2020 IEEE 27th International Conference on High Performance Computing, Data, and Analytics (HiPC 2020)

Virtual Conference 16 – 18 December 2020



IEEE Catalog Number: CFP20176-POD ISBN:

978-1-6654-4650-1

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:
 CFP20176-POD

 ISBN (Print-On-Demand):
 978-1-6654-4650-1

 ISBN (Online):
 978-1-6654-2292-5

ISSN: 1094-7256

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



2020 IEEE 27th International **Conference on High** Performance Computing, Data, and Analytics (HiPC) **HiPC 2020**

Table of Contents

Message from the General Co-chairs x
Message from the Program Chairs xii
HiPC 2020 Organization .xiv.
HiPC 2020 Technical Program Committee xvi
Keynote 1: Katherine Yelick xix
Keynote 2: Animashree Anandkumar xx
Keynote 3: Fabrizio Petrini xxi
Industry Sponsors xxii
HiPC 2020 Technical Program xxiii
Root Paper Soccion
Best Paper Session
SimGQ: Simultaneously Evaluating Iterative Graph Queries 1. Chengshuo Xu (University of California, Riverside, USA), Abbas Mazloumi (University of California, Riverside, USA), Xiaolin Jiang
(University of California, Riverside, USA), and Rajiv Gupta (University of California, Riverside, USA)
(University of California, Riverside, USA), and Rajiv Gupta

Session 1: Applications

Southern California, USA)

Performance Optimization and Scalability Analysis of the MGB Hydrological Model .31	
Exploring Task Parallelism for the Multilevel Fast Multipole Algorithm .41. Michael Lingg (Michigan State University, Array of Engineers), Stephen Hughey (Michigan State University, ARA), Doga Dikbayir (Michigan State University), Balasubramaniam Shanker (Michigan State University), and Hasan Metin Aktulga (Michigan State University)	•••
SparsePipe: Parallel Deep Learning for 3D Point Clouds .51 Keke Zhai (University of Florida), Pan He (University of Florida), Tania Banerjee (University of Florida), Anand Rangarajan (University of Florida), and Sanjay Ranka (University of Florida)	•••
HyPR: Hybrid Page Ranking on Evolving Graphs .62 Hemant Kumar Giri (Indian Institute of Information Technology Guwahati), Mridul Haque (Indian Institute of Information Technology Guwahati), and Dip Sankar Banerjee (Indian Institute of Information Technology Guwahati)	
Distributing Sparse Matrix/Graph Applications in Heterogeneous Clusters – an Experimental Study .72	
University of Berlin) Session 2: Scalable Data Science	
Processor Pipelining Method for Efficient Deep Neural Network Inference on Embedded Devices .82.	
Akshay Parashar (Samsung R&D Institute Bangalore, India), Arun Abraham (Samsung R&D Institute Bangalore, India), Deepak Chaudhary (Samsung R&D Institute Bangalore, India), and Vikram Nelvoy Rajendiran (Samsung R&D Institute Bangalore, India)	••
Avoiding Communication in Logistic Regression .91	
Aditya Devarakonda (Johns Hopkins University, USA) and James Demmel (University of California, Berkeley, USA)	
Aditya Devarakonda (Johns Hopkins University, USA) and James Demmel	••

Content-Defined Merkle Trees for Efficient Container Delivery .121
Model Checking as a Service using Dynamic Resource Scaling 131. Surya Teja Palavalasa (International Institute of Information Technology, Hyderabad), Yuvraj Singh (International Institute of Information Technology, Hyderabad), Adhish Singla (International Institute of Information Technology, Hyderabad), Suresh Purini (International Institute of Information Technology, Hyderabad), and Venkatesh Choppella (International Institute of Information Technology, Hyderabad)
Session 3: Algorithms
Parallel Hierarchical Clustering using Rank-Two Nonnegative Matrix Factorization .141
Pipelined Preconditioned Conjugate Gradient Methods for Distributed Memory Systems .151 Manasi Tiwari (Department of Computational and Data Sciences, Indian Institute of Science, Bengaluru) and Sathish Vadhiyar (Department of Computational and Data Sciences, Indian Institute of Science, Bengaluru)
Fair Allocation of Asymmetric Operations in Storage Systems .161. Thomas Keller (Rice University) and Peter Varman (Rice University)
A GPU Algorithm for Earliest Arrival Time Problem in Public Transport Networks 17.1
2D Static Resource Allocation for Compressed Linear Algebra and Communication Constraints .181 Olivier Beaumont (Inria Bordeaux Sud-Ouest and University of Bordeaux), Lionel Eyraud-Dubois (Inria Bordeaux Sud-Ouest and University of Bordeaux), and Mathieu Vérité (Inria Bordeaux Sud-Ouest and University of Bordeaux)
Algorithms for Preemptive Co-Scheduling of Kernels on GPUs .192. Lionel Eyraud-Dubois (Inria University of Bordeaux, France) and Cristiana Bentes (State University of Rio de Janeiro, Brazil)
Session 4: Runtime Systems
Understanding HPC Application I/O Behavior Using System Level Statistics 202. Arnab K. Paul (Virginia Tech, USA), Olaf Faaland (Lawrence Livermore National Laboratory, USA), Adam Moody (Lawrence Livermore National Laboratory, USA), Elsa Gonsiorowski (Lawrence Livermore National Laboratory, USA), Kathryn Mohror (Lawrence Livermore National Laboratory, USA), and Ali R. Butt (Virginia Tech, USA)

AMCilk: A Framework for Multiprogrammed Parallel Workloads .212
Extending SLURM for Dynamic Resource-Aware Adaptive Batch Scheduling .223
On the Marriage of Asynchronous Many Task Runtimes and Big Data: A Glance 233
Exposing Data Locality in HPC-Based Systems by using the HDFS Backend 243
PufferFish: NUMA-Aware Work-Stealing Library using Elastic Tasks .251
Design and Study of Elastic Recovery in HPC Applications 261. Kai Keller (Barcelona Supercomputing Center (BSC-CNS)), Konstantinos Parasyris (Lawrence Livermore National Laboratory (LLNL)), and Leonardo Bautista-Gomez (Barcelona Supercomputing Center (BSC-CNS))
Session 5: System Software and Architecture
Accelerating Force-Directed Graph Layout with Processing-in-Memory Architecture .27.1
Nonblocking Persistent Software Transactional Memory .283
GPU-FPtuner: Mixed-Precision Auto-Tuning for Floating-Point Applications on GPU .294
Batched Small Tensor-Matrix Multiplications on GPUs .305 Keke Zhai (University of Florida), Tania Banerjee (University of Florida), Adeesha Wijayasiri (University of Moratuwa), and Sanjay Ranka (University of Florida)
Temporal Based Intelligent LRU Cache Construction .315

Boosting LSTM Performance Through Dynamic Precision Selection .323	
Franyell Silfa (Polytechnic University of Catalonia), Jose Maria Arnau	
(Polytechnic University of Catalonia), and Antonio González	
(Polytechnic University of Catalonia)	
Author Index 335	