

2020 IEEE International Symposium on Workload Characterization (IISWC 2020)

**Beijing, China
27 – 29 October 2020**



**IEEE Catalog Number: CFP20236-POD
ISBN: 978-1-7281-7646-8**

**Copyright © 2020 by the Institute of Electrical and Electronics Engineers, Inc.
All Rights Reserved**

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

****** This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP20236-POD
ISBN (Print-On-Demand):	978-1-7281-7646-8
ISBN (Online):	978-1-7281-7645-1

Additional Copies of This Publication Are Available From:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: (845) 758-0400
Fax: (845) 758-2633
E-mail: curran@proceedings.com
Web: www.proceedings.com

CURRAN ASSOCIATES INC.
proceedings
.com

2020 IEEE International Symposium on Workload Characterization (IISWC) IISWC 2020

Table of Contents

Message from General Chairs	ix
Message from Program Chairs	x
Program Committee	xii
Organizing Committee	xiv
Steering Committee	xv
Keynotes	xvi

Session 1: Large-Scale Systems

Pocolo: Power Optimized Colocation in Power Constrained Environments	1
<i>Iyswarya Narayanan (The Pennsylvania State University), Adithya Kumar (The Pennsylvania State University), and Anand Sivasubramaniam (The Pennsylvania State University)</i>	
HetSim: Simulating Large-Scale Heterogeneous Systems using a Trace-Driven, Synchronization and Dependency-Aware Framework	13
<i>Subhankar Pal (University of Michigan, USA), Kuba Kaszyk (University of Edinburgh, UK), Siying Feng (University of Michigan, USA), Björn Franke (University of Edinburgh, UK), Murray Cole (University of Edinburgh, UK), Michael O'Boyle (University of Edinburgh, UK), Trevor Mudge (University of Michigan, USA), and Ronald G. Dreslinski (University of Michigan, USA)</i>	
HPC-MixPBench: An HPC Benchmark Suite for Mixed-Precision Analysis	25
<i>Konstantinos Parasyris (Lawrence Livermore National Laboratory, Livermore, CA, US), Ignacio Laguna (Lawrence Livermore National Laboratory, Livermore, CA, US), Harshitha Menon (Lawrence Livermore National Laboratory, Livermore, CA, US), Markus Schordan (Lawrence Livermore National Laboratory, Livermore, CA, US), Daniel Osei-Kuffuor (Lawrence Livermore National Laboratory, Livermore, CA, US), Giorgis Georgakoudis (Lawrence Livermore National Laboratory, Livermore, CA, US), Michael O. Lam (James Madison University: Harrisonburg, VA, US), and Tristan Vanderbruggen (Lawrence Livermore National Laboratory, Livermore, CA, US)</i>	
An In-Depth Analysis of Cloud Block Storage Workloads in Large-Scale Production	37
<i>Jinhong Li (The Chinese University of Hong Kong), Qiuping Wang (The Chinese University of Hong Kong), Patrick P. C. Lee (The Chinese University of Hong Kong), and Chao Shi (Alibaba Group)</i>	

Characterizing the Scale-Up Performance of Microservices Using TeaStore	48
<i>Sriyash Caculo (Advanced Micro Devices, India), Kanishka Lahiri (Advanced Micro Devices, India), and Subramaniam Kalambur (PES University)</i>	
MATCH: An MPI Fault Tolerance Benchmark Suite	60
<i>Luanzheng Guo (University of California, Merced), Giorgis Georgakoudis (Lawrence Livermore National Laboratory), Konstantinos Parasyris (Lawrence Livermore National Laboratory), Ignacio Laguna (Lawrence Livermore National Laboratory), and Dong Li (University of California, Merced)</i>	
CPU Microarchitectural Performance Characterization of Cloud Video Transcoding	72
<i>Yuhan Chen (University of Michigan, USA), Jingyuan Zhu (University of Michigan, USA), Tanvir Ahmed Khan (University of Michigan, USA), and Baris Kasikci (University of Michigan, USA)</i>	

Session 2: Benchmarking, Evaluation, and Simulation Methodologies

A Rigorous Benchmarking and Performance Analysis Methodology for Python Workloads	83
<i>Arthur Crapé (Ghent University, Belgium) and Lieven Eeckhout (Ghent University, Belgium)</i>	
Reliability Modeling of NISQ-Era Quantum Computers	94
<i>Ji Liu (North Carolina State University, USA) and Huiyang Zhou (North Carolina State University, USA)</i>	
CPI for Runtime Performance Measurement: The Good, the Bad, and the Ugly	106
<i>Li Yi (Intel Corporation), Cong Li (Intel Corporation), and Jianmei Guo (Alibaba Cloud)</i>	
High Frequency Performance Monitoring via Architectural Event Measurement	114
<i>Chutitep Woralert (Clarkson University), James Bruska (Independent Researcher), Chen Liu (Clarkson University), and Lok Yan (Air Force Research Lab)</i>	
Port or Shim? Stress Testing Application Performance on Intel SGX	123
<i>Aisha Hasan (Carnegie Mellon University in Qatar), Ryan Riley (Carnegie Mellon University in Qatar), and Dmitry Ponomarev (Binghamton University, USA)</i>	
Characterizing the Impact of Last-Level Cache Replacement Policies on Big-Data Workloads	134
<i>Alexandre Valentin Jamet (Barcelona Supercomputing Center), Lluç Alvarez (Barcelona Supercomputing Center; Universitat Politècnica de Catalunya), Daniel A. Jiménez (Texas A&M University), and Marc Casas (Barcelona Supercomputing Center)</i>	

Session 3: Machine Learning and Systems

AI on the Edge: Characterizing AI-Based IoT Applications Using Specialized Edge Architectures	145
<i>Qianlin Liang (University of Massachusetts, Amherst), Prashant Shenoy (University of Massachusetts, Amherst), and David Irwin (University of Massachusetts, Amherst)</i>	

Cross-Stack Workload Characterization of Deep Recommendation Systems	157
<i>Samuel Hsia (Harvard University, USA), Udit Gupta (Harvard University, USA; Facebook Inc., USA), Mark Wilkening (Harvard University, USA), Carole-Jean Wu (Facebook Inc., USA), Gu-Yeon Wei (Harvard University, USA), and David Brooks (Harvard University, USA)</i>	
A Case for Generalizable DNN Cost Models for Mobile Devices	169
<i>Vinod Ganesan (Indian Institute of Technology, Madras), Surya Selvam (Indian Institute of Technology, Madras), Sanchari Sen (Purdue University), Pratyush Kumar (Indian Institute of Technology, Madras), and Anand Raghunathan (Purdue University)</i>	
Empirical Analysis and Modeling of Compute Times of CNN Operations on AWS Cloud	181
<i>Ubaid Ullah Hafeez (Stony Brook University) and Anshul Gandhi (Stony Brook University)</i>	
A Sparse Tensor Benchmark Suite for CPUs and GPUs	193
<i>Jiajia Li (Pacific Northwest National Laboratory), Mahesh Lakshminarasimhan (University of Utah), Xiaolong Wu (Purdue University), Ang Li (Pacific Northwest National Laboratory), Catherine Olschanowsky (Boise State University), and Kevin Barker (Pacific Northwest National Laboratory)</i>	
Demystifying Power and Performance Bottlenecks in Autonomous Driving Systems	205
<i>Pedro Henrique Exenberger Becker (Universitat Politècnica de Catalunya, Spain), José María Arnau (Universitat Politècnica de Catalunya, Spain), and Antonio González (Universitat Politècnica de Catalunya, Spain)</i>	

Session 4: System Architecture and Applications

Evaluation of Graph Analytics Frameworks Using the GAP Benchmark Suite	216
<i>Ariful Azad (Indiana University), Mohsen Mahmoudi Aznavah (Texas A&M University), Scott Beamer (University of California, Santa Cruz), Mark Blanco (Carnegie Mellon University), Jinhao Chen (Texas A&M University), Luke D'Alessandro (Indiana University), Roshan Dathathri (The University of Texas at Austin), Tim Davis (Texas A&M University), Kevin Deweesé (University of Washington), Jesun Firoz (Pacific Northwest National Laboratory), Henry Gabb (Intel Corporation), Gurbinder Gill (The University of Texas at Austin), Balint Hegyi (Budapest University of Technology and Economics), Scott Kolodziej (Texas A&M University), Tze Meng Low (Carnegie Mellon University), Andrew Lumsdaine (University of Washington), Tugsbayasgalan Manlaibaatar (Massachusetts Institute of Technology), Timothy Mattson (Intel Corporation), Scott McMillan (Carnegie Mellon University), Ramesh Peri (Intel Corporation), Keshav Pingali (The University of Texas at Austin), Upasana Sridhar (Carnegie Mellon University), Gabor Szarnyas (Budapest University of Technology and Economics), Yunming Zhang (Massachusetts Institute of Technology), and Yongzhe Zhang (The Graduate University for Advanced Studies, SOKENDAI)</i>	

A Study of APIs for Graph Analytics Workloads	228
<i>Hochan Lee (The University of Texas at Austin), David Wong (Intel Corporation), Loc Hoang (The University of Texas at Austin), Roshan Dathathri (The University of Texas at Austin), Gurbinder Gill (The University of Texas at Austin), Vishwesh Jatala (The University of Texas at Austin), David Kuck (Intel Corporation), and Keshav Pingali (The University of Texas at Austin)</i>	
Vertex Reordering for Real-World Graphs and Applications: An Empirical Evaluation	240
<i>Reet Barik (Washington State University, USA), Marco Minutoli (Pacific Northwest National Laboratory, USA), Mahantesh Halappanavar (Pacific Northwest National Laboratory, USA; Washington State University, USA), Nathan R. Tallent (Pacific Northwest National Laboratory, USA), and Ananth Kalyanaraman (Washington State University, USA; Pacific Northwest National Laboratory, USA)</i>	
Scalable and Fast Lazy Persistency on GPUs	252
<i>Ardhi Wiratama Baskara Yudha (University of Central Florida), Keiji Kimura (Waseda University), Huiyang Zhou (North Carolina State University), and Yan Solihin (University of Central Florida)</i>	
Accelerating Number Theoretic Transformations for Bootstrappable Homomorphic Encryption on GPUs	264
<i>Sangpyo Kim (Seoul National University, South Korea), Wonkyung Jung (Seoul National University, South Korea), Jaiyoung Park (Seoul National University, South Korea), and Jung Ho Ahn (Seoul National University, South Korea)</i>	
Reconfigurable Accelerator Compute Hierarchy: A Case Study Using Content-Based Image Retrieval	276
<i>Nazanin Farahpour (University of California, Los Angeles), Yuchen Hao (University of California, Los Angeles), Zhenman Fang (Simon Fraser University, Canada), and Glenn Reinman (University of California, Los Angeles)</i>	
Selective Event Processing for Energy Efficient Mobile Gaming with SNIP	288
<i>Prasanna Venkatesh Rengasamy (Penn State), Haibo Zhang (Penn State), Shulin Zhao (Penn State), Anand Sivasubramaniam (Penn State), Mahmut T Kandemir (Penn State), and Chita R Das (Penn State)</i>	
Author Index	301