

Dazhao Cheng

- CONTACT** Department of Computer Science Phone: (704) 687-8381
University of North Carolina at Charlotte Email: dazhao.cheng@uncc.edu
Charlotte, NC 28223 USA Web: <http://webpages.uncc.edu/dcheng3/>
- RESEARCH INTERESTS** Big Data Processing and Analytics,
Cloud Computing, Sustainable Datacenters,
Distributed Systems and Middleware.
- WORK EXPERIENCE** **Tenure Track Assistant Professor** **August 2016 - Present**
Department of Computer Science, University of North Carolina at Charlotte, USA
- EDUCATION** Ph.D., Computer Science, University of Colorado, Colorado Springs, May 2016
M.S., Electrical and Computer Engineering, University of Science and Technology of China, 2009
B.S., Electrical and Computer Engineering, HeFei University of Technology, China, 2006
- PUBLICATIONS** **Journal Papers**
- Dazhao Cheng**, Jia Rao, Changjun Jiang, and Xiaobo Zhou, “Elastic Power-aware Resource Provisioning of Heterogeneous Workloads in Self-Sustainable Datacenters”, *IEEE Transactions on Computers (TC)*, in Preprints, 14 pages, April 2015.
 - Dazhao Cheng**, Yanfei Guo, Changjun Jiang and Xiaobo Zhou, “Self-tuning Batching with DVFS for Performance Improvement and Energy Efficiency in Internet Servers”, *ACM Transactions on Autonomous and Adaptive Systems (TAAS)*, Volume 10 Issue 1, 32 pages, March 2015.
 - Dazhao Cheng**, Jia Rao, Yanfei Guo, Changjun Jiang and Xiaobo Zhou, “Improving Performance of Heterogeneous MapReduce Clusters with Adaptive Task Tuning”, *IEEE Transactions on Parallel and Distributed Systems (TPDS)*, 14 pages, 2016.
 - Yanfei Guo, Jia Rao, **Dazhao Cheng**, Changjun Jiang and Xiaobo Zhou, “iShuffle: Improving Hadoop Performance with Shuffle-on-Write”, *IEEE Transactions on Parallel and Distributed Systems (TPDS)*, 14 pages, 2016.
- Conference Papers**
- Dazhao Cheng**, Palden Lama, Changjun Jiang, and Xiaobo Zhou, “Towards Energy Efficiency in Heterogeneous Hadoop Clusters by Adaptive Task Assignment”, in *Proceedings of the 35th IEEE International Conference on Distributed Computing Systems (ICDCS)*, 10 pages, Columbus, June 2015 (Acceptance rate 12.8%)
 - Dazhao Cheng**, Jia Rao, Changjun Jiang, and Xiaobo Zhou, “Resource and Deadline-aware Job Scheduling in Dynamic Hadoop Clusters”, in *Proceedings of the 29th IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, 10 pages, Hyderabad, May 2015 (Acceptance rate 22%)
 - Yanfei Guo, Jia Rao, **Dazhao Cheng**, Changjun Jiang, Cheng-Zhong Xu and Xiaobo Zhou, “StoreApp: A Shared Storage Appliance for Efficient and Scalable Virtualized Hadoop Clusters”,

in *Proceedings of the 34th IEEE Conference on Computer Communications (INFOCOM)*, 9 pages, Hong Kong, China, April 2015 (Acceptance rate 19%)

8. Dazhao Cheng, Jia Rao, Yanfei Guo and Xiaobo Zhou, “Improving MapReduce Performance in Heterogeneous Environments with Adaptive Task Tuning”, in *Proceedings of the 15th ACM/IFIP/USENIX International Conference on Middleware (Middleware)*, 12 pages, Bordeaux, France, December 2014 (Acceptance rate 18%)

9. Dazhao Cheng, Changjun Jiang, and Xiaobo Zhou, “Heterogeneity-aware Workload Placement and Migration in Distributed Sustainable Datacenters”, in *Proceedings of the 28th IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, 10 pages, Phoenix, May 2014 (Acceptance rate 21%)

10. Dazhao Cheng, Yanfei Guo and Xiaobo Zhou, “Self-tuning Batching with DVFS for Improving Performance and Energy Efficiency in Servers”, in *Proceedings of the 21th ACM/IEEE International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS)*, 10 pages, San Francisco, August 2013 (Acceptance rate 27%)

Posters and Patents

11. Dazhao Cheng and Xiaobo Zhou, “Dynamic Resource Provisioning and Workload Management in Sustainable Datacenters”, in *Proceedings of the 28th IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, PhD Forum, 2 pages, Phoenix, May 2014

12. Dazhao Cheng and Shengxiao Guan, etc. “Reset Time Detector of Elevator Buffer” Publication number: CN101580197 B, <http://www.google.com/patents/CN101580197B>, September, 2011

13. Dazhao Cheng Chi Chen, Ling Li and Shengxiao Guan, etc. “Proximity Sense Sensor Device for Contact State Test of Lift Buffer and Lift Car” Publication number: CN101654206 B, <http://www.google.com/patents/CN101654206B>, October, 2011

Papers under Peer Reviewing

14. Dazhao Cheng and Xiaobo Zhou, “Cross-platform Resource Scheduling for Spark-on-YARN”, under reviewing in *IEEE Transactions on Computers (TC)*.

RESEARCH
EXPERIENCE

University of Colorado

Distributed, Sustainable, and Cloud Computing Systems Lab

Supervisor: Xiaobo Zhou, Ph.D.

Research Assistant

August, 2011 - May 2016

Middleware Approaches to Predictable MapReduce Services in Multi-Tenant Clouds

Compared to in-house dedicated clusters, the deployment of MapReduce in datacenters and clouds often exhibits significant hardware and performance heterogeneity due to continuous server replacement and multi-tenant interferences. Co-sponsored by NSF grants CNS-1422119 “CSR: Moving MapReduce into the Cloud: Flexibility, Efficiency, and Elasticity”, I proposed a task tuning approach that allowed tasks to have different configurations, each optimized for the actual hardware capabilities, on heterogeneous nodes. Furthermore, I proposed a resource and deadline-aware MapReduce job scheduler that took the cluster resource availability into consideration when minimizing job deadline misses. In this area of research, I also proposed a heterogeneity-aware task assignment approach that aimed to minimize the overall energy consumption in a heterogeneous MapReduce cluster without sacrificing job performance.

Resource Scheduling and Power Management for Sustainable Cloud Computing

Sustainable cloud computing aims to adaptively operate datacenters efficiently and manage their

power demands in order to reduce carbon emissions without compromising the performance constraints of clients. Sponsored by NSF CAREER award CNS-0844983 and grant CNS-1217979, I developed an autonomic power-aware resource provisioning approach for heterogeneous workloads in sustainable datacenters that completely relied on renewable energy, which aimed to maximize the system performance and control the system power consumption with respect to green power supply. I further designed a heterogeneity-aware cloud workload placement and migration approach that aimed to maximize the overall system performance in distributed sustainable datacenters. In order to improve energy efficiency of datacenters, I designed a self-tuning request batching mechanism to simultaneously achieve performance improvement and energy efficiency in provisioning Internet services of datacenter servers.

Research Intern at Hewlett-Packard Labs
Systems Research Lab, Palo Alto, CA
 Mentor: Yuan Chen, Ph.D.

Research Associate
 June, 2015 - September, 2015

Data Dependency-aware Scheduling in Spark Streaming

As Spark Streaming had been increasingly adopted by industries for real time big data analysis, I found that the current job scheduler of Spark Streaming was inefficient while scheduling microbatch jobs of streaming workloads. In this project, I proposed and implemented a new scheduler for Spark Streaming workloads. It improved throughput, latency and energy efficiency by scheduling concurrent jobs and applying different scheduling policies based on their data dependencies.

PRESENTATIONS

1. Towards Energy Efficiency in Heterogeneous Hadoop Clusters by Adaptive Task Assignment; **ICDCS**, Columbus, June 2015.
2. Resource and Deadline-aware Job Scheduling in Dynamic Hadoop Clusters; **IPDPS**, Hyderabad, May 2015.
3. StoreApp: A Shared Storage Appliance for Efficient and Scalable Virtualized Hadoop Clusters; **INFOCOM**, Hong Kong, April 2015.
4. Heterogeneity-aware Workload Placement and Migration in Distributed Sustainable Datacenters; **IPDPS**, Phoenix, May 2014.
5. Self-tuning Batching with DVFS for Improving Performance and Energy Efficiency in Servers; **MASCOTS**, San Francisco, August 2013.

TEACHING EXPERIENCES

- Instructor of ITCS3166: Intro to Computer Networks in UNC Charlotte, Fall 2016
- Teaching Assistant of Principle of Computer Science and Operating Systems in UCCS

TEACHING INTERESTS

- Operating Systems, Computer Architecture, Data Structures and Algorithms
- Cloud Computing and Data Centers, Big Data Processing, Computer Networks
- Distributed and Parallel Systems, Feedback Control in Computer Systems

TECHNICAL SKILLS

- Big Data Skills: Hadoop, MapReduce, YARN, Spark, Spark Streaming
- Programming Languages: Java, C/C++, Python, Scala, Shell Script
- Virtualization Tools: VMware, Xen, Docker

PROFESSIONAL SERVICE

- Reviewer for IEEE Transactions on Cloud Computing
- Reviewer for IEEE Transactions on Parallel and Distributed Systems
- Reviewer for IEEE Conference on Computer Communications (INFOCOM'14, 15 and 16)
- Reviewer for USENIX International Conference on Autonomic Computing (ICAC'13, 14 and 15)