Qinglei Cao

qingleicao3@gmail.com St. Louis, MO, US www.qingleicao.com

RESEARCH INTERESTS

Parallel and distributed computing, Task-based runtime systems, Linear algebra algorithms, Extreme-scale domain applications, and Large-scale machine learning & deep learning

EDUCATION

The University of Tennessee, Knoxville (UTK), Computer Science — 2016 - 2022
PhD, High Performance Computing
Advisor: Dr. Jack Dongarra (Turing Award 2021)
Group Leader & Co-Advisor: Dr. George Bosilca

Ocean University of China (OUC), Computer Application Technology — 2013 - 2016
MS, Image Processing & Parallel Computing
Advisors: Dr. Yuntao Qian (Zhejiang University), Dr. Zhiqiang Wei (OUC)

Hunan University (HNU), Information and Computational Science — 2005 - 2009
BS, Mathematics

PROFESSIONAL EXPERIENCE

Department of Computer Science, Saint Louis University (SLU) — St. Louis, MO
Assistant Professor — 2023 - Present

Innovative Computer Laboratory (ICL), UTK — Knoxville, TN
Post-Doctoral Research Associate — 2023

Cerebras Systems, Inc. — Sunnyvale, CA
Member of Technical Staff — 2022 - 2023

Innovative Computer Laboratory (ICL), UTK — Knoxville, TN
Graduate Research Assistant — 2017 - 2022

Cerebras Systems, Inc. — Sunnyvale, CA
Summer Intern — 2021

Cadence Design Systems, Inc. — Austin, TX
Summer Intern — 2020

National University of Defense Technology (NUDT) — Changsha, China
HPC Software Developer & Research Scientist — 2010 - 2013

HONORS & AWARDS

- CRII Award, OAC, NSF, US — 2025
- 5,000 Node Hours on Polaris Supercomputer, Argonne National Laboratory, US — 2025
- ACM Gordon Bell Prize for Climate Modelling — 2024
- ACM Gordon Bell Prize, Finalist — 2024
- Subaward Funding, Oak Ridge National Laboratory, US — 2024
- 20,000 Node Hours on Frontier Supercomputer, Oak Ridge National Laboratory, US — 2024
- 206,780 Node Hours on Summit Supercomputer, Oak Ridge National Laboratory, US — 2024
- 20,000 Node Hours on Frontier Supercomputer, Oak Ridge National Laboratory, US — 2024
- ACM Gordon Bell Prize, Finalist — 2022

- ◇ 4,000,000 Node Hours on Fugaku Supercomputer, RIKEN, Japan 2022
- ◇ SIAM Student Travel Award 2021
- ◇ 40,000 Node Hours on Summit Supercomputer, Oak Ridge National Laboratory, US 2021
- ◇ Best Paper Award, CLUSTER 2020
- ◇ Graduate Student Senate (GSS) Travel Awards, UTK 2020
- ◇ Honor of Outstanding Graduates, OUC 2014
- ◇ Graduate Student Scholarship, OUC 2014
- ◇ Honor of Annual Advanced Worker, NUDT 2010, 2011
- ◇ Honor of Bronze Medal of TianHe-1A, NUDT 2010
- ◇ Outstanding Scholarship, HNU 2006, 2007
- ◇ Honor of Excellent Student Cadre, HNU 2006

PUBLICATIONS

◇ Conference & Journal Papers

1. Rabab Alomairy, Qinglei Cao, Hatem Ltaief and David Keyes. Scalable Low-Rank Solvers for Large-Scale 3D Mesh Deformation Using Global and Compact Support RBF Kernels. IEEE International Symposium on Cluster, Cloud, and Internet Computing (CCGRID, Scalable Computing Challenge Finalist), 2025 (Accepted)
2. Dali Wang, Chen Wang, Qinglei Cao, Jayesh Krishna, Danqing Wu, Weijian Zheng, Peter Schwartz, Fengming Yuan, Kathryn Mohror and Peter Thornton. Scaling Ultrahigh-Resolution E3SM Land Model for Leadership-Class Supercomputers. IEEE International Symposium on Cluster, Cloud, and Internet Computing (CCGRID, Scalable Computing Challenge Finalist), 2025 (Accepted)
3. Qiao Zhang, Rabab Alomairy, Dali Wang, Zhuowei Gu and Qinglei Cao. Leveraging Hardware-Aware Computation in Mixed-Precision Matrix Multiply: A Tile-Centric Approach. Workshop on Asynchronous Many-Task Systems and Applications (WAMTA), 2025 (Accepted)
4. Aurelien Bouteiller, Qinglei Cao, Joseph Schuchart and Thomas Herault. Comparing and Contrasting User and Runtime Directed Data Placement Strategies for Owner-Compute, Multi-Accelerator Distributed Task Based Scheduling. Workshop on Asynchronous Many-Task Systems and Applications (WAMTA), 2025 (Accepted)
5. Rabab Alomairy, Hatem Ltaief, Qinglei Cao and David Keyes. Analyzing Data Sparsity in Global and Compact Support Radial Basis Functions for 3D Unstructured Mesh Deformation. Workshop on Asynchronous Many-Task Systems and Applications (WAMTA), 2025 (Accepted)
6. Rabab Alomairy, Qinglei Cao, Hatem Ltaief, David E. Keyes, and Alan Edelman. Scalable Hamming Distance Computation Using Accelerated Matrix Transformations. ISC High Performance, 2025 (Accepted)
7. Wang, Dali, Peter Schwartz, Fengming Yuan, Franklin Eaglebarge, Danial Riccuito, Peter Thornton, Chris Layton, and Qinglei Cao. Portable Software Environment for Ultrahigh-Resolution ELM Development on GPUs. Journal of Computer and Communications, 2025
8. Aurelien Bouteiller, Thomas Herault, Qinglei Cao, Joseph Schuchart, and George Bosilca. PaRSEC: Scalability, Flexibility, and Hybrid Architecture Support for Task-based Applications in ECP. International Journal of High Performance Computing Applications (IJHPCA), 2025
9. Sameh Abdulah, Allison H. Baker, George Bosilca, Qinglei Cao, Stefano Castruccio, Marc G. Genton, David E. Keyes, Zubair Khalid, Hatem Ltaief, Yan Song, Georgiy L. Stenchikov, and Ying Sun. Boosting Earth System Model Outputs And Saving PetaBytes in their Storage Using Exascale Climate Emulators. International Conference for High Performance Computing, Networking, Storage and Analysis (SC, ACM Gordon Bell Prize for Climate Modelling), 2024 [PDF]

10. Hatem Ltaief, Rabab Alomairy, Qinglei Cao, Jie Ren, Lotfi Slim, Thorsten Kurth, Benedikt Dorschner, Salim Bougouffa, Rached Abdelkhalek, and David E. Keyes. Toward Capturing Genetic Epistasis From Multivariate Genome-Wide Association Studies Using Mixed-Precision Kernel Ridge Regression. International Conference for High Performance Computing, Networking, Storage and Analysis (SC, ACM Gordon Bell Prize Finalist), 2024 [PDF]

11. Shihui Song, Yafan Huang, Peng Jiang, Xiaodong Yu, Weijian Zheng, Sheng Di, Qinglei Cao, Yunhe Feng, Zhen Xie, and Franck Cappello. CereSZ: Enabling and Scaling Error-bounded Lossy Compression on Cerebras CS-2. International Symposium on High-Performance Parallel and Distributed Computing (HPDC), 2024 [PDF]

12. Kareem Shaik, Dali Wang, Weijian Zheng, Qinglei Cao, Heng Fan, Peter Schwartz, and Yunhe Feng. S3LLM : Large-Scale Scientific Software Understanding with LLMs using Source, Metadata, and Document. International Conference on Computational Science (ICCS), 2024 [PDF]

13. Qinglei Cao, Thomas Herault, Aurelien Bouteiller, Joseph Schuchart, and George Bosilca. Evaluating PaRSEC through Matrix Computations in Scientific Applications. Workshop on Asynchronous Many-Task Systems and Applications (WAMTA), 2024 [PDF]

14. Qinglei Cao, Sameh Abdulah, Hatem Ltaief, Marc G Genton, David E Keyes, and George Bosilca. Reducing Data Motion and Energy Consumption of Geospatial Modeling Applications Using Automated Precision Conversion. IEEE International Conference on Cluster Computing (CLUSTER), 2023 [PDF]

15. Qinglei Cao, Sameh Abdulah, Rabab Alomairy, Yu Pei, Pratik Nag, George Bosilca, Jack Dongarra, Marc G Genton, David E Keyes, Hatem Ltaief, and Ying Sun. Reshaping geostatistical modeling and prediction for extreme-scale environmental applications. International Conference for High Performance Computing, Networking, Storage and Analysis (SC, ACM Gordon Bell Prize Finalist), 2022 [PDF]

16. Qinglei Cao, George Bosilca, Nuria Losada, Wei Wu, Dong Zhong, and Jack Dongarra. Evaluating data redistribution in parsec. IEEE Transactions on Parallel and Distributed Systems (TPDS), 2022[PDF]

17. Sameh Abdulah, Qinglei Cao, Yu Pei, George Bosilca, Jack Dongarra, Marc G. Genton, David E. Keyes, Hatem Ltaief, and Ying Sun. Accelerating geostatistical modeling and prediction with mixed-precision computations: A high-productivity approach with parsec. IEEE Transactions on Parallel and Distributed Systems (TPDS), 2022 [PDF]

18. Qinglei Cao, Rabab Alomairy, Yu Pei, George Bosilca, Hatem Ltaief, David Keyes, and Jack Dongarra. A framework to exploit data sparsity in tile low-rank Cholesky factorization. IEEE International Parallel & Distributed Processing Symposium (IPDPS), 2022 [PDF]

19. Qinglei Cao, Yu Pei, Kadir Akbudak, George Bosilca, Hatem Ltaief, David Keyes, and Jack Dongarra. Leveraging parsec runtime support to tackle challenging 3d data-sparse matrix problems. IEEE International Parallel and Distributed Processing Symposium (IPDPS), 2021 [PDF]

20. Dong Zhong, Qinglei Cao, George Bosilca, and Jack Dongarra. Using long vector extensions for MPI reductions. Parallel Computing (PARCO), 2021 [PDF]

21. Yunhe Feng, Dong Zhong, Peng Sun, Weijian Zheng, Qinglei Cao, Xi Luo, and Zheng Lu. Micromobility in smart cities: A closer look at shared dockless e-scooters via big social data. IEEE International Conference on Communications (ICC), 2021 [PDF]

22. Qinglei Cao, George Bosilca, Wei Wu, Dong Zhong, Aurelien Bouteiller, and Jack Dongarra. Flexible data redistribution in a task-based runtime system. IEEE International Conference on Cluster Computing (CLUSTER), 2020 [PDF]

23. Qinglei Cao, Yu Pei, Kadir Akbudak, Aleksandr Mikhalev, George Bosilca, Hatem Ltaief, David Keyes, and Jack Dongarra. Extreme-scale task-based Cholesky factorization toward climate and weather prediction applications. ACM Platform for Advanced Scientific Computing Conference (PASC), 2020 [PDF]

24. Elliott Slaughter, Wei Wu, Yuankun Fu, Legend Brandenburg, Nicolai Garcia, Wilhem Kautz, Emily Marx, Kaleb S. Morris, Qinglei Cao, George Bosilca, Seema Mirchandaney, Wonchan Lee, Sean Treichler, Patrick McCormick, and Alex Aiken. Task bench: a parameterized benchmark for evaluating

parallel runtime performance. IEEE/ACM International Conference for High Performance Computing, Networking, Storage and Analysis (SC), 2020 [PDF]

25. Xi Luo, Wei Wu, George Bosilca, Yu Pei, Qinglei Cao, Thananon Patinyasakdikul, Dong Zhong, and Jack Dongarra. Han: a hierarchical autotuned collective communication framework. IEEE International Conference on Cluster Computing (CLUSTER, Best Paper), 2020 [PDF]
26. Dong Zhong, Qinglei Cao, George Bosilca, and Jack Dongarra. Using advanced vector extensions AVX-512 for MPI reductions. ACM European MPI Users' Group Meeting (EuroMPI), 2020 [PDF]
27. Dong Zhong, Pavel Shamis, Qinglei Cao, George Bosilca, Shinji Sumimoto, Kenichi Miura, and Jack Dongarra. Using ARM scalable vector extension to optimize OpenMPI. IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing (CCGRID), 2020 [PDF]
28. Yu Pei, Qinglei Cao, George Bosilca, Piotr Luszczek, Victor Eijkhout, and Jack Dongarra. Communication avoiding 2d stencil implementations over PaRSEC task-based runtime. IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW), 2020 [PDF]
29. Qinglei Cao, Yu Pei, Thomas Herault, Kadir Akbudak, Aleksandr Mikhalev, George Bosilca, Hatem Ltaief, David Keyes, and Jack Dongarra. Performance analysis of tile low-rank Cholesky factorization using parsec instrumentation tools. IEEE/ACM International Workshop on Programming and Performance Visualization Tools (ProTools at SC), 2019 [PDF]
30. Yan Yan, Jie Nie, Lei Huang, Zhen Li, Qinglei Cao, and Zhiqiang Wei. Exploring relationship between face and trustworthy impression using mid-level facial features. International Conference on Multimedia Modeling (MMM), 2016 [PDF]
31. Yan Yan, Jie Nie, Lei Huang, Zhen Li, Qinglei Cao, and Zhiqiang Wei. Is your first impression reliable? trustworthy analysis using facial traits in portraits. International Conference on Multimedia Modeling (MMM), 2015 [PDF]

◇ Abstracts & Posters

1. Wang, Dali, Peter Schwartz, Fengming Yuan, Daniel M. Ricciuto, Shih-Chieh Kao, Michele Thornton, Anthony Walker, Peter E. Thornton, Qinglei Cao, and Chen Wang. Kilometer-scale E3SM Land Model: Code Development, Deployment, Evaluation, and Applications. American Geophysical Union (AGU), 2024
2. Qinglei Cao, Yu Pei, Kadir Akbudak, Aleksandr Mikhalev, George Bosilca, Hatem Ltaief, David Keyes, and Jack Dongarra. Extreme-Scale Tile Low-Rank Cholesky Factorization Using the PaRSEC Task-Based Runtime. ACM Platform for Advanced Scientific Computing Conference (PASC), 2021

PROFESSIONAL ACTIVITIES

◇ Editorial Board

☞ American Journal of Computer Science and Technology 2023 - 2024

◇ Local Organizer

☞ Workshop on Asynchronous Many-Task Systems and Applications 2025

◇ Technical Program Chair

☞ Workshop on Asynchronous Many-Task Systems and Applications 2025

◇ Technical Program Committee

☞ International Conference on Parallel Processing (ICPP) 2025

☞ Intl Conference for High Performance Computing, Networking, Storage and Analysis (SC) 2025

☞ Asynchronous Many-Task systems for Exascale (AMTE) 2025

☞ International Conference on Parallel Processing (ICPP) 2024

- ☞ Association for the Advancement of Artificial Intelligence (AAAI) Undergraduate Consortium 2024
- ☞ International Supercomputing Conference (ISC) High Performance 2024
- ☞ International Parallel & Distributed Processing Symposium (IPDPS) 2024
- ☞ Workshop on HPC on Heterogeneous Hardware 2022, 2023
- ☞ Intl Conference for High Performance Computing, Networking, Storage and Analysis AD/AE (SC) 2021
- ☞ Intl Conference on Advances and Trends in Software Engineering (SOFTENG) 2021 - 2024

◇ **Conference & Journal Reviewer**

- ☞ Workshop on Asynchronous Many-Task Systems and Applications 2024
- ☞ Parallel Computing 2023
- ☞ Journal of Supercomputing 2023
- ☞ IEEE Transactions on Multimedia 2023
- ☞ ACM Transactions on Mathematical Software (TOMS) 2023
- ☞ International Conference on Emerging Information Security and Applications 2022
- ☞ Intl Conference for High Performance Computing, Networking, Storage and Analysis (SC) 2020, 2021
- ☞ PeerJ Computer Science 2021
- ☞ International Conference on Cluster Computing (CLUSTER) 2020
- ☞ International Conferences on High Performance Computing and Communications (HPCC) 2020, 2021

TEACHING EXPERIENCE

◇ **Lecturer**

- ☞ CSCI 2510 **Principles of Computing Systems** Spring 2025, SLU
- ☞ CSCI 4620/5620 **Distributed Computing** Fall 2024, SLU
- ☞ CSCI 4620/5620 **Distributed Computing** Spring 2024, SLU
- ☞ CSCI 4620/5620 **Distributed Computing** Fall 2023, SLU

◇ **Guest Lecturer**

- ☞ CSCE 5300 **Introduction to Big Data and Data Science** Spring 2025, UNT
- ☞ CSCI 5090 **Computer Science Colloquium** Fall 2024, SLU
- ☞ CSCI 5090 **Computer Science Colloquium** Fall 2023, SLU
- ☞ CSCE 5300 **Introduction to Big Data and Data Science** Fall 2023, UNT
- ☞ CSCE 5300 **Introduction to Big Data and Data Science** Spring 2023, UNT

◇ **Teaching Assistant**

- ☞ COSC 594 **Scientific Computing for Engineers** Spring 2018, UTK
- ☞ COSC 361 **Operating Systems** Spring 2017, UTK
- ☞ COSC 361 **Operating Systems** Fall 2016, UTK

PRESENTATIONS & TALKS

◇ **Talk**

- ☞ NVIDIA GTC (GPU Technology Conference) 2025
- ☞ SIAM Conference on Computational Science and Engineering 2025

- ☞ University of North Texas 2025
- ☞ Annual Workshop on Charm++ and Its Applications 2024
- ☞ Innovative Computer Laboratory (ICL) Lunch Talk 2019, 2020, 2021, 2022
- ☞ Joint Laboratory on Extreme Scale Computing Workshop (JLESC) 2021
- ☞ SIAM Conference on Computational Science and Engineering (CSE) 2021
- ☞ SIAM Conference on Parallel Processing for Scientific Computing (PP) 2020

◇ **Paper Presentation**

- ☞ Workshop on Asynchronous Many-Task Systems and Applications (WAMTA) 2024
- ☞ International Conference on Cluster Computing (CLUSTER) 2023
- ☞ International Parallel and Distributed Processing Symposium (IPDPS) 2021, 2022
- ☞ International Conference on Cluster Computing (CLUSTER) 2020
- ☞ Platform for Advanced Scientific Computing Conference (PASC) 2020
- ☞ International Workshop on Programming and Performance Visualization Tools (ProTools at SC) 2019

◇ **Poster**

- ☞ Joint Laboratory on Extreme Scale Computing Workshop (JLESC) 2020
- ☞ Platform for Advanced Scientific Computing Conference (PASC) 2020

♥ OPEN SOURCE CONTRIBUTIONS

- ◇ [**PaRSEC**]: Task-based runtime system, funded by Exascale Computing Project (ECP)
- ◇ [**DPLASMA**]: Leading implementation of a dense linear algebra package for distributed system
- ◇ [**HiCMA**]: Low-rank math library of exploiting the data sparsity of the matrix operator
- ◇ [**ExaGeostat**]: Parallel high performance unified framework for computational geostatistics

MEDIA COVERAGE

- ◇ 2024 Gordon Bell Prizes Awarded at SC24[ACM]
- ◇ Gordon Bell Prize for Climate Modelling Goes to Team for Exascale Emulator Breakthrough[HPCwire]
- ◇ Recipients of Prestigious Climate Modelling Prize Developed a Technique to Provide More Accurate and Detailed Climate Change Predictions[ACM]
- ◇ ACM Gordon Bell Prize for Climate Modelling[ACM]
- ◇ ACM Presents Winners of Gordon Bell Climate Modelling Prize[insideHPC]
- ◇ ORNL's Frontier Powers KAUST-Led Genome Study for Gordon Bell Prize Nomination[HPCwire]
- ◇ Bigger, Faster, Smarter Genetics Research[ORNL]
- ◇ Frontier Users' Exascale Climate Emulator Nominated for Gordon Bell Climate Prize[ORNL]
- ◇ Researchers Benchmark Nvidia's GH200 Supercomputing Chips[HPCwire]
- ◇ Gordon Bell Prize Finalists Develop Method for More Efficient Computing[AAAS Eurekalert][HLRS News]
- ◇ Inside the Gordon Bell Prize Finalist Projects[HPCwire]
- ◇ SC22 Unveils ACM Gordon Bell Prize Finalists[HPCwire]
- ◇ 2022 ACM Gordon Bell Prize Finalists Announced[Communications of the ACM]
- ◇ What's New in HPC Research: EXA2PRO, DQRA, and HiCMA-PaRSE Frameworks & More[HPCwire]

◇ KAUST Leverages Mixed Precision for Geospatial Data[HPCwire]

◇ Mixing Precision for Model Acceleration[Tech Xplore]

◇ Mixing It Up: Saudi Researchers Accelerate Environmental Models with Mixed Precision[Nvidia]

◇ 「富岳」を用いた3つの研究成果がゴードン・ベル賞ファイナリストに選出されました[RIKEN News]

Last updated: March 11, 2025