

## Raghunath Raja Chandrasekar, Ph.D.

---

CONTACT INFORMATION	6802 Telluride Trail Austin, TX 78749	+1-(614)-364-5080 rajachan (at) protonmail.com
RESEARCH INTERESTS	Designing highly-scalable systems; efficient storage, I/O and parallel file systems; resilience and fault-tolerance for High-Performance Computing (HPC) architectures; modern Non-Volatile Memory (NVM) architectures; energy-efficiency; profiling tools; parallel application performance optimization and tuning;	
EXPERIENCE	<b>Senior Software Engineer IV</b> Advanced Development Group - Storage R&D, Cray Inc.	July 2016 to Present
	<b>Software Engineer III</b> Storage R&D, Cray Inc.	Feb 2015 to July 2016
	<b>Research Scholar</b> Institute for Scientific Computing and Research, Lawrence Livermore National Laboratory	Summer 2012 & 2013
	<b>Graduate Research Assistant</b> Network-Based Computing Laboratory, The Ohio State University	Dec 2009 to Dec 2014
	<b>Research Assistant</b> Department of Information Technology, Sri Venkateswara College of Engineering Anna University	Aug 2007 to Mar 2009
EDUCATION	<b>The Ohio State University</b> , Columbus, OH, USA Ph.D., Computer Science and Engineering, 2014 <ul style="list-style-type: none"><li>Thesis: <i>Designing Scalable and Efficient I/O Middleware for Fault-Resilient High-Performance Computing Clusters</i></li></ul> <b>Anna University</b> , Chennai, India B.Tech., Information Technology, 2009	
PRIMARY PROJECTS	<b>Advanced Development at Cray</b> <ul style="list-style-type: none"><li>Developed prototype scale-out shared-nothing storage system with scaling and MTBF design targets of Exascale systems. Technologies used/evaluated: NoSQL DBs (Cassandra, RAMCloud, etc.); Hashing and distributed consensus algorithms; Distributed Object Stores (Ceph, Lustre, etc.); Persistent memory technologies (NVML suite, NVDIMM drivers, etc.); Flash technologies (SPDK, Open-Channel SSDs, NVMe); FUSE for userspace FS development; Low-latency fabric middleware (RDMA concepts, libfabric, UCX, etc.);</li><li>Developing core functionality for the Cray DataWarp Burst-Buffer product's <i>Single-Namespace Transparent Caching</i> capability. Technologies used/developed: Kernel VFS and WrapFS for stackable FS development; Cray Data Virtualization Service (DVS) for I/O scaling; Lustre networking API (LNet) for IPC; NVMe SSDs for metadata and object management; Workload manager (SLURM/PBS/MOAB) APIs to interact with batch jobs;</li><li>Developed solution for the Linux MD-RAID Write-hole problem. Technologies used/developed: Enhanced Linux RAID driver; Custom-developed kernel persistent memory driver; NVDIMM and NTB frameworks in the kernel for dirty-data caching and mirroring; Highly-dense dual-controller hardware platform;</li></ul>	2015-Present

## MVAPICH MPI

2009-2014

- A core developer in the MVAPICH MPI project, the software deliverables from which are being used by more than 2,000 organizations in over 70 countries, including many Top500 supercomputers
- Designed and implemented various Checkpoint-Restart modules and capabilities
- Implemented the support for MPI-3.0 MPI Tools Information Interface
- Designed and implemented job-startup related features
- Designed a performance-tuning framework that optimizes the multitude of thresholds and parameters used to control MVAPICH's runtime behavior across system architectures
- Involved in several major/minor release cycles (1.4 through 2.1)
- MVAPICH team's delegate in the MPI standardization forum

## Scalable Checkpoint-Restart (SCR)

2012-Present

- Designed and developed the CRUISE file system that leverages byte-addressable persistent memory for scalable and efficient in-memory checkpointing
- Integrated SCR with the MVAPICH MPI library to offer a hybrid, multi-level, transparent checkpointing scheme to HPC applications

## High-Performance File Systems

2009-2014

- Designed and implemented a user-space file system for pipelined data-staging using high-performance SSDs and Infiniband RDMA
- Designed a QoS-aware file system for that minimizes network contention
- Design and implementation of CRFS, a user-space file system for Checkpoint/Restart applications. This file system is being developed as an open-source package, and is integrated into the MVAPICH software stack
- Co-designed DUFFS, a file system that was developed to study the benefits of a distributed metadata processing layer for parallel file systems

## RESEARCH PUBLICATIONS

1. An Exploration into Object Storage for Exascale Supercomputers,  
**R. Rajachandrasekar**, L. Evans, and R. Wespetal, *Cray User Group Conference (To Appear)*, May 2017
2. Power-Check: An Energy-Efficient Checkpointing Framework for HPC Clusters,  
**R. Rajachandrasekar**, A. Venkatesh, K. Hamidouche, and D. K. Panda, *IEEE / ACM Int'l Symposium on Cluster, Cloud and Grid Computing*, May 2015
3. Understanding the Memory-Utilization of MPI Libraries: Challenges and Designs in Implementing the MPI\_T Interface,  
**R. Rajachandrasekar**, J. Perkins, K. Hamidouche, M. Arnold and D. K. Panda, *ACM European MPI Users' Group Meeting (EuroMPI/ASIA)*, Sept. 2014
4. MIC-Check: A Distributed Checkpointing Framework for the Intel Many Integrated Cores Architecture,  
**R. Rajachandrasekar**, S. Potluri, A. Venkatesh, K. Hamidouche, M. W. Rahman and D. K. Panda, *ACM Symposium on High-Performance Parallel and Distributed Computing*, June 2014
5. A 1 PB/s File System to Checkpoint Three Million MPI Tasks,  
**R. Rajachandrasekar**, A. Moody, K. Mohror and D. K. Panda, *ACM Symposium on High-Performance Parallel and Distributed Computing*, June 2013
6. Thinking Beyond the RAM Disk for In-Memory Checkpointing of HPC Applications,  
**R. Rajachandrasekar**, A. Moody, K. Mohror and D. K. Panda, *OSU Tech. Report (OSU-CISRC-1/13-TR02)*, Jan. 2013

7. Minimizing Network Contention in InfiniBand Clusters with a QoS-Aware Data-Staging Framework,  
**R. Rajachandrasekar**, J. Jaswani, H. Subramoni and D. K. Panda, *IEEE Cluster*, Sept. 2012
8. Monitoring and Predicting Hardware Failures in HPC Clusters with FTB-IPMI,  
**R. Rajachandrasekar**, X. Bessonon and D. K. Panda, *Workshop on System Management Techniques, Processes, and Services (SMTPS), held in conjunction with IPDPS '12*, May 2012
9. Can Checkpoint/Restart Mechanisms Benefit from Hierarchical Data Staging?,  
**R. Rajachandrasekar**, X. Ouyang, X. Bessonon, V. Meshram and D. K. Panda, *Workshop on Resiliency in High Performance Computing in Clusters, Clouds, and Grids (Resilience), held in conjunction with EuroPar*, Aug. 2011
10. High-Performance Design of YARN MapReduce on Modern HPC Clusters with Lustre and RDMA,  
M. W. Rahman, X. Lu, N. S. Islam, **R. Rajachandrasekar** and D. K. Panda, *IEEE Int'l Parallel and Distributed Processing Symposium (IPDPS)*, May. 2015
11. In-Memory I/O and Replication for HDFS with Memcached: Early Experiences,  
N. S. Islam, X. Lu, M. W. Rahman, **R. Rajachandrasekar** and D. K. Panda, *IEEE BigData*, Oct. 2014
12. MapReduce over Lustre: Can RDMA-based Approach Benefit?,  
M. W. Rahman, X. Lu, N. S. Islam, **R. Rajachandrasekar** and D. K. Panda, *Int'l European Conference on Parallel Processing (Euro-Par)*, Aug. 2014
13. High Performance Alltoall and Allgather designs for InfiniBand MIC Clusters,  
A. Venkatesh, S. Potluri, **R. Rajachandrasekar**, M. Luo, K. Hamidouche and D. K. Panda, *IEEE Int'l Parallel and Distributed Processing Symposium (IPDPS)*, May. 2014
14. High Performance RDMA-Based Design of HDFS over InfiniBand,  
N. S. Islam, M. W. Rahman, J. Jose, **R. Rajachandrasekar**, H. Wang, H. Subramoni, C. Murthy and D. K. Panda, *Int'l Conference on Supercomputing (SC '12)*, Nov. 2012
15. SSD Assisted Hybrid Memory to Accelerate Memcached over High Performance Networks,  
X. Ouyang, N. Islam, **R. Rajachandrasekar**, J. Jose, M. Luo, H. Wang and D. K. Panda, *Int'l Conference on Parallel Processing (ICPP)*, Sept. 2012
16. CRFS: A Lightweight User-Level Filesystem for Generic Checkpoint/Restart,  
X. Ouyang, **R. Rajachandrasekar**, X. Bessonon, Hao Wang, Jian Huang and D. K. Panda, *Int'l Conference on Parallel Processing (ICPP)*, Sept. 2011
17. Can a Decentralized Metadata Service Layer benefit Parallel Filesystems?,  
V. Meshram, X. Bessonon, X. Ouyang, **R. Rajachandrasekar**, R. P. Darbha and D. K. Panda, *Workshop on Interfaces and Architectures for Scientific Data Storage (IASDS), held in conjunction with Cluster '11*, Sept. 2011
18. High Performance Pipelined Process Migration with RDMA,  
X. Ouyang, **R. Rajachandrasekar**, X. Bessonon and D. K. Panda, *IEEE Int'l Symposium on Cluster, Cloud and Grid Computing (CCGrid)*, May. 2011
19. RDMA-Based Job Migration for MPI over InfiniBand,  
X. Ouyang, S. Marcarelli, **R. Rajachandrasekar** and D. K. Panda, *IEEE Cluster*, Sept. 2010

20. Job Submission to Grids using Mobile Device Interface,  
**R. Rajachandrasekar**, R. Nagarajan, G. Sridar and Sumathi Ganesan, *World Congress on Nature and Biologically Inspired Computing (NaBIC)*, Dec. 2009
21. Mobile Device Interface to Access Computational Grids,  
**R. Rajachandrasekar**, R. Nagarajan, G. Sridar and Sumathi Ganesan, *Int'l Conference on Advances in Recent Technologies in Communication and Computing (ARTCom)*, Oct. 2009

AWARDS	• Best Poster award, ACM HPDC	2014
	• Best Presentation finalist, ACM HPDC	2013
	• Best Paper nomination, ACM HPDC	2013
	• Best Paper nomination, ICPP	2012

Travel Awards

• Student Travel Grant, ACM HPDC, Vancouver, BC, Canada	2014
• Student Travel Grant, ACM HPDC, New York, USA	2013
• TPC's Student Travel Grant, IEEE Cluster, Beijing, China	2012
• Student Grant, USENIX FAST, San Jose, CA, USA	2012
• Student Volunteer Grant, IEEE/ACM Supercomputing, Seattle, WA, USA	2011
• Student Travel Grant, IEEE Cluster, Austin, TX, USA	2011

PROFESSIONAL SERVICE

Technical Program Committees Served

• Int'l Workshop on Extreme Scale Programming Models and Middleware (In conjunction with IEEE/ACM SC'16)	2016
• Int'l Conference for High Performance Computing, Networking, Storage and Analysis (SC)	2015
• Int'l Workshop on Fault Tolerant Systems (In conjunction with IEEE Cluster)	2015-2016

Reviewing Conference and Journal Publications

• Int'l Symposium on Cluster, Cloud and Grid Computing (CCGrid)	2014
• Int'l Parallel and Distributed Processing Symposium (IPDPS)	2011-2014
• Int'l Conference on Parallel and Distributed Systems (ICPADS)	2013
• Transactions on Parallel and Distributed Systems (TPDS)	2013
• Journal of Parallel and Distributed Computing (JPDC)	2012, 2017
• Int'l Conference for High Performance Computing, Networking, Storage and Analysis (SC)	2013
• Journal of Parallel and Distributed Computing (JPDC)	2011

Judging Research

• Intel International Science and Engineering Fair (ISEF) The Ohio Academy of Science	2014
• Edward F. Hayes Graduate Research Forum The Ohio State University	2012-2014
• Richard J. and Martha D. Denman Undergraduate Research Forum The Ohio State University	2012

CSE Delegate, Council of Graduate Students, The Ohio State University 2011-2012

PROFESSIONAL MEMBERSHIPS

- IEEE Member
- IEEE Computer Society Member
- ACM Member
- SIGHPC Member