Haibin Lai

Shenzhen | laihb2022@mail.sustech.edu.cn | | blog: www.haibinlaiblog.top | github.com/HaibinLai

About Me

I am a fourth-year undergraduate student at **Southern University of Science and Technology** (SUSTech), majoring in Computer Science. I am currently engaged in research in Parallel Computing, Machine Learning Systems, and GPU Computing at the SUSTech HPC Lab supervised by **Prof. Zhuozhao Li**.

Education

Southern University of Science and Technology, BS in Computer Science   Sept 2022 – Present
- GPA: 3.70/4.0; Member of Turing Class (designated for elite CS students at SUSTech)

Publication

1. **Haibin Lai**, Sicheng Zhou, Site Fan, Zhuozhao Li. *ParaCOSM: A Parallel Framework for Continuous Subgraph Matching*. In Proceedings of the 54th International Conference on Parallel Processing (ICPP '25), San Diego, CA, USA, 2025.
- Proposed **ParaCOSM**, a parallel framework enabling efficient continuous subgraph matching (CSM) on CPUs.
- Achieved $1.2\times$–$30.2\times$ speedups across diverse datasets, with up to two orders of magnitude faster execution.
- Released as open source at **GitHub**. **Paper Link**.

2. Wenyi Wang, Maxime Gonthier, **Haibin Lai**, Poornima Nookala, Haochen Pan, Ian Foster, Ioan Raicu, Kyle Chard. *Exploring Fine-Grained Parallelism in Dataflow Runtime Systems on Multi-Socket Many-Core Systems*. Supercomputing (SC '25) Poster Track, 2025.
- Profiling **taskflow** on Multi-Socket Many-Core systems. Taskflow is a a General-purpose Task-parallel Programming System. **Poster Link**. **Paper Link**.

Research & Engineering Projects

Utilizing Task-based Parallelism for LLM Inference on Many-Core Systems   March 2025 - Present
with **Prof. Kyle Chard** and **Prof. Ian Foster** *(Globus Lab, University of Chicago), and*
Prof. Ioan Raicu *(Illinois Institute of Technology)*
- Developed a suite of task-parallel LLM operators from `llama.cpp`, achieving over 80% throughput improvement on many-core CPU systems. Preparing a full research paper on task-parallel operator design and evaluation.
- **Target venue: HPDC'26**; preprint to appear on arXiv, Jan 2026;

Virtio-GPU Driver for Asterinas OS   Course Project, **Prof. Yinqian Zhang**   Nov 2024 – Jan 2025
- Developed a **virtio-gpu driver** in the Asterinas operating system, enabling user-space programs to access GPU functions via dedicated syscalls.
- Designed the driver integrated with QEMU virtio for para-virtualized GPU support.
- Technologies: Rust, QEMU, Para-virtualization.

Geochem Pi Development   Aug 2024 - March 2025
- **Contributor** of Geochemistry Pi — An open-source Python ML framework for data-driven geochemistry discovery on tabular data. Contributed pull requests enhancing functionality for geochemistry datasets.

Experience

Research Intern, UChicago Globus Lab   March 2025 - Sep 2025
- Supervised by *Prof. Kyle Chard*
- Boosting LLM inference on many cores CPU. Modifying `llama.cpp` inference engine.

Intern, Beijing Sunway World Technology Co., Ltd.   Aug 2023 - Sep 2023
- Studied the Laboratory Information Management System (LIMS) solutions, gaining insights into software workflows and architecture alongside the technical department.

Service

Student Assistant, SUSTech Center for Computational Science and Engineering Feb 2024 – Aug 2024
- Regularly participate in the operation and maintenance of SUSTech Qiming and Taiyi (Rank 127 on Top500, year 2018) Supercomputer. Conduct HPL/HPCG benchmark testing on the supercomputing cluster.

Deputy Class President, Turing Class, Computer Science Department, SUSTech Sep 2022 – Present
- Collaborated with class advisor on daily affairs and organized impactful events, including dialogues with Turing Award laureate Prof. Joseph Sifakis and thematic report sessions.

Selected Awards

8th APAC HPC-AI Student Competition APAC HPC-AI Nov 2025
- *Second Prize*
- Leader of Team SUSTech. Responsible for parallel optimization on NWChem HPC software and SGLang Deepseek inference

7th APAC HPC-AI Student Competition APAC HPC-AI Dec 2024
- *Third Prize*
- Leader of HPC Team. Responsible for parallel optimization on Hoomd-blue HPC software, using HPC-X communication library, and a better scheduling strategy for domain decompostion.

ASC Student Supercomputer Challenge ASC24 Apr 2024
- *First Prize; Group Competition Award*
- Responsible for LINPACK benchmark and parallel optimization of materials science calculations in the team, as well as Linux and network monitoring.

2025 Outstanding Student Award	Oct 2025
2024 Outstanding Student Award	Nov 2024
2023 Outstanding Student Award	Nov 2023
2022 President's Special Scholarship	Sep 2022

English Skills & Communication

TOEFL Nov 2025
- Score: 104/120; speaking 22, listening 30, reading 30, writing 22

CET6 Jun 2024
- Score: 649/710

Georgia Institute of Technology ASP Summer Program Jul 2023
- Participated in summer courses at Georgia Tech, enhancing communication and American experience.