ZONGZE LI

5801 S Ellis Ave, Chicago, IL 60637

WebPage Y zongzel@uchicago.edu

EDUCATION

University of Chicago

Chicago, IL, USA

Sep. 2024 - Present

Research Interests: Sys4AI, Heterogeneous Computing, HPC, AI Systems

ShanghaiTech University

Shanghai, China

B.Eng. in Computer Science, Advisor: Prof. Shu Yin & Rui Fan

Ph.D. in Computer Science, Advisor: Prof. Ce Zhang & Yanjing Li

Sep. 2020 - July 2024

Main Courses: Computer Architecture, Operating System, Parallel Computing, Database, NLP

RESEARCH EXPERIENCE

Gulliver - A Finer Grained Log-Structured PMEM File System

Research Intern, instructed by Shu Yin

Mar. 2023 – Nov. 2023 *Shanghai, China*

- Research kernel compilation, using suitable compilation options and auxiliary tools to enable the successful execution of the project prototype.
- Design an IOR testing plan and collaborate with team members to compare and assess the parallel access capabilities of heterogeneous file systems, such as Ext4, XFS, NOVA.

PowerInfer Mar. 2024 – June. 2024

Research Intern, instructed by Rui Fan

Shanghai, China

- Collaborated with SJTU-IPADS Lab to successfully migrate their PowerInfer project to AMD device platforms.
- Conducted comprehensive performance analysis on AMD architecture, identifying hotspots in memcopy between CPU and GPU, and implemented optimizations resulting in a xx% improvement in inference performance.

WORK EXPERIENCE

Architecture Design Intern

AMD Xilinx Department

Apr. 2023 - July 2024

Shanghai, China

- Designed and implemented full configuration environment based on the MI210 graphics card, including remote interface integration, and provided procedural documentation for internal remote access resources.
- Contributed to maintaining and developing the HACC-NUS supercomputing cluster, offering test cases for cluster testing and successfully training and inferring large models. Provided user-oriented improvement measures.
- Provided materials and guidance for the AMD 2024 Winter Camp and the 2024 Summer School courses. Assisted in deploying hardware for the SARI research group and supported the reform of Parallel Computing course at ShanghaiTech, offering technical and equipment support for course projects.
- Participated in the development of an open-source project for visualizing model training based on Unity, successfully bridging the interaction between simulation software and local hardware inference through network debugging. Deployed models for training completion.

Club Advisor Sep. 2022 – Dec. 2023

ShanghaiTech GeekPie HPC Club

Shanghai, China

• Develop GeekPie HPC team to participate in top tier student cluster competitions co-hosted with HPC conferences including ASC23, ISC23 and SC23, where students build a tiny cluster under a 3000W power constraint and accelerate a set of benchmarks and applications on it.

SERVICES

Operating System Course Aug. 2023 - Feb. 2024 Teaching Assistant

Computer Architecture Course

Teaching Assistant

Student Cluster Competition 2023

Advisor of the University Team

Shanghai, China

Mar. 2023 - July 2023 Shanghai, China

July 2023 - Nov. 2023

Denver, CO, USA

AWARDS

• ISC23, The Third Place - 2023

• Field Research, Outstanding Individual Award - 2022

SKILLS

Programming Languages: Python, C/C++, Matlab, CUDA, HIP, SQL, HTML(Not limited to any specific language)

System: Specialist in Performance Analysis, familiar with LLVM, MLIR, Gdb, Qemu, Docker

AI: Familiar with general knowledge of machine & deep learning(PyTorch), interested of Sys for ML/LLM