Jinpyo Kim

University of California, San Diego | jik066@ucsd.edu | +1-619-357-5244 | https://jinpyo-cs.github.io/

Education

University of California, San Diego San Diego, CA, US
Ph.D. Student in Computer Science, (Expected 2027)
Advisor: Professor Jishen Zhao

Sogang University Seoul, Korea
Master of Science in Computer Science and Engineering, 2016
Advisor: Professor Juho Kim

Sogang University Seoul, Korea
Bachelor of Engineering in Computer Science and Engineering, 2014

Research Interest

System Architecture for AI/ML Workloads

- System-level characterization and optimization of heterogeneous compute platforms
- Memory access and data placement optimization in memory-intensive workloads
- Energy-efficient computing leveraging emerging memory and near-data processing

Research Experience

Graduate Researcher, STABLE Lab UC San Diego

1. Heimdall: Cache-Coherent Heterogeneous Systems Benchmark Suite

- Developed and maintained **LLM Bench**, an open-source benchmarking suite for evaluating inference performance across **vLLM, llama.cpp and PyTorch based frameworks**.
- Designed experiments for **throughput, latency, and CPU memory offloading** focusing on heterogeneous memory environments (CPU DRAM, GPU HBM, and CXL memory).
- Conducted profiling studies using **perf, Nsight Systems**, and **AMD uProf** to characterize system-level bottlenecks

2. AlphaFold3 Workload Characterization

- Developed **AFSysBench**, an open-source benchmark suite for AlphaFold3 system-level performance analysis
- Characterized **MSA and inference bottlenecks** across CPU/GPU architectures and heterogeneous memory using system-level CPU/GPU profiling and I/O monitoring tools

Graduate Research Assistant, CAD & VLSI Research Lab Sogang University
Master's Thesis : Efficient Flash Cache Management in Online Transaction Processing Server
Project: Process variation and BTI-aware Static Timing Analysis for Samsung, 2014 – 2015.

Publications

1. Jinpyo Kim, Mingi Kwon, and Jishen Zhao.
"AlphaFold3 Workload Characterization: A Comprehensive Analysis of Bottlenecks and Performance Scaling" (AFSysBench benchmark suite)
IEEE International Symposium on Workload Characterization (IISWC), 2025. [Accepted].

2. Zixuan Wang, Suyash Mahar, Luyi Li, Jangseon Park, **Jinpyo Kim**, et al.
"The Hitchhiker's Guide to Programming and Optimizing Cache Coherent Heterogeneous Systems: CXL, NVLink-C2C, and AMD Infinity Fabric" (HEIMDALL benchmark suite)

Work Experience

Senior Embedded Software Engineer — **SK Hynix Inc., Korea**
Feb 2016 – Jul 2023

- Led team efforts on **power-off recovery feature development** during PE9110 project.
- Contributed to firmware development for enterprise SSDs, from **SATA SSD** to **PCIe Gen4 SSD**.
- Developed and optimized **Flash Translation Layer (FTL)** to improve performance and reliability.
- Performed **SSD performance benchmarking and analysis** to identify firmware-level bottlenecks.

Patents

U.S. 10,741,254: "Memory system and operating method thereof.", 2020
U.S. 10,860,227: "Memory controller, memory system having the same, and method of operating the same.", 2020
U.S. 11,269,528: "Data storage device, operation method thereof and controller therefor.", 2022
U.S. 11,307,942: "Memory system, memory controller and method for operating memory controller.", 2022
U.S. 11,404,137: "Memory system and operating method of memory system", 2022
U.S. 11,422,747: "Memory system and method for operating memory controller included therein.", 2022
U.S. 11,556,252: "Storage device and method of operating the same.", 2023
U.S. 11,593,006: "Data storage apparatus and method for managing valid data based on bitmap table.", 2023
U.S. 11,599,275: "Memory controller for controlling power loss recovery and method of operating the same.", 2023
U.S. 11,704,050: "Memory system for determining a memory area in which a journal is stored according to a number of free memory blocks.", 2023
U.S. 12,216,914: "Apparatus and method for power-loss data protection in a system.", 2025
U.S. 12,287,979: "Data storage apparatus and operating method thereof.", 2025

Honors & Awards

SK Hynix Ph.D. Fellowship Program in 2021
SK Hynix Industrial Scholarship in 2012

Technical Skills

LLM Benchmarking & Inference Frameworks: vLLM, SGLang, llama.cpp
Profiling & Benchmarking Tools: Linux perf, AMD uProf, NVIDIA Nsight Systems
Programming Languages: C/C++, Python, Java, Shell scripting (UNIX/Linux)
Platform: Linux, UNIX server
Debugger: TRACE32 (hardware debugger for embedded systems such as ARM architecture)

References

Prof. Jishen Zhao (advisor)
Associate Professor at University of California, San Diego, USA
Email: jzhao@ucsd.edu
Phone: 858-822-2449

Dr. Dongyoung Seo (co-worker)
Principal Embedded Software Engineer at Solidigm, Rancho Cordova, USA
Email: dongyoung.seo@gmail.com
Phone: 279-246-7172