Peizhi Liu

2233 Tech Drive
Mudd Hall, Room 3302
Evanston, IL 60208

peizhi.liu@u.northwestern.edu
+1 (248)-824-0359

RESEARCH INTERESTS

I am broadly interested in parallel/distributed computing systems. My current research is focused on enhancing the communication, performance, and programmability of these systems in light of heterogeneous hardware.

EDUCATION

Northwestern University, Evanston, IL

Ph.D. Student, Computer Science — Mar 2024 – *Present*
M.S. Student, Computer Science — Sep 2022 – Jun 2023
B.S. Student, Computer Science (Honors) — Sep 2019 – Jun 2023

RESEARCH EXPERIENCE

Graduate Student, Northwestern University — Mar 2024 – *Present*

Parameterized Collective Communication Library (PCCL) for GPU Clusters (In progress)
Advisor: Professor Peter Dinda

- Developing parameterized collective algorithms for GPU-GPU communication that are flexible and performant on modern HPC topologies. Received $25,000 in AWS resources and a total of 60,000 GPU Hours of compute time across three HPC clusters through the National Artificial Intelligence Research Resource (NAIRR) Pilot.
- Research Question – Can parameterized algorithms offer the flexibility and performance needed to achieve comparable performance as current state-of-the-art synthesized algorithms, but without the cost and scalability issues of synthesis?

Source to Specification Translation of Scientific Codebases with Large Language Models (In progress)
Mentors: Professor Andrew Crotty, Professor Christos Dimoulas, Professor Peter Dinda

- Proposed an LLM-based framework for translating low-level C and OpenMP code into Python. Collected and analyzed performance traces for the Flash-X to determine, as a proxy, relevant scientific functions for translation.
- Research Question – Can we use an LLM to transform a highly optimized hardware-specific codebase into a high-level, easy-to-understand specification?

Undergraduate Student, Northwestern University — Sep 2019 – Jun 2023

Generalized Collective Algorithms for the Exascale Era **[1]**
Mentors: Professor Peter Dinda and Dr. Michael Wilkins

- Developed parameterized versions of MPI All-Gather, All-Reduce, and Bcast collective algorithms to enable machine learning-based autotuning performance enhancements to account for dynamic non-programmatic factors of an HPC cluster.
- Achieved up to 4.5x speedup compared to original MPICH and native Cray MPICH implementations on Crusher, a testbed for the Frontier exascale supercomputer.

PDNS: A Fully Privacy-Preserving DNS **[2][Under Review]**
Mentors: Professor Aleksandar Kuzmanovic and Dr. Yunming Xiao

- Designed and implemented a privacy-preserving DNS recursive resolver by leveraging single-server private information retrieval (PIR). This allows users to retrieve DNS records anonymously without the recursive resolver knowing.
- Also developed and ran benchmarks for client and server-side query performance, including user webpage performance tests. Determined that PDNS offers significant privacy guarantees with acceptable performance.

Variation in Derived Extrasolar Planet Spectrum
Mentors: Professor Melville Ulmer and Dr. Cobi Rabinowitz

- Developed an end-to-end spectrum extraction pipeline using CHARIS Data Reduction pipeline and pyKLIP to explore variability in exoplanet spectra for HD 1160 b and Kappa And b. These techniques have the potential to be used for detecting solar flare events and the exoplanets' rotational velocity.
- Verified previously published Kappa And b spectrum and discovered substantial variability in the spectrum either due to physical, instrumental, or pipeline processing. Presented findings to other ISG students.

High School Student, Michigan State University Sep 2018 – Jun 2019

Utilizing Google Earth Engine to Retrieve the Devon Ice Cap's Equilibrium Line Altitude
Mentor: Professor Grant Gunn

- Computed and monitored the Equilibrium Line Altitude (ELA) of the Devon Ice Cap between 2015 and 2018 with remote sensing satellite imagery. Developed a novel technique that uses machine learning (K-Means) on Sentinel 1 and Landsat 8 imagery.
- The ELA of the Devon Ice Cap was found to have reached a maximum elevation of 1642 meters and the maximum percentage melt of the glacier was 92.57% in 2017. Won Grand Award at the Science and Engineering Fair of Metro Detroit (SEFMD) and presented results at the 2019 Intel International Science and Engineering Fair (ISEF). We were also interviewed by WKAR Public Media from Michigan State University for this work.

INDUSTRY EXPERIENCE

Software Development Engineer Intern, Amazon Jun 2022 – Sep 2022

Extending Log Filters and Cache for Recommendations Service Testing
Hosts: Christian Tan, Zimeng Zhang

- Worked in the Personalization organization to implement extensible test log filtering with DynamoDB to easily onboard new recommendations services without code changes. Addressed problems with insufficient service test data by designing a test data caching scheme with S3.
- Used build and pipeline systems to deploy changes onto production servers and to monitor service state for potential issues and bugs.

Junior Intern (Software Engineering), Robert Bosch LLC Jun 2021 – Aug 2021

Integrating Protocol Buffers in MCU Communications
Hosts: Syed Ahson, Barrington Brown

- Worked in the Perfectly Keyless team to integrate Protocol Buffers to be used for MCU cryptographic handshakes.
- Analyzed stack and MCU RAM usage of the protocol library to inform the customer of potential decision tradeoffs.

SELECTED PROJECTS

Accelerated Journal Bearing Hydrodynamic Lubrication Simulation on CUDA
Mentors: Professor Nikos Hardavellas and Professor Peter Dinda

- Developed a GPU-accelerated solution for modeling the tribology of journal bearings under hydrodynamic and non-Newtonian lubrication. GPU-specific optimizations include tiling, reduction, memory coalescing, and bank conflict avoidance.
- Saw over 9x speed-up in total simulation time over a comparable sequential CPU simulation for medium-sized meshes.

Benchmarking Overhead of Running Neural Networks in OP-TEE
Mentor: Professor Peter Dinda

- Investigated memory and performance limitations of deep learning in the Open Portable Trusted Execution Environment (OP-TEE), an implementation of Arm TrustZone running on an embedded device (Raspberry Pi 3b). Specifically focused on a layer-based partitioning technique to reduce the model's memory footprint.
- Discovered layer-based partitioning increases the maximum memory usage while only marginally decreasing average memory usage. Further experiments comparing the use of Secure Storage versus encrypted Shared Memory revealed significant performance improvements of the latter over the former.

PUBLICATIONS / DEMOS

[1] Wilkins, Michael, Hanming Wang, Peizhi Liu, Bangyen Pham, Yanfei Guo, Rajeev Thakur, Peter Dinda, and Nikos Hardavellas. "Generalized Collective Algorithms for the Exascale Era." In *2023 IEEE International Conference on Cluster Computing (CLUSTER)*, pp. 60-71. IEEE, 2023.

[2] Xiao, Yunming, Chenkai Weng, Ruijie Yu, Peizhi Liu, Matteo Varvello, and Aleksandar Kuzmanovic. "PDNS: A Fully Privacy-Preserving DNS." In *Proceedings of the ACM SIGCOMM 2023 Conference*, pp. 1182-1184. 2023.

TEACHING AND MENTORING EXPERIENCE

Teaching

Peer Mentor (Undergraduate TA), Northwestern University

Winter 2023	CS 336 – Design & Analysis of Algorithms	with Instructor Donald M. Stull
Winter 2022	CS 343 – Operating Systems	with Prof. Peter Dinda
Fall 2021	CS 336 – Design & Analysis of Algorithms	with Prof. Konstantin Makarychev

Teacher for AP Statistics Course

Taught weekly AP Statistics classes to a class of 11 high school students over the summer in 2021. Prepared slides, class material, and graded weekly homework assignments.

Teacher for Python Programming Course

Taught elementary and middle school students Python Programming as part of my church's summer camp program in 2023. For more advanced students, I also taught basic web development with HTML, CSS, and Javascript.

Mentoring

- Mentoring Sean Rhee (undergraduate student at Northwestern University) on the *Generalized Collective Algorithms for High-Performance GPU Clusters* project.

AWARDS AND HONORS

NAIRR Pilot Resource Award 2024 – 2025
- $25K in AWS resources, 60,000 GPU Hours across three HPCs

Dean's List, all quarters 2019 – 2023

National Merit Scholarship 2019 – 2023

Illinois Space Grant 2020
- Summer research experience with Prof. Melville Ulmer

Grand Award Winner at the Science and Engineering Fair of Metro Detroit 2019
- Participated in Intel ISEF 2019