Advanced and Sustainable Technologies
ASET
E-ISSN 2976-2294
Volume 3, No 2, December 2024 [51-60]

Flow Analysis in Cooling Channels for Graphics Processing Unit (GPU) Block

Mohd Izzat Aiman Adi[1*]

[1]Faculty of Mechanical Engineering & Technology, Universiti Malaysia Perlis (UniMAP), 02600 Arau, Perlis, Malaysia

Received 11 August 2024, Revised 4 September 2024, Accepted 8 September 2024

ABSTRACT

This study investigates the dynamics of fluid flow in the cooling channels of the graphics processing unit (GPU) blocks, focusing on the effects of varying inlet and outlet diameters as well as the shape of the pipes on the maximum pressure and temperature within the system. The primary objectives are: first, to evaluate the different inlet and outlet diameters (5 mm, 10 mm, and 15 mm) that impact peak pressure and temperature; second, to examine the influence of pipe designs on these parameters. Computational Fluid Dynamics (CFD) simulations were conducted to model and analyze these variations using SimFlow 4.0. The findings reveal significant correlations between the geometric configurations and the flow characteristics, with specific diameters and designs leading to notable changes in maximum temperature. These insights provide a deeper understanding of optimizing cooling channel designs for improved thermal management in GPUs. The study concludes by emphasizing the critical role of geometric parameters in the performance of cooling systems and suggesting directions for further research.

Keywords: Graphics Processing Unit, Simulation and modeling, Computational fluid dynamics.

1. INTRODUCTION

Effective thermal management is vital in modern computing systems, particularly with the growing demand for high-performance graphics processing units (GPUs) [1]. These GPUs, known for their intensive computational tasks, such as rendering graphics [2], machine learning [3], and scientific simulations [4], generate substantial heat. To ensure optimal performance and prevent overheating, it is crucial to implement efficient cooling mechanisms. These mechanisms maintain the GPU's performance and extend the hardware's lifespan by preventing thermal degradation. A key component in GPU cooling systems is the cooling block [5], which plays a critical role in thermal management. The cooling block facilitates the transfer of heat away from the GPU to the cooling medium, such as water or air. This process is essential because it directly affects the temperature regulation of the GPU, ensuring that it operates within safe temperature ranges.

Modern cooling blocks are designed with precision to maximize heat dissipation. They often feature complex designs with optimized internal structures that promote efficient fluid flow and heat transfer [6]. By enhancing the contact area between the GPU and the cooling medium, these blocks significantly improve the cooling system's overall efficiency. In addition, advancements in cooling block materials and manufacturing techniques have improved thermal conductivity and durability. High-quality materials like copper and aluminum [7] are commonly used for their excellent thermal properties, while innovative manufacturing methods enable the creation of more complex and compelling designs [8]. The cooling block is critical in GPU cooling systems,

*Corresponding author: mhd.izzataiman@gmail.com

and crucial for maintaining optimal thermal performance. As the demand for high-performance GPUs continues to rise, effective cooling solutions, particularly well-designed cooling blocks, become increasingly paramount.

Previous work on cooling systems and GPU blocks highlights significant advancements in the design and efficiency of these systems, particularly for high-density electronic devices. High-density cooling systems are specifically engineered to ensure efficient cooling of multiple electronic devices densely packed in a cooling bath [9]. These systems typically employ vertical and lateral inner barrier plates that divide the cooling bath into multiple accommodating sections, optimizing the cooling process. The coolant circulation involves a flow passage with a pump and a heat exchanger, where the coolant is continuously circulated, cooled by the heat exchanger, and reintroduced into the cooling bath, ensuring stable and efficient cooling performance. Components of such systems include a cooling bath with liquid coolant, a processor (which can be a CPU, GPU, or both), and a heat spreader [10]. Maintaining the coolant level to immerse all heat-generating components reduces evaporative losses. It prevents local boiling on the processor surface.

Heat exchangers play a critical role in these systems. With their placement options being flexible, they can be positioned inside or outside the cooling bath, above or below the liquid level, to optimize cooling efficiency and system configuration. Systems may use a single heat exchanger [11] for one cooling bath or multiple distributed exchangers [12], which can simplify the system's piping and overall configuration. The processors cooled by these systems can include CPUs, GPUs, or combinations of both, and the cooling system configurations are adaptable to various types and numbers of processors. Efficient heat management often involves direct cooling mechanisms [13], such as liquid coolant headers with multiple nozzles to distribute the coolant evenly across the processors, ensuring uniform cooling performance. The advantages of using liquid coolants in these systems are notable. The coolants are chosen for their low evaporative properties, which help maintain a stable coolant level over extended periods, reducing maintenance needs and improving overall reliability. Using forced circulation through headers with radially dispersed nozzles enhances the direct cooling effect [14], ensuring that all parts of the electronic devices, especially GPUs, are adequately cooled. These cooling system design and technology advancements significantly contribute to the performance and longevity of high-density electronic devices.

The primary objective of this project is to analyze the effects of varying inlet and outlet diameters on peak pressure and temperature in cooling channels, as well as to explore how pipe shape influences these parameters. The design of the GPU Block was created using the computer-aided design (CAD) software SOLIDWORKS [15]. The methodology involves the modeling with laminar flow assumptions and a steady-state solver, specifically the CHT Multi Region SIMPLE solver. The simulations utilize water as the fluid and aluminum as the plate material. It is anticipated that the results of the CFD simulations, visualized in Paraview to display temperature and pressure contours, will offer valuable insights into the performance of the GPU Block, thereby informing enhancements in cooling system design.

2. MATERIAL AND METHODS

This study employs SimFlow software [16] to import STL models and conduct simulations to understand the inlet and outlet diameter effects on maximum temperature and pressure. The model (Figure 1), distinguished by diameters of 5 mm, 10 mm, and 15 mm, is initially imported into SimFlow, where meshing properties are set with refinement to Min “1” and Max “1” for the cooler and pipe parts, while for the plate part, just setting a mesh geometry. The base mesh type is a box automatically adjusted to fit each model (6 meshed models are available in Figure 2). For the model and material point divisions, different values have been set according to the size and shape of the diameter. Refer to Table 1 for the details of the GPU Block. At axes X-, X+, Y-, Y+, Z-, and Z+ have been set as boundaries and all other surfaces as patches.

The simulation setup adopts a steady-state, compressible flow model with the CHT Multi Region SIMPLE solvers and employs laminar modeling. The thermophysical properties are water for the fluid and aluminum for the plate. Boundary conditions are configured with the inlet set as a velocity inlet with a surface normal fixed value and a reference value of 100.0 m/s. At the same time, the outlet is a pressure outlet with a fixed value of 100000 Pa and a type of velocity of zero gradient. Each simulation runs for 500 iterations on a dual-core CPU in serial mode. Post-simulation focuses on maximum pressure and temperature values extracted from the Paraview. The simulation is repeated for each model to comprehensively analyze how variations in inlet and outlet diameter impact the maximum pressure and temperature on flow distribution cooling channels of GPU Block.

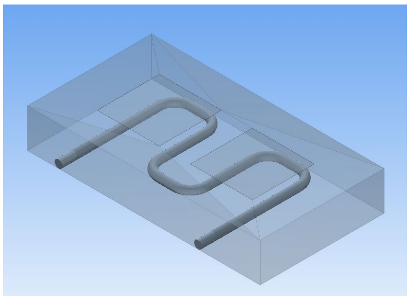
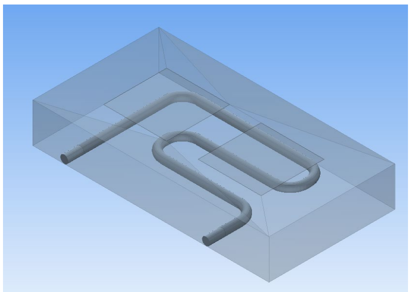
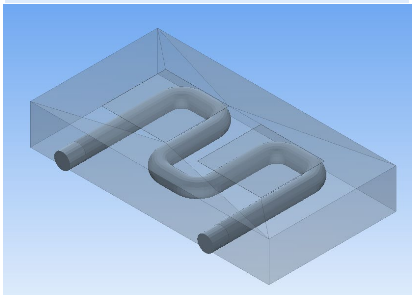
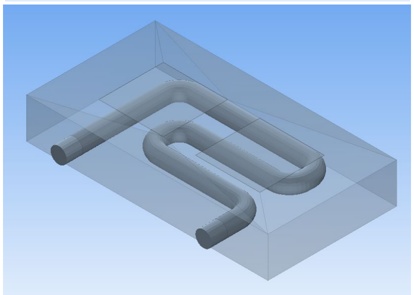
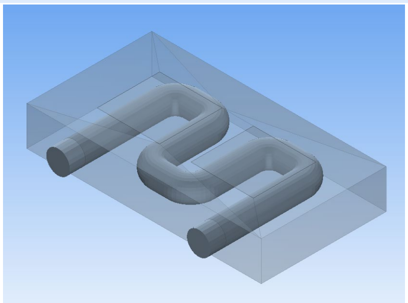
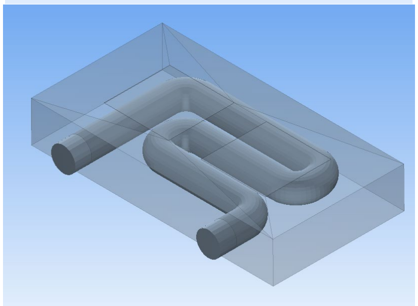
Size	**Design 1**	**Design 2**
5mm		
10 mm		
15 mm		

Figure 1: Size and shape of pipe GPU block.

Table 1: Meshing data for different designs and inlet-outlet diameters.

Base Mesh	Division	Design 1			Design 2		
		5 mm	**10 mm**	**15 mm**	**5 mm**	**10 mm**	**15 mm**
Material Point	Fluid	30, 55, 106	31, 65, 155.5	33, 117.5, 14	30.5, 55, 106	31, 76.5, 155.5	31, 141, 12
	Plate	70, 75, 102.5	77.5, 115, 155	78, 177.5, 2	75, 65, 105	77.5, 122, 155	80, 186, 13
Number of Cells		42 739	35 320	44 737	49 037	38 365	49 112
Number of Nodes		50 454	41 780	53 217	57 503	45 536	57 742
Fluid Cells		4 080	7 653	14 002	5 295	8 543	15 843
Fluid Nodes		6 962	10 666	19 734	8 489	12 135	22 602
Plate Cells		38 659	27 667	30 735	43 742	29 822	33 269
Plate Nodes		48 486	36 350	41 329	54 651	39 719	44 558
Skewness		3.4	1.36	11.23	1.86	6.33	5.3
Aspect ratio		12.63	7.34	8.42	12.03	8.03	8.59
Total Mesh size		85k	70k	89k	98k	76k	98k

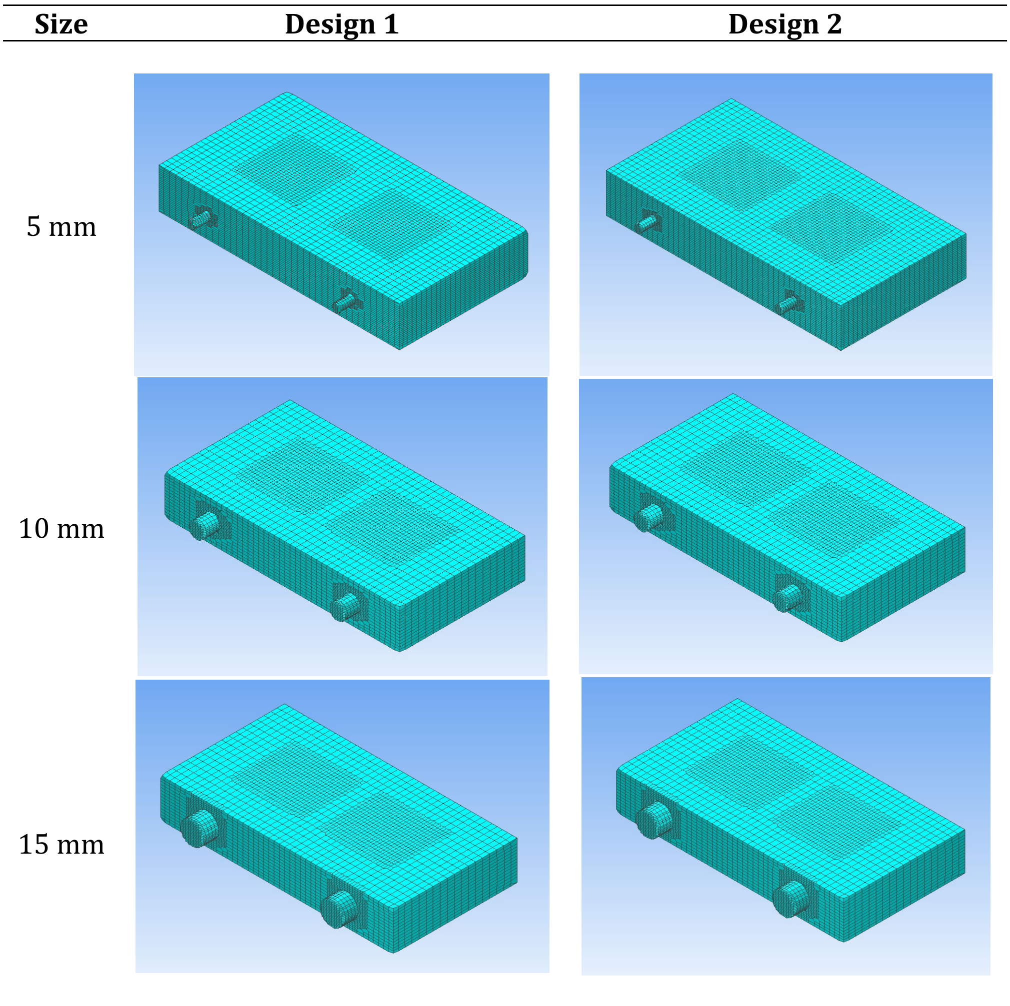


Figure 2: Meshed model for different designs and inlet-outlet diameters.

3. RESULTS AND DISCUSSION

The current study evaluates the impact of varying inlet and outlet diameters and pipe shapes on the peak pressure and temperature within GPU blocks' flow distribution cooling channels. The findings from the CFD simulations provide valuable insights into the optimization of cooling channel designs for improved thermal management in GPUs. The simulation results clearly correlate the inlet and outlet diameters and temperature within the cooling channels (Figures 3-4). Figures 5–8 show the temperature, pressure, and velocity contours of both GPU designs in the simulation. As the diameter increased from 5 mm to 15 mm, there was a significant rise in maximum pressure, with Design 1 showing an increase from 8.0e+06 Pa to 5.4e+07 Pa and Design 2 showing a rise from 8.0e+06 Pa to 5.0e+07 Pa. This indicates that larger diameters can handle higher pressure but may also increase stress on the cooling system components. Temperature variations were also notable. For Design 1, the temperature decreased from 400 K at 5 mm to 300 K at 10 mm and 15 mm (Figure 3). Design 2 exhibited consistent maximum temperatures of 300 K across all diameters (Figure 2). These results suggest that while larger diameters can effectively manage temperatures within an optimal range, the specific design of the pipe also plays a crucial role in thermal performance.

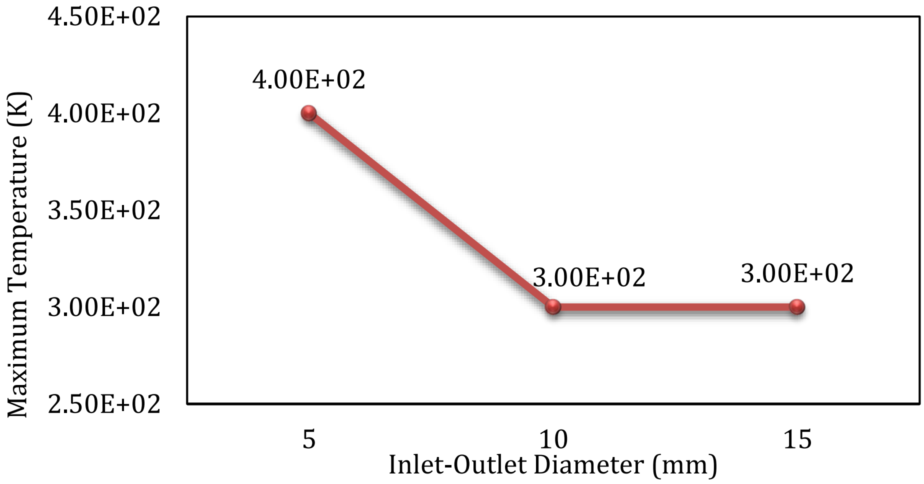


Figure 3: Maximum temperature for different inlet-outlet diameters using Design 1.

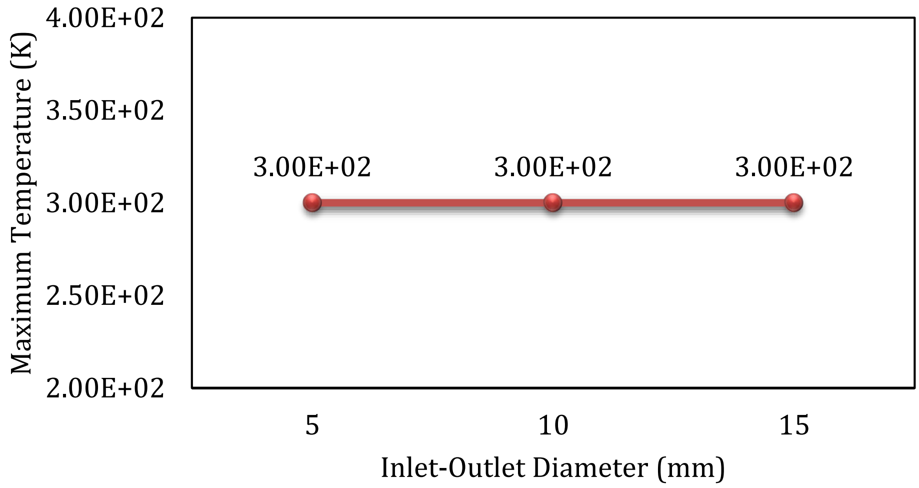


Figure 4: Maximum temperature for different inlet-outlet diameters using Design 2.

The study compared two pipe designs to determine their influence on pressure and temperature. Design 1 generally exhibited higher maximum pressures across all diameters than Design 2, indicating that the geometric design of the pipes significantly affects the flow characteristics. Figure 5 presents the temperature contours for Design 1 with three different inlet-outlet diameters (5 mm, 10 mm, and 15 mm). This result illustrates how the temperature distribution within the cooling channel changes with varying diameters. The contours show areas of higher and lower temperatures, allowing for a visual assessment of the cooling effectiveness at different sizes. Smaller diameters, like 5 mm, might show higher temperatures due to restricted flow, while larger diameters, such as 15 mm, could display a more uniform cooling distribution, indicating a more effective heat dissipation.

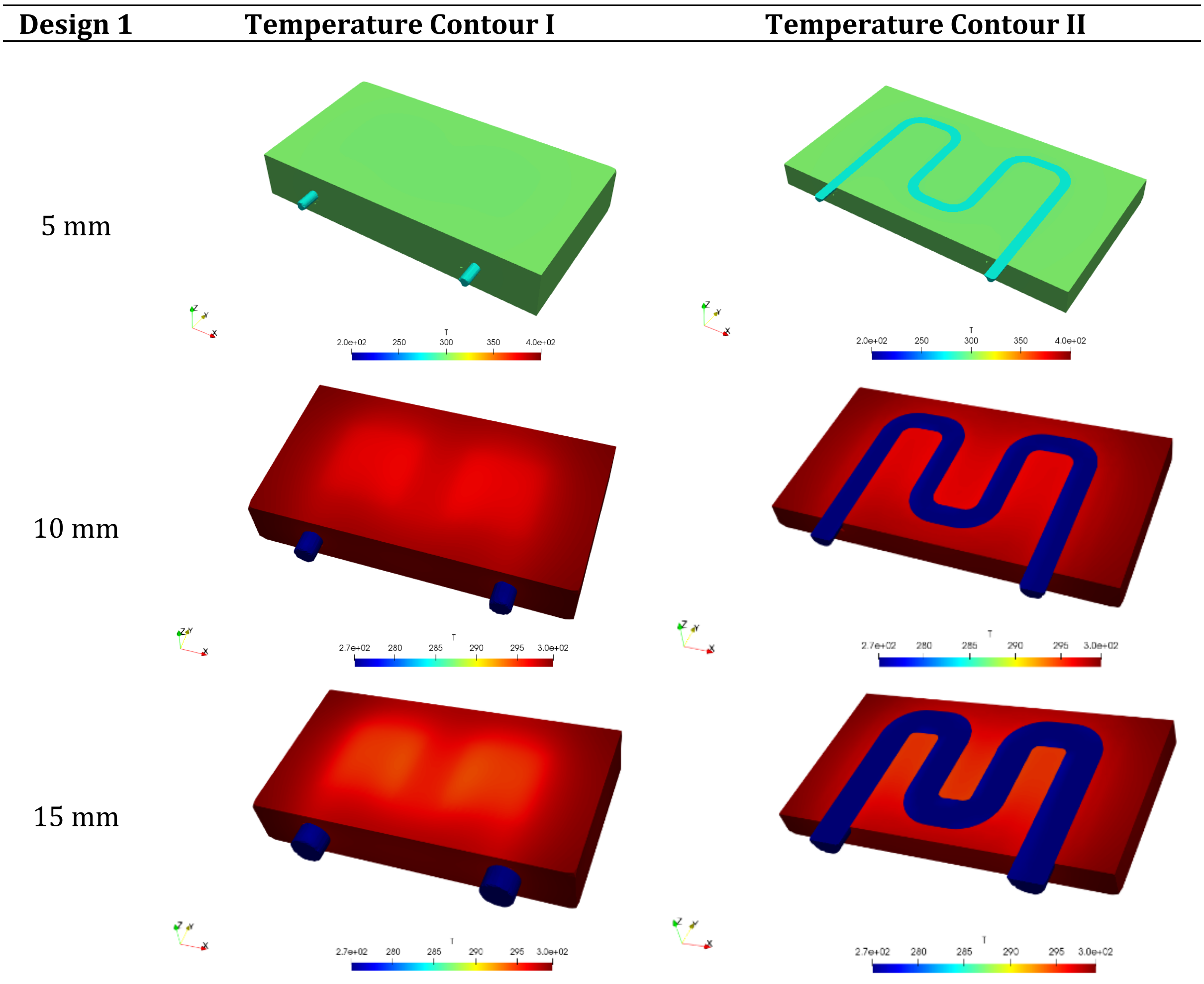


Figure 5: Temperature contour for different inlet-outlet diameters using Design 1.

Figure 6 shows the temperature contours for Design 2, using the three inlet-outlet diameters (5 mm, 10 mm, and 15 mm). Comparing this to Figure 5, the impact of the pipe design on temperature distribution becomes evident. The result highlights differences in thermal management between the two designs and how changes in the cooling block design influence the system's maximum temperature and overall thermal efficiency. Both designs maintained similar temperature profiles when 10 mm and 15 mm were used, suggesting that while the shape impacts pressure distribution, it may not significantly alter the thermal distribution. These findings have practical implications for designing and optimizing cooling channels in GPUs. Managing pressure and temperature can enhance the GPU components' performance and longevity. Designers can use these insights to choose appropriate diameters and designs based on specific cooling requirements and system constraints.

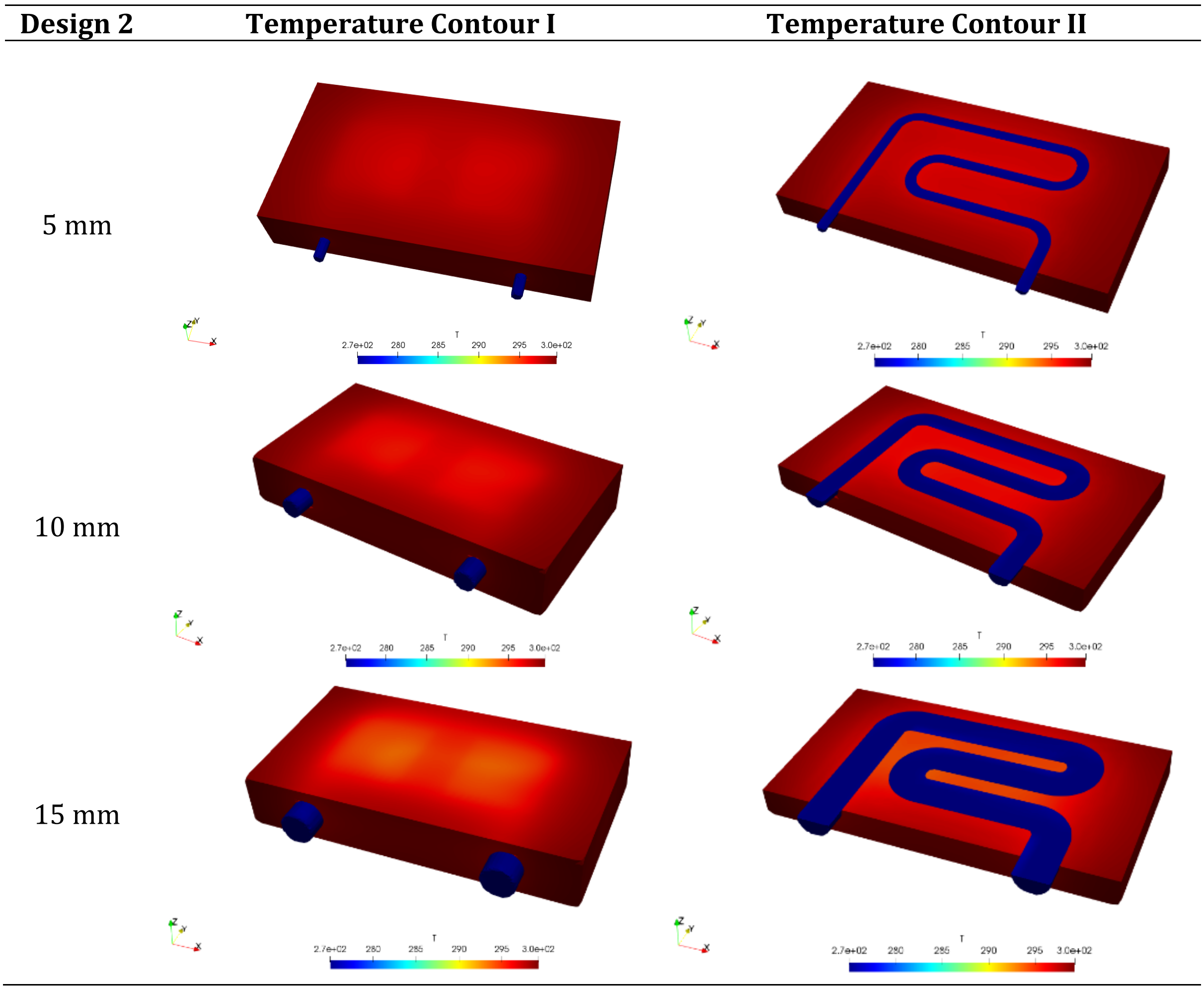


Figure 6: Temperature contour for different inlet-outlet diameters using Design 1.

Figure 7 illustrates the pressure and velocity contours for Design 1 with varying inlet-outlet diameters. The pressure contours indicate regions of high and low pressure within the cooling channel, which directly affects the fluid flow dynamics. Higher pressures might be observed at smaller diameters due to increased resistance, while larger diameters may result in lower pressures and more efficient flow. The velocity contours complement this by showing how fluid speed varies within the channel, with faster flows potentially occurring in narrower sections. Figure 8 shows the pressure and velocity contours for Design 2 with different inlet-outlet diameters, similar to Figure 7. This figure helps to compare the two designs' impact on pressure and velocity distributions. Differences between Design 1 and Design 2 in handling fluid dynamics can provide insights into optimizing cooling performance through design alterations. The contours might reveal that Design 2 offers a more balanced pressure distribution or consistent velocity profiles, suggesting potential advantages in specific operational scenarios.

From the fluid mechanics perspective, both designs show that smaller diameters (5 mm) lead to higher pressure drops and increased fluid velocities due to greater flow resistance, which can raise energy consumption and risk flow instability. Conversely, larger diameters (10 mm and 15 mm) promote lower pressure drops and more uniform velocity profiles, enhancing cooling efficiency and stability. The design comparison reveals that Design 2 may offer better flow management, with more consistent pressure and velocity distributions, suggesting an optimized internal channel geometry that reduces flow disruptions. These insights underscore the importance of optimizing geometric design and diameter selection to improve GPU cooling performance.

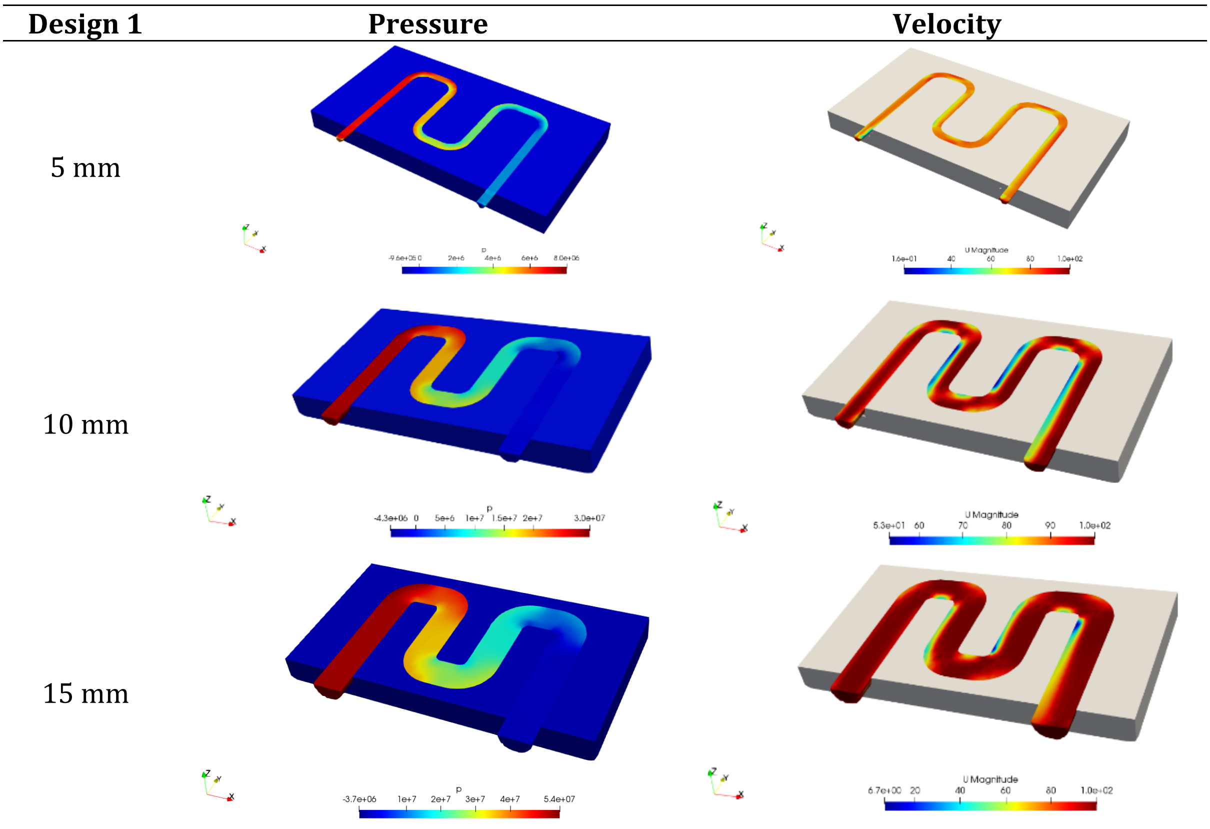


Figure 7: Pressure and velocity contour for different inlet-outlet diameters using Design 1.

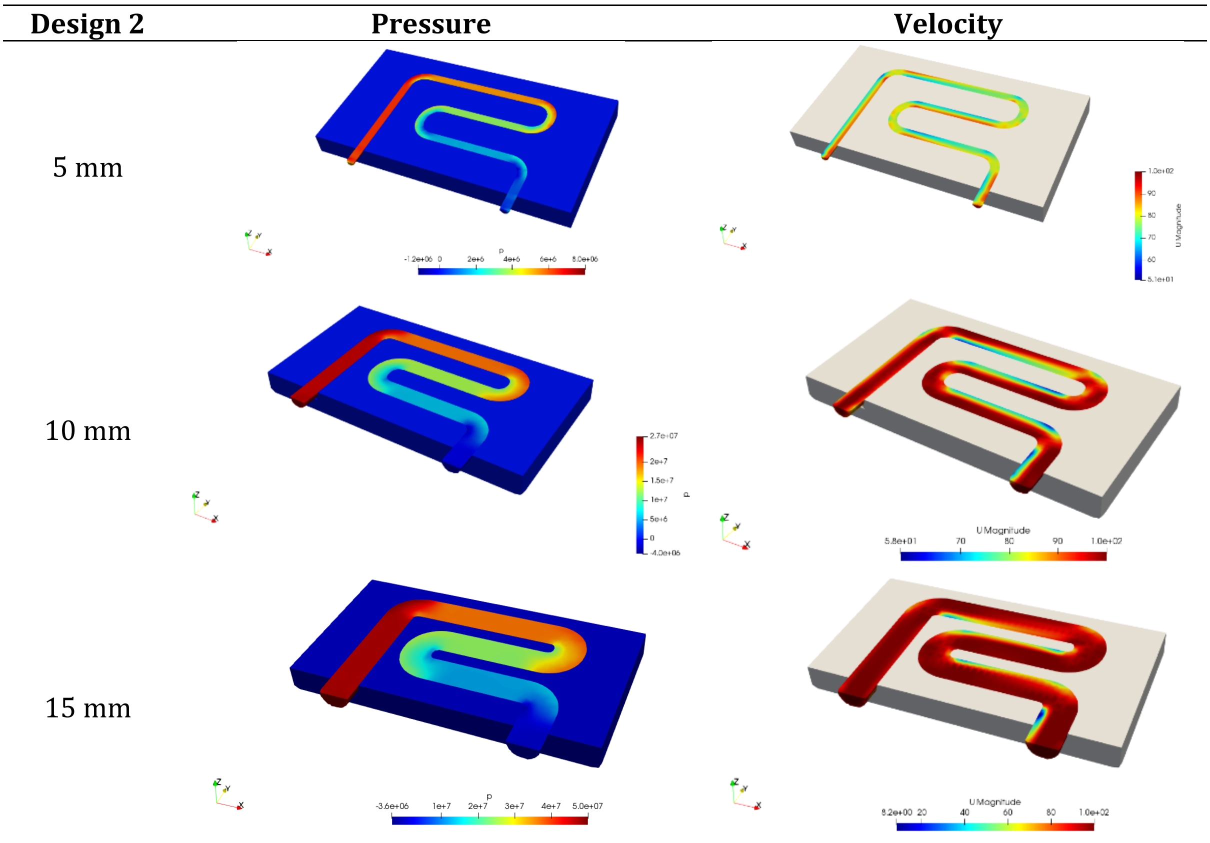


Figure 8: Pressure and velocity contour for different inlet-outlet diameters using Design 2.

4. CONCLUSION

This current study emphasized the impact of the inlet and outlet diameters and pipe design on the performance of cooling channels in GPU blocks. The findings reveal that increasing the diameter leads to higher maximum pressures but helps maintain lower temperatures, with diameters of 10 mm and 15 mm proving particularly effective in optimizing thermal performance without significantly increasing pressure. Additionally, the geometric design of the pipes plays a crucial role in shaping the pressure distribution within the cooling channels, with Design 1 exhibiting higher pressures than Design 2. However, both designs achieved similar temperature profiles. These insights provide valuable guidance for optimizing cooling channel designs to enhance thermal management and overall GPU performance. Future research should explore a broader range of geometric configurations and materials to improve cooling efficiency. Overall, the study highlights the importance of carefully considering geometric parameters and design configurations in developing effective cooling systems for high-performance electronic devices.

ACKNOWLEDGEMENTS

The author thanks the Faculty of Mechanical Engineering & Technology for their invaluable support and resources throughout this research. Special thanks to the coordinator of the CFD course for his guidance and mentorship.

REFERENCES

[1] Kong, J., Chung, S. W., & Skadron, K. Recent thermal management techniques for microprocessors. ACM Computing Surveys (CSUR), vol 44, issue 3 (2012) pp. 1-42.
[2] Owens, J. D., Houston, M., Luebke, D., Green, S., Stone, J. E., & Phillips, J. C. GPU computing. Proceedings of the IEEE, vol 96, issue 5 (2008) pp. 879-899.
[3] Carvalho, P., Clua, E., Paes, A., Bentes, C., Lopes, B., & Drummond, L. M. D. A. Using machine learning techniques to analyze the performance of concurrent kernel execution on GPUs. Future Generation Computer Systems, vol 113, (2020) pp. 528-540.
[4] Hawick, K. A., Leist, A., & Playne, D. P. Mixing multi-core CPUs and GPUs for scientific simulation software. Res. Lett. Inf. Math. Sci, vol 14, (2009) pp.25-77.
[5] Siricharoenpanich, A., Wiriyasart, S., & Naphon, P. Study on the thermal dissipation performance of GPU cooling system with nanofluid as coolant. Case Studies in Thermal Engineering, vol 25, (2021) p. 100904.
[6] Feng, S., Kamat, A. M., & Pei, Y. Design and fabrication of conformal cooling channels in molds: Review and progress updates. International Journal of Heat and Mass Transfer, vol 171, (2021) p. 121082.
[7] Górecki, K., & Posobkiewicz, K. Cooling systems of power semiconductor devices—A Review. Energies, vol 15, issue 13 (2022) p. 4566.
[8] Sahan, R. A., Mohammed, R. K., Xia, A., & Pang, Y. F. Advanced Liquid Cooling Technology Evaluation for High Power CPUs and GPUs. In International Electronic Packaging Technical Conference and Exhibition, vol 44625, (2011) pp. 311-318).
[9] Mancin, S., Zilio, C., Righetti, G., & Rossetto, L. Mini Vapor Cycle System for high density electronic cooling applications. International journal of refrigeration, vol 36, issue 4 (2013) pp. 1191-1202.
[10] Liu, C., & Yu, H. Experimental investigations on heat transfer characteristics of direct contact liquid cooling for CPU. Buildings, vol 12, issue 7 (2022) p. 913.
[11] Alam, T., & Kim, M. H. A comprehensive review on single phase heat transfer enhancement techniques in heat exchanger applications. Renewable and Sustainable Energy Reviews, vol 81, (2018) pp. 813-839.

[12] Xia, G., Zhuang, D., Ding, G., Lu, J., Han, W., & Qi, H. A distributed parameter model for multi-row separated heat pipe with micro-channel heat exchangers. Applied Thermal Engineering, vol 182, (2021) p. 116113.
[13] Kong, R., Zhang, H., Tang, M., Zou, H., Tian, C., & Ding, T. Enhancing data center cooling efficiency and ability: A comprehensive review of direct liquid cooling technologies. Energy, vol 308, (2024) p. 132846.
[14] Vaghetto, R. Experimental Study of the Thermal-Hydraulic Phenomena in the Reactor Cavity Cooling System and Analysis of the Effects of Graphite Dispersion (Doctoral dissertation, Texas A & M University) (2012).
[15] Zamri, M. Z. Multi-Objective Optimization of Absorber Piston Rod Design Parameter using Response Surface Methodology. Advanced and Sustainable Technologies (ASET), vol 1, issue 2 (2022) pp. 20-29
[16] Paramasivan, S. Analysis of Tesla CyberTruck Speed on the Velocity and Pressure Distribution Using SimFlow Software. Advanced and Sustainable Technologies (ASET), vol 3, issue 1 (2024) pp. 71-79.

Conflict of interest statement: The author declares no conflict of interest.

Author contributions statement: Conceptualization; Methodology; Software; Formal Analysis; Investigation; Writing & Editing, Mohd Izzat Aiman Adi.