Sarat Sreepathi

sarat@sarats.com +91 80744.32976
Hyderabad, India

sarats.com
linkedin.com/in/sarats

Summary

Sarat Sreepathi is a Computer Scientist with 20 years of experience working on the world's fastest supercomputers.
As Performance Coordinator for the Energy Exascale Earth System Model (E3SM) project (U.S. Department of Energy's flagship climate modeling effort - $30M/year), he led an interdisciplinary team of scientists across national laboratories to focus on computational performance and application readiness on exascale (10^{18} numerical operations per second) supercomputers.

- Initiated and coordinated effort that won the inaugural Gordon Bell Prize for Climate Modelling from the Association for Computing Machinery (ACM), the most prestigious team award in Supercomputing.
- Formulated and executed GPU strategy for large teams (100+ scientists) in multiple domains (climate, nuclear fusion) over the past decade across various GPU vendors and architecture generations (NVIDIA K20X/V100/A100, AMD MI100/MI250X, Intel ATS/PVC). Led teams on GPU code development and performance optimization.
- Coordinated co-design activities with scientific teams, system software developers and hardware vendors (Cray/HPE, Intel, AMD, NVIDIA, IBM) as part of U.S. Exascale Computing project.
- Demonstrated track record in strategic planning, securing research funding, and supercomputer allocations through proposal writing.
- Experience in managing people across organizational boundaries. Proven ability to bridge disciplines and motivating teams including members outside direct chain of command to successfully execute strategic high-visibility efforts.
- Extensive experience in the design and development of efficient parallel scientific applications and scalable algorithms on multiple generations of leadership class supercomputers, including on Titan ($\sim 18k$ NVIDIA Kepler K20X GPUs), Summit ($\sim 27k$ NVIDIA V100 GPUs), Perlmutter ($\sim 7k$ NVIDIA A100 GPUs), Frontier ($\sim 75k$ AMD MI250X GPUs) and Aurora ($\sim 63k$ Intel PVC GPUs), presently the second and third fastest supercomputers respectively.

Employment History

Jun 2013 – Aug 2025	Computer Scientist, Oak Ridge National Laboratory, USA. (Largest national laboratory for open science in the U.S.) Computational Earth Sciences, Computational Sciences and Engineering Division (2021 - 2025). Future Technologies Group, Computer Science and Mathematics Division (2013-2020).
Oct 2014 – Oct 2017	Adjunct Assistant Professor, CCEE Department, North Carolina State University, USA.

Employment History (continued)

Jan 2013 – Jun 2013	Postdoctoral Research Associate, North Carolina State University, USA.
Jan 2007 – Dec 2012	Research Assistant, North Carolina State University, USA.
Feb 2009 – Jul 2009	Research Fellow, Blue Brain Project, École Polytechnique Fédérale de Lausanne, Switzerland.
Aug 2005 – Sep 2005	Research Intern, Oak Ridge National Laboratory, USA.
May 2005 – Aug 2005	Intern, Microsoft, USA.

Leadership Activities

2024 – 2025	Member, U.S. Interagency Council on Advancing Meteorological Services (ICAMS) High Performance Computing (HPC) team (White House Initiative)
2024 – 2028	High Performance Computing Lead, Working Group on Numerical Experimentation (WGNE) within World Climate Research Programme (WCRP)/World Meteorological Organization (WMO).
2022 – 2025	Member, SciDAC (Scientific Discovery through Advanced Computing) Coordination Committee
2020 – 2024	Member, National Energy Research Scientific Computing Center (NERSC) User Group Executive Committee
2020 – 2021	Chair, Oak Ridge Leadership Computing Facility User Group Executive Board

Education

2007 – 2012	**Ph.D., Computer Science** North Carolina State University, Raleigh, USA. Thesis: *Optimus: A Scalable Parallel Metaheuristic Optimization Framework With Environmental Engineering Applications.*
2004 – 2006	**M.S., Computer Science** North Carolina State University, Raleigh, USA. Thesis: *Cyberinfrastructure for Contamination Source Characterization in Water Distribution Systems.*

Awards

2023	**Inaugural Gordon Bell Prize for Climate Modelling, Association for Computing Machinery (ACM)**, Team award for "The Simple Cloud-Resolving E3SM Atmosphere Model Running on the Frontier Exascale System".
2021	**Senior member** of Association for Computing Machinery (ACM) and Institute of Electrical and Electronics Engineers (IEEE).
2015	**Significant Event Award, ORNL**, Energy Exascale Earth System Model (`E3SM`) v1.0 open source model release.
2012	**Gold Medal**, ACM Graduate Student Research Competition Supercomputing 2012 conference.

Recent Invited Talks

Sept 2025
- "Reflections on the Road to Exascale" at Center for Development of Advanced Computing (CDAC), Pune, India.
- "Climate Modeling in the Exascale Era" at Indian Institue of Tropical Meteorology (IITM), Pune, India.
- "Reflections on the Road to Exascale" at Indian Institute of Technology (IIT) Bombay, India.

Nov 2025
- "GPU Programming models for Weather and Climate Modeling", World Climate Research Programme meeting (WGNE40), China Meteorological Administration, Beijing, China.

Key Projects

- **Energy Exascale Earth System Model** (E3SM)
  E3SM is the U.S. Departent of Energy's state of the art coupled earth system model that combines detailed simulation of the earth's atmosphere, ocean, ice, land and rivers to predict changes in the coming decades to inform energy strategies, climate change planning and mitigation efforts. E3SM is designed to target ultra high-resolution scales and deliver exceptional performance on the world's fastest supercomputers.

  I am the Performance Coordinator/Lead for the project, managing a group of 15 computational and domain scientists spread across 8 national labs to focus on application readiness, model development, and computational performance on exascale supercomputers. I have been deeply involved in the strategic planning of the project and led execution of our exascale strategy. We have developed several model components targeting GPUs using Kokkos and OpenACC/OpenMP offload. Furthermore, we have optimized for efficient GPU utilization through coalesced memory accesses, kernel size (fission and fusion as appropriate), GPU-GPU communication etc.

  Our team has developed a global cloud resolving model operating at 3 km resolution that obtained a record setting performance of 1.26 simulated years per day (simulated more than an year's worth of climate using a single day of computation) on Frontier, the first exascale supercomputer. This work has been awarded the Association for Computing Machinery's 2023 Gordon Bell Special Prize for Climate Modeling.

  Additionally, I developed a comprehensive framework called Performance Analytics for Computational Experiments (PACE) to capture model provenance and performance data to inform performance optimization efforts.

- **Exascale Computing Project**
  The Exascale Computing Project (ECP) was the largest software R&D project ($1.8B) undertaken by the US Department of Energy (2016–2024) with the mission to accelerate delivery of a capable exascale computing ecosystem. As a Senior researcher with the climate and nuclear fusion sub-projects, I initiated pathfinding efforts on programming models for GPU architectures and worked on the development of next generation models. We have developed multiple versions of a cloud model using various GPU programming models starting with directives (OpenACC, OpenMP) and rewrote in C++ using template based approaches (YAKL, Kokkos) for performance portability. Performance was evaluated on multiple GPU architectures spanning pre-exascale (V100, A100) and exascale (MI250X, Intel PVC) machines.

  I have led co-design activities in collaboration with hardware vendors (communicating workload characteristics, evaluating prototype hardware etc.), system software teams (compiler features, evaluation, bugfixes) and standards bodies (programming model specifications, benchmark development). Furthermore, I have helped define and achieve the success metrics (figure of merit) that essentially amounted to a 50x improvement in model performance on exascale supercomputers.

- **SciDAC (Scientific Discovery Through Advanced Computing)**
  Topology-aware communication optimization for large scientific applications: Developed techniques for mapping processes to processor cores at runtime to minimize communication overhead by taking advantage of network topology.
- **Optimization Methods for Universal Simulators (Optimus)**
  Designed and developed Optimus, a parallel metaheuristic optimization framework that effectively scaled to a quarter of a million cores on the Jaguar supercomputer at ORNL. Designed a novel swarm intelligence technique, TAPSO (Topology Aware Particle Swarm Optimization) for network based optimization problems with applications to Water Distribution Systems problems. This was my dissertation research that was awarded the ACM SRC Gold Medal at Supercomputing 2012.
- **Blue Brain Project**
  Designed scalable multi-objective optimization algorithms for neuron modeling on the BlueGene/L supercomputer. This project had since then evolved into the Human Brain Project that strove to reverse-engineer the human brain through detailed simulations on supercomputers.

Technical Skills

Programming Languages	C, C++, Fortran, Python
Parallel Programming	MPI, OpenACC, OpenMP, Kokkos, CUDA, HIP, SYCL/DPC++
Supercomputing Architectures	Nvidia Tesla/V100/A100, AMD MI250x and Intel PVC GPUs, Fujitsu A64FX, Intel KNL/Xeon Phi, Cray XT4/XK6, IBM BlueGene/L/Q, Intel Itanium, Cray X1E vector supercomputers.

Selected Funding

- **Co-PI**, Energy Exascale Earth System Model (E3SM) - $30M/year, U.S. Department of Energy (DOE). Phase-3 (April 2023 – Present), Phase-2 (July 2018 – July 2021).
- **Co-PI**, Scientific Discovery Through Advanced Computing - ImPACTS project, $260k/year, DOE Advanced Scientific Computing Research (ASCR).
- **Senior Personnel**, Exascale Computing Project (ECP) (2017-2024) E3SM-MMF (Multiscale Modeling Framework): $2.5M/year, XGC (Nuclear Fusion): $2.3M/year, DOE Office Of Science.
- **Co-PI**, Cyber-Enabled Water and Energy Systems Sustainability Utilizing Climate Information (2014-2017) - $1.2M, National Science Foundation (NSF).
- **Co-PI**, DOE Innovative and Novel Computational Impact on Theory and Experiment (INCITE) supercomputer allocations for nuclear fusion (2021-2023) and climate (2018-2025) projects.

Selected Synergistic Activities

- Technical reviewer, Parallel Computing Journal, IEEE Transactions on Evolutionary Computing, IEEE Transactions on Parallel and Distributed Systems, Journal of Parallel and Distributed Computing, Journal of Water and Climate Change, Geoscientific Model Development Journal, Journal of Open Source Software.
- Panel Member, Computing and Computational Sciences Directorate Advisory Committee (ORNL) Panel on Accelerated Node Architectures, 2021.
- Primary Convener and Session Chair of AGU Fall Meeting 2021 session(s) on “Accelerating Earth System Predictability: Advances in High Performance Computing, Numerical modeling, Artificial Intelligence and Machine Learning”

- Program Committee, International Conference of Parallel Processing (ICPP), 2019.
- Program Committee, Supercomputing Conference, 2017.
- Reviewer, DOE Small Business Innovation Research (SBIR) Phase-I and Phase-II proposals, 2014.

Selected Publications

https://scholar.google.com/citations?hl=en&user=9WaY8iAAAAAJ&sortby=pubdate

- Carmin, J., Elbert, O., Giraldo, F., Govett, M., Harris, L., Hauser, T., McCarren, D., Mouallem, J., Olsen, M., Ringler, T., **Sreepathi**, **S.**, & Taylor, M. (2025). *Assessment of ESM Readiness Level for Exascale HPC, Report from the Interagency Council for Advancing Meteorological Services Implementation Team - High Performance Computing* (tech. rep.). Sandia National Laboratories (SNL-NM), Albuquerque, NM (United States). https://www.osti.gov/servlets/purl/2566651
- Donahue, A. S., Caldwell, P. M., Bertagna, L., Beydoun, H., Bogenschutz, P. A., Bradley, A. M., Clevenger, T. C., Foucar, J., Golaz, C., Guba, O., Hannah, W., Hillman, B. R., Johnson, J. N., Keen, N., Lin, W., Singh, B., Sreepathi, S., Taylor, M. A., Tian, J., … Zhang, Y. (2024). To Exascale and Beyond—The Simple Cloud-Resolving E3SM Atmosphere Model (SCREAM), a Performance Portable Global Atmosphere Model for Cloud-Resolving Scales **(Journal Cover/Highlight)**. *Journal of Advances in Modeling Earth Systems*, *16*(7), e2024MS004314. https://doi.org/https://doi.org/10.1029/2024MS004314
- Ikuyajolu, O. J., Van Roekel, L., Brus, S. R., Thomas, E. E., Deng, Y., & **Sreepathi**, **S.** (2023). Porting the wavewatch iii (v6.07) wave action source terms to GPU **(Highlight Paper)**. *Geoscientific Model Development*, *16*(4), 1445–1458. https://doi.org/10.5194/gmd-16-1445-2023
- Malaya, N., Messer, B., Glenski, J., Georgiadou, A., Lietz, J., Gottiparthi, K., Day, M., Chen, J., Rood, J., Esclapez, L., White III, J., Jansen, G. R., Curtis, N., Nichols, S., Kurzak, J., Chalmers, N., Freitag, C., Bauman, P., Fanfarillo, A., … Joubert, W. (2023). Experiences readying applications for exascale. *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. https://doi.org/10.1145/3581784.3607065
- Taylor, M., Caldwell, P. M., Bertagna, L., Clevenger, C., Donahue, A., Foucar, J., Guba, O., Hillman, B., Keen, N., Krishna, J., Norman, M., **Sreepathi**, **S.**, Terai, C., White, J. B., Salinger, A. G., McCoy, R. B., Leung, L.-y. R., Bader, D. C., & Wu, D. (2023). The Simple Cloud-Resolving E3SM Atmosphere Model Running on the Frontier Exascale System **(ACM Gordon Bell Prize for Climate Modeling Winner)**. *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. https://doi.org/10.1145/3581784.3627044
- Evans, T. M., Norman, M. R., Bader, D. C., Eldred, C., Hannah, W. M., Hillman, B. R., Jones, C. R., Lee, J. M., Leung, L., Lyngaas, I., Pressel, K. G., **Sreepathi**, **S.**, Taylor, M. A., & Yuan, X. (2022). Unprecedented cloud resolution in a GPU-enabled full-physics atmospheric climate simulation on OLCF's summit supercomputer. *Int. J. High Perform. Comput. Appl.*, *36*(1), 93–105. https://doi.org/10.1177/10943420211027539
- **Sreepathi**, **S.**, & Taylor, M. (2021). Early Evaluation of Fugaku A64FX Architecture Using Climate Workloads. *2021 IEEE International Conference on Cluster Computing (CLUSTER)*, 719–727. https://doi.org/10.1109/Cluster48925.2021.00107
- **Sreepathi**, **S.**, Kumar, J., Mills, R. T., Hoffman, F. M., Sripathi, V., & Hargrove, W. W. (2017). Parallel multivariate spatio-temporal clustering of large ecological datasets on hybrid supercomputers. *2017 IEEE International Conference on Cluster Computing (CLUSTER)*, 267–277.
- **Sreepathi**, **S.**, D'Azevedo, E., Philip, B., & Worley, P. (2016). Communication characterization and optimization of applications using topology-aware task mapping on large supercomputers. *Proceedings of the 7th ACM/SPEC on International Conference on Performance Engineering*, 225–236. https://doi.org/10.1145/2851553.2851575