

2017 46th International Conference on Parallel Processing (ICPP 2017)

**Bristol, United Kingdom
14-17 August 2017**



**IEEE Catalog Number: CFP17127-POD
ISBN: 978-1-5386-1043-5**

**Copyright © 2017 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:	CFP17127-POD
ISBN (Print-On-Demand):	978-1-5386-1043-5
ISBN (Online):	978-1-5386-1042-8
ISSN:	0190-3918

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

2017 46th International Conference on Parallel Processing

ICPP 2017

Table of Contents

Message from the General Co-Chairs	xii
Message from the Program Co-Chairs	xiii
Organizing Committee	xv
Program Committee	xvii
Reviewers	xxi

Highlighted Papers (S1-T1)

Preparing HPC Applications for the Exascale Era: A Decoupling Strategy	1
<i>Ivy Bo Peng, Roberto Gioiosa, Gokcen Kestor, Erwin Laure, and Stefano Markidis</i>	
An Efficient, Distributed Stochastic Gradient Descent Algorithm for Deep-Learning Applications	11
<i>Guojing Cong, Onkar Bhardwaj, and Minwei Feng</i>	
Large-Scale Parallelization of Smoothed Particle Hydrodynamics Method on Heterogeneous Cluster	21
<i>Yingrui Wang, Leisheng Li, and Rong Tian</i>	

Graph Analytics and ML (S2-T1)

Boosting the Efficiency of HPCG and Graph500 with Near-Data Processing	31
<i>Erik Vermij, Leandro Fiorin, Christoph Hagleitner, and Koen Bertels</i>	
GCN: GPU-Based Cube CNN Framework for Hyperspectral Image Classification	41
<i>Han Dong, Tao Li, Jiabing Leng, Lingyan Kong, and Gang Bai</i>	
Nearly Balanced Work Partitioning for Heterogeneous Algorithms	50
<i>Mallipeddi Hardhik, Dip Sankar Banerjee, Kiran Raj Ramamoorthy, Kishore Kothapalli, and Kannan Srinathan</i>	

Enhancing Programming Runtime Systems (S2-T2)

GLTO: On the Adequacy of Lightweight Thread Approaches for OpenMP Implementations	60
<i>Adrián Castelló, Sangmin Seo, Rafael Mayo, Pavan Balaji, Enrique S. Quintana-Ortí, and Antonio J. Peña</i>	
Locality-Aware Dynamic Task Graph Scheduling	70
<i>Jordyn Maglalang, Sriram Krishnamoorthy, and Kunal Agrawal</i>	
Practical Experience with Transactional Lock Elision	81
<i>Tingzhe Zhou, Pante A Zardoshti, and Michael Spear</i>	

Linear Algebra Algorithms (S2-T3)

Variable-Size Batched LU for Small Matrices and Its Integration into Block-Jacobi Preconditioning	91
<i>Hartwig Anzt, Jack Dongarra, Goran Flegar, and Enrique S. Quintana-Ortí</i>	
High-Performance and Memory-Saving Sparse General Matrix-Matrix Multiplication for NVIDIA Pascal GPU	101
<i>Yusuke Nagasaka, Akira Nukada, and Satoshi Matsuoka</i>	
Sparse Tensor Factorization on Many-Core Processors with High-Bandwidth Memory	111
<i>Shaden Smith, Jongsoo Park, and George Karypis</i>	

Data and Networks (S3-T1)

Efficient Data Sharing on Heterogeneous Systems	121
<i>Víctor García-Flores, Eduard Ayguade, and Antonio J. Peña</i>	
HyPPI NoC: Bringing Hybrid Plasmonics to an Opto-Electronic Network-on-Chip	131
<i>Vikram K. Narayana, Shuai Sun, Armin Mehrabian, Volker J. Sorger, and Tarek El-Ghazawi</i>	
ES2: Aiming at an Optimal Virtual I/O Event Path	141
<i>Xiaokang Hu, Wang Zhang, Jian Li, Ruhui Ma, Feng Wu, and Haibing Guan</i>	

GPU & Runtime Systems (S3-T2)

MPI-GDS: High Performance MPI Designs with GPUDirect-aSync for CPU-GPU Control Flow Decoupling	151
<i>Akshay Venkatesh, Khaled Hamidouche, Sreeram Potluri, Davide Rosetti, Ching-Hsiang Chu, and Dhabaleswar K. Panda</i>	

Efficient and Scalable Multi-Source Streaming Broadcast on GPU Clusters for Deep Learning	161
<i>Ching-Hsiang Chu, Xiaoyi Lu, Ammar A. Awan, Hari Subramoni, Jahanzeb Hashmi, Bracy Elton, and Dhabaleswar K. Panda</i>	
Overlapping Data Transfers with Computation on GPU with Tiles	171
<i>Burak Bastem, Didem Unat, Weiqun Zhang, Ann Almgren, and John Shalf</i>	

Graphs and Networks (S3-T3)

Accelerating Graph Analytics by Utilising the Memory Locality of Graph Partitioning	181
<i>Jiawen Sun, Hans Vandierendonck, and Dimitrios S. Nikolopoulos</i>	
Parallel Algorithms for the Computation of Cycles in Relative Neighborhood Graphs	191
<i>Hari Sundar and Parmeshwar Khurd</i>	
High Performance Query Processing for Web Scale RDF Data Using BSP Style Communication and Balanced Distribution	201
<i>Minho Bae, Junho Eum, Donghoon Kim, and Sangyoon Oh</i>	

Storage (S4-T1)

OptiMatch: Enabling an Optimal Match between Green Power and Various Workloads for Renewable-Energy Powered Storage Systems	211
<i>Xiaoyang Qu, Jiguang Wan, Fengguang Song, Xiaozhao Zhuang, Fei Wu, and Changsheng Xie</i>	
Favorable Block First: A Comprehensive Cache Scheme to Accelerate Partial Stripe Recovery of Triple Disk Failure Tolerant Arrays	221
<i>Luyu Li, Houxiang Ji, Chentao Wu, Jie Li, and Minyi Guo</i>	
Non-Sequential Striping for Distributed Storage Systems with Different Redundancy Schemes	231
<i>Yanwen Xie, Dan Feng, and Fang Wang</i>	

IO & Cloud (S4-T2)

Predicting Response Latency Percentiles for Cloud Object Storage Systems	241
<i>Yi Su, Dan Feng, Yu Hua, and Zhan Shi</i>	
WA-Dataspaces: Exploring the Data Staging Abstractions for Wide-Area Distributed Scientific Workflows	251
<i>Mehmet Fatih Aktas, Javier Diaz-Montes, Ivan Roderio, and Manish Parashar</i>	
Scalable Write Allocation in the WAFL File System	261
<i>Matthew Curtis-Maury, Ram Kesavan, and Mrinal K. Bhattacharjee</i>	

Numerical Applications (S4-T3)

Parallel Construction of Simultaneous Deterministic Finite Automata on Shared-Memory Multicores	271
<i>Minyoung Jung, Jinwoo Park, Johann Blieberger, and Bernd Burgstaller</i>	
Parallel Reconstruction of Three Dimensional Magnetohydrodynamic Equilibria in Plasma Confinement Devices	282
<i>Sudip K. Seal, Mark R. Cianciosa, Steven P. Hirshman, Andreas Wingen, Robert S. Wilcox, and Ezekial A. Unterberg</i>	
Performance Analysis and Optimization of Sparse Matrix-Vector Multiplication on Modern Multi- and Many-Core Processors	292
<i>Athena Elafrou, Georgios Goumas, and Nectarios Koziris</i>	

Networks (S5-T1)

Network Aware Multi-User Computation Partitioning in Mobile Edge Clouds	302
<i>Lei Yang, Jiannong Cao, Zhenyu Wang, and Weigang Wu</i>	
Fading-Resistant Link Scheduling in Wireless Networks	312
<i>Chenxi Qiu and Haiying Shen</i>	
Order/Radix Problem: Towards Low End-to-End Latency Interconnection Networks	322
<i>Ryota Yasudo, Michihiro Koibuchi, Koji Nakano, Hiroki Matsutani, and Hideharu Amano</i>	

Cloud Scheduling (S5-T2)

A Dynamic Resource Controller for a Lambda Architecture	332
<i>MohammadReza HoseinyFarahabady, Javid Taheri, Zahir Tari, and Albert Y. Zomaya</i>	
CELIA: Cost-Time Performance of Elastic Applications on Cloud	342
<i>Sunimal Rathnayake, Dumitrel Loghin, and Yong Meng Teo</i>	
The Cloud as an OpenMP Offloading Device	352
<i>Hervé Yviquel and Guido Araújo</i>	

GPU Applications (S5-T3)

Simple and Fast Parallel Algorithms for the Voronoi Map and the Euclidean Distance Map, with GPU Implementations	362
<i>Takumi Honda, Shinnosuke Yamamoto, Hiroaki Honda, Koji Nakano, and Yasuaki Ito</i>	

High-Performance Recommender System Training Using Co-Clustering on CPU/GPU Clusters	372
<i>Kubilay Atasu, Thomas Parnell, Celestine Dünner, Michail Vlachos, and Haralampos Pozidis</i>	

Exploiting GPUs for Fast Force-Directed Visualization of Large-Scale Networks	382
<i>Govert G. Brinkmann, Kristian F.D. Rietveld, and Frank W. Takes</i>	

Data and IO (S6-T1)

A Coflow-Based Co-Optimization Framework for High-Performance Data Analytics	392
<i>Long Cheng, Ying Wang, Yulong Pei, and Dick Epema</i>	
PDS: An I/O-Efficient Scaling Scheme for Parity Declustered Data Layout	402
<i>Zhipeng Li, Yinlong Xu, Yongkun Li, Chengjin Tian, and Youhui Bai</i>	
Data Caching in Next Generation Mobile Cloud Services, Online vs. Off-Line	412
<i>Yang Wang, Shuibing He, Xiaopeng Fan, Chengzhong Xu, Joseph Culberson, and Joseph Horton</i>	

Computation Optimization (S6-T2)

Towards Highly Efficient DGEMM on the Emerging SW26010 Many-Core Processor	422
<i>Lijuan Jiang, Chao Yang, Yulong Ao, Wanwang Yin, Wenjing Ma, Qiao Sun, Fangfang Liu, Rongfen Lin, and Peng Zhang</i>	
Optimizations of Two Compute-Bound Scientific Kernels on the SW26010 Many-Core Processor	432
<i>James Lin, Zhigeng Xu, Akira Nukada, Naoya Maruyama, and Satoshi Matsuoka</i>	
Bitslice Vectors: A Software Approach to Customizable Data Precision on Processors with SIMD Extensions	442
<i>Shixiong Xu and David Gregg</i>	

Data Analytics (S6-T3)

Runtime Data Layout Scheduling for Machine Learning Dataset	452
<i>Yang You and James Demmel</i>	
A Machine Learning Approach for Efficient Parallel Simulation of Beam Dynamics on GPUs	462
<i>Kamesh Arumugam, Desh Ranjan, Mohammad Zubair, Baša Terzić, Alexander Godunov, and Tunazzina Islam</i>	

Multiple Pattern Matching for Network Security Applications: Acceleration through Vectorization	472
<i>Charalampos Stylianopoulos, Magnus Almgren, Olaf Landsiedel, and Marina Papatriantafilou</i>	

Graph Algorithms (S7-T1)

Parallel Space-Time Kernel Density Estimation	483
<i>Erik Saule, Dinesh Panchananam, Alexander Hohl, Wenwu Tang, and Eric Delmelle</i>	
Parallel Algorithm for Single-Source Earliest-Arrival Problem in Temporal Graphs	493
<i>Peng Ni, Masatoshi Hanai, Wen Jun Tan, Chen Wang, and Wentong Cai</i>	
Greed Is Good: Parallel Algorithms for Bipartite-Graph Partial Coloring on Multicore Architectures	503
<i>Mustafa Kemal Taş, Kamer Kaya, and Erik Saule</i>	

Performance & Power Tuning for Heterogeneous Platforms (S7-T2)

A Scalable Hierarchical Semi-Separable Library for Heterogeneous Clusters	513
<i>Isuru Dilanka Fernando, Sanath Jayasena, Milinda Fernando, and Hari Sundar</i>	
Autotuning GPU Kernels via Static and Predictive Analysis	523
<i>Robert Lim, Boyana Norris, and Allen Malony</i>	
A Pareto Framework for Data Analytics on Heterogeneous Systems: Implications for Green Energy Usage and Performance	533
<i>Aniket Chakrabarti, Srinivasan Parthasarathy, and Christopher Stewart</i>	

Various Parallel Algorithms (S8-T1)

Scheduling Independent Tasks in Parallel under Power Constraints	543
<i>Ayham Kassab, Jean-Marc Nicod, Laurent Philippe, and Veronika Rehn-Sonigo</i>	
A Novel Minimum Time Parallel 2-D Discrete Wavelet Transform Algorithm for General Purpose Processors	553
<i>Eduardo Moscoso Rubino, Alberto Jose Alvares, Raul Marin Prades, and Pedro Sanz Valero</i>	
A Parallel TSP-Based Algorithm for Balanced Graph Partitioning	563
<i>Harshvardhan Das and Subodh Kumar</i>	

Resilience & Power Aware Scheduling (S8-T2)

E-Storm: Replication-Based State Management in Distributed Stream Processing Systems	571
<i>Xunyun Liu, Aaron Harwood, Shanika Karunasekera, Benjamin Rubinstein, and Rajkumar Buyya</i>	
Resilience for Stencil Computations with Latent Errors	581
<i>Aiman Fang, Aurélien Cavelan, Yves Robert, and Andrew A. Chien</i>	
Application-Aware Power Coordination on Power Bounded NUMA Multicore Systems	591
<i>Rong Ge, Pengfei Zou, and Xizhou Feng</i>	
Author Index	601