Rong Shi

CONTACT INFORMATION

Department of Computer Science and Engineering
The Ohio State University
2015 Neil Avenue
Columbus, Ohio 43210, USA

Office: Dreese Lab 190
Phone: (614) 397-2394
E-mail: rshi007@gmail.com
https://vdr007.github.io

SUMMARY OF QUALIFICATIONS

- Broad knowledge and experience on distributed system and parallel programming models
- 3+ years experience on design and implementation of distributed systems
- 2+ years experience on co-designing applications with parallel programming models
- Excellent team worker with strong communication skills

EDUCATION

The Ohio State University Columbus, OH

Ph.D. Computer Science and Engineering, 09/2011 – 08/2018

University of Electronic Science and Technology of China Chengdu, China

M.S. Computer Science and Engineering, 07/2011
B.S. Computer Science and Engineering, 06/2007

EXPERIENCE

Software Engineering Intern, Google Inc., Mountain View, CA 05/2017 – 08/2017

Worked on designing and implementing Location Extension Database (LEDB) that aggregates data from Google My Business, F1 Adwords and other sources, and serve multiple users and pipelines. All flume pipelines will be executed daily to generate snapshots and refresh LEDB dashboard metrics and statistics. The LEDB was launched in production.

Graduate Research Assistant, Advisor: Yang Wang 01/2015 – present

- **Evaluating System Scalability Bottlenecks by Workload Extrapolation**

  Designed and developed PatternMiner tool to identify and predict workload patterns for large-scale system. Built simulator to play extrapolated workload and emulate large-scale cluster with a few machines. For evaluation, applied approach to HDFS NameNode and YARN Resource Manager facilitated identifying performance bugs and bottlenecks.

- **Cheap and Available State Machine Replication (SMR)**

  Designed a general approach to reduce the cost of asynchronous SMR protocols while maintaining their high availability. Applied ideas to Paxos and built ThriftyPaxos from scratch in Java. Built a remoteHashMap benchmark, and exported the ThriftyPaxos to replicate database system H2. ThriftyPaxos achieved higher throughput and similar availability, yet with fewer replicas.

Graduate Research Assistant, Advisor: D.K. Panda 09/2012 – 11/2014

- **Message Passing Library MVAPICH2** (**0.4+ million** downloads, 2,875 organizations used)

  Involved in the design, development, testing, release and maintenance of MVAPICH2 software stacks. Participated in the design and implementation of MPI runtime (MVAPICH2-GDR) for GPU clusters. Proposed efficient data movement approach for GPU clusters using techniques like GPUDirect RDMA, pipelining and Fastcopy.

- **Hybrid High Performance Linpack (HPL) benchmark over Heterogeneous Clusters**

  Designed and implemented the hybrid HPL benchmark (the yardstick to rank the Top500 supercomputers) with C and CUDA using two-level adaptive workload scheduling and communication-aware process grid reordering to gain performance on heterogeneous clusters.

Selected Course Projects

Topic prediction on Reuters Dataset 08/2014 – 12/2014

Coded in Python to preprocess the raw data into refined dataset composed of feature vectors, build classifier models to predict the topic of each article, such as Naive Bayes and KNN.

Lisp Interpreter Simulation 01/2012 – 03/2012

Implemented a lisp interpreter composed of lexer, parser and evaluator using Java. Lexer analyzed input commands as tokens, parser built a parse tree given a stream of tokens, and then evaluator evaluated all functions and expressions.

Computer Skills

- Programming Languages: Java, C, Python, C++, SQL, Shell scripting
- Big Data Processing Frameworks: Hadoop, Spark
- Parallel Programming Models and Libraries: MPI, OpenMP, CUDA
- Database Management System: MySQL, PostgreSQL, Hive

Honors and Awards

Student Travel Grants: SOSP 2017, OSDI 2016, SOSP 2015, Cluster 2013

Best Student Paper Award: Cluster 2013

Selected Publications

Rong Shi, Yifan Gan, Yang Wang, Evaluating Scalability Bottlenecks by Workload Extrapolation, IEEE MASCOTS 2018.

Rong Shi, Yang Wang, Cheap and Available State Machine Replication, USENIX ATC 2016.

Rong Shi, Sreeram Potluri, Khaled Hamidouche, Mingzhe Li, Davide Rossetti and D. K. Panda, Designing Efficient Small Message Transfer Mechanism for Inter-node MPI Communication on InfiniBand GPU Clusters, Conference on High Performance Computing (HiPC'14), Goa, India, 2014.

Rong Shi, Xiaoyi Lu, Sreeram Potluri, Khaled Hamidouche, Jie Zhang, and D. K. Panda, HAND: A Hybrid Approach to Accelerate Non-contiguous Data Movement using MPI Datatypes on GPU Clusters, International Conference on Parallel Processing (ICPP'14), Minneapolis, USA, 2014.

Rong Shi, Sreeram Potluri, Khaled Hamidouche, Xiaoyi Lu, Karen Tomko and D. K. Panda, A Scalable and Portable Approach to Accelerate Hybrid HPL on Heterogeneous CPU-GPU Clusters. Cluster'13, Indianapolis, USA, 2013.

Professional Experience

- External reviewer of international conferences: HiPC 2018, HiPC 2017, IPDPS 2014, Cluster 2013
- ACM Student Member

Teaching Experience

Teaching Assistant at the Ohio State University, 09/2011 – 05/2014

- Instructor for CSE 2451: Advanced C Programming, Spring 2014.
- Lab instructor for CSE 2111: Spreadsheets and Databases, Autumn 2012.