

2016 IEEE International Conference on Cluster Computing (CLUSTER 2016)

**Taipei, Taiwan
12 – 16 September 2016**



**IEEE Catalog Number: CFP16235-POD
ISBN: 978-1-5090-3654-7**

**Copyright © 2016 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 publication is a representation of what appears in the IEEE Digital Libraries. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP16235-POD
ISBN (Print-On-Demand):	978-1-5090-3654-7
ISBN (Online):	978-1-5090-3653-0
ISSN:	1552-5244

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

2016 IEEE International Conference on Cluster Computing

CLUSTER 2016

Table of Contents

Message from the General Co-Chairs.....	xii
Message from the Program Chair.....	xiii
FTS 2016 Workshop Welcome Message.....	xv
Organizing Committee.....	xvi
Technical Program Committee.....	xvii
Steering Committee.....	xx
Poster Committee.....	xxi
External Reviewers.....	xxii
Keynote 1.....	xxiv
Keynote 2.....	xxv
Keynote 3.....	xxvi
Sponsors and Organizers.....	xxvii

Best Paper Nominees (Paper Session 1)

Adaptive and Dynamic Design for MPI Tag Matching.....	1
<i>M. Bayatpour, H. Subramoni, S. Chakraborty, and D. K. Panda</i>	
A Low Disk-Bound Transaction Logging System for In-memory Distributed Data Stores	11
<i>Dayal Dilli, Kenneth B. Kent, Yang Wang, and Chengzhong Xu</i>	
Realizing Out-of-Core Stencil Computations Using Multi-tier Memory Hierarchy on GPGPU Clusters	21
<i>Toshio Endo</i>	
Parallel DTFE Surface Density Field Reconstruction	30
<i>Esteban Rangel, Nan Li, Salman Habib, Tom Peterka, Ankit Agrawal, Wei-Keng Liao, and Alok Choudhary</i>	

Network Performance (Paper Session 2)

Evaluation of Topology-Aware Broadcast Algorithms for Dragonfly Networks	40
<i>Matthieu Dorier, Misbah Mubarak, Rob Ross, Jianping Kelvin Li, Christopher D. Carothers, and Kwa-Liu Ma</i>	
(SAI) Stalled, Active and Idle: Characterizing Power and Performance of Large-Scale Dragonfly Networks	50
<i>Taylor Groves, Ryan E. Grant, Scott Hemmer, Simon Hammond, Michael Levenhagen, and Dorian C. Arnold</i>	
Compiler-Assisted Overlapping of Communication and Computation in MPI Applications	60
<i>Jichi Guo, Qing Yi, Jiayuan Meng, Junchao Zhang, and Pavan Balaji</i>	

Virtualization (Paper Session 3)

vProbe: Scheduling Virtual Machines on NUMA Systems	70
<i>Song Wu, Huahua Sun, Like Zhou, Qingtian Gan, and Hai Jin</i>	
GLAP: Distributed Dynamic Workload Consolidation through Gossip-Based Learning	80
<i>Mansour Khelghatdoust, Vincent Gramoli, and Daniel Sun</i>	
CORP: Cooperative Opportunistic Resource Provisioning for Short-Lived Jobs in Cloud Systems	90
<i>Jinwei Liu, Haiying Shen, and Liuhua Chen</i>	

Memory Optimization (Paper Session 4)

Towards Resource Disaggregation — Memory Scavenging for Scientific Workloads	100
<i>Alexandru Uta, Ana-Maria Oprescu, and Thilo Kielmann</i>	
CHOPPER: Optimizing Data Partitioning for In-memory Data Analytics Frameworks	110
<i>Arnab Kumar Paul, Wenjie Zhuang, Luna Xu, Min Li, M. Mustafa Rafique, and Ali R. Butt</i>	
Improving Collective I/O Performance Using Non-volatile Memory Devices	120
<i>Giuseppe Congiu, Sai Narasimhamurthy, Tim Süß, and André Brinkmann</i>	

Poster Papers

SMARTPARTITION: Efficient Partitioning for Natural Graphs	130
<i>Chengfei Zhang, Yiming Zhang, Dongsheng Li, Jia Li, and Minne Li</i>	
Accelerating I/O Performance of SVM on HDFS	132
<i>Mao Ye, Jun Wang, Jiangling Yin, and Xuhong Zhang</i>	
Energy and Performance Efficient Underloading Detection Algorithm of Virtual Machines in Cloud Data Centers	134
<i>Juan Fang, Lifu Zhou, Xiaoting Hao, Min Cai, and Xingtian Ren</i>	

Machine Status Prediction for Dynamic and Heterogenous Cloud Environment	136
<i>Jinliang Xu, Ao Zhou, Shangguang Wang, Qibo Sun, Jinglin Li, and Fangchun Yang</i>	
Design and Analysis of Fault Tolerance Mechanisms for Big Data Transfers	138
<i>Preethika Kasu, Youngjae Kim, Sungyong Park, Scott Atchley, and Geoffroy R. Vallée</i>	
Time Optimization Modeling for Big Data Placement and Analysis for Geo-Distributed Data Centers	140
<i>Awais Khan, Muhammad Attique, Tae-Sun Chung, and Youngjae Kim</i>	
Enhancing Performance of Large-Scale Electronic Structure Calculations with Many-Core Computing	142
<i>Hoon Ryu and Yosang Jeong</i>	
Reduced-Precision Floating-Point Formats on GPUs for High Performance and Energy Efficient Computation	144
<i>Daichi Mukunoki and Toshiyuki Imamura</i>	
Themis: A Scalable Performance Evaluation Framework for Virtualized Datacenter	146
<i>Zhengmin Li, Di Zhang, Xinran Liu, Bin Sun, Zhicheng Yao, and Xiufeng Sui</i>	
Conflict Prediction-Based Transaction Execution for Transactional Memory in Multi-core In-memory Databases	148
<i>Min Yoon, Moon-Hwan Kang, Yeon-Woo Jang, and Jae-Woo Chang</i>	
Skyline Service Selection Based on QoS Prediction	150
<i>Yan Guo and Shangguang Wang</i>	
Minimizing CMT Miss Penalty in Selective Page-Level Address Mapping Table	152
<i>Ronnie Mativenga, Joon-Young Paik, Junghee Lee, Tae-Sun Chung, and Youngjae Kim</i>	
Efficient Semantic-Aware Coflow Scheduling for Data-Parallel Jobs	154
<i>Ziyang Li, Yiming Zhang, Yunxiang Zhao, and Dongsheng Li</i>	
Quick Eviction of Virtual Machines through Proactive Snapshots	156
<i>Dinuni Fernando, Hardik Bagdi, Yaohui Hu, Ping Yang, Kartik Gopalan, Charles Kamhoua, and Kevin Kwiatt</i>	
Spatial Locality Aware, Fast, and Scalable SLINK Algorithm for Commodity Clusters	158
<i>Poonam Goyal, Sonal Kumari, Sumit Sharma, Vivek Kishore, Navneet Goyal, and Sundar S. Balasubramaniam</i>	
High Throughput Log-Based Replication for Many Small In-Memory Objects	160
<i>Kevin Beineke, Stefan Nothaas, and Michael Schöttner</i>	
TwinPCG: Dual Thread Redundancy with forward Recovery for Preconditioned Conjugate Gradient Methods	162
<i>Kiril Dichev and Dimitrios S. Nikolopoulos</i>	
Optimizing Locality by Topology-Aware Placement for a Task Based Programming Model	164
<i>Jens Gustedt, Emmanuel Jeannot, and Farouk Mansouri</i>	

SSDUP: An Efficient SSD Write Buffer Using Pipeline	166
<i>Ming Li, Xuanhua Shi, Wei Liu, Hai Jin, and Yong Chen</i>	
streamingRPHash: Random Projection Clustering of High-Dimensional Data in a MapReduce Framework	168
<i>Jacob Franklin, Samuel Wenke, Sadiq Quasem, Lee A. Carraher, and Philip A. Wilsey</i>	
SuperGlue: Standardizing Glue Components for HPC Workflows	170
<i>Jay Lofstead, Alexis Champsaur, Jai Dayal, Matthew Wolf, and Greg Eisenhauer</i>	

Large-Scale System Analysis and Modeling: Combined Paper/Panel Session (Paper Session 5)

Fast Multi-parameter Performance Modeling	172
<i>Alexandru Calotoiu, David Beckinsale, Christopher W. Earl, Torsten Hoefler, Ian Karlin, Martin Schulz, and Felix Wolf</i>	
Active Learning in Performance Analysis	182
<i>Dmitry Duplyakin, Jed Brown, and Robert Ricci</i>	
A Model for Weak Scaling to Many GPUs at the Basis of the Linpack Benchmark	192
<i>David Rohr, Jan de Cuveland, and Volker Lindenstruth</i>	
When Amdahl Meets Young/Daly	203
<i>Aurélien Cavelan, Jiafan Li, Yves Robert, and Hongyang Sun</i>	

Energy and Resilience (Paper Session 6)

A Case for Criticality Models in Exascale Systems	213
<i>Brian Kocoloski, Leonardo Piga, Wei Huang, Indrani Paul, and John Lange</i>	
Dynamically Building Energy Proportional Data Centers with Heterogeneous Computing Resources	217
<i>Violaine Villebonnet, Georges Da Costa, Laurent Lefevre, Jean-Marc Pierson, and Patricia Stolf</i>	
On Energy Proportionality and Time-Energy Performance of Heterogeneous Clusters	221
<i>Lavanya Ramapantulu, Dumitrel Loghin, and Yong Meng Teo</i>	
Unequal Failure Protection Coding Technology for Cloud Storage Systems	231
<i>Yupeng Hu, Yonghe Liu, Wenjia Li, Nong Xiao, Zheng Qin, and Shu Yin</i>	

Resource Management (Paper Session 7)

Probabilistic Network-Aware Task Placement for MapReduce Scheduling	241
<i>Haiying Shen, Ankur Sarker, Lei Yu, and Feng Deng</i>	
Tiresias: Low-Overhead Sample Based Scheduling with Task Hopping	251
<i>Chunliang Hao, Jie Shen, Heng Zhang, Yanjun Wu, and Mingshu Li</i>	
Machine Learning Predictions of Runtime and IO Traffic on High-End Clusters	255
<i>Ryan McKenna, Stephen Herbein, Adam Moody, Todd Gamblin, and Michela Taufer</i>	

Exploring Plan-Based Scheduling for Large-Scale Computing Systems	259
<i>Xingwu Zhang, Zhou Zhou, Xu Yang, Zhiling Lan, and Jia Wang</i>	

In-Situ Applications (Paper Session 8)

Adaptive Performance-Constrained In Situ Visualization of Atmospheric Simulations	269
<i>Matthieu Dorier, Robert Sisneros, Leonardo Bautista Gomez, Tom Peterka, Leigh Orf, Lokman Rahmani, Gabriel Antoniu, and Luc Bougé</i>	
Bredala: Semantic Data Redistribution for In Situ Applications	279
<i>Matthieu Dreher and Tom Peterka</i>	

I/O Optimization (Paper Session 9)

Exploring Data Migration for Future Deep-Memory Many-Core Systems	289
<i>Swann Perarnau, Judicael A. Zounmevo, Balazs Geroft, Kamil Iskra, and Pete Beckman</i>	
GraphMeta: A Graph-Based Engine for Managing Large-Scale HPC Rich Metadata	298
<i>Dong Dai, Yong Chen, Philip Carns, John Jenkins, Wei Zhang, and Robert Ross</i>	
Extending SSD Lifetime with Persistent In-Memory Metadata Management	308
<i>Qingsong Wei, Cheng Chen, Mingdi Xue, Chundong Wang, and Jun Yang</i>	
FlashStager: Improving the Performance of SSD-Based Data Staging Systems via Write Redirection	312
<i>Xuechen Zhang, Fang Zheng, Karsten Schwan, and Matthew Wolf</i>	
Performance Optimization for All Flash Scale-Out Storage	316
<i>Myoungwon Oh, Jugwan Eom, Jungyeon Yoon, Jae Yeun Yun, Seungmin Kim, and Heon Y. Yeom</i>	

Parallel Applications (Paper Session 10)

Distributed Parallel #SAT Solving	326
<i>Jan Burchard, Tobias Schubert, and Bernd Becker</i>	
Smart-MLlib: A High-Performance Machine-Learning Library	336
<i>David Siegal, Jia Guo, and Gagan Agrawal</i>	
Distributed Bayesian Probabilistic Matrix Factorization	346
<i>Tom Vander Aa, Imen Chakroun, and Tom Haber</i>	
A Software-Defined Storage for Workflow Applications	350
<i>Samer Al-Kiswany and Matei Ripeanu</i>	
VOLAP: A Scalable Distributed System for Real-Time OLAP with High Velocity Data	354
<i>Frank Dehne, David Robillard, Andrew Rau-Chaplin, and Neil Burke</i>	

Data Analytics (Paper Session 11)

Horme: Random Access Big Data Analytics	364
<i>Guangchen Ruan and Beth Plale</i>	
In-cache MapReduce: Leverage Tiling to Boost Temporal Locality-Sensitive MapReduce Computations	374
<i>Daniel Magro and Hervé Paulino</i>	
Intra-host Rate Control with Centralized Approach	384
<i>Zhuang Wang, Ke Liu, Yifan Shen, Jack Y. B. Lee, Mingyu Chen, and Lixin Zhang</i>	
Fast Big Data Analysis in Geo-Distributed Cloud	388
<i>Yue Li, Laiping Zhao, Chenzhou Cui, and Ce Yu</i>	

Optimized Checkpointing (Paper Session 12)

A Lightweight Causal Message Logging Protocol to Lower Fault Tolerance Overhead	392
<i>Jin-Min Yang</i>	
Design and Implementation for Checkpointing of Distributed Resources Using Process-Level Virtualization	402
<i>Kapil Arya, Rohan Garg, Artem Y. Polyakov, and Gene Cooperman</i>	
Deduplication Potential of HPC Applications' Checkpoints	413
<i>Jürgen Kaiser, Ramy Gad, Tim Süß, Federico Padua, Lars Nagel, and André Brinkmann</i>	

Big Data (Paper Session 13)

A Comparative Survey of the HPC and Big Data Paradigms: Analysis and Experiments	423
<i>HamidReza Asaadi, Dounia Khaldi, and Barbara Chapman</i>	
Spark Versus Flink: Understanding Performance in Big Data Analytics Frameworks	433
<i>Ovidiu-Cristian Marcu, Alexandru Costan, Gabriel Antoniu, and María S. Pérez-Hernández</i>	
Results of a Model for Hadoop YARN MapReduce Tasks	443
<i>Thomas C. Bressoud and Qiuyi Tang</i>	
Application-Assisted Writeback for Hadoop Clusters	447
<i>Jungi Jeong, Daewoo Lee, and Seungryoul Maeng</i>	
Custody: Towards Data-Aware Resource Sharing in Cloud-Based Big Data Processing	451
<i>Shiyao Ma, Jingjie Jiang, Bo Li, and Baochun Li</i>	

Threads: Heterogeneity and Tasking (Paper Session 14)

ARCS: Adaptive Runtime Configuration Selection for Power-Constrained OpenMP Applications	461
<i>Md Abdullah Shahneous Bari, Nicholas Chaimov, Abid M. Malik, Kevin A. Huck, Barbara Chapman, Allen D. Malony, and Osman Sarood</i>	
A Review of Lightweight Thread Approaches for High Performance Computing	471
<i>Adrián Castelló, Antonio J. Peña, Sangmin Seo, Rafael Mayo, Pavan Balaji, and Enrique S. Quintana-Ortí</i>	
Directive-Based Pipelining Extension for OpenMP	481
<i>Xuewen Cui, Thomas R. W. Scogland, Bronis R. de Supinski, and Wu-Chun Feng</i>	
Serving More GPU Jobs, with Low Penalty, Using Remote GPU Execution and Migration	485
<i>Pak Markthub, Akihiro Nomura, and Satoshi Matsuoka</i>	
VarySched: A Framework for Variable Scheduling in Heterogeneous Environments	489
<i>Tim Süß, Nils Döring, Ramy Gad, Lars Nagel, André Brinkmann, Dustin Feld, Thomas Soddemann, and Stefan Lankes</i>	
Extending the Roofline Model for Asynchronous Many-Task Runtimes	493
<i>Joshua D. Suetterlein, Joshua Landwehr, Andres Marquez, Joseph Manzano, and Guang R. Gao</i>	

FTS 2016 — The Second International Workshop on Fault Tolerant Systems

FTS 2016 Workshop Keynote Speech	497
<i>David Abramson</i>	
A Runtime Heuristic to Selectively Replicate Tasks for Application-Specific Reliability Targets	498
<i>Omer Subasi, Gulay Yalcin, Ferad Zyulkyarov, Osman Unsal, and Jesus Labarta</i>	
TwinPCG: Dual Thread Redundancy with Forward Recovery for Preconditioned Conjugate Gradient Methods	506
<i>Kiril Dichev and Dimitrios S. Nikolopoulos</i>	
An ABFT Scheme Based on Communication Characteristics	515
<i>Upama Kabir and Dhrubajyoti Goswami</i>	
Separation Kernel Robustness Testing: The XtratuM Case Study	524
<i>S. Gixti, N. Sammut, M. Hernek, E. Carrascosa, M. Masmano, and A. Crespo</i>	

Additional Paper

MBL: A Multi-Stage Bufferless High-radix Router	532
<i>Wenxiang Yang, Dezun Dong, Jingyue Zhao, and Cunlu Li</i>	
Author Index	534