Curriculum Vitae

Sathish S. Vadhiyar

Professor, Department of Computational and Data Sciences (CDS)
Chair, Supercomputer Education and Research Centre (SERC)
Indian Institute of Science (IISc)
Bangalore- 560012
India.
Phone: +91-80-22932615 **Email:** vss@iisc.ac.in
URL: `http://cds.iisc.ac.in/faculty/vss`

1 Research Interests and Areas

My research interests are in high performance computing (HPC) in general, and building middleware and runtime systems for parallel applications and systems, in particular.

- **HPC Runtime Systems / Application Frameworks**: Our lab works on building runtime systems for HPC applications on both accelerator and general HPC systems. We primarily focus on irregular applications including graph applications, N-Body simulations, Molecular Dynamics (MD), and Adaptive Mesh Refinement (AMR) applications. We have also worked with applications in climate science and visualization in collaboration with researchers working in these areas.
  - Accelerator systems: Techniques for load balancing, hybrid executions utilizing both CPU and GPU cores, and efficient executions of different programming models on these systems.
  - Large-scale systems: Scalability studies, tools and techniques, processor allocation, mapping and remapping strategies on HPC network topologies.
- **Middleware**:
  - Job Scheduling: Our work is on analyzing job submissions, developing prediction strategies and scheduling techniques that use predictions for efficient job management and utilization of production supercomputer systems.
  - Fault Tolerance: For exascale applications using checkpointing and process replication.

Besides these, we are also eager to learn about and providing solutions for efficient executions, good speedups, and in general remarkable reduction in turnaround times of large time-consuming Industrial applications.

2 Education

- **Ph.D:** May 2003, Department of Computer Science, University of Tennessee, U.S.A.
  Dissertation: *A Preemption-Based Meta-Scheduling System for Distributed Computing*
  Advisor: Dr. Jack Dongarra
- **M.S.:** May 1999, Department of Computer Science, Clemson University, South Carolina, U.S.A.
  Project: *Static Assignment of Multithreaded Systems*
  Advisor: Dr. Harold Grossman
- **B.E.:** May 1997, Computer Science and Engineering, Thiagarajar College of Engineering, Madurai, India
  Project: *Anti-missile simulation software based on neural networks*

3 Experience

- **March 2021 - present:** Professor, Department of Computational and Data Sciences, Indian Institute of Science, Bangalore, India.
- **November 2009 - March 2021:** Associate Professor, Department of Computational and Data Sciences (erstwhile Supercomputer Education and Research Centre till December 2015), Indian Institute of Science, Bangalore, India.
- **January 2018 - present:** Chair, Supercomputer Education and Research Centre, Indian Institute of Science, Bangalore, India.
- **June 2013:** JT Oden Faculty Fellow Lab of Prof. Keshav Pingali, The Institute for Computational and Engineering Sciences (ICES), University of Texas, Austin, USA
- **July - August 2010:** Visitor, Parallel Programming Lab (Lab of Prof. Laxmikant Kale), Computer Science Department, University of Illinois, Urbana-Champaign, USA
- **November 2003 - November 2009:** Assistant Professor, Supercomputer Education and Research Centre, Indian Institute of Science, Bangalore, India.
- **May 2003 - August 2003:** Senior Research Associate, Innovative Computing Laboratory, Department of Computer Science, Knoxville, Tennessee, U.S.A.
- **May 1999 - May 2003:** Graduate Research Assistant, Innovative Computing Laboratory, Department of Computer Science, Knoxville, Tennessee, U.S.A.

4 Honors, Awards

- Best paper award in IEEE International Conference on High Performance Computing, Data, and Analytics (HiPC), 2023.
- Outstanding Reviewer Status for Journal of Parallel and Distributed Computing, May 2015.
- NVIDIA Innovation Award. For the project "GPU-enabled Efficient Executions of Radiation Calculations on GPUs.". 2013.

- Best student paper nominee for our student's paper in IEEE/ACM SC 2012.
- Yahoo! Faculty Research Award, 2011.
- Indian National Academy of Engineering (INAE) Young Engineer Award, 2009.
- University of Tennessee citation for Extraordinary Professional Promise, 2003.

5 Publications

5.1 Papers in Refereed Journals

1. Anwesha Bhowmick, Sathish Vadhiyar, Varun PV. Scalable Multi-node Multi-GPU Louvain Community Detection Algorithm for Heterogeneous Architectures. *Concurrency and Computation: Practice and Experience*, Volume 34(17), 2022.

2. Manasi Tiwari, Sathish Vadhiyar. Pipelined Preconditioned Conjugate Gradient Methods for Real and Complex Linear Systems for Distributed Memory Architectures. *Journal of Parallel and Distributed Computing*, Volume 163, Pages 147–155, 2022.

3. Rintu Panja, Sathish S. Vadhiyar. HyPar:A Divide-and-conquer Model for Hybrid CPUGPU Graph Processing. *Journal of Parallel and Distributed Computing*, Volume 132, Pages 8–20, 2019.

4. Prakash Murali, Sathish Vadhiyar. Metascheduling of HPC Jobs in Day-Ahead Electricity Markets. *IEEE Transactions on Parallel and Distributed Systems.* Volume 29(3), March 2018.

5. Amlesh Kashyap, Sathish Vadhiyar, PN Vinayachandran, Ravi Nanjundiah. Asynchronous and Synchronous Models of Executions on Intel® Xeon Phi™ Coprocessor Systems for High Performance of Long Wave Radiation Calculations in Atmosphere models. *Journal of Parallel and Distributed Computing*, Vol 102, pages 199-212, 2017.

6. Prakash Murali, Sathish Vadhiyar. Qespera: An Adaptive Framework for Prediction of Queue Waiting Times in Supercomputer Systems. *Concurrency and Computation: Practice and Experience*, Vol 28/9, pp 2685-2710, June 2016.

7. Cijo George, Sathish Vadhiyar. Fault Tolerance on Large Scale Systems using Adaptive Process Replication. *IEEE Transactions on Computers*, Vol 64/8, pp 2213-2225, August 2015.

8. Hari Raghavan, Sathish Vadhiyar. Adaptive Executions of Hyperbolic Block-Structured AMR Applications on GPU Systems. *International Journal of High Performance Computing Applications (IJHPCA)*, Vol 29/2, pp 135-153, 2015.

9. Hari Raghavan, Sathish Vadhiyar. Efficient Asynchronous Executions of AMR Computations and Visualization on a GPU System. *Journal of Parallel and Distributed Computing (JPDC)*, Vol 73/6, pp 866-875, 2013.

10. Sivagama Sundari M., Sathish Vadhiyar, Ravi Nanjundiah. Large Improvements in Application Throughput of Long-running Multi-Component Applications using Batch Grids. *Concurrency and Computation: Practice & Experience*, Vol 24/15, pp 1775-1791, 2012.

11. Sivagama Sundari M., Sathish Vadhiyar, Ravi Nanjundiah. Adaptive Executions of Multi-Physics Coupled Applications on Grids. *Journal of Grid Computing*, Vol 9/4, pp 455-478, 2011.

12. H.A. Sanjay and Sathish Vadhiyar. Strategies for Rescheduling Tightly-Coupled Parallel Applications in Multi-Cluster Grids. *Journal of Grid Computing*, Vol 9/3, pp 379-403, 2011.

13. Sivagama Sundari M., Sathish Vadhiyar, Ravi Nanjundiah. Grids with Multiple Batch Systems for Performance Enhancement of Multi-Component and Parameter Sweep Parallel Applications. *Future Generation Computer Systems*, Vol 26/2, pp 217-227, 2010.

14. H.A. Sanjay and Sathish Vadhiyar. A Strategy for Scheduling Tightly-Coupled Parallel Applications on Clusters. *Concurrency and Computation: Practice & Experience*, Vol 21/18, pp 2491-2517, 2009.

15. Sivagama Sundari M., Sathish Vadhiyar, Ravi Nanjundiah. Dynamic Component Extension: a Strategy for Performance Improvement in Multicomponent Applications. *International Journal of High Performance Computing Applications*, Vol 23/1 pp 84-98, 2009.

16. Yadnyesh Joshi and Sathish Vadhiyar. Analysis of DNA Sequence Transformations on Grids. *Journal of Parallel and Distributed Computing*, Vol 69/1, pp 80-90, 2009.

17. H.A. Sanjay and Sathish Vadhiyar. Performance Modeling of Parallel Applications for Grid Scheduling. *Journal of Parallel and Distributed Computing*, vol 68/8 pp 1135-1145, 2008.

18. Sandip Tikar and Sathish Vadhiyar. Efficient Reuse of Replicated Parallel Data Segments in Computational Grids. *Future Generation Computer Systems*, vol 24/7 pp 644-657, 2008.

19. J. Dongarra, G. Bosilca, Z. Chen, V. Eijkhout, G. Fagg, E. Fuentes, J. Langou, P. Luszczek, J. Pjesivac-Grbovic, K. Seymour, H. You, and Sathish Vadhiyar. Self-adapting Numerical Software (SANS) Effort. *IBM Journal of System Research and Development. Exploratory Systems Research*, volume 50, Number 2/3, pp. 223-238, 2006.

20. Sathish Vadhiyar and J. Dongarra. Self adaptivity in Grid computing. *Concurrency and Computation: Practice and Experience*, volume 17, Issue 2-4, pp. 235-257, Special Issue: Grid Performance . Issue Edited by John Gurd, Tony Hey, Juri Papay, Graham Riley. Copyright ©. John Wiley & Sons, Ltd. Feb 7, 2005.

21. Sathish Vadhiyar and Jack Dongarra. GrADSolve - A Grid-based RPC system for Parallel Computing with Application-level Scheduling. *Journal of Parallel and Distributed Computing*, volume 64, pp. 774-783, 2004.

22. Sathish Vadhiyar, Graham Fagg and Jack Dongarra. Toward an Accurate Model for Collective Communications. *International Journal of High Performance Computing Applications*, volume 18, Number 1, Spring 2004, pp 159-166, ISSN 1094-3420.

23. Sathish Vadhiyar and Jack Dongarra. SRS - A Framework for Developing Malleable and Migratable Parallel Applications for Distributed Systems. *Parallel Processing Letters*, vol. 13, number 2, pp. 291-312, June 2003.

24. M. Beck and D. Arnold and A. Bassi and F. Berman and H. Casanova and J. Dongarra and T. Moore and G. Obertelli and J. Plank and M. Swany and S. Vadhiyar and R. Wolski. Middleware for the use of Storage in Communication. *Parallel Computing*, Volume 28 , Issue 12, pages 1773 - 1787, 2002.

25. A. Petitet, S. Blackford, J. Dongarra, B. Ellis, G. Fagg, K. Roche and Sathish Vadhiyar. Numerical Libraries and The Grid: The GrADS Experiments with ScaLAPACK. *Journal of High Performance Applications and Supercomputing*, Vol. 15, number 4, pp. 359-374, Winter 2001.

5.2 Papers at Refereed Conferences and Workshops

26. Koushik Sen, Sathish Vadhiyar, PN Vinayachandran. Strategies for Fast I/O Throughput in Large-scale Climate Modeling Applications. In the proceedings of *IEEE International Conference on High Performance Computing, Data, and Analytics (HiPC) 2023*, Goa, India, December 2023. **Best paper award.**

27. Anwesh Panda, Sathish Vadhiyar. Dynamic Strategies for High Performance Training of Knowledge Graph Embeddings. In the proceedings of *International Conference on Parallel Processing (ICPP)*, pages 38:1-38:10, Virtual, August/September 2022.

28. Anoop Kumar, Vibha Balaji, MA Chandrashekar, Ambedkar Dukkipati, Sathish Vadhiyar. Graph Convolutional Neural Networks for Alzheimer's Classification with Transfer Learning and HPC Methods. In the proceedings of the IPDPS Workshops on HiCOMB (High Performance Computational Biology), Virtual, May 2022.

29. Manasi Tiwari, Sathish Vadhiyar. Strategies for Efficient Execution of Pipelined Conjugate Gradient Method on GPU Systems. In the proceedings of the ISC Workshops on HPC on Heterogeneous Hardware, Hamburg, Germany, May/June 2022.

30. Manasi Tiwari, Sathish Vadhiyar. Pipelined Preconditioned s-step Conjugate Gradient Methods for Distributed Memory Systems. In the proceedings of *IEEE Cluster 2021*, Virtual, September 2021.

31. Manasi Tiwari, Sathish Vadhiyar. Pipelined Preconditioned Conjugate Gradient Methods for Distributed Memory Systems. In the proceedings of *IEEE International Conference on High Performance Computing, Data, and Analytics (HiPC) 2020*, pages 151-160, 2020.

32. Preeti Malakar, Vijay Natarajan, Sathish S. Vadhiyar. Adaptive and Efficient Transfer for Online Remote Visualization of Critical Weather Applications. In the proceedings of *International Conference on Computational Science (ICCS)*, pages 674-693, 2020.

33. Anwesha Bhowmik, Sathish Vadhiyar. HyDetect: A Hybrid CPU-GPU Algorithm for Community Detection. *IEEE International Conference on High Performance Computing, Data, and Analytics (HiPC)*, pages 2-11, 2019.

34. Udit Gupta, Sathish Vadhiyar. Fast and Accurate Learning of Knowledge Graph Embeddings at Scale. *IEEE International Conference on High Performance Computing, Data, and Analytics (HiPC)*, pages 173-182, 2019.

35. Rintu Panja, Sathish Vadhiyar. A Multi-Node Multi-Device Parallel Boruvka's MST Algorithm. In the proceedings of the *International Conference on Parallel Processing (ICPP)*, Eugene, USA, August 2018.

36. Srinivasan Ramesh, Sathish Vadhiyar, P. N. Vinayachandran, Ravi S. Nanjundiah. Deep and Shallow Convection Calculations in Atmosphere Models on Intel Xeon Phi Coprocessor Systems. In the proceedings of the *High Performance Computing Commmunications (HPCC)*, December 2017, Bangkok, Thailand. Acceptance Rate: 36

37. Vineetha Kondameedi, Sathish Vadhiyar. Adaptive Hybrid Queue Configuration for Supercomputer Systems. In the proceedings of the 17th *IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid)*, May 2017, Madrid, Spain.

38. Aketh TM, Sathish Vadhiyar, PN Vinayachandran, Ravi Nanjundiah. High Performance Horizontal Diffusion Calculations in Ocean Models on Intel® Xeon Phi™ Coprocessor Systems. In the Proceedings of the *International Conference on High Performance Computing (HiPC)*, December 2016, Hyderabad, India.

39. Prakash Murali, Sathish Vadhiyar. Metascheduling of HPC Jobs in Day-Ahead Electricity Markets. In the Proceedings of the *International Conference on High Performance Computing (HiPC)*, December 2015, Bangalore, India.

40. Anirudh Jayakumar, Prakash Murali, Sathish Vadhiyar. Matching Application Signatures for Performance Predictions using a Single Execution. In the Proceedings of the *IEEE International Parallel and Distributed Processing Symposium (IPDPS)* 2015, Hyderabad, India.

41. Rajath Kumar, Sathish Vadhiyar. Prediction of Queue Waiting Times for Metascheduling on Parallel Batch Systems. In the proceedings of the *Workshop on Job Scheduling Strategies for Parallel Processing (JSSPP)*, May 2014, Phoenix, USA, appeared in Lecture Notes in Computer Science (LNCS), Volume 8828, pages 108-128, 2015.

42. Anurag Murty, Vijay Natarajan, Sathish Vadhiyar. Efficient Homology Computations on Multicore and Manycore Systems. In the proceedings of the *International Conference on High Performance Computing (HiPC)*, December 2013, Bangalore, India.

43. Sai Kiran Korwar, Sathish Vadhiyar, Ravi Nanjundiah. GPU-enabled Efficient Executions of Radiation Calculations in Climate Modeling. In the proceedings of the *International Conference on High Performance Computing (HiPC)*, December 2013, Bangalore, India.

44. Preeti Malakar, Vijay Natarajan, Sathish Vadhiyar, Ravi Nanjundiah. A Diffusion-Based Processor Reallocation Strategy for Tracking Multiple Dynamically Varying Weather Phenomena. In the proceedings of the *International Conference on Parallel Processing (ICPP)*, October 2013, Lyon, France.

45. R. Vasudevan, Sathish Vadhiyar, Laxmikant Kale. G-Charm: An Adaptive Runtime System for Message-Driven Parallel Applications on Hybrid Systems. In the proceedings of *International Conference on Supercomputing (ICS)*, pp 349-358, June 2013, Eugene, Oregon, USA.

46. Preeti Malakar, Thomas George, Sameer Kumar, Rashmi Mittal, Vijay Natarajan, Yogish Sabharwal, Vaibhav Saxena, Sathish Vadhiyar. A Divide and Conquer Strategy for Scaling Weather Simulations with Multiple Regions of Interest. In the proceedings of *IEEE/ACM Supercomputing conference, SC 2012*, November 2012, Salt Lake City, Utah, USA. **Best student paper nominee**.

47. Cijo George, Sathish Vadhiyar. AdFT:An Adaptive Framework for Fault Tolerance on Large Scale Systems using Application Malleability. In the proceedings of *International Conference on Computational Science (ICCS)*, June 2012, Omaha, Nebraska, USA, pp 166-175.

48. Rajath Kumar, Sathish Vadhiyar. Identifying Quick Starters: Towards an Integrated Framework for Efficient Predictions of Queue Waiting Times of Batch Parallel Jobs. In the Workshop on *Job Scheduling Strategies for Parallel Processing (JSSPP)*, May 2012, Shanghai, China.

49. Preeti Malakar, Vijay Natarajan, Sathish Vadhiyar. InSt: An Integrated Steering Framework for Critical Weather Applications. In the proceedings of *International Conference on Computational Science (ICCS)*, June 2011, Singapore, pp 116-125.

50. Preeti Malakar, Vijay Natarajan, Sathish Vadhiyar. An Adaptive Framework for Simulation and Online Remote Visualization of Critical Climate Applications in Resource-constrained Environments. In the proceedings of *IEEE/ACM Supercomputing conference, SC 2010*, November 2010, New Orleans, USA.

51. Priyank Raj Khatariya and Sathish Vadhiyar. Phylogenetic Predictions on Grids. In *Proceedings of the 5th IEEE International Conference of e-Science*, Oxford, UK, December 2009.

52. Rakhi Gupta and Sathish Vadhiyar. An Efficient MPI_Allgather for Grids. *In proceedings of High Performance Distributed Computing (HPDC)*, pp 169-178, Monterey, California, USA, June 2007.

53. Jay Yagnik, H. A. Sanjay and Sathish Vadhiyar. Performance Modeling based on Multidimensional Surface Learning for Performance Predictions of Parallel Applications in Non-Dedicated Environments. *In proceedings of 35th International Conference on Parallel Processing (ICPP)*, pp 513-520, Columbus, Ohio, USA, August 2006.

54. Rakhi Gupta and Sathish Vadhiyar. Application-Oriented Adaptive MPI_Bcast for Grids. *In proceedings of 20th IEEE International Parallel & Distributed Processing Symposium (IPDPS)*, Rhodes Island, Greece, April 2006.

55. Sathish Vadhiyar, Jack Dongarra and Asim Yarkhan. GrADSolve - RPC for High Performance Computing on the Grid. *Euro-Par 2003, 9th International Euro-Par Conference*, Proceedings, Springer, LCNS 2790, p. 394-403, August 26 -29, 2003.

56. Sathish Vadhiyar and Jack Dongarra. A Performance Oriented Migration Framework for the Grid. *Proceedings of The 3rd IEEE/ACM International Symposium on Cluster Computing and the Grid (CCGrid 2003)*, pp 130-137, Tokyo, Japan, May 2003.

57. Sathish Vadhiyar and Jack Dongarra. A Metascheduler for the Grid. *Proceedings of the 11th IEEE International Symposium on High Performance Distributed Computing*, pp 343-351, Edinburgh, Scotland, July, 2002.

58. K. Kennedy, M. Mazina, J. Mellor-Crummey, K. Cooper, L. Torczon, F. Berman, A. Chien, H. Dail, O. Sievert, D. Angulo, I. Foster, D. Gannon, L. Johnsson, K. Kasselman, R. Aydt, D. Reed, J. Dongarra, Sathish Vadhiyar and R. Wolski. Toward a Framework for Preparing and Executing Adaptive Grid Programs. *Proceedings of NSF Next Generation Systems Program Workshop (International Parallel and Distributed Processing Symposium 2002)*, Fort Lauderdale, Florida, USA, 2002.

59. A. Petitet, S. Blackford, J. Dongarra, B. Ellis, G. Fagg, K. Roche and S. Vadhiyar. Numerical Libraries and The Grid: The GrADS Experiments with ScaLAPACK. *Proceedings of Supercomputing 2001 (SC)*, Denver, Colorado, 2001.

60. Sathish Vadhiyar, G. Fagg and J. Dongarra. Performance Modeling For Self Adapting Collective Communications for MPI. *Proceedings of Los Alamos Computer Science Institute (LACSI) Symposium*, Santa Fe, USA, October 2001.

61. M. Beck, D. Arnold, A. Bassi, F. Berman, H. Casanova, J. Dongarra, T. Moore, G. Obertelli, J. Plank, M. Swany, Sathish Vadhiyar, R. Wolski. Logistical Computing and Internetworking: Middleware for the Use of Storage in Communication. *Third Annual International Workshop on Active Middleware Services (AMS)*, San Francisco, CA, August, 2001.

62. Sathish Vadhiyar, Graham Fagg and Jack Dongarra. Towards an Accurate Model for Collective Communications. *In Proceedings of International Conference on Computational Science - ICCS 2001*, San Francisco, CA. Lecture Notes in Computer Science, Vol. 2073 (Berlin: Springer Verlag), pp 41-50, 2001.

63. Graham Fagg, Sathish Vadhiyar and Jack Dongarra. ACCT: Automatic Collective Communications Tuning. *Proceedings of the 7th European PVM/MPI Users' Group Meeting on Recent Advances in Parallel Virtual Machine and Message Passing Interface*, pages 354-362, Hungary, 2000.

64. Sathish Vadhiyar, Graham Fagg and Jack Dongarra. Automatically Tuned Collective Communications. *Proceedings of Supercomputing 2000*, Dallas, Texas, 2000.

65. Sathish Vadhiyar and Harold Grossman. Static Assignment of Multithreaded Systems. *ACM Southeast Regional Conference* 1999.

5.3 Book Chapters

66. H. Dail, O. Sievert, F. Berman, H. Casanova, A. YarKhan, Sathish Vadhiyar, J. Dongarra, C. Liu, L. Yang, D. Angulo, I. Foster. Scheduling in the Grid Application Development Software Project. *in Grid Resource Management: State of the art and future trends, Edited by Jarek Nabrzyski, Jennifer Schopf and Jan Weglarz, Kluwer Academic Publisher, 2003, ISBN 1-4020-7575-8.*

67. Sudesh Agrawal, Jack Dongarra, Keith Seymour, and Sathish Vadhiyar. NetSolve: Past, Present, and Future - a Look at a Grid Enabled Server. *in Grid Computing: Making the Global Infrastructure a Reality, Edited by F. Berman, G. Fox, and A. Hey, Wiley Publisher, 2003, ISBN 0-470-85319-0.*

5.4 Technical Reports

68. Sarthak Joshi, Sathish Vadhiyar. PartRePer-MPI: Combining Fault Tolerance and Performance for MPI Applications. CoRR abs/2310.16370, 2023.

69. Manasi Tiwari, Sathish Vadhiyar. Efficient executions of Pipelined Conjugate Gradient Method on Heterogeneous Architectures. CoRR abs/2105.06176 2021.

70. Swetha Hariharan, Prakash Murali, Abhishek Pasari, Sathish Vadhiyar. End-to-End Predictions-Based Resource Management Framework for Supercomputer Jobs. CoRR abs/2008.08292. 2020.

71. Ashirbad Mishra, Sathish Vadhiyar, Rupesh Nasre, Keshav Pingali. A Fine-Grained Hybrid CPU-GPU Algorithm for Betweenness Centrality Computations. CoRR abs/2008.05718. 2020.

72. Vasudevan Rengasamy, Sathish Vadhiyar. Strategies for Efficient Executions of Irregular Message-Driven Parallel Applications on GPU Systems. CoRR abs/2008.05712. 2020.

73. Ananya Raval, Rupesh Nasre, Vivek Kumar, R. Vasudevan, Sathish Vadhiyar, Keshav Pingali: Dynamic Load Balancing Strategies for Graph Applications on GPUs. CoRR abs/1711.00231 (2017).

74. K. Raghavendra, Sathish S. Vadhiyar: Determination of Checkpointing Intervals for Malleable Applications. CoRR abs/1711.00270 (2017).

75. Srinivasan Ramesh, Sathish Vadhiyar, Ravi S. Nanjundiah, P. N. Vinayachandran: Deep and Shallow convections in Atmosphere Models on Intel Xeon Phi Coprocessor Systems. CoRR abs/1711.00289 (2017).

76. Prakash Murali, Sathish Vadhiyar: Metascheduling of HPC Jobs in Day-Ahead Electricity Markets. CoRR abs/1712.10201 2017.

77. Sathish Vadhiyar, Graham Fagg and Jack Dongarra. Toward an Accurate Model for Collective Communications. *Computer Science Department, University of Tennessee, ut-cs-05-550*, 2005.

78. Typically a team of 8-10 including S Vadhiyar. User's guide to NetSolve, versions 1,1.4,1.4.1,2.0. Computer Science Department, University of Tennessee, Technical Reports.

5.5 Refereed Conference Short Papers and Posters

79. Sarthak Joshi, Sathish Vadhiyar. Combining Checkpoint/Restart and Replication for Fault Tolerance with High Performance. *Student Research Symposium poster, International Conference on High Performance Computing (HiPC)*, Bangalore, India, 2024.

80. Chaitanya Vinod Patil, Sathish Vadhiyar. Balanced and Efficient Distribution of Coarse and Nested Grids in Regional Ocean Modeling System. *Student Research Symposium poster, International Conference on High Performance Computing (HiPC)*, Bangalore, India, 2024.

81. Manasi Tiwari, Sathish Vadhiyar. Communication Overlapping Pipelined Conjugate Gradients for Distributed Memory Systems and Heterogeneous Architectures. In the proceedings of *Euro-Par Workshops*, pages 535-539, August 2021.

82. Surendra Varma Pericherla, Sathish Vadhiyar. High Performance and Enhanced Scalability for Parallel Applications using MPI-3's non-blocking Collectives. In the proceedings of the International Conference on Computational Science (ICCS), Zurich, Switzerland, 2017.

83. Rajath Prasad, Sathish Vadhiyar. Scheduling Strategies for Multi-Physics Applications. *Student Research Symposium poster, International Conference on High Performance Computing (HiPC)*, Bangalore, India, 2011.

84. Archana Venkatesh, Sathish Vadhiyar. Asynchronous Parallelism for Molecular Dynamics on GPU cores. *Student Research Symposium poster, International Conference on High Performance Computing (HiPC)*, Bangalore, India, 2011.

85. Rajath Prasad, Sathish Vadhiyar. Scheduling Strategies for Multi-Physics Applications. *Poster. IBM I-CARE*, 2011.

86. Archana Venkatesh, Sathish Vadhiyar. Asynchronous Parallelism for Molecular Dynamics on GPU cores. *Poster. IBM I-CARE*, 2011.

87. Sivagama Sundari M., Sathish Vadhiyar, Ravi Nanjundiah. Morco: Middleware Framework for Long-running Multi-component Applications on Batch Grids. Short paper. In the proceedings of the *ACM High Performance Distributed Computing (HPDC)*, pp 328-331, June 2010, Chicago, USA.

88. Preeti Malakar, Vijay Natarajan, Sathish Vadhiyar and Ravi Nanjundiah. An Integrated Simulation and Visualization Framework for Tracking Cyclone Aila. *Student Research Symposium poster, International Conference on High Performance Computing (HiPC)*, Kochi, India, 2009. **Best Paper Award**.

89. Sivagama Sundari M., Sathish Vadhiyar and Ravi Nanjundiah. Middleware for Long-Running Applications on Batch Grids. *Student Research Symposium poster, International Conference on High Performance Computing (HiPC)*, Kochi, India, 2009.

90. Preeti Malakar, Vijay Natarajan, Sathish Vadhiyar and Ravi Nanjundiah. An Integrated Simulation and Visualization Framework for Tracking Cyclone Aila. *Student poster in ATIP First Workshop on HPC in India: Indigenous Hardware, Software, and Infrastructure Research*, held in conjunction with IEEE/ACM Supercomputing (SC), Portland, Oregon, USA, 2009.

91. Sivagama Sundari M., Sathish Vadhiyar and Ravi Nanjundiah. Middleware for Long-running Applications on Batch Grids. *Student poster in ATIP First Workshop on HPC in India: Indigenous Hardware, Software, and Infrastructure Research*, held in conjunction with IEEE/ACM Supercomputing (SC), Portland, Oregon, USA, 2009.

92. Sivagama Sundari M., Sathish Vadhiyar and Ravi Nanjundiah. Executing Long-running Multi-component Applications on Batch Grids. *Student Research Symposium poster, International Conference on High Performance Computing (HiPC)*, Bangalore, India, 2008.

93. H. A. Sanjay and Sathish Vadhiyar. Performance Modeling Based Scheduling and Rescheduling of Parallel Applications on Computational Grids. *International Conference on High Performance Computing (HiPC)*, Goa, India, 2007.

94. Sivagama Sundari M., Sathish Vadhiyar and Ravi Nanjundiah. Coupled Climate Models on Grids. *International Conference on High Performance Computing (HiPC)*, Goa, India, 2007.

6 Software Developed in Our Group

- PIPECG2. A piplelined Preconditioned Conjugate Gradient (PCG) solver that reduces global synchronizations to one per two iterations. The software has been integrated into widely-used PETSc (`https://www.mcs.anl.gov/petsc/`) software for scientific computations. Available at
`https://gitlab.com/petsc/petsc/-/blob/master/src/ksp/ksp/impls/cg/pipecg2/pipecg2.c`.

- Fast_ANN. Fast Scalable Approximate Nearest Neighbor Search for High-dimensional Data. Available at
`https://github.com/renshyam/fast\_ann`.

- HyDetect. A Hybrid CPU-GPU Community Detection Package. Available at
`https://github.com/marslabiisc/HyDetect`.

- MCMF-Metascheduler-Predictors. A software for predictions of loads, electricity costs and metascheduling in HPC systems. Available at `https://github.com/MARS-CDS-IISc/mcmf-metaschedulerpredictors`.

7 Sponsored Projects

- Rs. 2,60,46,000 (US $314,290). August 2022-July 2025. Shell Technologies, India. *A Comprehensive and Integrated Framework for ML/DL Applications.*

- Rs. 1,92,93,320 (US $232,807). March 2022-March 2025. Investigator along with two other PIs. Department of Science and Technology (DST), SERB Core Research Grant. *High Performance Simulations and Visualizations for High-Resolution Climate Modeling.*

- Rs. 1,50,00,000 (US $208,335). August 2019-March 2021. Investigator, along with six other PIs. IISc Institute of Eminence (IoE) Research Grant. *An Integrated HPC Stack for Extreme-scale Computing.*

- Rs. 10,00,000 (US $13,900). Investigator, along with one other PI. July 2019-December 2020. IISc Pratiksha Trust Research Grant. *Deep Unsupervised Learning and HPC Frameworks for Large-scale Study of Brain Networks.*

- Rs. 6,00,000 (US $8,340). Siemens Technologies. May 2019-December 2020. it Neural Network Model Compression Techniques.

- Rs. 44,67,554 (US $70,915). Investigator, along with two other PIs. Intel Corp. Intel Parallel Computing Centre (IPCC) for Modeling Monsoons and Tropical Climate. June 2014 - June 2019.

- Rs. 10,17,500 (US $15,650). Centre for Development of Advanced Computing (CDAC). June 2013 - March 2014. *A Framework for Efficient Executions of Irregular Applications on Hybrid Systems.*

- A Kepler Card. NVIDIA CUDA Research Centre (CRC). January - December 2013.

- Rs. 28,36,800 (US $43,650). Department of Science and Technology. December 2012 - June 2016. *A Robust Middleware for Job management in Supercomputer systems.*

- Rs. 13,90,500 (US $24500). Centre for Development of Advanced Computing (CDAC). August 2010 - March 2012. With Dr. Vijay Natarajan, CSA, IISc and Prof. Ravi Nanjundiah, CAOS, IISc. *Middleware for Online Remote Visualization of Weather Applications.*

- Rs. 5,56,500 (US $13000). European Commission. October 2010 - September 2011. With Prof. Ravi Nanjundiah, CAOS, IISc. *Sustainable e-infrastructures across Europe and India.*

- Rs. 2,25,000 (US $2000). Yahoo! India. Consultancy project. May 2008 - July 2008. *Exploration of Research Challenges and Methods in Yahoo!'s Hadoop Infrastructure.*

- Rs. 38,30,000 (US $76,600). Ministry of Information Technology. May 2007 - May 2010. With Prof. Ravi Nanjundiah, CAOS, IISc. *Coupled climate models on grids.*

- Rs. 22,28,400 (US $44,568). Department of Science and Technology. July 2006 - Jan. 2010. *A Checkpointing infrastructure for parallel scientific applications on computational grids.*

- Rs. 8,00,000 (US $16,000). Indian Institute of Science. Part (2A) Grant (252/SERC). November 2004 - December 2005. With Prof. R. Govindarajan, SERC, IISc. *Connecting 2 clusters.*
- Rs. 15,22,593 (US $35,000). Microsoft Research. February 2005 - March 2006. *Parallel Numerical Applications as Web Services.*
- Rs. 12,00,000 (US $24,000). Indian Institute of Science. Part (2A) Special Grant (45/SERC). January 2003 - January 2004. *Development of Grid Computing.*

8 Invited Talks and Lectures

8.1 Invited Talks

1. Divide-and-Conquer Paradigm for Asynchronous CPU-GPU Computations in Graph and AI/ML Applications and NSM R&D. Keynote talk, Student Research Symposium, HiPC, December 2024.
2. Divide-and-Conquer Paradigm for Asynchronous CPU-GPU Computations in Graph and AI/ML Applications. 2nd Workshop on Present and Future Computing Systems. CSA, IISc, November 2024.
3. Pipelined Preconditioned Conjugate Gradient Methods for Distributed Memory Architectures. NSM HPC Research Week, Mathematics, IIT Madras, November 2023.
4. Divide-and-Conquer Paradigm for Asynchronous CPU-GPU Computations in Graph and AI/ML Applications. NSM HPC Research Week, Computer Science, IIT Madras, November 2023.
5. Essential Ingredients for Scalable Solutions - Asynchronous Methods, Multi-node Hybrid CPU-GPU Executions, Harnessing Domain Knowledge, Shell-IISc workshop, November 2021.
6. Parallel Algorithms for Sorting, Search and Graph Applications. FDP program in NMIT, Bangalore, September 2021.
7. Towards Preparing for the Exascale Era: Developing Bulk-Synchronous Avoiding Algorithms. IIT Indore NSM workshop, July 2021.
8. Bulk Synchronization Avoiding Algorithms in Graph Applications and Iterative Solvers. Adena workshop, October 2020.
9. Heterogeneous Computing for Graph Algorithms. IndoSys workshop, July 2019.
10. Metascheduling. Garuda NKN Meet. NIAS, September 2016.
11. Metascheduling of HPC Jobs in Day-Ahead Electricity Markets. Jawaharlal Nehru University (JNU), Delhi, November 2015.
12. Optimization Strategies for GPU Computing. National Institute of Technology (NIT), Trichy, India. January 2015.
13. Hybrid Executions. Keynote Talk, Nitte Meenakshi Institute of Technology (NMIT), Bangalore, India. August 2014.
14. Parallel Programming Basics and Areas for Research Projects. Tutorial, Indo-Sys Workshop, Bangalore India. June 2014.

15. Outreach talks for the department: RVCE, MSRIT, BMSCE, 2014.

16. GPU Enabled Efficient Executions of Radiation Calculations in Climate Modeling. NVIDIA Professors' Meet, Bangalore, India. December 2013.

17. Fault Tolerance for Exascale Systems using Adaptive Process Replication. US-India workshop on High Performance Computing, Systems and Applications, Supercomputer Education and Research Centre, Indian Institute of Science, Bangalore, India. December 2013.

18. Partitioning and Repartitioning Strategies for Weather Applications on Torus Systems. Garuda-NKN Partners Meet, Bangalore, India. July 2013.

19. Overview of Research in GARL. Prof. Pingali's group, The Institute for Computational Engineering and Sciences (ICES), University of Texas, Austin, USA, June 2013.

20. Grid Middleware Research for Parallel Computing. BMSIT, Bangalore, India. March 2013.

21. 3 Parallel Programming Models - MPI, OpenMP, CUDA. MS Ramiah Institute of Technology (MSRIT), Bangalore, India. March 2013.

22. Middleware Research in Parallel Computing. Workshop in Computational Science (WCP). Supercomputer Education and Research Centre, Indian Institute of Science, Bangalore, India. February 2013.

23. Research in Parallel Computing. Workshop in Computational Science (WCP). Supercomputer Education and Research Centre, Indian Institute of Science, Bangalore, India. February 2012.

24. A Comprehensive Framework for Real Time Visualization and Steering of Critical Weather Applications. Garuda-NKN Partners Meet. July 2011.

25. Research in Parallel Computing. Workshop in Computational Science (WCP). Supercomputer Education and Research Centre, Indian Institute of Science, Bangalore, India. February 2011.

26. Tracking and Steering Critical Weather Events on Garuda. CDAC Round-Table Meeting on Weather and Climate Modeling, Centre for Development of Advanced Computing (CDAC), Bangalore, India. February 2011.

27. Middleware Frameworks for Adaptive Executions and Visualizations of Climate and Weather Applications on Grids. Hadoop India Summit, Indian Institute of Science, Bangalore, India. February 2011.

28. Scheduling and Rescheduling in Grid. Workshop on Grid Computing, Madras Institute of Technology (MIT), Chennai, India. February 2011.

29. Coupled Climate Models on Grids. EU-India Grid2 Workshop on Climate Change, Indian Institute of Technology (IIT) Delhi, India. December 2010.

30. Middleware Framework for Online Remote Visualization of Cyclone Events using Garuda. *5th CDAC Garuda Partners Meet*, Bangalore, India. May 2010.

31. Middleware Frameworks for Climate Simulations on Multiple Parallel Systems, *IBM Software Lab, Bangalore, India.* April 2010.

32. Biology on Clouds and Climate Science on Grids, *Yahoo! India, Bangalore, India.* February 2010.

33. Grid Middleware for High Performance Computing, *Indian National Academy of Engineering (INAE) Annual Convention, Kalpakkam, India.* December 2009.

34. Grid Middleware for High Performance Computing, *ATIP First Workshop on HPC in India: Indigenous Hardware, Software, and Infrastructure Research, held in conjunction with IEEE/ACM Supercomputing (SC) conference, Portland, Oregon, USA*. November 2009.

35. Top Supercomputers-India, *Centre for Development of Advanced Computing (CDAC), Bangalore, India.* October 2009.

36. Top Supercomputers-India, *International Conference on High Performance Computing (HiPC), Bangalore, India.* December 2008.

37. Towards Uniform Map-Reduce Communications. *Yahoo!, Bangalore, India.* October 2008.

38. High Performance Computing in Grids. *Computational Research Lab (CRL) Ltd., Pune, India.* March 2008.

39. Grid Scheduling, Rescheduling and Applications for Garuda. *Third Garuda Partners Meet, Bangalore, India.* March 2008.

40. Performance Modeling of Parallel Applications for Grids. *Wipro Technologies, Bangalore, India.* January 2008.

41. GrADSolve: A Performance Oriented Grid System & Grid Activities in GARL. *Supercomputer Education and Research Centre (SERC), Indian Institute of Science (IISc), Bangalore, India.* April 2005.

42. GrADSolve Resource Management System and Grid Research Activities in SERC, IISc. *Tata Institute of Fundamental Research (TIFR), Mumbai, India.* January 2005.

43. Self Adaptive Scientific Applications in Computational Grids. *IBM, Bangalore, India.* November 2004.

44. GrADSolve - Grid Resource Management System. *Centre for Development of Advanced Computing (CDAC), Bangalore, India.* August 2004.

45. Rescheduling, Redistribution and Towards Dynamic Tuning of Grid Policies. *Indo-UK eScience Workshop, IIT, Delhi, India.* February 2004.

46. Application-Level Scheduling and Metascheduling for the Grid. *Indo-UK eScience Workshop, IIT, Delhi, India.* February 2004.

47. GrADSolve - Grid Resource Management System. *Tata Institute of Fundamental Research (TIFR), Indian Institute of Science (IISc), Bangalore, India.* February 2004.

8.2 Invited Lectures

1. Parallel Algorithms and Applications. Faculty Development Program on High Performance Computing. *IIT Mandi, India.* November 2018.

2. Parallel Algorithms and Applications. National Supercomputing Mission (NSM) Faculty Development Program. *IISc, Bangalore, India.* September 2011.

3. HPC and Parallel Computing. *Centre for Development of Advanced Computing (CDAC), Bangalore, India.* January 2011.

4. Parallel and Distributed Computing. *Workshop on Large Scale Computing for SAP Research Group, Bangalore in Indian Institute of Science, India.* September 2010.

5. Grid Computing. Lecture. *Vellore Institute of Technology, India.* July 2009.

6. Parallel Programming. Lecture. *Vellore Institute of Technology, India.* July 2009.

7. Writing Parallel Programs. Lecture. *Supercomputer Education and Research Centre (SERC), Indian Institute of Science (IISc), Bangalore, India.* February 2009.

8. Parallel Programming for Distributed Memory Machines. Lecture. *Atria Institute of Technology, Bangalore, India.* August 2007.

9 Professional Activities

9.1 Editorial Board and Membership in Professional Bodies

- Associate Editor, IEEE Transactions on Parallel and Distributed Systems. April 2014-April 2018. Extended for a second 2-year term in 2016 in recognition of "your ongoing hard work and dedication to the journal".

- Chair, ACM Senior Member Award Committee. 2018.

- Member, ACM Senior Member Award Committee. March 2015-December 2017.

- Senior Member, IEEE.

- Member, ACM.

9.2 Chair for Conferences

- Session Chair, Data Science - Scalable Algorithms and Analytics, International Conference on High Performance Computing, HiPC, 2023.

- Program Co-Chair, HPC, International Conference on High Performance Computing, HiPC, 2022.

- Vice Chair - System Software track, International Conference on High Performance Computing, HiPC, 2021.

- Session Chair, Applications, High Performance Computing Communications (HPCC), 2017.

- Session Chair, Algorithms, International Conference on Supercomputing (ICS), 2013.

- Chair, Student Affairs Committee, 15th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming (PPoPP), 2010.
- Track Chair, Parallel/Grid Computing Systems track, 17th International Conference on Advanced Computing and Communications (ADCOM), 2009.
- Chair, Session on Garuda usage policies. 3rd Garuda Partners Meet, Bangalore, India. 2008.
- Chair, Session on Performance, 3rd IEEE International Conference on e-Science and Grid Computing, 2007.
- Tutorial Chair, 3rd IEEE International Conference on e-Science and Grid Computing, 2007.

9.3 Program Committee

- IEEE International Parallel and Distributed Processing Symposium (IPDPS): 2007, 2014, 2017-2019, 2021, 2023.
- IEEE Cluster conference, 2014, 2015, 2022.
- International Conference on High Performance Computing, HiPC: 2008, 2010-17, 2020.
- International Conference on Parallel Programming, ICPP: 2009, 2017, 2020.
- International Supercomputing Conference, ISC: 2019.
- Heterogeneous Computing Workshop (HCW): 2016.
- IEEE/ACM Supercomputing (SC) conference Posters, 2015.
- IEEE/ACM Supercomputing (SC) conference Birds-of-a-Feather (BoF), 2014.
- International Symposium on Cluster, Cloud and Grid Computing, CCGrid: 2006, 2010, 2011, 2014.
- IEEE International Conference on High Performance Computing and Communications, HPCC: 2005, 2009, 2011-2013.
- Student Research Symposium, International Conference on High Performance Computing, HiPC: 2009-2012.
- International Conference on Advanced Computing and Communications (ADCOM), 2009.
- Workshop on Scheduling and Resource Management for Parallel and Distributed Systems, SRMPDS: 2007, 2008.
- IEEE International Conference on e-Science and Grid Computing: 2005, 2007.
- Joint Workshop on High-Performance Grid Computing and High-Level Parallel Programming Models, HIPS-HPGC: 2005.

9.4 Reviewer

- Parallel Computing journal (ParCo): 2013, 2015, 2017.
- Concurrency & Computing: Practice & Experience (CPE), 2014, 2017.
- ACM Computing Surveys: 2011, 2016
- Journal of Supercomputing: 2014, 2016.
- Journal of Parallel and Distributed Computing, JPDC: 2004, 2005, 2008-2015.
- IEEE Transactions on Computers (TC), 2015.
- International Journal of High Performance Computing Applications (IJHPCA), 2014.
- IEEE Transactions on Cloud Computing, TCC: 2013.
- IEEE Transactions on Parallel and Distributed Systems, TPDS: 2010-2013.
- ACM Transactions on Computers: 2011, 2012.
- Future Generation Computer Systems, FGCS: 2007, 2008.
- ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming, PPoPP: 2008.
- EuroPar: 2005.
- Journal of Grid Computing: 2005.
- International Conference on High Performance Computing, HiPC: 2004.

9.5 Organizer

- Co-organized Indo-US Workshop on Computational Science and Data Engineering (CS&DE), held in conjunction with International Conference on High Performance Computing (HiPC), 2014.

10 Graduate Students

1. Chaitanya Vinod Patil, PhD, *January 2023 – date.*
2. R Akhila, PhD, *January 2023 – date.*
3. Sarthak Joshi, PhD, *August 2021 – date.*
4. Manasi Tiwari, PhD, *August 2017 – January 2023.*
5. Surendra Pericherla, Ph.D, *August 2015 – date.*
6. Preeti Malakar, Ph.D, *August 2008 – April 2014.*
7. Sivagama Sundari, Ph.D, *August 2005 – July 2011.*
8. H.A. Sanjay, Ph.D, *August 2005 – June 2009.*

9. Anwesha Bhowmick, M.Tech (Research), *August 2017 – January 2020.*
10. Soham Halder, M.Tech (Research), *August 2016 – July 2019.*
11. Rintu Panja, M.Sc (Engg.), *August 2015 – January 2018.*
12. Ashirbad Mishra, M.Sc (Engg.), *August 2013 – December 2016.*
13. Prakash Murali, M.Sc (Engg.), *August 2012 – April 2015.*
14. R. Vasudevan, M.Sc (Engg.), *August 2011 – December 2014.*
15. Cijo George, M.Sc (Engg.), *August 2010 – March 2013.*
16. Hari Raghavan, M.Sc (Engg.), *August 2010 – February 2013.*
17. Rajath Kumar, M.Sc (Engg.), *August 2010 – April 2013.*
18. Yadnyesh Joshi, M.Sc (Engg.), *August 2005 – November 2007.*
19. Rakhi Gupta, M.Sc (Engg.), *August 2004 – March 2007.*
20. Sandip Tikar, M.Sc (Engg.), *January 2004 – December 2006.*
21. Tanya Gautam, M.Tech *August 2021 – present.*
22. Arun Joshua Kennedy, M.Tech *August 2021 – present.*
23. Kevin Mahesh Kuriakose, M.Tech *August 2023 – present.*
24. Ruchil Prajapati, M.Tech *August 2021 – 2023.*
25. Aditya Kumar Patel, M.Tech *August 2020 – 2022.*
26. Anwesh Panda, M.Tech *August 2019 – July 2021.*
27. Varun PV, M.Tech *August 2019 – July 2021.*
28. Renga Bashyam KG, M.Tech *August 2018 – July 2020.*
29. Koushik Sen, M.Tech *August 2018 – July 2020.*
30. Mayank Gupta, M.Tech *August 2017 – July 2019.*
31. Udit Gupta, M.Tech *August 2017 – July 2019.*
32. Siddhant Panda, M.Tech *August 2016 – July 2018.*
33. Abhishek Upperwal, M.Tech *August 2016 – July 2018.*
34. Iresh Agrawal, M.Tech *August 2016 – July 2018.*
35. Ponnezhil Dass, M.Tech *August 2015 – July 2017.*
36. Vineetha K, M.Tech, *August 2014 – 2016.*
37. Sabyasachi Sahoo, M.Tech, *August 2014 – 2016.*

38. Anurag Murty, M.Tech, *August 2011 – 2013*.
39. K Saikiran Murty, M.Tech, *August 2011 – 2013*.
40. K Vivek Murty, M.Tech, *August 2011 – 2013*.
41. Gowthami Manogna Gottipati, M.Tech, *August 2010 – 2012*.
42. T. Santanu, M.Tech, *August 2010 – 2012*.
43. Sharat Chandra Racha, M.Tech, *August 2010 – 2012*.
44. Jitender Singh, M.Tech, *August 2008 – July 2010*.
45. J. Sivabhavani, M.E., *August 2008 – July 2010*.
46. Priyank Raj, M.Tech, *August 2007 – July 2009*.
47. Karthikeyan Raman, M.Tech, *August 2006 – July 2008*.
48. Antoine Henry, Student Intern from INSA, Lyon, France, *January 2007 – July 2007*.

11 Teaching

1. Indian Institute of Science, SE 295/DS 295, *Parallel Programming*: Spring 2004 – date.
2. Indian Institute of Science, DS 221, *Introduction to Scalable Systems*: Fall 2017 – date.
3. Indian Institute of Science, SE 286, *Data Structures and Programming*: Fall 2014.
4. Indian Institute of Science, SE 292, *High Performance Computing*: Fall 2009-2013, 2015-2016.
5. Indian Institute of Science, SE 293, *Topics in Grid Computing*: Fall 2004-2007.

Research Statement

I am the founder-convener of Middleware and Runtime Systems Lab (MARS) first in Supercomputer Education and Research Centre (SERC) and now in Department of Computational and Data Sciences (CDS), Indian Institute of Science (IISc). The lab, established in 2004, has been working on the following themes during the review period.

1. HPC (High Performance Computing) Application Frameworks for Graph and ML/DL Applications

This line of research involves developing runtime systems for efficient execution of applications on HPC systems. Our primary target has been multiple nodes of GPU systems and on applications with irregular computation, communication and memory access characteristics including graph applications, adaptive mesh refinement(AMR), N-Body simulations, Molecular Dynamics applications and sparse matrix solvers.

Our lab has been heavily working on developing frameworks for executions of graph applications on GPU systems with the objective of efficient and simultaneous harnessing of the power of both the CPU and the GPU cores. Most of the existing works on graph applications are CPU-only or GPU-only strategies. A hybrid strategy involving computations on both the CPU and GPU cores can help to explore large graphs and harness the power of all the devices. The few hybrid frameworks that exist employ the popular Bulk Synchronous Processing (BSP) model across the devices. We have convincingly shown in our works that BSP models involve large communication times and under-utilization of the devices. In an important work, we proposed a divide-and-conquer model (DC) for single-node GPU systems, called *HyPar* that fills this essential gap by promoting large-scale asynchronism. Our framework employs the strategy of partitioning the graph into two parts for the CPU and GPU, and invoking the original graph problem on the two devices for completely independent processing. The individual results on the two devices are then merged and post processed. We were able to successfully demonstrate our divide-and-conquer model to applications with different levels of complexities including applications that are directly amenable for the model, applications that may not seem amenable but whose algorithmic properties can be utilized to adopt to the model, applications whose independent results will have to be refined to obtain correct results and applications where rethinking of the fundamental algorithms is required. In each of these cases, our method yielded 24-92% performance benefits over the popular and widely prevalent BSP-only approaches. We extended this work to multiple nodes using a novel merging algorithm that avoids the single-node space complexity bottlenecks of the usual merging strategies, thereby truly facilitating exploration of very large graphs that cannot be accommodated in a single node.

In our recent work funded by Shell Technologies, India, we have been developing a comprehensive and integrated HPC framework for ML/DL applications. The project includes developing 1) asynchronous and communication minimizing strategies for large-scale AI/ML frameworks including mitigating and eliminating performance bottlenecks in current frameworks and proposing new models of parallelism for ML/DL frameworks, 2) strategies for high performance I/O for ML and DL workloads and 3) checkpoint-restart and fault tolerance frameworks for ML/DL problems. We have proposed a novel model of parallelism for ML/DL applications that overcomes the inherent limitations of existing models. The existing models of parallelism have inherent limitations in terms of model sizes that can be accommodated within the available memory and/or large communication and synchronization bottlenecks that limit their scalability for large-scale DNNs. In our work, we developed a divide-and-conquer model for asynchronous training of large-scale Deep Neural Networks

(DNNs) on multiple devices. Our work forms multiple small models from a given large model for asynchronous executions using the variational dropouts of the neurons. Our experiments with up to 16 GPUs on large-scale models and data sets show that our framework comprehensively outperforms the default data parallelism model in PyTorch, providing up to about 10% higher accuracy values and up to 2.5X speedup.

2. Middleware for Supercomputers and HPC Grids

Supercomputer installation sites and funding agencies for the sites have been increasingly facing the challenges of highly prohibitive and ever-increasing operational costs. With the growing demand for exascale computation, the power consumption and operational costs of these systems are expected to increase super-linearly over the years. Dynamic electricity price markets have become popular in US and many European countries, catering to large industries and manufacturing units. In a futuristic work with disruptive techniques, we proposed a first design of metascheduling in HPC grid systems that exploits the electricity price differences across states and countries in day-ahead electricity markets to execute jobs with minimum electricity cost and response times. The novelty is in the formulation of the scheduling problem as a Minimum Cost Maximum Flow problem and leveraging queue waiting time and electricity price predictions to accurately estimate the cost of job execution at a system. Experiments showed that our approach can potentially save $167K in annual electricity cost while obtaining 25% reduction in average response time compared to a baseline strategy.

Another important challenge in supercomputer centers is to determine configuration of job submission queues. In a widely prevalent and inefficient practice, system administrators manually choose these queue configurations based on their prior knowledge or experience, that may not be well-defined. The chosen configurations in general tend to be static over a period of time irrespective of the workload fluctuations, and thus cannot yield good utilization and response times under all workloads. We proposed a dynamic scheme for setting queue configurations, a non-linear programming (NLP) formulation for partitioning and mapping of nodes to the queues and a hybrid partitioned-nonpartitioned scheme for allocating processors to the jobs submitted to the queues. Our simulation results for our Institute's supercomputer system with 35,000+ CPU cores based on real workloads showed that our hybrid scheme gave up to 74% reduction in queue waiting times and up to 12% higher utilizations than static queue configurations.

3. High Performance Climate Modeling

Our lab has a long history of successful collaborations with climate modeling and visualization researchers to provide large-scale performance and acceleration of climate and weather models and in-situ visualization of the modeling output. We had developed strategies for providing high performance to climate models using computational grids of HPC batch systems. Our work on novel combination of performance prediction, processor allocation methods and topology-aware mapping of multiple regions of interests, was nominated as **best student paper** in the prestigious IEEE/ACM SC 2012 conference.

In another line of work, we developed methods for high performance of climate modeling routines on accelerators including GPUs and Intel Xeon Phis. The techniques include a novel asynchronous execution strategy, loop rearrangement and vectorization for accelerating different climate modeling routines, namely, long-wave radiation calculations in atmosphere models, horizontal diffusion computations in ocean models and convection calculations. Our works on Intel Xeon Phi Coprocessors were part of a project designated as the Intel Parallel Computing Centre (IPCC) on climate modeling, the first such centre in India.

In our current project funded by DST's SERB-Core Grant, we have been developing various strategies for integrated HPC simulations and visualization for large-scale ocean models, including the finest 500 metre resolutions that are attempted first-time in India for the Indian monsoon region. Our work includes profiling and identification of computational and I/O bottlenecks, developing hybrid CPU-GPU execution models for time-consuming ocean modeling routines, and developing I/O optimizations. Our work related to development of I/O optimization strategies for ocean models including transformation of independent to collective I/O, and selective writing strategies based on domain knowledge won the **best paper award** in the reputed HiPC conference in 2023.

The works by my group on accelerating climate models provided significant benefits to climate modeling scientists including executions of the costly long-wave radiation calculations that compute the absorptivities for gases from the default frequency of once every twelve simulated hours to the desirable frequency of once every simulated hour to adequately capture water vapor concentrations. Our works also provided 17.4-23% reduction in execution times for the ocean models which are typically major bottlenecks in climate models since ocean processes are slower relative to other components and where even small reductions in execution times translate to large savings in wall clock times, thus enabling faster analyses of temperature and salinity fields in oceans. We also showed that with the help of domain knowledge, it is possible to identify slow varying phases in the climate models and perform approximate computations for these phases by offloading these computations to the accelerators and allow the CPU to execute the other phases asynchronously with the GPU computations. The CPU will proceed with its computations making use of the earlier results of these slow-varying phases and will use the new results when the GPU completes its computations. This was the first such work that harnessed the CPUs and GPUs using the domain knowledge on slow-varying phases and has far-reaching impact to also include domains other than climate modeling.

4. Fault Tolerance for Exascale Systems

As we have entered Exascale computing, the faults in high-performance systems are expected to increase considerably. Our lab has long been developing methods for fault tolerance for HPC applications. In an important work during this review period, we developed adaptive strategies that use the promising process replication strategy for fault tolerance for exascale systems. We had developed strategies that combine failure prediction and adaptive process replication to replicate only a small subset (1%) of the total number of processes. This subset is changed at every interval based on the likelihood of failure of the processes in the next interval. By means of extensive simulations with difference scenarios of projected exascale systems, we showed that the adaptive partial replication of processes provided application efficiency in excess of 75%. In our current work, we have developed a novel fault-tolerant MPI library that adopts partial replication of some of the launched MPI processes in order to provide resilience from failures. The novelty of our work is that it combines both fault tolerance and high performance, due to the use of communication protocols in the native MPI library that is generally fine-tuned for specific HPC platforms. Overall, this work is expected to stand the test of some time into the future.

5. Performance Modeling, Predictions and Scalability Characterization

Performance modeling, predictions and scalability characterization of parallel applications are essential and useful for various purposes including identifying performance bottlenecks, projections for future systems and tuning applications and algorithms. We developed a prediction framework that used only a single small scale application execution for performance predictions of large problem

sizes and processors by matching application execution signatures with reference kernels in terms of instruction mix. Our research provided an important insight that the fundamental routines in many of the large-scale applications may exhibit similar coding and execution patterns. A different line of work is to predict the scalability trends using only small-scale runs for even those applications which exhibit performance bottlenecks only at scale. We developed techniques that use modeling the number of invocations of time-consuming functions using small-scale executions and predict scalability trends of the application using the model. With executions on only up to 512 cores, our techniques were able to predict scalability trends for up to 16K cores for large scientific applications with correlation coefficients of minimum 97%.

Leadership in HPC:
Chair of a Supercomputer Center and Investigator in National Supercomputing Mission (NSM)

Since 2018, I have been the Chair of Supercomputer Education and Research Centre (SERC), IISc, a premier supercomputer center in the country. During this tenure of the past six years, important developments have taken place with excellent support from our administration and good help from our team. This includes the installation of a NSM machine called Param Pravega, the largest supercomputer system in an Indian academic Institution, with tremendous support from C-DAC. This involved careful planning of the data center infrastructure including cooling and power components over a sustained period. Our team also procured a small-scale cluster for simulations and a machine for AI/ML/DL applications, during this period. Our center has also been organizing numerous training programs both at the basic and advanced level for practitioners across the country. In this period, we also helped established a stable workforce for managing the center.

I have also been the investigator from Indian Institute of Science in the National Supercomputing Mission (NSM), a major initiative of Government of India. Under this project, IISc in coordination with JNCASR had issued Call for Proposals (CFP) and granted 104 research projects in HPC areas for the PIs across the country. This is the first time that HPC-related projects been coordinated at such a large scale in the country. These projects gave major impetus to large-scale HPC research in the country, resulting in about 260 research papers in reputed journals and conferences, training of about 300 personnel including PhDs, PostDocs, Masters, Undergraduates and research staff in the HPC projects, generation of about 43 software in different domains, and receipt of significant awards including the ACM Gordon Bell Prize. As part of the projects, we had also sanctioned 13 workshops to evoke interest among the students.