

2022 IEEE 29th International Conference on High Performance Computing, Data, and Analytics (HiPC 2022)

**Bengaluru, India
18 – 21 December 2022**



**IEEE Catalog Number: CFP22176-POD
ISBN: 978-1-6654-9424-3**

**Copyright © 2022 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: | CFP22176-POD |
| ISBN (Print-On-Demand): | 978-1-6654-9424-3 |
| ISBN (Online): | 978-1-6654-9423-6 |
| 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

CURRAN ASSOCIATES INC.
proceedings
.com

2022 IEEE 29th International Conference on High Performance Computing, Data, and Analytics (HiPC) **HiPC 2022**

Table of Contents

| | |
|--|-------|
| Message from the HiPC 2022 General Co-Chairs | x |
| Message from the HiPC 2022 Program Chairs | xiii |
| HiPC 2022 Organization | xv |
| HiPC 2022 Steering Committee | xvii |
| HiPC 2022 Technical Program Committee | xviii |
| Keynote 1: Paolo Lenne | xxii |
| Keynote 2: P Sadayappan | xxiii |
| Keynote 3: Jack Dongarra | xxiv |
| Keynote 4: Per Stenstrom | xxv |

Technical Session 1: Neural Networks

| | |
|---|----|
| Split-Knit Convolution: Enabling Dense Evaluation of Transpose and Dilated Convolutions on GPUs | 1 |
| <i>Arjun Menon Vadakkeveedu (Indian Institute of Technology Madras, India), Debabrata Mandal (AI & Advanced Computing Lab (ACL), KLA-Tencor Software India Pvt. Ltd., India), Pradeep Ramachandran (AI & Advanced Computing Lab (ACL), KLA-Tencor Software India Pvt. Ltd., India), and Nitin Chandrachoodan (Indian Institute of Technology Madras, India)</i> | |
| Low-Latency Mini-Batch GNN Inference on CPU-FPGA Heterogeneous Platform | 11 |
| <i>Bingyi Zhang (University of Southern California, USA), Hanqing Zeng (Meta AI, USA), and Viktor Prasanna (University of Southern California, USA)</i> | |
| Accelerating Broadcast Communication with GPU Compression for Deep Learning Workloads | 22 |
| <i>Qinghua Zhou (The Ohio State University, USA), Quentin Anthony (The Ohio State University, USA), Aamir Shafi (The Ohio State University, USA), Hari Subramoni (The Ohio State University, USA), and Dhableswar K. Panda (The Ohio State University, USA)</i> | |
| AccDP: Accelerated Data-Parallel Distributed DNN Training for Modern GPU-Based HPC Clusters | 32 |
| <i>Nawras Alnaasan (The Ohio State University, USA), Arpan Jain (The Ohio State University, USA), Aamir Shafi (The Ohio State University, USA), Hari Subramoni (The Ohio State University, USA), and Dhableswar K. Panda (The Ohio State University, USA)</i> | |

| | |
|---|----|
| Joint Partitioning and Sampling Algorithm for Scaling Graph Neural Network | 42 |
| <i>Manohar Lal Das (Indian Institute of Technology, India), Vishwesh Jatala (Indian Institute of Technology, India), and Gagan Raj Gupta (Indian Institute of Technology, India)</i> | |
| Building a Performance Model for Deep Learning Recommendation Model Training on GPUs | 48 |
| <i>Zhongyi Lin (University of California, Davis, USA), Louis Feng (Meta Platforms, Inc, USA), Ehsan K. Ardestani (Meta Platforms, Inc, USA), Jaewon Lee (Meta Platforms, Inc, USA), John Lundell (Meta Platforms, Inc, USA), Changkyu Kim (Meta Platforms, Inc, USA), Arun Kejarawal (Meta Platforms, Inc, USA), and John D. Owens (University of California, Davis, USA)</i> | |

Technical Session 2: HPC Architecture and Communication

| | |
|--|----|
| Accelerating Prefix Scan with In-network Computing on Intel PIUMA | 59 |
| <i>Kartik Lakhota (Intel Labs), Fabrizio Petrini (Intel Labs), Rajgopal Kannan (U.S. Army Research Lab), and Viktor Prasanna (University of Southern California)</i> | |
| memwalkd : Accelerating Key-Value Stores using Page Table Walkers | 69 |
| <i>Ravi Shreyas Anupindi (Microsoft Research, India), Swaroop Kotni (Rubrik India Pvt. Ltd, India), and Arkaprava Basu (Indian Institute of Science, India)</i> | |
| Energy Consumption Evaluation of Optane DC Persistent Memory for Indexing Data Structures | 75 |
| <i>Manolis Katsaragakis (National Technical University of Athens(NTUA), Greece; KU Leuven(KUL), Belgium), Lazaros Papadopoulos (National Technical University of Athens(NTUA), Greece), Christos Baloukas (National Technical University of Athens(NTUA), Greece), Verena Kantere (National Technical University of Athens(NTUA), Greece), Francky Catthoor (KU Leuven (Belgium); IMEC, Leuven (Belgium)), and Dimitrios Soudris (National Technical University of Athens(NTUA), Greece)</i> | |
| LDT: Lightweight Dirty Tracking of Memory Pages for x86 Systems | 85 |
| <i>Rohit Singh (Indian Institute of Technology Kanpur, India), Arun KP (Indian Institute of Technology Kanpur, India), and Debadatta Mishra (Indian Institute of Technology Kanpur, India)</i> | |
| Designing Efficient Pipelined Communication Schemes using Compression in MPI Libraries | 95 |
| <i>Bharath Ramesh (The Ohio State University, USA), Qinghua Zhou (The Ohio State University, USA), Aamir Shafi (The Ohio State University, USA), Mustafa Abduljabbar (The Ohio State University, USA), Hari Subramoni (The Ohio State University, USA), and Dhabaleswar K. Panda (The Ohio State University, USA)</i> | |

| | |
|--|-----|
| Efficient Personalized and Non-Personalized Alltoall Communication for Modern Multi-HCA GPU-Based Clusters | 100 |
| <i>Kaushik Kandadi Suresh (The Ohio State University, USA), Akshay Paniraja Guptha (The Ohio State University, USA), Benjamin Michalowicz (The Ohio State University, USA), Bharath Ramesh (The Ohio State University, USA), Mustafa Abduljabbar (The Ohio State University, USA), Aamir Shafi (The Ohio State University, USA), Hari Subramoni (The Ohio State University, USA), and Dhabaleswar Panda (The Ohio State University, USA)</i> | |

Technical Session 3: HPC Algorithms and Solvers

| | |
|--|-----|
| High-Performance Truss Analytics in Arkouda | 105 |
| <i>Zhihui Du (New Jersey Institute of Technology, USA), Joseph Patchett (New Jersey Institute of Technology, USA), Oliver Alvarado Rodriguez (New Jersey Institute of Technology, USA), Fuhuan Li (New Jersey Institute of Technology, USA), and David A. Bader (New Jersey Institute of Technology, USA)</i> | |
| Parallel Vertex Color Update on Large Dynamic Networks | 115 |
| <i>Arindam Khanda (Missouri University of Science and Technology, USA), Sanjukta Bhowmick (University of North Texas, USA), Xin Liang (Missouri University of Science and Technology, USA), and Sajal K. Das (Missouri University of Science and Technology, USA)</i> | |
| IMpart: A Partitioning-Based Parallel Approach to Accelerate Influence Maximization | 125 |
| <i>Reet Barik (Washington State University), Marco Minutoli (Pacific Northwest National Laboratory), Mahantesh Halappanavar (Pacific Northwest National Laboratory), and Ananth Kalyanaraman (Washington State University)</i> | |
| Leveraging GPU Tensor Cores for Double Precision Euclidean Distance Calculations | 135 |
| <i>Benoit Gallet (Northern Arizona University, USA) and Michael Gowanlock (Northern Arizona University, USA)</i> | |
| A Portable Sparse Solver Framework for Large Matrices on Heterogeneous Architectures | 145 |
| <i>Fazlay Rabbi (Michigan State University, USA), Christopher S. Daley (Lawrence Berkeley National Laboratory, USA), Umit V. Catalyurek (Amazon Web Services and Georgia Institute of Technology, USA), and Hasan Metin Aktulga (Michigan State University, USA)</i> | |
| Performance Analysis of GPU Accelerated Meshfree q-LSKUM Solvers in Fortran, C, Python, and Julia | 156 |
| <i>Nischay Ram Mamidi (Birla Institute of Technology and Science Pilani - Hyderabad Campus, India), Dhruv Saxena (Birla Institute of Technology and Science Pilani - Hyderabad Campus, India), Kumar Prasun (Courant Institute of Mathematical Sciences, USA), Anil Nemili (Birla Institute of Technology and Science Pilani - Hyderabad Campus, India), Bharatkumar Sharma (NVIDIA, India), and Suresh Deshpande (FASc, FNAE Fellow, India)</i> | |

Technical Session 4: High Performance and Data Science Applications

| | |
|---|-----|
| A Deep Learning-Based In Situ Analysis Framework for Tropical Cyclogenesis Prediction | 166 |
| <i>Abir Mukherjee (Citrix R&D India Pvt. Ltd., India) and Preeti Malakar (Indian Institute of Technology Kanpur, India)</i> | |
| HiBGT: High-Performance Bayesian Group Testing for COVID-19 | 176 |
| <i>Weicong Chen (Case Western Reserve University, USA), Curtis Tatsuoka (University of Pittsburgh, USA), and Xiaoyi Lu (University of California Merced, USA)</i> | |
| Customer Churn Prediction in Telecommunications Industry Based on Conditional Wasserstein GAN | 186 |
| <i>Chang Su (Chongqing University of Posts and Telecommunications, China), Linglin Wei (Chongqing University of Posts and Telecommunications, China), and Xianzhong Xie (Chongqing University of Posts and Telecommunications, China)</i> | |
| A Real-Time Flood Inundation Prediction on SX-Aurora TSUBASA | 192 |
| <i>Yoichi Shimomura (Tohoku University, Japan), Akihiro Musa (Tohoku University, Japan), Yoshihiko Sato (NEC Solution Innovators, Japan), Atsuhiko Konja (Mitsui Consultants Co., Ltd, Japan), Guoqing Cui (Mitsui Consultants Co., Ltd, Japan), Rei Aoyagi (Tohoku University, Japan), Keichi Takahashi (Tohoku University, Japan), and Hiroyuki Takizawa (Tohoku University, Japan)</i> | |
| Precise Parallel FEM-Based Interactive Cutting Simulation of Deformable Bodies | 198 |
| <i>Harshvardhan Das (Indian Institute of Technology, India), Suraj Kumar (Indian Institute of Technology, India), and Subodh Kumar (Indian Institute of Technology, India)</i> | |
| Scaling the SOO Global Blackbox Optimizer on a 128-Core Architecture | 204 |
| <i>David Redon (University of Lille, France), Bilel Derbel (University of Lille, Inria, France), and Pierre Fortin (University of Lille, France)</i> | |

Technical Session 5: HPC System Software and Libraries

| | |
|--|-----|
| A GPU-Accelerated Data Transformation Framework Rooted in Pushdown Transducers | 215 |
| <i>Tri Nguyen (North Carolina State University, USA) and Michela Becchi (North Carolina State University, USA)</i> | |
| An Algorithmic and Software Pipeline for Very Large Scale Scientific Data Compression with Error Guarantees | 226 |
| <i>Tania Banerjee (University of Florida, USA), Jong Choi (Oak Ridge National Laboratory, USA), Jaemoon Lee (University of Florida, USA), Qian Gong (Oak Ridge National Laboratory, USA), and Ruonan Wang (Oak Ridge National Laboratory, USA)</i> | |
| COMPROF and COMPLACE: Shared-Memory Communication Profiling and Automated Thread Placement via Dynamic Binary Instrumentation | 236 |
| <i>Ryan Kirkpatrick (University of St Andrews, UK), Christopher Brown (University of St Andrews, UK), and Vladimir Janjic (University of Dundee, UK)</i> | |

| | |
|--|-----|
| LuxIO: Intelligent Resource Provisioning and Auto-Configuration for Storage Services | 246 |
| <i>Keith Bateman (Illinois Institute of Technology), Neeraj Rajesh (Illinois Institute of Technology), Jaime Cernuda Garcia (Illinois Institute of Technology), Luke Logan (Illinois Institute of Technology), Jie Ye (Illinois Institute of Technology), Stephen Herbein (Lawrence Livermore National Laboratory), Anthony Kougkas (Illinois Institute of Technology), and Xian-He Sun (Illinois Institute of Technology)</i> | |
| IRIS-BLAS: Towards a Performance Portable and Heterogeneous BLAS Library | 256 |
| <i>Narasinga Rao Miniskar (Oak Ridge National Laboratory, USA), Mohammad Alaul Haque Monil (Oak Ridge National Laboratory, USA), Pedro Valero-Lara (Oak Ridge National Laboratory, USA), Frank Liu (Oak Ridge National Laboratory, USA), and Jeffery S. Vetter (Oak Ridge National Laboratory, USA)</i> | |
| Towards Efficient Cache Allocation for High-Frequency Checkpointing | 262 |
| <i>Avinash Maurya (Rochester Institute of Technology, USA), Bogdan Nicolae (Argonne National Laboratory, USA), M. Mustafa Rafique (Rochester Institute of Technology, USA), Amr M. Elsayed (Brightskies Technologies, Egypt), Thierry Tonellot (Saudi Aramco, Saudi Arabia), and Franck Cappello (Argonne National Laboratory, USA)</i> | |

Technical Session 6: Data Science Methods

| | |
|--|------------|
| 1-bit LAMB: Communication Efficient Large-Scale Large-Batch Training with LAMB's Convergence Speed | 272 |
| <i>Conglong Li (Microsoft), Ammar Ahmad Awan (Microsoft), Hanlin Tang (University of Rochester), Samyam Rajbhandari (Microsoft), and Yuxiong He (Microsoft)</i> | |
| Input Feature Pruning for Accelerating GNN Inference on Heterogeneous Platforms | 282 |
| <i>Jason Yik (Harvard University), Sanmukh R. Kuppannagari (Case Western Reserve University), Hanqing Zeng (Meta AI), and Viktor K. Prasanna (University of Southern California)</i> | |
| Provenance-Based Workflow Diagnostics using Program Specification | 292 |
| <i>Yuta Nakamura (DePaul University, USA), Tanu Malik (DePaul University, USA), Iyad Kanj (DePaul University, USA), and Ashish Gehani (SRI International, USA)</i> | |
| EECAAP: Efficient Edge-Computing Based Anonymous Authentication Protocol for IoV | 302 |
| <i>Himani Sikarwar (Indian Institute of Technology(IIT), India) and Debasis Das (Indian Institute of Technology(IIT), India)</i> | |
| Author Index | 309 |