

ANIRBAN MANDAL

Renaissance Computing Institute(RENCI) Email: anirban@renci.org
University of North Carolina, Chapel Hill Phone: 713.443.1847 (cell)
100 Europa Drive, Suite 540 Fax: 919.445.9669
Chapel Hill, NC 27517 <http://www.renci.org/~anirban/>

INTERESTS

Research and development in the area of provisioning, scheduling, fault-tolerance and performance analysis for high-performance, parallel and distributed computing systems.

EDUCATION

Ph.D. in Computer Science, May 2006

Rice University, Houston, TX

Thesis: Toward a Tool for Scheduling Application Workflows onto Distributed Grid Systems

Advisor: Dr. Ken Kennedy

M.S. in Computer Science, November 2002

Rice University, Houston, TX

Thesis: Mapping HPF onto the Grid

Advisor: Dr. Ken Kennedy

B.Tech. in Computer Science and Engineering, May 2000

Indian Institute of Technology, Mumbai, India

Thesis: Design and Implementation of a Java JIT Compiler

Advisor: Dr. D. M. Dhamdhare

RESEARCH EXPERIENCE

Research Scientist at RENCi, UNC-Chapel Hill

July 2006 - present

Serve as a **co-PI** for UNC for a **NSF** Software Development for CyberInfrastructure (SDCI) project in collaboration with Duke University and University of Massachusetts, Amherst. Served as institutional **PI** for the VGrADS (Virtual Grid Application Development Software) project, a **NSF** sponsored collaborative research project across seven institutions. Develop research agenda and track research activity.

Research on distributed grid computing, cloud computing and next-generation networking. Develop tools and research software for multi-disciplinary collaborative projects, lead multi-institutional development efforts, review and write scientific papers and research proposals, represent RENCi at key national meetings and assist in hiring personnel. Highlights of research and technical activities -

- **Scheduling Scientific Workflows with Reliability for Grid Computing** : Research on scheduling algorithms for grid computing. Evaluated effectiveness of combined workflow scheduling and fault-tolerance approaches by analyzing their impact on the reliability of workflow execution, workflow performance and resource usage under different reliability models, failure prediction accuracies and workflow application

types.

- **Fault Tolerance and Recovery for Scientific Workflows** : Designed and implemented algorithms for over-provisioning and migration for scientific workflow components. It is integrated with a Fault Tolerance and Recovery web-service, which is a part of the *Linked Environment for Atmospheric Discovery* (LEAD) production portal.
- **GENI - Global Environment for Network Innovations** : Worked on developing tools to support next-generation network experiments by exposing resource management and resource discovery APIs for provisioning network and edge substrates. Also, developed a central registry for the ORCA - a control framework to provision virtual networked systems via secure and distributed management of heterogeneous resources over federated substrate sites. The registry offers an xml-rpc interface to the users and ORCA agents to offer easy publishing and discovery capabilities.
- **Memory Performance Analysis for Multi-core Architectures** : Did empirical evaluation of multi-core memory concurrency for variety of multi-socket, multi-core architectures and developed memory performance models. Also, investigated performance consistency on multi-socket AMD Systems.
- **CyberInfrastructure Evaluation** : Developed scaling and performance models for application benchmarks on the IBM Blue-Gene/L platform to evaluate peta-scale systems. Investigated I/O benchmarks for large-scale parallel systems.
- **Open Science Grid** : Assisted atmospheric scientists to run models on distributed resources using the *Open Science Grid* infrastructure and grid middleware.

Graduate Research Assistant, Rice University

May 2001 - May 2006

Advisor: Dr. Ken Kennedy

Investigated techniques for automatic, yet efficient application level scheduling for large-scale scientific applications in order to map them onto distributed Grid like systems in the context of the *Virtual Grid Application Development Software* (VGrADS) project.

- **Heuristic Workflow Scheduling using performance models** : Developed heuristic scheduling algorithms for in-advance scheduling of scientific application workflows. Implemented the scheduler that efficiently scheduled workflows from a bio-imaging application (EMAN) and meso-scale weather prediction application (LEAD). The makespan for the workflows dramatically improved from the ones obtained using existing scheduling strategies.
- **Scheduling onto Batch Queue Systems**: Extended workflow scheduling to incorporate batch queue wait time predictions to map workflows to supercomputers with batch queue front ends. Joint work with Dr. Daniel Nurmi and Dr. Richard Wolski.
- **Scalable Grid Application Scheduling Strategies**: Investigated scalability of workflow scheduling. Designed a scheduling algorithm to schedule workflows onto abstract resource classes like clusters. Have also investigated an approach of decoupled resource selection and scheduling to address the scaling problem.
- **HPF Mapper**: Developed strategies for efficiently mapping tightly-coupled parallel HPF applications onto the Grid. Building the mapper entailed program analysis using task-graphs to discover the communication structure of the program and mapping processes to processors to minimize wide-area communication cost.

Graduate Intern, USC/Information Sciences Institute, CA May 2003 - Aug 2003
Supervisors: Dr. Carl Kesselman and Dr. Ewa Deelman

Extended the *Pegasus* Grid Workflow management system by incorporating scheduling smarts in the system. Investigated the value of whole-workflow scheduling over dynamic task based scheduling approaches.

PUBLICATIONS

Book Chapter

R. J. Fowler, T. Gamblin, G. Kandaswamy, A. Mandal, A. K. Porterfield, L. Ramakrishnan and D. A. Reed, "Challenges of Scale: When All Computing Becomes Grid Computing," in Lucio Grandinetti, Eds., *High Performance Computing and Grids in Action*, Chapter, IOS Press, Amsterdam, 2008.

Refereed Articles

A. Mandal, Y. Xin, I. Baldine, P. Ruth, C. Heermann, J. Chase, V. Orlikowski and A. Yumerefendi, "Provisioning and Evaluating Multi-domain Networked Clouds for Hadoop-based Applications," in *Proceedings of the 3rd International Conference on Cloud Computing Technologies and Science 2011 (IEEE Cloudcom '11)*, Athens, Greece, December 2011, to appear.

Y. Xin, I. Baldine, A. Mandal, C. Heermann, J. Chase and A. Yumerefendi, "Embedding Virtual Topologies in Networked Clouds," in *Proceedings of the 6th International Conference on Future Internet Technologies 2011 (CFI '11)*, Seoul, Korea, June 2011.

I. Baldine, Y. Xin, A. Mandal, C. Heermann, J. Chase, V. Marupadi, A. Yumerefendi and D. Irwin, "Autonomic Cloud Network Orchestration: A GENI Perspective", in *Proceedings of the 2nd IEEE International Workshop on Management of Emerging Networks and Services (IEEE MENS '10)* in conjunction with GLOBECOM'10, Miami, FL, December 2010.

A. Mandal, R. Fowler and A. Porterfield, "Modeling Memory Concurrency for Multi-Socket Multi-Core Systems", in *Proceedings of the IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS'10)*, pp. 66-75, White Plains, NY, March 2010.

L. Ramakrishnan, D. Nurmi, A. Mandal, C. Koelbel, D. Gannon, T. M. Huang, Y. S. Kee, G. Obertelli, K. Thyagaraja, R. Wolski, A. Yarkhan and D. Zagorodnov, "VGrADS: Enabling e-Science Workflows on Grids and Clouds with Fault Tolerance," in *Proceedings of the IEEE/ACM International Conference for High Performance Computing, Networking, Storage and Analysis (SC)*, Portland, OR, November 2009.

R. Zhang, A. Mandal, C. Koelbel and K. Cooper, "Combined Fault Tolerance and Scheduling Techniques for Workflow Applications on Computational Grids," in *Proceedings of the*

IEEE/ACM International Symposium on Cluster Computing and the Grid (CCGrid) , pp. 244-251, Shanghai, China, May 2009.

G. Kandaswamy, A. Mandal, and D. A. Reed, “Fault Tolerance and Recovery of Scientific Workflows on Computational Grids,” in *Proceedings of the IEEE International Symposium on Cluster Computing and the Grid (CCGrid)* , pp. 777-782, Lyon, France, May 2008.

D. Nurmi, A. Mandal, J. Brevik, K. C. Koelbel, R. Wolski and K. Kennedy, “Evaluation of a Workflow Scheduler Using Integrated Performance Modeling and Batch Queue Wait Time Prediction,” in *Proceedings of the IEEE/ACM International Conference for High Performance Computing, Networking, Storage and Analysis (SC)* , pp. 29-, Tampa, FL, November 2006.

R. Zhang, A. Mandal, H. Casanova, A. Chien, Y. Kee, K. Kennedy and C. Koelbel, “Scalable Grid Application Scheduling with Virtual Grids via Decoupled Resource Selection and Scheduling,” in *Proceedings of the IEEE International Symposium on Cluster Computing and the Grid (CCGrid)* , Vol 1, pp. 568-575, Singapore, May 2006.

A. Mandal, K. Kennedy, C. Koelbel, G. Marin, J. Mellor-Crummey, B. Liu and L. Johnsson, “Scheduling Strategies for Mapping Application Workflows onto the Grid,” in *Proceedings of the IEEE International Symposium on High Performance Distributed Computing (HPDC)* , pp. 125-134, Raleigh, NC, July 2005.

J. Blythe, S. Jain, E. Deelman, Y. Gil, K. Vahi, A. Mandal and K. Kennedy, “Task Scheduling Strategies for Workflow-based Applications in Grids.” in *Proceedings of the IEEE International Symposium on Cluster Computing and the Grid (CCGrid)* , Vol 2, pp. 759-767, Cardiff, UK, May 2005.

A. Mandal, A. Dasgupta, K. Kennedy, M. Mazina, C. Koelbel, G. Marin, K. Cooper, J. Mellor-Crummey, B. Liu and L. Johnsson, “Scheduling Workflow Applications in GrADS,” in *Proceedings of the IEEE International Symposium on Cluster Computing and the Grid (CCGrid)* , pp. 790-797, Chicago, IL, April 2004.

F. Berman, H. Casanova, A. Chien, K. Cooper, H. Dail, A. Dasgupta, W. Deng, J. Dongarra, L. Johnsson, K. Kennedy, C. Koelbel, B. Liu, X. Liu, A. Mandal, G. Marin, M. Mazina, J. Mellor-Crummey, C. Mendes, A. Olugbile, M. Patel, D. Reed, Z. Shi, O. Sievert, H. Xia and A. YarKhan, “New Grid Scheduling and Rescheduling Methods in the GrADS Project,” in *International Journal of Parallel Programming (IJPP)* , Volume 33(2-3):pp. 209-229, 2005.

Technical Reports, Posters and Dissertations

A. Mandal, M. Y. Lim, A. Porterfield and R. Fowler, “Implications for Applications and Compilers of Multi-core Memory Concurrency,” in Poster at *International Workshop on Languages and Compilers for Parallel Computing (LCPC'10)* , Houston, TX, October 2010.

A. Porterfield, R. Fowler, A. Mandal and M. Y. Lim, “Empirical Evaluation of Multi-Core

Memory Concurrency, Available as Technical Report TR-09-01, RENC1, North Carolina, Jan. 2009.

A. Porterfield, R. J. Fowler, A. Mandal and M. Y. Lim, "Performance Consistency on Multi-socket AMD Opteron Systems," Available as Technical Report TR-08-07, RENC1, North Carolina, Dec. 2008.

A. Mandal, "Toward a Tool for Scheduling Application Workflows onto Distributed Grid Systems", PhD Thesis, Available as Technical Report TR06-875, Computer Science, Rice University, May 2006.

A. Mandal, "Mapping HPF onto the Grid", M.S. Thesis, Available as Technical Report TR03-417, Computer Science, Rice University, Nov. 2003.

GRANTS

"Missing Link: Connecting Eucalyptus Clouds with Multi-Layer Networks," National Science Foundation (NSF) SDCI grant, Co-PI for RENC1, UNC - Chapel Hill, Sep 2010 - Aug 2013, \$295,000.

"Orchestrating Distributed Resource Ensembles for Petascale Science," Department of Energy (DoE) ASCR Next Generation Networking grant, Senior Personnel, Sep 2010 - Aug 2013, \$509,100.

"Virtual Grid Application Development Software (VGrADS)," NSF large ITR grant, PI for UNC subcontract, Dec 2007 - Sep 2009, managed \$225,000.

PROFESSIONAL ACTIVITIES

Proposal panels and program committees

Invited to and served on **National Science Foundation** Proposal Review Panels (March 2009, May 2009).

Program Committee member for Grid Computing Environments (GCE'09), held in conjunction with SC'09, Portland.

Journal article reviews

Reviewer for journal, *IEEE Transactions on Parallel and Distributed Systems (TPDS)* (2007-2009).

Reviewer for journal, *Journal of Systems and Software* (2010).

Reviewer for journal, *Parallel Computing (PARCO)* (2006).

Reviewer for journal, *Future Generation Computer Systems (FGCS)* (2005-2007).

Reviewer for journal published from China, *Journal of Computer Science and Technology* (2007).

Reviewer for journal, *Journal of Parallel and Distributed Computing (JPDC)* (2004-2006).

Reviewer for journal, *Journal of Grid Computing (JOGC)* (2003).

Reviewer for journal, *Concurrency and Computation: Practice and Experience* (2001, 2004).

Conference article reviews

Reviewer for international conferences: International Parallel and Distributed Processing Symposium (IPDPS) (2003, 2005, 2010), Supercomputing (SC) (2004), Euro-Par (2004, 2006, 2008), Middleware for Grid Computing (MGC) (2004) & Los Alamos Computer Science Institute (LACSI) Symposium (2004)

PRESENTATIONS AND DEMONSTRATIONS

“Modeling Memory Concurrency for Multi-Socket Multi-Core Systems” - Presentation at International Symposium on Performance Analysis of Systems and Software (ISPASS 2010), White Plains, NY, March 2010.

“Fault Tolerance and Recovery of Scientific Workflows on Computational Grids” - Presentation at International Symposium on Cluster Computing and the Grid (CCGrid 2008), Lyon, France, May 2008.

“Scheduling Strategies for Mapping Application Workflows onto the Grid” - Presentation at International Symposium on High Performance Distributed Computing (HPDC 2005), Raleigh, NC, July 2005.

“Virtual Grid Execution System: Fault Tolerance Planning and Run-time Rescheduling of Scientific Workflows” - Presentation at Supercomputing (SC) (2007), Reno, NV, November 2007.

“Application Scheduling with Batch-queued Resources” - Demonstration at Supercomputing (SC) (2005), Seattle, WA, November 2005.

“EMAN Workflow Scheduling in GrADS/VGrADS” - Demonstrations at Supercomputing (SC) (2003-2004)

TEACHING EXPERIENCE

Teaching Assistant, Rice University

Aug 2000 - Dec 2002

Served as a teaching assistant for five semesters for advanced undergraduate courses such as Operating Systems, Compiler Construction, Design & Analysis of Algorithms and an introductory course on Principles of Computing and Programming. Conducted laboratory sessions, assisted students with projects and graded exams and assignments. Mentored visiting undergraduate summer students and junior graduate students.

COURSE HIGHLIGHTS

Advanced Compilation for High-Performance Computers, High-Performance Computer Architecture, Operations Research, Advanced Compiler Construction, Distributed Systems

Seminar, Design and Analysis of Algorithms, Operating Systems

REFERENCES

Available on request.