

Zhonghao Chen

University of Florida, Gainesville, FL, 32608
E-mail: zh.chen@ufl.edu | Tel: (209)–504–3779
Website: <https://diogeneschcn.github.io/>

Education

University of Florida, Gainesville, FL, U.S.
Ph.D. in ECE, Advisor: Dr. Xiaoyi Lu

Jan. 2026 -

University of California, Merced, Merced, CA, U.S.
Ph.D. in EECS, Advisor: Dr. Xiaoyi Lu
(Transferring to University of Florida in Jan. 2026)

Aug 2024 - Dec 2025

Tianjin University, Tianjin, Tianjin, China
B.E. (with Honors), Computer Science and Technology

September 2020 - July 2024

Research Interests

High-performance computing and networking; Runtime and communication systems; Scalable training and inference of large language models; Distributed and federated learning systems; AI–HPC co-design and convergence for efficient large-scale machine learning.

Research Experience

- **Reproduction and Characterization of Large Reasoning Models on HPC**

University of California, Merced, Advisor: Dr. Xiaoyi Lu

December 2024 -

- Investigate inference and distillation performance of large reasoning models (e.g., dense Llama-70B, MoE DeepSeek-R1-671B, Kimi-K2-1T) on HPC-scale GPU clusters (NVIDIA, AMD, Intel, etc.) and interconnects (Slingshot, InfiniBand).
- Employ parallelisms including Data, Tensor, Pipeline, and Expert, together with prefill-decode disaggregation, KV-cache transfer engines etc., to scale model deployment up to 3,840 GPUs.
- Analyze pipeline imbalance, communication overhead, and KV-cache bottlenecks limiting throughput and concurrency; identify optimal GPU and cluster configurations for scalable inference and provide system-level insights for AI–HPC co-design.
- Using Python, C/C++, and CUDA for building high-performance distributed ML systems, MPI/NCCL for communication, DeepSpeed/Megatron, vLLM, VerL, Ray for LLM parallelism, and PyTorch Profiler, Nsight system for performance profiling.
- Ongoing project; partial results published at SC '25.

- **Scalable and Resilient Federated Learning**

University of California, Merced, PI: Dr. Xiaoyi Lu

July 2024 -

- Design a scalable and fault-tolerant modeling and simulation framework for complex federated learning (FL) workflows across cross-silo and cross-device scenarios under diverse networking conditions.
- Develop a discrete-event simulator on SimGrid that abstracts FL training, communication, and aggregation into lightweight data-plane events.
- Build a user-space, MPI-level long-haul RDMA simulator with rate limiting and distance-based propagation delay modeling to compare RDMA and TCP/IP for geo-distributed FL workloads, offering practical guidance for real-world long-haul deployment.
- Compare RDMA and TCP/IP for FL workloads, and verify on real-world long-haul RDMA from Berkley, CA to Chicago, IL upon ESNt testbed
- Ongoing project; partial results published at SEC '25, SC '25.

Selected Publications

- [1] [ICML'26] Jin Lee, **Zhonghao Chen**, Xuhang He, Robert Underwood, Bogdan Nicolae, Franck Cappello, Xiaoyi Lu, Sheng Di, and Zheng Zhang. "SPARe: Stacked Parallelism with Adaptive Reordering for Fault-Tolerant LLM Pretraining Systems with 100k+ GPUs". In: *Proceedings of the 43rd International Conference on Machine Learning*. ICML'26. Vienna, Austria: JMLR.org, 2026.
- [2] [CAIS'26] **Zhonghao Chen**, Yuke Li, and Xiaoyi Lu. "FedMECA: Scalable Federated Learning via Memory-Efficient and Concurrent Aggregation". In: *Proceedings of the 1st ACM conference on AI and Agentic Systems*. CAIS'26. San Jose, California: Association for Computing Machinery, 2026.
- [3] [HPDC'26] Yuke Li, **Zhonghao Chen**, and Xiaoyi Lu. "When RDMA Goes Long-Haul: Characterization, Modeling, and Verbs-Level Emulation with Implications for Federated Learning". In: *Proceedings of the 35th International Symposium on High-Performance Parallel and Distributed Computing*. HPDC '26. New York, NY, USA: Association for Computing Machinery, 2026.
- [4] [SEC'25] **Zhonghao Chen**, Weicong Chen, Duo Zhang, Kibaek Kim, Guanpeng Li, Sheng Di, and Xiaoyi Lu. "FedDES: Discrete Event Based Performance Simulation for Federated Learning Systems". In: *The Tenth ACM/IEEE Symposium on Edge Computing (SEC)*. IEEE. 2025.
- [5] [SC'25] (**Best Research Poster Candidate**) **Zhonghao Chen**, Yuke Li, Duo Zhang, and Xiaoyi Lu. "Can Long-Haul RDMA Benefit Federated Learning?" In: *SC25: International Conference for High Performance Computing, Networking, Storage and Analysis*. IEEE. 2025.
- [6] [SC'25] Adam Weingram, Duo Zhang, **Zhonghao Chen**, Hao Qi, and Xiaoyi Lu. "HPC-R1: Characterizing R1-like Large Reasoning Models on HPC". In: *SC25: International Conference for High Performance Computing, Networking, Storage and Analysis*. IEEE. 2025.
- [7] Jitao Bai, **Zhonghao Chen**, Simiao Zhang, and Jiahe Liang. "Building form optimization for renewable energy-economic utility of flexible solar cells as building integrated photovoltaics". In: *Sustainable Cities and Society* 133 (2025), p. 106819. ISSN: 2210-6707. DOI: <https://doi.org/10.1016/j.scs.2025.106819>. URL: <https://www.sciencedirect.com/science/article/pii/S2210670725006924>.
- [8] Jitao Bai, Keyong Yang, **Zhonghao Chen**, Jiahe Liang, Simiao Zhang, and Yu Diao. "Geometry and Material Criteria for Low-Carbon Design of I/H-Beams in Sustainable Steel Structures Considering Both Mechanical Properties and Carbon Emissions". In: *Materials* 18.21 (2025). ISSN: 1996-1944. DOI: 10.3390/ma18214930. URL: <https://www.mdpi.com/1996-1944/18/21/4930>.
- [9] Ying Zhang, Jitao Bai, Yu Diao, **Zhonghao Chen**, Chu Wang, Kun Yang, Zeng Gao, and Huajie Wei. "Risk and Energy Based Optimization for Fire Monitoring System in Utility Tunnel Using Cellular Automata". In: *Sustainability* 16.11 (2024), p. 4717.

Professional Experience

1. **NXP Semiconductors**, Tianjin, Tianjin, China
Software Engineer Intern - GUI Development *April 2023 - June 2023*
2. **Tianjin University Software Studio(TWT)**, Tianjin, Tianjin, China
Software Engineer Intern - iOS Backend Development *November 2020 - May 2021*

Honors, Awards and Activities

- Student Travel Grant Q-CORE @ QCE 2025 *August 2025*
- B.E. with honors *July 2024*
- Talent Student in Scientific and Technological Research *September 2023*
- The 7th Undergraduate Integrated Circuits Innovation and Entrepreneurship Competition *June 2023*
- Member, Undergraduate Academic Committee *September 2020 - July 2021*