(630) 779-5671
Bay Area, California, USA
liuzhengchun@gmail.com

Zhengchun Liu

Senior Applied Scientist, PhD

lzhengchun.github.io
github.com/lzhengchun
linkedin.com/in/liuzhengchun

Zhengchun has 10+ years of experience in Research and Development, specializing in Data Infrastructure, Machine Learning, and High-Performance Computing systems. Since 2022, he has been serving as a Senior Applied Scientist at AWS, where he Leads the Technical development of ML capabilities in Redshift, focusing on enabling autonomous capabilities for data warehouse. Before this, he was a Computer Scientist at Argonne National Laboratory and the University of Chicago from 2016 to 2024 where he led or co-led 10+ NSF/DOE funded research projects, securing over $12M in funding. His expertise spans technical leadership, innovative problem-solving, and advancing cutting-edge technologies in his field.

Research & Engineering Experience

Amazon Web Service (AWS), AI Labs **Aug. 2022 — present**
Senior Applied Scientist
- Tech Lead (Initiate, Design & Review) of Machine Learning model for AI-driven Autonomous data warehouse.
- The ML Models developed support billions of decisions for Scaling, Scheduling and Optimization every day.
- Drive predictive analytics of customer behavior, and launch data-driven initiatives for customer pain points.
- Lead scientist developing the Agentic AI system that powers analytic SQL generation in Amazon Ads Agent.

Argonne National Laboratory, Data Science and Learning division **Aug. 2019 — Aug. 2024**
Sabbatical leave (Aug. 2022), Computer Scientist (Tenured, May 2022), Assistant Computer Scientist (Sep. 2019)
- Serve as a Principal Investigator (PI) or co-PI of 10+ NSF/DOE research projects totaling more than $12M of funding.
- Lead a team conducting research on Artificial Intelligence for Science (AI4S), and HPC/Systems for AI4S applications.
- Lead author of more than 20 Publications in prestigious venues on Data Infrastructure, HPC and AI4S.

University of Chicago, Computation Institute **Mar. 2018 — Aug. 2019**
Research Scientist
- Senior Personnel or co-PI of research projects on the architecture of Autonomous science ecosystems.
- Drive the research of performance modeling and characterization of scientific applications on supercomputers.

Argonne National Laboratory, Mathematics and Computer Science division **Sep. 2016 — Mar. 2018**
Postdoctoral Researcher
- Led research on High Performance File Transfer over Wide-Area Science Network.
- Key Personnel of a project on Modeling, Simulation and Optimization of scientific workflows over distributed infra.

Freelancer **May 2010 — Nov. 2013**
Embedded Software Engineer *(Part-time); Xi'an FengLiTong Electronic, China* *May 2010 — Sep. 2013*
- Board Support Package development for μC/OS-II real-time operating system.
- Designed and Led the implementation of an event streaming service for Vehicle Event Data Recorder (connected IoT devices).

Founder*; Embedded System Technology Consulting & Outsourcing LLC, China* *Sep. 2010 — Nov. 2013*
- 10+ Consulting and/or Outsourcing contracts, three prototypes went into massive production.

Selected Awards, Grants & Honors

- `Best Paper` award $\times$5; XLOOP@SC'21, FGCS'21, MLN'19, MLN'18, and TRIDENTCOM'18 — 2021, 2019, 2018
- `Impact Argonne/Pacesetter` award $\times$3; for notable achievement in `Innovation` — Nov. 2019, May 2020, July 2022
- `Top Winner` of the Technology Challenge at the ACM/IEEE Supercomputing Conference (SC'19) — Nov. 2019

Education

- **PhD., Computer Science**, Autonomous University of Barcelona, Spain — Jul. 2016
  - `With Honors(Cum Laude); Extraordinary Doctorate Award(1 out of 5 thesis).`
- **MSc., Guidance, Navigation and Control**, Northwestern Polytechnical University, China — Apr. 2013
- **BSc., Aircraft Manufacturing Engineering**, Northwestern Polytechnical University, China — Jun. 2010
  - `National Scholarship` (award to top 0.2% of outstanding undergraduates) $\times$2, 2008 and 2009.
  - `Champion` of the Chinese Robot Competition (RoboCup 2008), dancing session. Dec. 2008

Activities

- `Editorship`: Journal of Future Generation Computer Systems (FGCS), 2020-2023, Impact Factor: 7.2.
- `Workshop Co-Chair`: SRMPDS 2017, 2018, 2019, 2020; AI-Science'19.
- `Technical Program Committee`: ICDS'19; DAAC'17-19; DLS'19; ICDCS'20; SC'20; HPCC'20; HiPC'22; SC'23; IPDPS'23,24.
- `Journal Reviewer`: MDPI Sensors; Elsevier FGCS, JOCS; IEEE Access, TPDS.

Selected Publications

* denotes student/post-doc supervised. **Full list**: Google Scholar

Papers on Machine Learning and Systems

1. Alexander van Renen, Dominik Horn, Pascal Pfeil, Kapil Eknath Vaidya, Wenjian Dong, Murali Narayanaswamy, **Zhengchun Liu**, Gaurav Saxena, Andreas Kipf, Tim Kraska. ***Why TPC is not enough: An analysis of the Amazon Redshift fleet***. VLDB'24.
2. Vikram Nathan, Vikramank Singh, **Zhengchun Liu**, Mohammad Rahman, Andreas Kipf, Dominik Horn, Davide Pagano, Gaurav Saxena, Balakrishnan Narayanaswamy, Tim Kraska. ***Intelligent Scaling in Amazon Redshif***. SIGMOD'24.
3. Ziniu Wu*, Ryan Marcus, **Zhengchun Liu**, Parimarjan Negi, Vikram Nathan, Pascal Pfeil, Gaurav Saxena, Mohammad Rahman, Balakrishnan Narayanaswamy, Tim Kraska. ***Stage: Query Execution Time Prediction in Amazon Redshift***. SIGMOD'24.
4. **Zhengchun Liu**, Rajkumar Kettimuthu, Michael E. Papka, Ian Foster. ***FreeTrain: A Framework to Utilize Unused Supercomputer Nodes for Training Neural Networks***. IEEE/ACM International Symposium in Cluster, Cloud, and Grid Computing (CCGrid'23).
5. Ahsan Ali*, Hemant Sharma, Rajkumar Kettimuthu, Peter Kenesei, Dennis Trujillo, Antonino Miceli, Ian Foster, Ryan Coffee, Jana Thayer, **Zhengchun Liu**. ***fairDMS: Rapid Model Training by Data and Model Reuse***. 2022 IEEE International Conference on Cluster Computing. arXiv:2204.09805
6. Joaquin Chung, Wojciech Zacherek, AJ Wisniewski, **Zhengchun Liu**, Tekin Bicer, Rajkumar Kettimuthu and Ian Foster. ***SciStream: Architecture and Toolkit for Data Streaming between Federated Science Instruments***. ACM HPDC'2022.
7. **Zhengchun Liu**, Ahsan Ali*, Peter Kenesei, Antonino Miceli, Hemant Sharma, Nicholas Schwarz, Dennis Trujillo, Hyunseung Yoo, Ryan Coffee, Ryan Herbst, Jana Thayer, Chun Hong Yoon, Ian Foster. ***Bridge Data Center AI Systems with Edge Computing for Actionable Information Retrieval***. XLOOP@SC'21. ***Best Paper Awarded***.
8. **Zhengchun Liu**, Rajkumar Kettimuthu, Joaquin Chung, Rachana Ananthakrishnan, Michael Link, Ian Foster. ***Design and Evaluation of a Simple Data Interface for Efficient Data Transfer Across Diverse Storage***. ACM Transactions on Modeling and Performance Evaluation of Computing Systems, 2021 Vol. 6, No.1.
9. Papadimitriou, George, Cong Wang, Karan Vahi, Rafael Ferreir da Silva, Anirban Mandal, **Zhengchun Liu**, Rajiv Mayania, Mats Rynge, Mariam Kiran, Vickie E. Lynch, Rajkumar Kettimuthu, Ewa Deelman, Jeffrey S. Vetter, Ian Foster. ***End-to-End Online Performance Data Capture and Analysis for Scientific Workflows***. Future Generation Computer Systems, Vol. 117, 2021, Pages 387-400, ***2021 Best Paper Award***.
10. Tirthak Patel, **Zhengchun Liu**, Rajkumar Kettimuthu, Paul Rich, Bill Allcock, Devesh Tiwari. ***Job Characteristics on Large-Scale Systems: Long-Term Analysis, Quantification, and Implications***. The International Conference for High Performance Computing, Networking, Storage, and Analysis (SC'20).
11. **Zhengchun Liu**, Ryan Lewis*, Rajkumar Kettimuthu, Kevin Harms, Philip Carns, Nageswara Rao, Ian Foster and Michael Papka. ***Characterization and Identification of HPC Applications at Leadership Computing Facility***. International Conference on Supercomputing (ICS'20).
12. Yuanlai Liu*, **Zhengchun Liu**, Rajkumar Kettimuthu, Nageswara Rao, Zizhong Chen and Ian Foster. ***Data transfer between scientific facilities - bottleneck analysis, insights and optimizations***. IEEE/ACM International Symposium in Cluster, Cloud, and Grid Computing (CCGrid'19).
13. **Zhengchun Liu**, Rajkumar Kettimuthu, Prasanna Balaprakash, Nageswara S. V. Rao and Ian Foster. ***Building a Wide-Area Data Transfer Performance Predictor: An Empirical Study***. International Conference on Machine Learning for Networking (MLN'18).
14. **Zhengchun Liu**, Rajkumar Kettimuthu, Ian Foster, Peter H. Beckman. ***Towards a Smart Data Transfer Node***. Future Generation Computer Systems, 2018(89), Pages 10—18.
15. Rajkumar Kettimuthu, **Zhengchun Liu**, David Wheeler, Ian Foster, Katrin Heitmann, Franck Cappello. ***Transferring a Petabyte in a Day***. Future Generation Computer Systems, 2018(88).
16. **Zhengchun Liu**, Rajkumar Kettimuthu, Ian Foster and Yuanlai Liu. ***A comprehensive study of wide area data movement at a scientific computing facility***. IEEE 38th International Conference on Distributed Computing Systems (SNTA@ICDCS'18).
17. Rajkumar Kettimuthu, **Zhengchun Liu**, Ian Foster, Peter H. Beckman, Alex Sim, John Wu, Wei-keng Liao, Qiao Kang, Ankit Agrawal, and Alok Choudhary. 2018. ***Toward Autonomic Science Infrastructure: Architecture, Limitations, and Open Issues***. The 1st Autonomous Infrastructure for Science workshop (AI-Science@HPDC'18).
18. **Zhengchun Liu**, Rajkumar Kettimuthu, Ian Foster and Nageswara S.V. Rao. ***Cross-geography Scientific Data Transfer Trends and User Behavior Patterns***. International Symposium on High-Performance Parallel and Distributed Computing (HPDC'18).
19. **Zhengchun Liu**, Rajkumar Kettimuthu, Sven Leyffer, Prashant Palkar and Ian Foster. ***A mathematical programming and simulation based framework to evaluate cyberinfrastructure design choices***. IEEE International Conference on eScience.
20. **Zhengchun Liu**, Prasanna Balaprakash, Rajkumar Kettimuthu and Ian Foster. ***Explaining Wide Area Data Transfer Performance***. International Symposium on High-Performance Parallel and Distributed Computing (HPDC'17), 167-178.

Papers on AI/HPC for Science

1. Anakha V Babu, Tao Zhou, Saugat Kandel, Tekin Bicer, **Zhengchun Liu**, William Judge, Daniel J Ching, Yi Jiang, Sinisa Veseli, Steven Henke, Ryan Chard, Yudong Yao, Ekaterina Sirazitdinova, Geetika Gupta, Martin V Holt, Ian T Foster, Antonino Miceli and Mathew J Cherukara. ***Deep Learning at the Edge Enables Real-time Streaming Ptychographic Imaging***. Nature Communications 14, 7059 (2023).
2. Lipeng Wan, Jieyang Chen, Xin Liang, Ana Gainaru, Qian Gong, Qing Liu, Ben Whitney, Joy Arulraj, **Zhengchun Liu**, Ian Foster, Scott Klasky. ***RAPIDS: Reconciling Availability, Accuracy, and Performance in Managing Geo-Distributed Scientific Data***. International Symposium on High-Performance Parallel and Distributed Computing (HPDC'23).
3. Petro Junior Milan, Hongqian Rong, Craig Michaud, Naoufal Layad, **Zhengchun Liu**, Ryan Coffee. ***Enabling real-time adaptation of machine learning models at x-ray Free Electron Laser facilities with high-speed training optimized computational hardware***. Frontiers in Physics, Volume 10 - 2022.
4. Mike Kraus, Naoufal Layad, **Zhengchun Liu**, Ryan Coffee. ***EdgeAI: Machine learning via direct attached accelerator for streaming data processing at high shot rate x-ray free-electron lasers***. Frontiers in Physics, Volume 10 - 2022.
5. **Zhengchun Liu**, Rajkumar Kettimuthu, Ian Foster. ***Masked Sinogram Model with Transformer for ill-Posed Computed Tomography Reconstruction: a Preliminary Study***. arXiv:2209.01356
6. **Zhengchun Liu**, Hemant Sharma, Jun-Sang Park, Peter Kenesei, Antonino Miceli, Jonathan Almer, Rajkumar Kettimuthu and Ian Foster. ***BraggNN: Fast X-ray Bragg Peak Analysis Using Deep Learning***. IUCrJ, Vol. 9, No. 1, 2022.
7. Aniket Tekawade, **Zhengchun Liu**, Peter Kenesei, Tekin Bicer, Francesco De Carlo, Rajkumar Kettimuthu, Ian Foster. ***3D Autoencoders For Feature Extraction In X-ray Tomography***. 2021 IEEE International Conference on Image Processing.
8. Jiali Wang, **Zhengchun Liu**, Ian Foster, Won Chang, Rajkumar Kettimuthu, Rao Kotamarthi. ***Fast and accurate learned multiresolution dynamical downscaling for precipitation***. journal of Geoscientific Model Development.
9. Selin Aslan, **Zhengchun Liu**, Viktor Nikitin, Tekin Bicer, Sven Leyffer, Doga Gursoy. ***Joint Ptycho-Tomography with Deep Generative Priors***. Machine Learning Science and Technology, 2021, Vol. 2, No.4
10. Ziling Wu*, Tekin Bicer, **Zhengchun Liu**, Vincent De Andrade, Yunhui Zhu, Ian T. Foster. ***Deep Learning-based Low-dose Tomography Reconstruction with Hybrid-dose Measurements***. AI4S@SC'20.
11. **Zhengchun Liu**, Tekin Bicer, Rajkumar Kettimuthu, Doga Gursoy, Francesco De Carlo and Ian Foster. ***TomoGAN: Low-Dose Synchrotron X-Ray Tomography with Generative Adversarial Networks***. Optical Society of America A, 2020, Vol. 37, No. 2.
12. Vibhatha Abeykoon*, **Zhengchun Liu**, Tekin Bicer, Rajkumar Kettimuthu, Geoffrey Fox and Ian Foster. ***Scientific Image Restoration Anywhere***. XLOOP @SC'19.
13. **Zhengchun Liu**, Tekin Bicer, Rajkumar Kettimuthu and Ian Foster. ***Deep Learning Accelerated Light Source Experiments***. IEEE/ACM Deep Learning on Supercomputers DLS@SC'19.
14. **Zhengchun Liu**, Dolores Rexachs, Francisco Epelde, and Emilio Luque. ***A simulation and optimization based method for calibrating agent-based emergency department models under data scarcity***. Computers & Industrial Engineering, 2017.
15. Xueping Zhu, **Zhengchun Liu** and Jun Yang. ***Model of Collaborative UAV Swarm Toward Coordination and Control Mechanisms Study***. 2015 International Conference on Computational Science (ICCS'15), Vol 51, 493-502.
16. **Zhengchun Liu**, Eduardo Cabrera, Manel Taboada, Francisco Epelde, Dolores Rexachs and Emilio Luque. ***Quantitative Evaluation of Decision Effects in the Management of Emergency Department Problems***. International Conference on Computational Science (ICCS'15), Vol 51, Pages 433-442.
17. **Zhengchun Liu**, Eduardo Cabrera, Dolores Rexachs and Emilio Luque. ***A Generalized Agent-Based Model to Simulate Emergency Departments***. The 6th International Conference on Advances in System Simulation (SIMUL'14).

– ***Last updated: 2025***