Contents

Topical Section Headings: Benchmarking

On Big Data Benchmarking ...................................................... 3
  Rui Han, Xiaoyi Lu, and Jiangtao Xu

A Micro-benchmark Suite for Evaluating Hadoop MapReduce on High-Performance Networks .............................................. 19

MemTest: A Novel Benchmark for In-memory Database ............. 34
  Qiangqiang Kang, Cheqing Jin, Zhao Zhang, and Aoying Zhou

DSIMBench: A Benchmark for Microarray Data Using R .......... 47
  Shicai Wang, Ioannis Pandis, Ibrahim Emam, David Johnson, Florian Guitton, Axel Oehmichen, and Yike Guo

A Benchmark to Evaluate Mobile Video Upload to Cloud Infrastructures ................................................................. 57
  Afsin Akdogan, Hien To, Seon Ho Kim, and Cyrus Shahabi

Benchmarking Replication and Consistency Strategies in Cloud Serving Databases: HBase and Cassandra .................. 71
  Huajin Wang, Jianhui Li, Haiming Zhang, and Yuanchun Zhou

Topical Section Headings: Workload Characterization

I/O Characterization of Big Data Workloads in Data Centers .......... 85
  Fengfeng Pan, Yinliang Yue, Jin Xiong, and Daxiang Hao

Characterizing Workload of Web Applications on Virtualized Servers .... 98
  Xiajun Wang, Song Huang, Song Fu, and Krishna Kavi

Topical Section Headings: Performance Optimization and Evaluation

Performance Benefits of DataMPI: A Case Study with BigDataBench ...... 111
  Fan Liang, Chen Feng, Xiaoyi Lu, and Zhiwei Xu

InvarNet-X: A Comprehensive Invariant Based Approach for Performance Diagnosis in Big Data Platform .................. 124
  Pengfei Chen, Yong Qi, Di Hou, and Huachong Sun
Tuning Hadoop Map Slot Value Using CPU Metric .......................... 141
  Kamal Kc and Vincent W. Freeh

A Study of SQL-on-Hadoop Systems .......................... 154
  Yueguo Chen, Xiongpai Qin, Haoqiong Bian, Jun Chen, Zhaoan Dong,
  Xiaoyong Du, Yanjie Gao, Dehai Liu, Jiaheng Lu, and Huijie Zhang

Predoop: Preempting Reduce Task for Job Execution Accelerations. .... 167
  Yi Liang, Yufeng Wang, Minglu Fan, Chen Zhang, and Yuqing Zhu

Record Placement Based on Data Skew Using Solid State Drives .... 181
  Jun Suzuki, Shivaram Venkataraman, Sameer Agarwal,
  Michael Franklin, and Ion Stoica

Efficient HTTP Based I/O on Very Large Datasets for High Performance
  Computing with the Libdavix Library .......................... 194
  Adrien Devresse and Fabrizio Furano

Topical Section Headings: Emerging Hardware

Exploring Opportunities for Non-volatile Memories in Big Data Applications ... 209
  Wei Wei, Dejun Jiang, Jin Xiong, and Mingyu Chen

Author Index .......................................................... 221
Big Data Benchmarks, Performance Optimization, and Emerging Hardware
4th and 5th Workshops, BPOE 2014, Salt Lake City, USA, March 1, 2014 and Hangzhou, China, September 5, 2014, Revised Selected Papers
Zhan, J.; Han, R.; Weng, C. (Eds.)
2014, X, 221 p. 100 illus., Softcover
ISBN: 978-3-319-13020-0