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. **NVIDIA** Santa Clara, CA
   cuPyNumeric *Jun. 2025 –Sep. 2025*
   Multi-GPU Multi-Node numerical method development and interoperability with PyTorch Distributed.
   Compiler Optimization for OpenMP Offload *Jun. 2024 –Sep. 2024*
   Optimized OpenMP target offload in SPEChpc 2021 on GH200/H200 GPUs, achieving up to 10× speedup without source changes; work adopted by RWTH Aachen University and merged into SPEChpc 1.1.9.
2. **USC ISI / Argonne National Laboratory** Los Angeles, CA / Lemont, IL
   Scientific Workflow Applications on Resilient Metasystem *Jan. 2024 –Present*
   Designed DGRO, a Q-learning + GNN-based topology protocol to reduce network diameter in geo-distributed systems; deployed a resilient single-hop gossip-based failure detector across 20+ global sites on FABRIC testbed.
3. **UCR / Lawrence Berkeley National Laboratory** Riverside, CA
   ABFT-Enhanced GPU Kernels for Scientific Computing *Sep. 2022 –Present*
   Developed ABFT GEMM and FFT pipelines outperforming cuBLAS/cuFFT, and proposed the first GPU-resident fault-tolerant K-means framework with low-overhead resilience for exascale scientific workloads.
4. **Columbia University / AI4Finance Foundation** New York, NY
   ElegantRL: Parallel Deep Reinforcement Learning Library *Aug. 2021 –Jul. 2022*
   Contributed multi-agent RL algorithms to ElegantRL (4k GitHub stars).

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 **[SC'25]** **Shixun Wu***, Jinwen Pan*, Jinyang Liu, Jiannan Tian, Ziwei Qiu, Jiajun Huang, Kai Zhao, Xin Liang, Sheng Di, Zizhong Chen, and Franck Cappello. "Boosting Scientific Error-Bounded Lossy Compression through Optimized Synergistic Lossy-Lossless Orchestration." *2025 SC25: International Conference for High Performance Computing, Networking, Storage and Analysis.*

C2 **[SC'25]** **Shixun Wu**, Yujia Zhai, Hairui Zhao, Huliang Dai, Yue Zhu, Haiyang Hu, and Zizhong Chen. "TurboFNO: High-Performance Fourier Neural Operator with Fused FFT-GEMM-iFFT" *2025 SC25: International Conference for High Performance Computing, Networking, Storage and Analysis.*

C3 **[SC'25]** Huangliang Dai, **Shixun Wu**, Hairui Zhao, Jiajun Huang, Zizhe Jian, Yue Zhu, Haiyang Hu, and Zizhong Chen. "FT-Transformer: Resilient and Reliable Transformer with End-to-End Fault Tolerant Attention." *2025 SC25: International Conference for High Performance Computing, Networking, Storage and Analysis.*

C4 **[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]

C5 **[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]

C6 **[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]

C7 **[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]

C8 **[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."*2024 IEEE International Symposium on Parallel and Distributed Processing (IPDPS).* [paper]

C9 **[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." *Proceedings of the ACM on Management of Data 2, no. 1 (2024): 1-27* [paper]

C10 **[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" *2023 IEEE International Conference on Big Data (BigData)*[paper]

C11 **[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.* [paper]

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).* [paper]

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" [paper]

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* [paper]

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" [paper]

PROFESSIONAL SERVICES

Reviewer: Parallel Computing (2025), GPGPU Workshop (2025), ICDCS (2025), Computing Surveys (2023), Journal of Intelligent & Fuzzy Systems (2023), ICS (2025), CCGrid (2025), IPDPS (2023, 2024), IEEE Open Journal of the Computer Society (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

Teaching Assistant, University of California, Riverside: CS161 (Design & Architecture of Computer Systems) in Fall 2024 and Spring 2024; CS160 (Concurrent Programming and Parallel Systems) in 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***