Shixun Wu

- Email: swu264@ucr.edu
- Phone: +1 909-836-7100
- Contact: Winston Chung Hall 459, Riverside, CA, 92507
- Website: https://www.shixun404.com
- Google Scholar

EDUCATION

University of California, Riverside ***Sep. 2022 - Present***
Ph.D Candidate in Computer Science, advised by Dr. Zizhong Chen

Columbia University ***Sep. 2020 - May 2022***
M.S. in Electrical Engineering

Peking University ***Sep. 2016 - Jul. 2020***
B.S. in Computer Science
B.S. in Economics(Double Major)

RESEARCH EXPERIENCE

1. **USC ISI / Argonne National Laboratory** Los Angles, CA / Lemont, IL
   Scientific Workflow Applications on Resilient Metasystem Jan. 2024 - Present
   Mentors: Dr. Franck Cappello, Dr. Sheng Di, Dr. Krishnan Raghavan (ANL); Dr. Ewa Deelman (USC ISI).
   - Designed a Q-learning + GNN-based topology protocol (DGRO) that reduces network diameter by optimizing virtual rings over heterogeneous, failure-prone systems.
   - Implemented a single-hop gossip-based failure detector, resilient to network jitter and churn, enabling decentralized membership monitoring across 20+ globally distributed sites.
   - DGRO on FABRIC testbed spanning Japan, Europe, Hawaii, and 15+ U.S. locations, demonstrating fast convergence and robustness at international scale.
2. **UCR / Lawrence Berkeley National Laboratory** Riverside, CA
   Data-driven Exascale Control of Optically Driven Excitations in Chemical and Material Systems Sep. 2022 - Present
   Mentor: Dr. Zizhong Chen
   - Designed and implemented in-kernel ABFT GEMM using tensor cores, achieving higher performance than **cuBLAS** while ensuring fault detection and correction under soft errors.
   - Developed a fully GPU-resident ABFT FFT pipeline, outperforming **cuFFT**, and enabling error-resilient spectral analysis in scientific simulations.
   - Proposed the first ABFT-enabled K-means clustering framework on GPUs, exceeding **cuML** performance with integrated resilience support.
   - Innovated lightweight, low-overhead in-kernel fault tolerance mechanisms across linear algebra and ML workloads, demonstrating resilience-performance co-design in exascale systems.
3. **Nvidia** Santa Clara, CA
   Compiler Optimization for OpenMP Target Offload on Heterogeneous GPU Architectures Jun. 2024 - Sep. 2024
   Mentor: Dr. David Appelhans
   - Investigated performance bottlenecks of OpenMP target offload in SPEChpc 2021 on GH200/H200 GPUs.
   - Developed compiler/runtime optimizations achieving up to 10× speedup without source code changes.
   - Analyzed OpenMP vs. OpenACC performance and contributed optimized version to SPEChpc 1.1.9.
   - Work adopted by RWTH Aachen University, demonstrating both research impact and practical relevance.
4. **Columbia University / AI4Finance Foundation**, New York, NY
   ElegantRL: Massively Parallel Deep Reinforcement Learning Library *Aug. 2021 - Jul. 2022*
   Mentors: Dr. Xiaoyang Liu, Dr. Xiaodong Wang
   - Develop multi-agent reinforcement learning algorithms in ElegantRL, a RL library with 4k starts on GitHub.
   - Co-leader of ElegantRL_Solver, a high-performance solver, outperforming Gurobi for dense cases in MaxCut Problem.

PEER REVIEWED PUBLICATIONS

- My research interest lies in **High-Performance Computing**, **Reinforcement Learning**, **Fault Tolerance & Resilience Deep Learning Systems**, **Parallel, Distributed & Heterogeneous Systems**, **Lossy Compression & Data Management**.
- 10 peer-reviewed papers (4 papers as the 1st author), 4 papers in submission, 2 workshop, 1 poster.
- 1st Author publications: PPoPP, SC, ICS, Cluster. Co-first authors are marked with *.

CONFERENCE PAPERS

C1 **[PPoPP'25] Shixun Wu**, Yujia Zhai, Jinyang Liu, Jiajun Huang, Zizhe Jian, Sheng Di, Franck Cappello, Zizhong Chen. "TurboFFT: Co-Designed High-Performance and Fault-Tolerant Fast Fourier Transform on GPUs",*ACM Symposium on Principles and Practice of Parallel Programming (PPoPP), 2025.* [paper]

C2 **[Cluster'24] Shixun Wu***, Yitong Ding*, Yujia Zhai, Jinyang Liu, Jiajun Huang, Zizhe Jian, Huangliang Dai, Sheng Di, Bryan Wong, Zizhong Chen, and Franck Cappello. "FT K-means: A High-Performance K-means on GPU with Fault Tolerance.", 2024 IEEE International Conference on Cluster Computing (CLUSTER). [paper]

C3 **[SC'24]** Jinyang Liu*, Jiannan Tian*, **Shixun Wu***, Sheng Di, Boyuan Zhang, Robert Underwood, Yafan Huang, Jiajun Huang, Kai Zhao, Guanpeng Li, Dingwen Tao, Zizhong Chen, and Franck Cappello. "cuSZ-I: High-Fidelity Error-Bounded Lossy Compression for Scientific Data on GPUs." *2024 SC24: International Conference for High Performance Computing, Networking, Storage and Analysis.* [paper]

C4 **[ICS'23] Shixun Wu***, Yujia Zhai*, Jinyang Liu, Jiajun Huang, Zizhe Jian, Bryan Wong, Zizhong Chen. "Anatomy of High-Performance GEMM with Online Fault Tolerance on GPUs." *The 37th ACM International Conference on Supercomputing*, Orlando, FL, USA, June 21–23, 2023. [paper]

C5 **[IPDPS'24]** Zizhe Jian, Sheng Di, Jinyang Liu, Kai Zhao, Xin Liang, Haiying Xu, Robert Underwood, **Shixun Wu**, Jiajun Huang, Zizhong Chen, and Franck Cappello. "CliZ: Optimizing Lossy Compression for Climate Datasets with Adaptive Fine-tuned Data Prediction."

C6 **[SIGMOD'24]** Jinyang Liu, Sheng Di, Kai Zhao, Xin Liang, Sian Jin, Zizhe Jian, Jiajun Huang, **Shixun Wu**, Zizhong Chen, Franck Cappello. "High-performance Effective Scientific Error-bounded Lossy Compression with Auto-tuned Multi-component Interpolation." *2023 IEEE International Conference on Big Data (BigData)*

C7 **[BigData'23]** Jiajun Huang, Jinyang Liu, Sheng Di, Yujia Zhai, Zizhe Jian, **Shixun Wu**, Kai Zhao, Zizhong Chen, Yanfei Guo, Franck Cappello. "Exploring Wavelet Transform Usages for Error-bounded Scientific Data Compression" *Companion of the 2023 International Conference on Management of Data.*

C8 **[Allerton'23]** Jeremy Johnston, Xiaoyang Liu, **Shixun Wu**, Xiaodong Wang. "Downlink beamforming optimization via deep learning" *2023 59th Annual Allerton Conference on Communication, Control, and Computing.*

JOURNAL PAPERS

J1 **[IJHPCA'25]** Balaprakash Prasanna, Krishnan Raghavan, Franck Cappello, Ewa Deelman, Anirban Mandal, Hongwei Jin, Imtiaz Mahmud, Komal Thareja, **Shixun Wu**, Pawel Zuk, Mariam Kiran, Zizhong Chen, Sheng Di, Kesheng Wu. "SWARM: Reimagining Scientific Workflow Management Systems in a Distributed World",*International Journal of High Performance Computing Applications, 2025.* [paper]

J2 **[TSP'23]** Jeremy Johnston, Xiaoyang Liu, **Shixun Wu**, Xiaodong Wang. "A Curriculum Learning Approach to Optimization with Application to Downlink Beamforming." *IEEE Transactions on Signal Processing (2023).*

POSTER PAPERS

P1 **[HPDC'23] Shixun Wu***, Yujia Zhai*, Jiajun Huang, Zizhe Jian, Zizhong Chen. "FT-GEMM: A Fault Tolerant High Performance GEMM Implementation on x86 CPUs."*The 32nd ACM International Symposium on High-Performance Parallel and Distributed Computing*, Orlando, FL, USA, June 21–23, 2023.[poster]

WORKSHOP PAPERS

W1 **[SC'24]** Ewa Deelman, Prasanna Balaprakash, Mariam Kiran, Anirban Mandal, Krishnan Raghavan, Sheng Di, Franck Cappello, John Wu, Zizhong Chen, **Shixun Wu**, Hongwei Jin, Cong Wang, Imtiaz Mahmud, Komal Thareja, Erik Scott, Pawel Zuk, Aiden Hamade. "SWARM: Scientific Workflow Applications on Resilient Metasystem"

W2 **[ICLR'23]** Xiaoyang Liu, Zechu Li, **Shixun Wu**, Xiaodong Wang. "Stationary Deep Reinforcement Learning With Quantum K-Spin Hamiltonian Regularization." *ICLR 2023 Workshop on Physics for Machine Learning. 2023*

ARXIV (IN SUBMISSION)

A1 **Shixun Wu**, Sheng Di, Krishnan Ragahvan, Kesheng Wu, Ewa Deelman, Franck Cappello "DGRO: Diameter-Guided Ring Optimization for Integrated Research Infrastructure Membership"

A2 Huangliang Dai, **Shixun Wu**, Zizhong Chen "FT-Transformer"

A3 **Shixun Wu**, Zizhong Chen "TurboFNO: High-Performance Fourier Neural Operator with Fused FFT-GEMM-iFFT on GPU"

A4 **Shixun Wu***, Jinyang Liu*, Jiannan Tian*, Jinwen Pan*, Sheng Di, Zizhong Chen, and Franck Cappello "Optimizing GPU-Based Error-Bounded Lossy Compression with Advanced Interpolation and Synergistic Lossy-Lossless Scheme"

PROFESSION SERVICES

1. Reviewer, Parallel Computing, 2025
2. SubReviewer, International Conference on Supercomputing (ICS), 2025
3. SubReviewer, International Symposium on Cluster, Cloud and Internet Computing (CCCGrid), 2025
4. Reviewer, Workshop on General Purpose Processing Using GPU (GPGPU), 2025
5. Reviewer, International Conference on Distributed Computing Systems (ICDCS), 2025
6. SubReviewer, International Parallel and Distributed Processing Symposium (IPDPS), 2024
7. SubReviewer, International Parallel and Distributed Processing Symposium (IPDPS), 2023
8. Reviewer, Computing Surveys, 2023
9. Reviewer, Journal of Intelligent & Fuzzy Systems, 2023

PROJECTS INVOLVED

1. DOE SWARM: Scientific Workflow Applications on Resilient Metasystem
2. DOE DECODE: Data-driven Exascale Control of Optically Driven Excitations in Chemical and Material Systems

TEACHING

1. Teaching Assistant, CS161: Design & Architecture of Computer Systems, University of California, Riverside, Riverside, CA, Fall 2024.
2. Teaching Assistant, CS161: Design & Architecture of Computer Systems, University of California, Riverside, Riverside, CA, Spring 2024.
3. Teaching Assistant, CS160: Concurrent Programming and Parallel Systems, University of California, Riverside, Riverside, CA, Fall 2023.

HONORS AND AWARDS

- NSF PPoPP Travel Grant ***Feb, 2025***
- Outstanding Teaching Award, University of California, Riverside ***May, 2024***
- Third Prize at UCRPC, University of California, Riverside, ***Nov, 2023***
- Distinguished Dean's Fellowship, University of California, Riverside, ***Sep, 2022***
- Second Prize in PKU ACM in ***2017, 2018***
- PKU May 4th Scholarship. ***2017***