Euijun Chung

Email: euijun@gatech.edu
LinkedIn: euijun-chung
Github: github.com/ejchung0406

Introduction

Euijun Chung is a 2nd-year CS PhD student at Georgia Tech, focusing on **GPU architecture, system and simulations** for machine learning. His research interests lie on cycle-level GPU simulation and performance modeling, GPU memory safety, GPU + SSD architecture co-design, and Multi-GPU systems.

Education

Georgia Institute of Technology — Atlanta, GA, USA
Ph.D. in Computer Science — Jan. 2024 – Present
Advisor: Hyesoon Kim

Korea Advanced Institute of Science and Technology (KAIST) — Daejeon, Korea
B.S. Major in Electrical Engineering, Minor in Mathematical Sciences — Feb. 2018 – Feb. 2024
GPA: 4.05/4.30 (*Summa Cum Laude*)

Experience

AMD RAD (Research and Advanced Development) — Santa Clara, CA, USA
Research Associate at Advanced Memory Team — May. 2025 – Aug. 2025

Proposed a near-zero-cost algorithmic improvement for future chiplet-based GPU architectures, significantly reducing remote HBM traffic during end-to-end LLM inference and training. Modified PyTorch API, `ATen`'s tensor operators, and memory allocation in the `c10` library for implementation and evaluation.

Research Projects

Georgia Tech HPArch Lab — Atlanta, GA, USA
Graduate Research Assistant (Advisor: Hyesoon Kim) — Jan. 2023 – Present

- **GPU Simulation & Performance Modeling**: Proposed statistical kernel-sampling methodology for accelerating cycle-level GPU simulations on ML/LLM workloads [1], [2], achieving scalable microarchitecture simulations with $< 1\%$ sampling error. First-author publication at MICRO 2025.
- **GPU Memory Safety**: Designed and evaluated LMI [3], a novel fine-grained hardware bounds-checking solution for GPUs with under 1% performance overhead in HPC and ML benchmarks. Second-author publication at HPCA 2025.
- **GPU-SSD architecture co-design**: Developed a GPU-SSD integrated simulator [link] for evaluating adaptive GPU block scheduling and address mapping policies on co-designed architectures [4].
- **Open-source GPU Simulator**: Added SASS-level tracer for Macsim (cycle-level GPU simulator) to run workloads on latest CUDA versions [link], utilized NVBit and CUDA Runtime APIs.
- **Vortex 2.0**: Contributed to the cache subsystem of the next-generation open-source RISC-V GPU.

KAIST INALab — Daejeon, Korea
Undergraduate Research Assistant (Advisor: Dongsu Han) — Jul. 2021 – Aug. 2022

- **Scene-clustered Superresolution Training**: Developed a content-aware video delivery and upscaling methodology using scene-level classification and fine-tuning, achieving + 5.8 dB PSNR improvement over prior methods with the same resource usage, using ONNX and TensorRT.

Publications

[1] **Euijun Chung**, Seonjin Na, Sung Ha Kang, and Hyesoon Kim, "Swift and trustworthy large-scale GPU simulation with fine-grained error modeling and hierarchical clustering", in *2025 58th IEEE/ACM International Symposium on Microarchitecture* (*MICRO*), IEEE, 2025.

[2] **Euijun Chung**, Seonjin Na, and Hyesoon Kim, "Allegro: GPU simulation acceleration for machine learning workloads", in *Machine Learning for Computer Architecture and Systems 2024* (*MLArchSys, co-located with ISCA*), 2024.

[3] Jaewon Lee, **Euijun Chung**, Saurabh Singh, Seonjin Na, Yonghae Kim, Jaekyu Lee, and Hyesoon Kim, "Let-me-in: (still) employing in-pointer bounds metadata for fine-grained GPU memory safety", in *2025 IEEE International Symposium on High-Performance Computer Architecture* (*HPCA*), 2025.

[4] Xueyang Liu, Seonjin Na, **Euijun Chung**, Jiashen Cao, Jing Yang, and Hyesoon Kim, "Contention-aware GPU thread block scheduler for efficient GPU-SSD", *IEEE Computer Architecture Letters*, 2025.

[5] Huanzhi Pu, Rishabh Ravi, Shinnung Jeong, Udit Subramanya, **Euijun Chung**, Jisheng Zhao, Chihyo Ahn, and Hyesoon Kim, "Hardware vs. software implementation of warp-level features in vortex RISC-V GPU", in *OSSMPIC - Open Source Solutions for Massively Parallel Integrated Circuits* (*Co-located with DATE*), 2025.

[6] Myoung Jae Lee and **Euijun Chung**, "Experimental analysis on the 0 dimensional plasma model in an inductively coupled plasma (ICP)", in *2016 New Physics: Sae Mulli*, 66:1183–1189.

Teaching

- **Teaching Assistant for CS 8803 - GPU Hardware & Software** Summer 2024, Fall 2025
  Designed and developed a **lightweight GPU architecture simulator** for two programming assignments, demonstrating microarchitectural concepts such as warp scheduling and tensor core utilization.

  Developed a course project implementing **prefill and decode FlashAttention-2 kernels in CUDA**, and integrated the custom kernels with Hugging Face LLMs for performance evaluation and analysis.
- **Tutor in Freshman Tutoring Program** Fall 2021, Spring 2022, Fall 2022, Fall 2023
  Tutored Calculus II (Vector Calculus) to KAIST freshmen through weekly lectures and office hours.

Skills

- **Programming**: C/C++, CUDA, Python, C#
- **Machine learning frameworks**: cuBLAS, Pytorch (`ATen` and `c10`), vLLM
- **Tools and simulators**: NVBit, Macsim, ASTRA-Sim, gem5, MATLAB, ARM Mbed, Unity, LabWindows/CVI

Scholarships and Honors

- **Gem5 Bootcamp** Attendee with Full Travel Grant at UC Davis. Jul.–Aug. 2024
- **5th uArch Workshop at ISCA 2023** Full Grant Recipient Jun. 2023
- **KOSAF** (Korea Student Aid Foundation) National Science & Technology Scholarship 2022 – 2024
  Awarded scholarship for being an outstanding undergraduate student in engineering.
- **Dean's List** for KAIST EE: Awarded to students with top 3% academic performance. Fall 2022

Extracurricular Activities & Experiences

- **2nd place in Susquehanna Brainteaser Battle** at Georgia Tech Fall 2024
- **AI Competition for Agricultural Commodity Price Prediction** Fall 2022
  Participated in Nongnet AI price prediction competition utilizing a comprehensive 10-year agricultural transaction database. Achieved a top 13% ranking out of 69 participating teams.
- **Republic of Korea Air Force (ROKAF)** Aug. 2019 – May 2021