

Contents

Benchmarking

| | |
|---|----|
| Revisiting Benchmarking Principles and Methodologies for Big Data Benchmarking | 3 |
| <i>Liutao Zhao, Wanling Gao, and Yi Jin</i> | |
| BigDataBench-MT: A Benchmark Tool for Generating Realistic Mixed Data Center Workloads | 10 |
| <i>Rui Han, Shulin Zhan, Chenrong Shao, Junwei Wang, Lizy K. John, Jiangtao Xu, Gang Lu, and Lei Wang</i> | |

Benchmarking and Workload Characterization

| | |
|--|----|
| Towards a Big Data Benchmarking and Demonstration Suite for the Online Social Network Era with Realistic Workloads and Live Data | 25 |
| <i>Rui Zhang, Irene Manotas, Min Li, and Dean Hildebrand</i> | |
| On Statistical Characteristics of Real-Life Knowledge Graphs. | 37 |
| <i>Wenliang Cheng, Chengyu Wang, Bing Xiao, Weining Qian, and Aoying Zhou</i> | |
| Mbench: Benchmarking a Multicore Operating System Using Mixed Workloads | 50 |
| <i>Gang Lu, Xinlong Lin, and Runlin Zhou</i> | |

Performance Optimization and Evaluation

| | |
|--|----|
| Evolution from Shark to Spark SQL: Preliminary Analysis and Qualitative Evaluation | 67 |
| <i>Xinhui Tian, Gang Lu, Xiexuan Zhou, and Jingwei Li</i> | |
| How Data Volume Affects Spark Based Data Analytics on a Scale-up Server | 81 |
| <i>Ahsan Javed Awan, Mats Brorsson, Vladimir Vlassov, and Eduard Ayguade</i> | |
| An Optimal Reduce Placement Algorithm for Data Skew Based on Sampling | 93 |
| <i>Zhuo Tang, Wen Ma, Rui Li, Kenli Li, and Keqin Li</i> | |

| | |
|--|-----|
| AAA: A Massive Data Acquisition Approach in Large-Scale System Monitoring | 107 |
| <i>Runlin Zhou, Yanfei Lv, Dongjin Fan, Hong Zhang, and Chungu Zhu</i> | |

Emerging Hardware

| | |
|--|-----|
| A Plugin-Based Approach to Exploit RDMA Benefits for Apache and Enterprise HDFS. | 119 |
| <i>Adithya Bhat, Nusrat Sharmin Islam, Xiaoyi Lu, Md. Wasi-ur-Rahman, Dipti Shankar, and Dhabaleswar K. (DK) Panda</i> | |

| | |
|--|-----|
| Stream-Based Lossless Data Compression Hardware Using Adaptive Frequency Table Management | 133 |
| <i>Shinichi Yamagiwa, Koichi Marumo, and Hiroshi Sakamoto</i> | |

| | |
|-------------------------------|-----|
| Author Index | 147 |
|-------------------------------|-----|

Big Data Benchmarks, Performance Optimization, and
Emerging Hardware

6th Workshop, BPOE 2015, Kohala, HI, USA, August 31 -
September 4, 2015. Revised Selected Papers

Zhan, J.; Han, R.; Zicari, R.V. (Eds.)

2016, X, 147 p. 53 illus. in color., Softcover

ISBN: 978-3-319-29005-8