PAVAN BALAJI

Mathematics and Computer Science (MCS) Division, Argonne National Laboratory (ANL)
Phone (O): 630-252-3017 **Fax:** 630-252-5986
Email: balaji@mcs.anl.gov **Home Page:** http://www.mcs.anl.gov/∼balaji

EDUCATION

Ph.D., Computer Science and Engg., Ohio State University	06/2006
M.S., Computer Science and Engg., Ohio State University	06/2003
B.Tech., Computer Science and Engg., Indian Institute of Technology (IIT), Madras	08/2001

AFFILIATION AND RESEARCH INTERESTS

I'm currently a post-doctoral researcher in the Mathematics and Computer Science (MCS) division of the Argonne National Laboratory (ANL). My research broadly falls into the general areas of parallel and distributed computing. Specifically, I focus on three areas:

- Network-based parallel and distributed computing including high-speed interconnects (InfiniBand, 10-Gigabit Ethernet, Myrinet), offloaded protocol stacks (TCP offload engines, RDMA offload engines) and IP-based protocols (iWARP, TCP/IP, UDP/IP).
- Parallel programming models and other middleware (MPI, sockets, file-systems, web-based data-center middleware).
- Various aspects of job scheduling and resource management (QoS, scheduling of parameter sweep applications, energy-aware scheduling).

AWARDS AND ACCOMPLISHMENTS

1. *Los Alamos Director's Technical Achievement Award for Summer Interns*, Los Alamos National Laboratory, 2005.
2. *Outstanding Research Award of Excellence*, Computer Science and Engineering, Ohio State University, 2005.

PROFESSIONAL EXPERIENCE

1. **Argonne National Laboratory, Post-doctoral Researcher**, Argonne, IL, Aug '06 to present
   Supervisor: Dr. Rusty Lusk
   Responsibilites: Designing advanced communication mechanisms for MPICH2, an implementation of the Message Passing Interface (MPI)
2. **Ohio State University, Graduate Research Assistant**, Columbus, Ohio, Jan '02 to Jun '06
   Advisor: Prof. D. K. Panda
   Responsibilities: High Performance Communication Support for Sockets-based Applications over High-speed Networks
3. **Los Alamos National Laboratory, Summer Intern**, New Mexico, Mar '05 to Aug '05
   Team Lead: Dr. Wu-chun Feng
   Project: High Performance Sockets evaluation and extensions for multi-Gigabit Networks
4. **Intel Corporation, Summer Intern**, Austin, Texas, Jun '03 to Sep '03
   Manager: Dr. Scott Hahn
   Project: Evaluating 10-Gigabit Network Interfaces with TCP/IP and InfiniBand
5. **Ohio State University, Graduate Teaching Assistant**, Columbus, Ohio, Mar '03 to Jun '03
   Faculty: Prof. Sandra Mamrak
   Course: Advanced Operating Systems Laboratory

6. **Intel Corporation, Summer Intern**, Austin, Texas, Jun '02 to Sep '02
   Manager: Dr. Hemal V. Shah
   Project: Direct Data Placement over TCP/IP: A Software Prototype

7. **Ohio State University, Graduate Teaching Assistant**, Columbus, Ohio, Sep '01 to Dec '01
   Faculty: Prof. P. Sadayappan
   Courses: Introduction to High Performance Computing; Introduction to VLSI Designing

8. **Network Based Computing Laboratory, System Administrator**, Columbus, Ohio, Sep '02 to Jun '06
   Faculty: Prof. D. K. Panda
   Responsibilities: System maintenance and trouble shooting for high-end research equipment

9. **Indian Institute of Technology (IIT), Head Teaching Assistant**, Madras, India, Jan '00 to Dec '00
   Faculty: Prof. R. Pandurangan (Head of the Department)
   Responsibilities: In charge of a class of over 300 students and coordinating other teaching assistants

10. **Indian Institute of Technology (IIT), Teaching Assistant**, Madras, India, Jul '99 to Dec '00
    Faculty: Prof. C. Sivaramamurthy and Prof. R. Pandurangan (Head of the Department)
    Responsibilities: Taught an introductory course in Computers to a class of about 40 students

11. **ENIAC, Lecturer**, India, Lecturer, May '99 to Jul '99
    Courses: Introduction to Programming in C; Introduction to the AIX Environment

PUBLICATIONS

Refereed Journal Articles

1. **P. Balaji**, W. Feng and D. K. Panda. "Bridging the Ethernet-Ethernot Performance Gap". *IEEE Micro Journal; Special Issue on High-Performance Interconnects*, pp. 24-40, Vol. 26, Issue 3, 2006.

2. H. -W. Jin, **P. Balaji**, C. Yoo, J. -Y. Choi and D. K. Panda. "Exploiting NIC Architectural Support for Enhancing IP based Protocols on High Performance Networks". *Special Issue of the Journal of Parallel and Distributed Computing (JPDC) on Design and Performance of Networks for Super-, Cluster- and Grid-Computing*, pp. 1348-1365, Vol. 65, Issue 11, 2005.

3. M. Islam, **P. Balaji**, P. Sadayappan and D. K. Panda. "QoPS: A QoS based scheme for Parallel Job Scheduling (extended journal version)". *IEEE Springer LNCS Journal Series*, pp. 252-268, Vol. 2862, 2003.

4. M. Islam, **P. Balaji**, P. Sadayappan and D. K. Panda. "Towards Quality of Service Guarantees in Job Scheduling (extended journal version)". *International Journal of High Performance Computing and Networking (JHPCN)*; to appear.

5. **P. Balaji**, G. Sabin and P. Sadayappan. "Opportune Job Shredding: An Efficient Approach for Scheduling Parameter Sweep Applications (extended journal version)". *International Journal for High Performance Computing Applications (JHPCA)*; to appear.

Refereed Conference Publications

6. **P. Balaji**, W. Feng, Q. Gao, R. Noronha, W. Yu and D.K. Panda. "Head-to-TOE Comparison for High Performance Sockets over Protocol Offload Engines". *IEEE International Conference on Cluster Computing (Cluster)*, Sep 26th to 30th, 2005, Boston, Massachusetts.

7. W. Feng, **P. Balaji**, C. Baron, L. N. Bhuyan and D. K. Panda. "Performance Characterization of a 10-Gigabit Ethernet TOE". *IEEE International Symposium on High Performance Interconnects (HotI)*, Aug 17th to 19th, 2005, Palo Alto, California.

8. **P. Balaji**, K. Vaidyanathan, S. Narravula, H. -W. Jin and D. K. Panda. "On the Provision of Prioritization and Soft QoS in Dynamically Reconfigurable Shared Data-Centers over InfiniBand". *IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS)*, Mar 20th to 22nd, 2005, Austin, Texas.

9. S. Narravula, **P. Balaji**, K. Vaidyanathan, H. -W. Jin and D. K. Panda. "Architecture for Caching Responses with Multiple Dynamic Dependencies in Multi-Tier Data-Centers over InfiniBand". *IEEE/ACM International Symposium on Cluster Computing and the Grid (CCGrid)*, May 9th to 12th, 2005, Cardiff, UK.

10. M. Islam, **P. Balaji**, P. Sadayappan and D. K. Panda. "Towards Provision of Quality of Service Guarantees in Job Scheduling". *IEEE International Conference on Cluster Computing (Cluster)*, Sep 20th to Sep 23rd, 2004, San Diego, California.

11. **P. Balaji**, S. Narravula, K. Vaidyanathan, S. Krishnamoorthy, J. Wu and D. K. Panda. "Sockets Direct Protocol over InfiniBand in Clusters: Is it Beneficial?". *IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS)*, Mar 10th to 12th, 2004, Austin, Texas.

12. R. Kurian, **P. Balaji** and P. Sadayappan. "Opportune Job Shredding: An Efficient Approach for Scheduling Parameter Sweep Applications". *Los Alamos Computer Science Institute (LACSI) Symposium*, Oct 12th to 14th, 2003, Santa Fe, New Mexico.

13. **P. Balaji**, J. Wu, T. Kurc, U. Catalyurek, D. K. Panda and J. Saltz. "Impact of High Performance Sockets on Data Intensive Applications". *IEEE International Symposium on High Performance Distributed Computing (HPDC)*, Jun 22nd to 24th, 2003, Seattle, Washington.

14. R. Gupta, **P. Balaji**, J. Nieplocha and D. K. Panda. "Efficient Collective Operations using Remote Memory Operations on VIA-based Clusters". *IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, Apr 22nd to 26th, 2003, Nice, France.

15. **P. Balaji**, P. Shivam, P. Wyckoff and D. K. Panda. "High Performance User-level Sockets over Gigabit Ethernet". *IEEE International Conference on Cluster Computing (Cluster)*, Sep 23rd to Sep 26th, 2002, Chicago, Illinois.

Refereed Workshop Publications

16. **P. Balaji**, S. Bhagvat, H. -W. Jin and D. K. Panda. "Asynchronous Zero-copy Communication for Synchronous Sockets in the Sockets Direct Protocol (SDP) over InfiniBand". *Workshop on Communication Architecture for Clusters (CAC); in conjunction with the IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, Apr 25th, 2006, Rhodes Island, Greece.

17. V. Viswanathan, **P. Balaji**, W. Feng, J. Leigh, D. K. Panda. "A Case for UDP Offload Engines in LambdaGrids". *Workshop on Protocols for Fast Long-Distance Networks (PFLDnet)*, Feb 2nd and 3rd, 2006, Nara, Japan.

18. **P. Balaji**, H. -W. Jin, K. Vaidyanathan and D. K. Panda. "Supporting RDMA capable network compatibility and features for regular network adapters". *Workshop on Remote Direct Memory Access (RDMA): Applications, Implementations and Techniques (RAIT); in conjunction with IEEE International conference on Cluster Computing (Cluster)*, Sep 26th, 2005, Boston, Massachusetts.

19. H. -W. Jin, S. Narravula, G. Brown, K. Vaidyanathan, **P. Balaji** and D. K. Panda. "RDMA over IP Networks: A Study with the Ammasso Gigabit Ethernet NIC". *Workshop on High Performance Interconnects for Distributed Computing (HPI-DC); in conjunction with IEEE International Symposium on High Performance Distributed Computing (HPDC)*, Jul 24th, 2005, Research Triangle Park, North Carolina.

20. K. Vaidyanathan, **P. Balaji**, H. -W. Jin and D. K. Panda. "Workload driven analysis of File Systems in Shared Multi-Tier Data-Centers over InfiniBand". *Workshop on Computer Architecture Evaluation using Commercial Workloads (CAECW); in conjunction with IEEE International Symposium on High Performance Computer Architecture (HPCA)*, Feb 12th, 2005, San Francisco, California.

21. **P. Balaji**, H. V. Shah and D. K. Panda. "Sockets vs. RDMA Interface over 10-Gigabit Networks: An In depth Analysis of the Memory Traffic Bottleneck". *Workshop on Remote Direct Memory Access (RDMA): Applications, Implementations and Technologies (RAIT); in conjunction with IEEE International Conference on Cluster Computing (Cluster)*, Sep 20th, 2004, San Diego, California.

22. **P. Balaji**, K. Vaidyanathan, S. Narravula, S. Krishnamoorthy, H. -W. Jin and D. K. Panda. "Exploiting Remote Memory Operations to Design Efficient Reconfiguration for Shared Data-Centers over InfiniBand". *Workshop on Remote Direct Memory Access (RDMA): Applications, Implementations and Technologies (RAIT); in conjunction with IEEE International Conference on Cluster Computing (Cluster)*, Sep 20th, 2004, San Diego, California.

23. S. Narravula, **P. Balaji**, K. Vaidyanathan, S. Krishnamoorthy, J. Wu and D. K. Panda. "Supporting Strong Coherency for Active Caches in Multi-Tier Data-Centers over InfiniBand". *Workshop on System Area Networks (SAN); in conjunction with IEEE International Symposium on High Performance Computer Architecture (HPCA)*, Feb 14th, 2004, Madrid, Spain.

24. M. Islam, **P. Balaji**, P. Sadayappan and D. K. Panda. "QoPS: A QoS based scheme for Parallel Job Scheduling". *Job Scheduling Strategies for Parallel Processing (JSSPP) workshop; in conjunction with IEEE International Symposium on High Performance Distributed Computing (HPDC)*, Jun 24th, 2003, Seattle, Washington.

Invited Papers

25. **P. Balaji**, K. Vaidyanathan, S. Narravula, H. -W. Jin and D. K. Panda. "Designing Next Generation Data-centers with Advanced Communication Protcols and Systems Services". *Workshop on the National Science Foundation Next Generation Software (NSFNGS) Program; in conjunction with the IEEE International Parallel and Distributed Processing Symposium(IPDPS)*, Apr 25th, 2006. Rhodes Island, Greece.

Technical Reports

26. K. Vaidyanathan, **P. Balaji,** J. Wu, H. -W. Jin and D. K. Panda. "An Architectural Study of Cluster-based Multi-tier Data-Centers". *Technical Report, OSU-CISRC-5/04-TR25, Ohio State University*.

27. S. Krishnamoorthy, **P. Balaji,** K. Vaidyanathan, H. -W. Jin and D. K. Panda. "Dynamic Reconfigurability Support for providing Soft Quality of Service Guarantees in Multi-Tier Data-Centers over InfiniBand". *Technical Report, OSU-CISRC-2/04-TR10, Ohio State University*.

RESEARCH GRANTS, DONATIONS AND LOANERS

1. *"Achieving Complete WAN compatibility with Myrinet 10-G Network Adapters"*. Co-PI (with W. Feng) on an equipment donation from Myricom Inc., 09/2006 onward (permanent grant), $20,000.

2. *"Designing Efficient System Software Services for iWARP-enabled Clusters"*. PI on an equipment loaner from NetEffect Inc., 10/2006 to 03/2007, $45,000.

PRESENTATIONS

Invited Talks

1. *"SDP and Extended Sockets: A TOE and iWARP Perspective"*. NetEffect Inc., June 19th, 2006, Austin, TX.

2. *"TCP and iWARP Offload Engines: Challenges and Pitfalls"*. NetEffect Inc., June 19th, 2006, Austin, TX.

3. *"Designing High-performance Communication Support for Sockets-based Applications"*. Hewlett-Packard (HP), May 23rd, 2006, Palo Alto, CA.

4. *"Enhancing the Performance and Compatibility of Sockets over High-Speed Networks"*. QLogic Corporation (formerly Pathscale Incorporation), May 5th, 2006, Mountain View, CA.

5. *"Designing High Performance Communication Support for Sockets-based Applications over System-area and Wide-area Networks"*. Argonne National Laboratory (ANL), Mar 24th, 2006, Argonne, Illinois.

6. *"Comparing 10-Gigabit Ethernet with InfiniBand and Myrinet for High Performance Sockets over Protocol Offload Engines"*. IEEE International conference on Supercomputing (SC), Nov 16th, 2005, Seattle, Washington.

Conference and Workshop Presentations

7. *"Asynchronous Zero-copy Communication for Synchronous Sockets in the Sockets Direct Protocol (SDP) over InfiniBand"*. Workshop on Communication Architecture for Clusters (CAC); held in conjunction with the IEEE International Parallel and Distributed Processing Symposium (IPDPS), Apr 25th, Rhodes Island, Greece

8. *"Head-to-TOE Comparison for High Performance Sockets over Protocol Offload Engines"*. IEEE International Conference on Cluster Computing (Cluster), Sep 29th, 2005, Boston, Massachusetts.

9. *"Supporting RDMA Capable Network Compatibility and Features for Regular Network Adapters"*. Workshop on Remote Direct Memory Access (RDMA): Applications, Implementations and Techniques (RAIT); held in conjunction with the IEEE International conference on Cluster Computing (Cluster), Sep 26th, 2005, Boston, Massachusetts.

10. *"Performance Characterization of a 10-Gigabit Ethernet TOE"*. IEEE International Symposium on High Performance Interconnects (HotI), Aug 18th, 2005, Palo Alto, California.

11. *"Sockets vs. RDMA Interface over 10-Gigabit Networks: An In depth Analysis of the Memory Traffic Bottleneck"*. Workshop on Remote Direct Memory Access (RDMA): Applications, Implementations and Technologies (RAIT); held in conjunction with the IEEE International Conference on Cluster Computing (Cluster), Sep 20th, 2004, San Diego, California.

12. *"Exploiting Remote Memory Operations to Design Efficient Reconfiguration for Shared Data-Centers over InfiniBand"*. Workshop on Remote Direct Memory Access (RDMA): Applications, Implementations and Technologies (RAIT); held in conjunction with the IEEE International Conference on Cluster Computing (Cluster), Sep 20th, 2004, San Diego, California.

13. *"Towards Provision of Quality of Service Guarantees in Job Scheduling"*. IEEE International Conference on Cluster Computing (Cluster), Sep 22nd, 2004, San Diego, California.

14. *"Sockets Direct Protocol over InfiniBand in Clusters: Is it Beneficial?"*. IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS), Mar 10th, 2004, Austin, Texas.

15. *"Impact of High Performance Sockets on Data Intensive Applications"*. IEEE International Symposium on High Performance Distributed Computing (HPDC), Jun 22nd, 2003, Seattle, Washington.

16. *"Efficient Collective Operations using Remote Memory Operations on VIA-based Clusters"*. IEEE International Parallel and Distributed Processing Symposium (IPDPS), Apr 24th, 2003, Nice, France.

17. *"High Performance User-level Sockets over Gigabit Ethernet"*. IEEE International Conference on Cluster Computing, Sep 25th, 2002, Chicago, Illinois.

Tutorials

18. D. K. Panda, **P. Balaji** and S. Sur. "Designing Next Generation Clusters, File/Storage Systems and Data-Centers with InfiniBand". *IEEE International Conference on Supercomputing (SC)*, 13th Nov, 2005, Seattle, Washington

Demo/Project Showcase

19. *"mpiBLAST on the GreenGene Distributed Supercomputer: Sequencing the NT database against the NT database (An NT-complete problem)"*. In conjunction with the Los Alamos National Laboratory at the IEEE International Conference on Supercomputing (SC), Nov 14th to 17th, 2005, Seattle, Washington.

20. *"InfiniBand: Performance and Implications on Next Generation Applications"*. Exhibition Demo at the 10th Annual Coalition for National Science Funding Science Exhibition and Reception, Jun 23rd, 2003, Capitol Hill, Washington DC.

PROFESSIONAL ACTIVITIES

Session Chair Positions

1. IEEE International Conference on Computer Communications and Networks (ICCCN '06)

Program Committee Member Positions

1. International Conference on Information Science and Security (ICISS '07)

2. IEEE International Conference on Computer Communications and Networks (ICCCN '06)

Awards Committees

1. Departmental Awards Committee, Computer Science and Engineering, Ohio State University, 2006.

Technical Referee Positions

1. Transactions of Parallel and Distributed Systems (2006)
2. Transactions on Computers (2006)
3. Electronics and Telecommunications Research Institute Journal (2006)
4. Journal of Grid Computing (2005)
5. Journal of Parallel and Distributed Computing (2004, 2005)
6. IEEE International Conference on Supercomputing (SC' 05)
7. IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS '05)
8. IEEE Symposium on High-Performance Computer Architecture (HPCA '05)
9. IEEE Symposium on High Performance Distributed Computing (HPDC '04 and HPDC '05)
10. IEEE International Parallel and Distributed Processing Symposium (IPDPS '04, IPDPS '05 and IPDPS '06)
11. IEEE International Conference on Cluster Computing (2003, 2004 and 2005)
12. Workshop on Communication Architecture for Clusters (CAC '03, CAC '04 and CAC '05)
13. Workshop on RDMA: Applications, Implementations and Technologies (RAIT '04 and RAIT '05)

Professional Memberships

1. Member, IEEE Computer Society
2. Member, ACM Computer Society

REFERENCES

Available on Request