Bench 2018 Program

Conference Program

Day 1 (Mon, Dec 10th, 2018)

8:30-8:35 Opening Remarks (Dr. Chen Zheng)

8:35-9:00 **Prof. Jianfeng Zhan, Institute of Computing Technology, CAS**

Benchmarking Opportunities and Challenges: present and future of BenchCouncil

9:05-10:05 ***Keynote I: Big Data Benchmarking: Applications and Systems***

Prof. Geoffrey Fox, Indiana University, APS and ACM Fellow.

Abstract: We classify applications and identify key features that affect performance. We relate these to features in the hardware and software used in Big Data Analysis. We suggest issues in benchmarking that emerge from this study and work done in BDEC2 -- https://www.exascale.org/bdec/- on identifying consensus around a shared cyberinfrastructure for science in the data saturated world that is now emerging.

Bio: Geoffrey Charles Fox (https://www.engineering.indiana.edu/, http://www.dsc.soic.indiana.edu/, gcf@indiana.edu)

Fox received a Ph.D. in Theoretical Physics from Cambridge University where he was Senior Wrangler. He is now a distinguished professor of Engineering, Computing, and Physics at Indiana University where he is director of the Digital Science Center. He previously held positions at Caltech, Syracuse University, and Florida State University after being a postdoc at the Institute for Advanced Study at Princeton, Lawrence Berkeley Laboratory, and Peterhouse College Cambridge. He has supervised the Ph.D. of 73 students and published around 1300 papers (over 500 with at least ten citations) in physics and computing with an hindex of 78 and over 35000 citations. He is a Fellow of APS (Physics) and ACM (Computing) and works on the interdisciplinary interface between computing and applications. Current work is in Biology, Pathology, Sensor Clouds and Ice-sheet Science, Image processing, Deep Learning and Particle Physics. His architecture work is built around High-performance computing enhanced Software Defined Big Data Systems on Clouds and Clusters. The analytics focuses on scalable parallel machine learning. He is an expert on streaming data and robot-cloud interactions. He is involved in several projects to enhance the capabilities of Minority Serving Institutions. He has experience in online education and its use in MOOCs for areas like Data and Computational Science.

10:00-10:20 coffee Break

10:20-10:40 **Session 1: Cloud I**

10:40-12:10 ***BenchCouncil: Benchmarking proposal***

10:40-11:10 **Dr. Wanling Gao, ICT, CAS**

DataMotif: A Benchmark Proposal for Big Data and AI

Bio: Wangling Gao is a Ph.D candidate in computer science at the Institute of Computing Technology, Chinese Academy of Sciences and University of Chinese Academy of Sciences. Her research interests focus on big data benchmark and big data analytics. She received her B.S. degree in 2012 from Huazhong University of Science and Technology.

11:10-11:40 **Prof. Xiaoyi Lu, The Ohio State University**

A Benchmark proposal for Deep Learning Benchmarks

Bio: Dr. Xiaoyi Lu is a Research Scientist in the Department of Computer Science and Engineering at the Ohio State University, USA. His current research interests include high performance interconnects and protocols, Big Data, Deep Learning, Hadoop/Spark/Memcached/TensorFlow Ecosystem, Parallel Computing Models (MPI/PGAS), Virtualization, and Cloud Computing. He has published over 90 papers in International journals and conferences related to these research areas. He has been actively involved in various professional activities (PC Co-Chair, PC Member, Invited Reviewer) in academic journals and conferences. Recently, Dr. Lu is leading the research and development of RDMA-based accelerations for Apache Hadoop, Spark, HBase, and Memcached, and OSU HiBD micro-benchmarks, which are publicly available from (http://hibd.cse.ohio-state.edu). These libraries are currently being used by more than 275 organizations from 34 countries. More than 25,500 downloads of these libraries have taken place from the project site. He is a core member of the MVAPICH2 (High-Performance MPI over InfiniBand, Omni-Path, Ethernet/iWARP, and RoCE) project and he is leading the research and development of MVAPICH2-Virt (high-performance and scalable MPI for hypervisor and container based HPC cloud). He is a member of IEEE and ACM. More details about Dr. Lu are available at http://web.cse.ohio-state.edu/~lu.932.

11:40-12:10 **Discussion**

12:00 Lunch

13:30-14:50 ***BenchCouncil: Benchmarking proposal***

13:30-14:00 **Dr. Chen Zheng, ICT, CAS**

A Benchmark proposal for Datacenter Computing

Bio: Chen Zheng is a post doc researcher at the Institute of Computing Technology, Chinese Academy of Sciences and University of Chinese Academy of Sciences. His research focuses on Operating System, Virtualization, benchmarks, and data center workload characterization. He received his PHD degree in 2017 from Institute of Computing Technology in China.

14:00-14:30 **Prof. Weining Qian, East China Normal University**

PeakBench: A Benchmark Proposal for Scalable Transaction Processing

Bio: Prof. Weining Qian is the Dean and a Professor at School of Data Science and Engineering (DaSE), East China Normal University. He was a Lecturer at Department of Computer Science and Engineering, Fudan University, from 2004 to 2006. He got my Ph.D. from Fudan University. His research interests include scalable transaction processing, and management & mining of massive datasets.

14:30-14:50 **Discussion**

14:50 **Session 2: Best Paper Session I**

15:30 Coffee Break

15:50 **Session 3: Best Paper Session II**

16:40 **Session 4: Big Data**

17:40 **Session 5: Modeling and Prediction**

18:20 End

Day 2 (Tue, Dec 11th, 2018)

8:35 Opening and Welcome

8:45-9:45 ***Keynote II:***

MLPerf: The Vision Behind an ML Benchmark Suite for Measuring the Performance of ML Software Frameworks, ML Hardware Accelerators, and ML Cloud and Edge Platforms

Prof. Vijay Janapa Reddi, Harvard University

Bio: Vijay Janapa Reddi is currently an Associate Professor in the John A. Paulson School of Engineering and Applied Sciences (SEAS) at Harvard University. His research interests include computer architecture and software design to enhance mobile and high-performance computing systems,

specifically focusing on always-on computing and end-user experience for mobile devices and energy efficiency and reliability for heterogeneous system architectures. Dr. Janapa Reddi is a recipient of multiple awards, including the National Academy of Engineering (NAE) Gilbreth Lecturer Honor (2016), IEEE TCCA Young Computer Architect Award (2016), Intel Early Career Award (2013), Google Faculty Research Awards (2012, 2013, 2015, 2017), Best Paper at the 2005 International Symposium on Microarchitecture, Best Paper at the 2009 International Symposium on High Performance Computer Architecture, and IEEE's Top Picks in Computer Architecture awards (2006, 2010, 2011, 2016, 2017). Beyond his technical research contributions, Dr. Janapa Reddi is passionate about STEM education. He is responsible for the Austin Independent School District's "hands-on" computer science (HaCS) program, which teaches 6th- and 7th-grade students programming and the general principles that govern a computing system using open-source electronic prototyping platforms. He received a BS in computer engineering from Santa Clara University, an MS in electrical and computer engineering from the University of Colorado at Boulder, and a PhD in computer science from Harvard University.

9:45-10:05 coffee Break

10:05-10:35 **Invited talk: Dr. Arne Berre, SINTEF Digital**

Benchmarking for Digital Platforms with Big Data, IoT, AI, Cloud, HPC and CyberSecurity

Abstract: The European Big Data Value Association (BDVA, www.bdva.eu) has created a reference model for big data and digital platforms that identifies a number of technical areas relevant for benchmarking and standardisation within Big data and AI, such as data management, data processing, data protection, data analytics with machine learning and AI, and data visualisation.

The reference model shows the relationships to related boundary areas and organisations, such IoT with the Alliance of Internet of Things Innovation (AIOTI, www.aioti.eu), Cloud with the Networked Software and Services Initiative (NESSI, www.nessi-europe.com) , HPC with the European Technology Platform for High Performance Computing (ETP4HPC, www.etp4hpc.eu) and CyberSecurity with the European CyberSecurity Organisation (ECSO, www.ecs-org.eu)

The European Horizon 2020 project DataBench www.databench.eu aims at supporting a Benchmarking Community of existing benchmarks and benchmarking initiatives that together can be a foundation for both business oriented and technical benchmarks within the digital platforms areas of Big Data and AI.

Bio: Dr. Arne Berre is working with Digital Platforms and Systems Interoperability, with focus on Big Data and Cloud/HPC processing support for Analytics/AI/Machine Learning. He is Chief Scientist at SINTEF Digital, Department for Software and Service Innovation, Group for Smart Data in Oslo, Norway and the leader of European BDVA (Big Data Value Association) TF6 Technical Priorities, GEMINI Center for Big Data with SINTEF, NTNU and the University of Oslo. He is the Norwegian representative to the Big Data group within ISO JTC1/SC42 Artificial Intelligence. He was responsible for the HyperModel Benchmark as one of the earlier benchmarks for object-oriented and graph-oriented databases.

10:35-11:55 ***BenchCouncil: Benchmarking proposal***

10:35-11:05 **Prof. Yueguo Chen, Renmin University of China**

TS-benchmark: a benchmark proposal for time series databases

Bio: PhD., Professor of Information School, Renmin University of China. His research interests include big data analytics and semantic search. He has published more than 30 papers on top conferences and journals such as SIGMOD, SIGIR, ICDE, AAAI, IEEE TKDE, and WWW.

11:05-11:35 **Prof. Zhiyuan Chen, Prof. Jianwu Wang, University of Maryland, Baltimore County**

A Benchmark proposal for large-scale and high-speed spatiotemporal data processing and analytic

Bio: Dr. Zhiyuan Chen is an associate professor at information systems department, UMBC. He has a PhD in Computer Science from Cornell University. His research interests include big data processing and performance tuning, data navigation and visualization, privacy preserving data mining, database compression, and cyber security. He has published over 40 research articles in peer reviewed journals and conferences and has managed a number of research projects funded by NSF, Navy, IBM, and MITRE.

11:35-11:55 **Discussion**

11:55-12:20 **BenchCouncil Open Meeting: Formation of working groups**

12:20 Lunch

13:30-14:10 **Session 6: Algorithms and Implementations**

14:10-15:10 **Session 7: Cloud II**

15:10-17:30 **Tutorial: BigDataBench Tutorial: a scalable and unified Big Data and AI benchmark suite**

15:10 **Introduction of BigDataBench 4.0 & Benchmarking Methodology**

16:10 Coffee Break

16:30 **How to use BigDataBench 4.0**

17:00 **Best Paper Award and Closing the symposium**

17:15 End

Program Details

Length of presentations (including Q&A):

Keynotes: 60 minutes

Benchmark Proposal: 30 minutes

Invited talk: 30 minutes

Best Paper Candidates: 25 minutes

Regular Papers: 20 minutes

Session 1: Cloud I

Monday (Dec 10th) 9:40-10:40 (Including a Coffee Break)

Benchmarking VM Startup Time in the Cloud

Samiha Islam Abrita, Moumita Sarker, Faheem Abrar, Muhammad Abdullah Adnan

Bangladesh University of Engineering and Technology, Bangladesh

University of Saskatchewan, Canada

Session 2: Best Paper Session I

Monday (Dec 10th), 14:30-15:30

Session Chair:

DCMIX: Generating Mixed Workloads for the Cloud Data Center

XingWang Xiong, Lei Wang, WanLing Gao, Chen Zheng, Yu Wen, Yi Liang

Institute of Computing Technology, Chinese Academy of Sciences, China

College of Computer Science, Beijing University of Technology, China

EC-Bench: Benchmarking Onload and Offload Erasure Coders on Modern Hardware Architectures

Haiyang Shi, Xiaoyi Lu, Dhabaleswar K. (DK) Panda

The Ohio State University, United States

Session 3: Best Paper Session II

Monday (Dec 10th), 15:50-16:40

Testing Raft-replicated Database Systems

Guohao Ding, Weining Qian, Peng Cai, Tianze Pang, Qiong Zhao

East China Normal University, China

Bank of Communications, China

Machine-Learning Based Spark and Hadoop Workload Classification Using Container Performance Patterns

Mikhail Genkin, Frank Dehne, Pablo Navarro, Siyu Zhou

Carleton University, Canada

Session 4: Big Data

Monday (Dec 10th), 16:40-17:40

Benchmarking for Transaction Processing Database Systems in Big Data Era

Chunxi Zhang, Yuming Li, Rong Zhang, Weining Qian, Aoying Zhou

EAST CHINA NORMAL UNIVERSITY, China

UMDISW: A Universal Multi-Domain Intelligent Scientific Workflow Framework for the Whole Life Cycle of Scientific Data

Qi Sun, Yue Liu, Wenjie Tian, Yike Guo, and Bocheng Li

School of Computer Engineering and Science, Shanghai University, China

IBDB: A Benchmark Suite for Industrial Big Data System

Yijian Cheng, Mengqian Cheng, Hao Ge, Yuhe Guo, Yuanzhe Hao, Xiaoguang Sun, Xiongpai Qin, Wei Lu, and Yueguo Chen

Renmin University of China, Key Laboratory of Data Engineering and Knowledge Engineering, China

Session 5: Modeling and Prediction

Monday (Dec 10th), 17:40-18:20

Power Characterization of Memory Intensive Applications: Analysis and Implications

Yeliang Qiu, Congfeng Jiang, Yumei Wang, Youhuizi Li, and Dongyang Ou

Hangzhou Dianzi University, China

Multi-USVs coordinated detection in marine environment based on deep reinforcement learning

Ruiying Li, Rui Wang, Kai Li, Xiaohui Hu, and Haichang Li

Institute of Software, Chinese Academy of Sciences, China

Session 6: Algorithms and Implementations

Tuesday (Dec 11th), 13:30-14:10

Benchmarking SpMV on Many-core Platforms

Biwei Xie, Zhen Jia

Institute of Computing Technology, Chinese Academy of Sciences, China

Department of Computer Science, Princeton University, USA

Benchmarking Parallel Implementations of K-Means Cloud Type Clustering from Satellite Data

Carlos Barajas, Pei Guo, Lipi Mukherjee, Susan Hoban, Jianwu Wang, Daeho Jin, Aryya Gangopadhyay, and Matthias Gobbert

University of Maryland, Baltimore County, United States

GESTAR, USRA, and NASA GSFC, United States

Session 7: Cloud II

Tuesday (Dec 11th), 14:10-15:10

An Open Source Cloud-based NoSQL and NewSQL Database Benchmarking Platform for IoT Data

Arjun Pandya, Chaitanya Kulkarni, Kunal Mali, and Jianwu Wang

University of Maryland, Baltimore County, USA

Scalability Evaluation of Big Data Processing Services in Clouds

Wei Huang, Congfeng Jiang, Zujie Ren, and Jian Wan

Hangzhou Dianzi University, China

Zhejiang University of Sciences and Technology, China

PAIE: A Personal Activity Intelligence Estimator in the Cloud

Yingjie shi, Fang Du, Yanyan Zhang

Beijing Institute of Fashion Technology, China

Ningxia University, China

Chinese Academy of Social Sciences, China