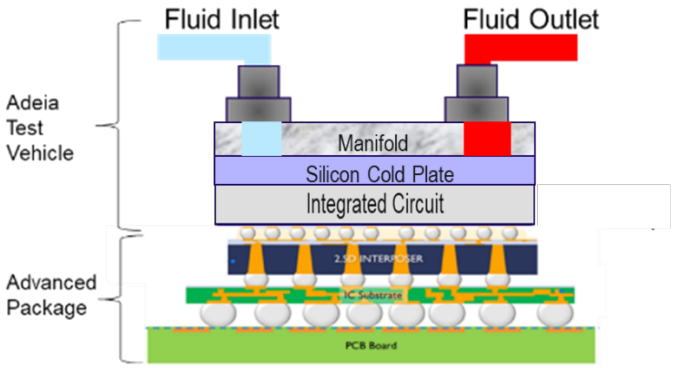


Figure 3. Cross-sectional view of the experimental test vehicle showing manifold, silicon microchannel cold plate, molded MCM, and PCB stack.

The coolant used in all experiments is deionized water. Flow rate is regulated using a calibrated pump and flow meter, and inlet coolant temperature is maintained constant for all test cases to ensure repeatability. Differential pressure across the cold plate is measured using pressure sensors located at the inlet and outlet manifolds.

Temperature sensors are embedded within the devices at selected locations, including regions proximal to known hot spots in the XPU. These sensors provide junction-adjacent temperature readings during operation, enabling spatially resolved thermal characterization of the MCM.

Prior to testing, all temperature sensors are calibrated using a controlled thermal environment to minimize systematic error. Pressure sensors and flow meters are similarly calibrated to ensure accurate measurement of hydraulic performance. During testing, power is applied to the devices according to the prescribed power map, and steady-state conditions are reached before data acquisition. Steady state is defined by stable temperature readings over a sustained time window, after which temperature and pressure data are averaged to reduce the influence of measurement noise.

Repeatability is assessed by comparing steady-state measurements across multiple runs under identical operating conditions. Each experimental condition was repeated at least three times to ensure measurement consistency. The variation in both temperature and pressure drop across repeated runs was within 2% for all monitored quantities. Reported values represent the arithmetic average of these repeated measurements. This repeatability confirms the stability of the test platform and the robustness of the measured trends. The resulting variation provides a practical estimate of experimental scatter, which is used to contextualize simulation-to-test discrepancies reported in Section VI. The experimental setup described above therefore enables direct validation of the numerical flow partitioning methodology under realistic packaging and thermal boundary conditions.

VI. Results and Discussion

The thermal and hydraulic performance of the proposed flow partitioning strategies is evaluated using both numerical simulation and experimental measurement. Results are presented to compare baseline uniform-flow operation, alternative cold plate layouts, and the optimized partitioned-flow design. Experimental data obtained using the test vehicle described in Section V are used to validate the numerical predictions and to assess the effectiveness of the proposed approach.

First, the influence of cold plate layout and coolant footprint is evaluated using the design options in Fig. 4 and summarized in Table II. Next, the optimized partitioned-flow configuration is discussed using the results in Table III to highlight thermal–hydraulic tradeoffs. Finally, simulation predictions are compared against experimental measurements (Table IV) to validate the modeling approach and confirm the effectiveness of flow partitioning in practice.

A. Impact of Flow Partitioning on Thermal Performance

Optimized flow partitioning significantly reduces peak junction temperature while simultaneously lowering pressure drop at constant total flow rate. Compared to uniform-flow operation, redistributing coolant flow toward high-power regions increases local convective heat transfer where it is most needed, while avoiding unnecessary flow through low-power regions.

This targeted flow delivery improves thermal uniformity across the MCM and enables more efficient utilization of available coolant. The results highlight that peak temperature reduction is achieved not by increasing total flow rate, but by improving the spatial matching between coolant delivery and power dissipation.

The observed temperature reduction is primarily driven by redistributing coolant flow to increase local Reynolds number and convective heat transfer coefficient in regions aligned with dominant hot spots. At the same time, reducing flow in low-power regions decreases unnecessary shear and viscous losses, contributing to the reduced pressure drop. Overall, flow partitioning effectively reshapes the thermal–hydraulic resistance network of the cold plate to better match the spatial power dissipation profile.

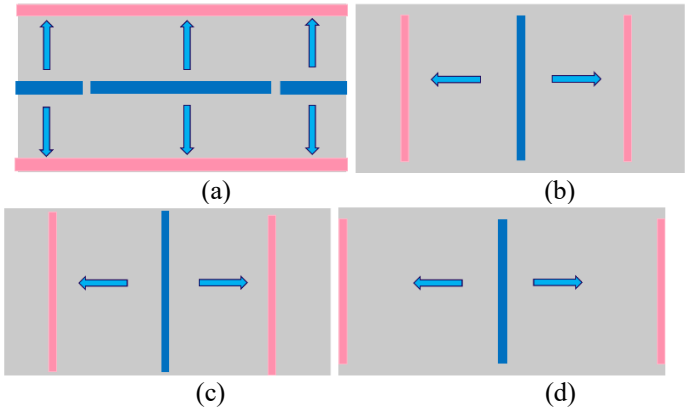


Figure 4. Cold plate layout configurations evaluated under constant total volumetric flow rate: (a) Separate cold plates for XPU and HBMs (uniform-flow baseline); (b) Single cold plate with split-flow distribution; (c) Split-flow configuration with extended vertical coolant coverage; (d) Split-flow configuration with extended horizontal coolant coverage. Arrows indicate inlet-to-outlet flow direction.

Table II. Temperature ($T_{j,max}$) and Pressure Drop (ΔP) Results for Layout Options

Layout Option	$T_{j,max}$ (°C)	ΔP (psi)
a (baseline)	90.8	25
b	69.6	71
c	70.6	60
d	68.2	93

As shown in Table II, option (d) yields the lowest peak temperature but incurs the highest pressure drop, making it impractical. Option (b) provides a more balanced alternative; however, its pressure drop remains significantly higher than that of the uniform-flow design (option (a)). Consequently, further effort is devoted to optimizing flow distribution to identify a thermal–hydraulic sweet spot between options (a) and (b).

The results indicate an inherent coupling between thermal performance and hydraulic resistance, consistent with prior studies [4], [7]. Accordingly, the optimization objective is defined as minimizing hydraulic cost under fixed volumetric

flow rate while preserving substantial peak temperature reduction.

To fine-tune regional flow allocation, flow dividers are introduced as barriers to separate flows and achieve appropriate flow distribution. These dividers are implemented in both the cold plate and the manifold. After extensive simulation DOE, the optimized layout shown in Fig. 5 achieves an improved balance between temperature reduction and pressure drop.

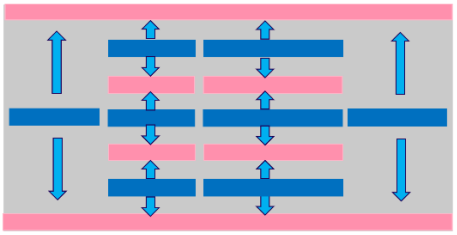


Figure 5. Optimized partitioned-flow cold plate configuration showing regional flow segmentation within the XPU area. Internal flow dividers and manifold routing features are used to redistribute coolant toward high-power regions while maintaining uniform microchannel geometry. The design achieves a balanced thermal–hydraulic operating point under constant total volumetric flow rate.

Compared with options (a) and (b), the XPU portion of the cold plate is divided into multiple regions, enabling targeted flow distributions that better match local power dissipation. Higher coolant flow is directed toward regions containing dominant hot spots, while lower flow is supplied to regions with reduced thermal demand. In parallel, the manifold design is refined to reduce hydraulic resistance while supporting the optimized regional flow rates. The reduction in pressure drop relative to the uniform-flow baseline arises from both flow redistribution and geometric refinement of the manifold. In the baseline configuration, coolant traverses the full channel length across all regions regardless of local thermal demand, resulting in unnecessary viscous dissipation in low-power areas. In contrast, the optimized partitioned layout shortens the effective flow path in selected regions, reduces coolant coverage in thermally non-critical areas, and lowers average velocity where high convective heat transfer is not required. These effects collectively reduce frictional and minor losses, yielding a lower overall hydraulic resistance despite the introduction of flow dividers. Because pressure drop scales approximately with the square of velocity in laminar microchannel flow when accounting for developing and manifold losses, reducing velocity in low-demand regions produces a disproportionate reduction in hydraulic penalty. The resulting thermal and hydraulic performance is summarized in Table III.

TABLE III. EXPANDED RESULTS WITH OPTIMIZED LAYOUT OPTION

Layout Option	T_j,max (°C)	ΔP (psi)
a	90.8	25
b	69.6	71
c	70.6	60
d	68.2	93
Optimized	78.4	8

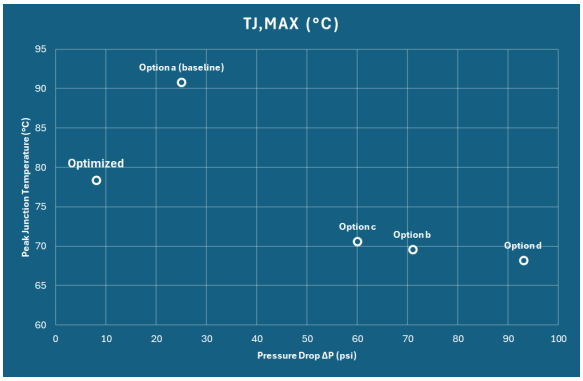


Figure 6. Peak junction temperature versus absolute pressure drop for evaluated configurations

Figure 6 presents peak junction temperature as a function of absolute pressure drop for all evaluated configurations. The distribution of points highlights the thermal–hydraulic trade space, with temperature-minimizing layouts located at higher pressure drop and the optimized design positioned in a lower-pressure region while maintaining significant temperature improvement.

It should be noted that the optimized configuration does not minimize the peak junction temperature in isolation. All layouts are evaluated at constant total volumetric flow rate; therefore, differences in pressure drop directly translate to differences in required pumping power. While layout (b) achieves a lower absolute peak temperature, it incurs substantially higher pressure drop. The optimized configuration is therefore selected to minimize pumping power while preserving a substantial reduction in peak temperature relative to the baseline.

These results emphasize the practical importance of balancing thermal performance with hydraulic cost. While temperature-minimizing layouts incur substantial pressure penalties, the optimized partitioned design achieves meaningful temperature reduction at significantly lower pumping requirements. Such balance is critical for system-level integration, where cooling overhead directly affects overall energy efficiency and infrastructure scaling.

For experimental validation, a dedicated test vehicle is utilized using the same molded package architecture. Temperature sensors are embedded within the devices at selected locations, enabling direct measurement of junction-adjacent temperatures during operation. These embedded sensors provide spatially resolved temperature data for comparison with simulation predictions. For pressure drop validation, pressure sensors located at the inlet and outlet of the manifold measure differential pressure across the cold plate at multiple flow rates, as shown in Fig. 7.

TABLE IV. SIMULATED VS. MEASURED TEMPERATURE COMPARISONS

Monitor #	Simulated (°C)	Measured (°C)	Differences	
1	72.4	73.3	-0.9	-3%
2	73.3	74.6	-1.3	-4%
3	74.7	73.1	1.6	5%
4	72.0	73.6	-1.6	-5%

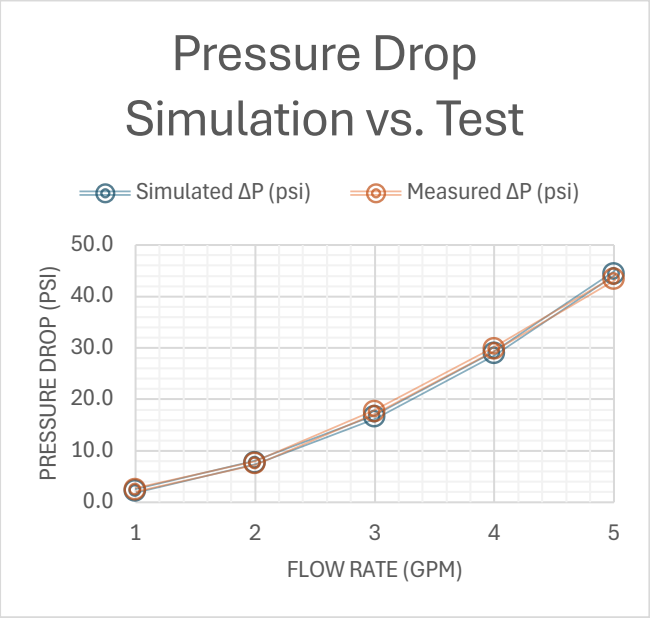


Figure 7. Simulated vs measured pressure drop plot

Agreement between simulation and experimental measurements is evaluated using temperature data collected at multiple sensor locations. As shown in Table IV, discrepancies between simulated and measured temperatures are well within 10% across all monitored points. Considering temperature sensor tolerances, the agreement is quite acceptable.

Agreement between simulated and measured pressure drop values is within 5% for all evaluated configurations, as shown iner Fig. 7. This validates the hydraulic accuracy of the numerical model, including manifold and divider-induced losses.

Sources of experimental uncertainty include sensor placement tolerance, sensor calibration accuracy, and variations in coolant inlet temperature and flow rate. Despite these factors, the observed agreement demonstrates that the conjugate heat transfer model accurately captures the dominant thermal and hydraulic behavior of the system.

The validated model is therefore suitable for evaluating alternative flow-partitioning strategies and for guiding design decisions in similar high-power multi-chip modules.

B. Pumping Power Trade-Offs

The optimized design achieves approximately 70% lower pressure drop than temperature-minimizing split-flow configurations, translating directly to reduced pumping power under constant total flow rate. Beyond lowering energy consumption, reduced hydraulic resistance provides increased system-level design margin, enabling smaller pumps, improved flow stability, and greater tolerance to manufacturing variation. These benefits underscore that practical cooling optimization must consider hydraulic efficiency alongside thermal performance, particularly in high-power computing environments where infrastructure scalability and operational cost are critical.

C. Design Implications

The results suggest that effective thermal management of non-uniform power dissipation does not require increased coolant flow rate, but rather improved spatial control of flow delivery. From a design perspective, this enables higher power operation within existing pumping and cooling infrastructure, reducing system-level cost and complexity.

D. Design Trade Space Considerations

The results illustrate that optimal cooling performance exists within a multidimensional design space defined by thermal performance, pressure drop, manufacturability, and system constraints. Designs that aggressively minimize peak junction temperature tend to incur higher hydraulic penalties, while overly conservative designs sacrifice thermal margin. The optimized partitioned configuration represents a compromise within this space, balancing these competing objectives.

Such trade space considerations are increasingly important as cooling solutions are integrated at the system level, where power delivery, pumping capacity, and reliability constraints must be considered alongside thermal performance.

E. Scalability and Generality

Although the present study focuses on a specific five-chip molded MCM, the proposed flow partitioning strategy is general and applicable to a wide range of package architectures. The methodology relies on matching coolant delivery to spatial power distribution rather than on specific channel geometries, making it well suited for future devices with higher power densities, heterogeneous integration, and evolving power maps.

As power densities continue to increase in heterogeneous integration platforms, manifold-aware and partitioned-flow cold plate concepts are expected to play an increasingly important role in improving system-level energy efficiency and thermal reliability [9]. The present methodology provides a manufacturable pathway for implementing such strategies in silicon-based cold plates.

The approach can be extended to systems with different chip counts, asymmetric layouts, and multiple hot spot locations, provided that the power distribution is known a priori.

F. Limitations and Future Work

The present study focuses on demonstrating the thermal benefits of flow partitioning at the cold plate and system level, rather than providing a detailed treatment of manifold design methodology. While the optimized layouts achieve up to a 70% reduction in pressure drop, the specific geometric optimization strategies employed within the manifold are not discussed in detail here.

Future work will focus on a systematic investigation of manifold design parameters, including inlet and outlet geometry, flow divider placement, and regional hydraulic resistance tuning. These studies will aim to establish general design guidelines for achieving low-pressure-drop flow partitioning across a broader range of power maps and package configurations. In addition, extension of the present approach to multiple hot spot scenarios and dynamic power distributions will be explored.

VII. Conclusions

This work demonstrates that optimized flow partitioning using a manufacturable uniform-channel silicon microchannel cold plate can significantly improve thermal–hydraulic performance in high-power multi-chip modules implemented as molded packages. By redistributing coolant flow to match non-uniform power dissipation, the proposed approach reduces peak junction temperature by more than 10 °C while substantially lowering pressure drop. The combination of simulation and experimental validation confirms that flow partitioning provides a practical and scalable cooling strategy for next-generation electronic systems. At the system level, the proposed flow partitioning approach enables improved thermal performance without increasing total coolant flow rate, reducing both pumping power and overall energy consumption. This capability is particularly relevant for high-power computing systems, where thermal efficiency, power delivery, and cooling overhead increasingly constrain performance scaling.

References

[1] J. Kim, J. H. Shin, S. Sohn, and S. H. Yoon, “Analysis of non-uniform flow distribution in parallel micro-channels,” *J. Mech. Sci. Technol.*, vol. 33, no. 8, pp. 3859–3864, Aug. 2019.

[2] Y. S. Ong and K. Ku Shaari, “CFD investigation of the effect of manifold and microchannel ratio on the hydrodynamic performance of microchannel heat sink,” *SN Appl. Sci.*, vol. 2, Art. no. 1199, Jun. 2020.

[3] W. Li, Z. Zhao, S. Li, and D. Liu, “Analysis of a parallel microchannel heat sink with tailored flow maldistribution under non-uniform heat sources,” *Appl. Therm. Eng.*, vol. 282, Art. no. 128804, Jan. 2026.

[4] D. Wang, H. Song, G. Wang, Y. Yang, S. Wang, and S. Xiang, “Optimal arrangements of inlet and outlet in topology liquid-cooled microchannel heat sink based on multi-objective optimization,” *Int. J. Therm. Sci.*, vol. 209, Art. no. 109552, Mar. 2025.

[5] J. Wilhite and C. Kharangate, “Fluid and thermal analysis of a manifold microchannel heat sink,” in *Proc. Thermal and Fluids Analysis Workshop (TFAWS)*, Aug. 26–30, 2024.

[6] J. Zhou, J. Chen, Q. Wang, X. Xie, P. Guan, and H. Zheng, “Numerical study on optimization of manifold microchannel heat sink,” *Energies*, vol. 18, no. 22, Art. no. 5883, 2025.

[7] B. Cao and Z. Wu, “Microchannel heat sinks for hotspot thermal management: achieving minimal pressure drop and maximal thermal performance,” *Int. J. Heat Mass Transfer*, vol. 236, pt. 2, Art. no. 126411, Jan. 2025.

[8] R. Kumar, G. Singh, and D. Mikielewicz, “Numerical study on mitigation of flow maldistribution in parallel microchannel heat sink: channels variable width versus variable height approach,” *J. Electron. Packag.*, vol. 141, no. 2, Art. no. 021009, Jun. 2019.

[9] Z. Wu, G. Zhang, S. Lu, P. Leng, Y. Yu, J. Deng, and W. Huang, “A comprehensive review of cold plate liquid cooling technology for data centers,” *Chem. Eng. Sci.*, vol. 310, Art. no. 121525, May 2025.

[10] D. B. Tuckerman and R. F. W. Pease, “High-performance heat sinking for VLSI,” *IEEE Electron Device Lett.*, vol. 2, no. 5, pp. 126–129, May 1981.

[11] S. G. Kandlikar, “High flux heat removal with microchannels—A roadmap of challenges and opportunities,” *Heat Transfer Eng.*, vol. 26, no. 8, pp. 5–14, 2005.

[12] K. S. Ryu, S. J. Kim, and H. H. Bau, “Thermal performance of microchannel heat sinks subject to non-uniform heating,” *J. Heat Transfer*, vol. 125, no. 1, pp. 75–83, Feb. 2003.

[13] A. Bar-Cohen and P. Wang, “Thermal management of on-chip hot spots and 3D chip stacks,” *J. Heat Transfer*, vol. 134, no. 5, Art. no. 051017, 2012.

[14] R. Zhang, G. Fountain, K. M. Bang, L. Mirkarimi, S. Sadiq, and A. Avellan, “Revolutionary thermal solutions for hot chips,” in *Proc. IEEE ITherm*, Dallas, TX, USA, May 2025.