

2016 16th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid 2016)

**Cartagena, Colombia
16-19 May 2016**



**IEEE Catalog Number: CFP16276-POD
ISBN: 978-1-5090-2454-4**

**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:	CFP16276-POD
ISBN (Print-On-Demand):	978-1-5090-2454-4
ISBN (Online):	978-1-5090-2453-7

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 16th IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing

CCGrid 2016

Table of Contents

Message from the CCGrid 2016 General Chairs.....	xiv
Message from the CCGrid 2016 Program Chairs.....	xvi
Organizing Committee.....	xvii
Technical Program Committee.....	xx
List of Reviewers.....	xxv
Keynotes.....	xxx

Regular Papers

Scheduling and Resource Management

Automatic Communication Optimization of Parallel Applications in Public Clouds	1
<i>Emmanuel D. Carreño, Matthias Diener, Eduardo H. M. Cruz, and Philippe O. A. Navaux</i>	
Tyrex: Size-Based Resource Allocation in MapReduce Frameworks	11
<i>Bogdan Ghiț and Dick Epema</i>	
Demand-Aware Power Management for Power-Constrained HPC Systems	21
<i>Thang Cao, Yuan He, and Masaaki Kondo</i>	
Landrush: Rethinking In-Situ Analysis for GPGPU Workflows	32
<i>Anshuman Goswami, Yuan Tian, Karsten Schwan, Fang Zheng, Jeffrey Young, Matthew Wolf, Greg Eisenhauer, and Scott Klasky</i>	
Service Level and Performance Aware Dynamic Resource Allocation in Overbooked Data Centers	42
<i>Luis Tomás, Ewnetu Bayuh Lakew, and Erik Elmroth</i>	

DieHard: Reliable Scheduling to Survive Correlated Failures in Cloud Data Centers	52
<i>Mina Sedaghat, Eddie Wadbro, John Wilkes, Sara De Luna, Oleg Seleznev, and Erik Elmroth</i>	
SHMEMPMI — Shared Memory Based PMI for Improved Performance and Scalability	60
<i>Sourav Chakraborty, Hari Subramoni, Jonathan Perkins, and Dhabaleswar K. (DK) Panda</i>	
DiBA: Distributed Power Budget Allocation for Large-Scale Computing Clusters	70
<i>Masoud Badiei, Xin Zhan, Reza Azimi, Sherief Reda, and Na Li</i>	
KOALA-F: A Resource Manager for Scheduling Frameworks in Clusters	80
<i>Aleksandra Kuzmanovska, Rudolf H. Mak, and Dick Epema</i>	
Elastic Partition Placement for Non-stationary Graph Algorithms	90
<i>Ravikant Dindokar and Yogesh Simmhan</i>	
In-Memory Caching Orchestration for Hadoop	94
<i>Jaewon Kwak, Eunji Hwang, Tae-Kyung Yoo, Beomseok Nam, and Young-Ri Choi</i>	
Increasing the Performance of Data Centers by Combining Remote GPU Virtualization with Slurm	98
<i>Sergio Iserte, Javier Prades, Carlos Reaño, and Federico Silla</i>	
Scheduling In-Situ Analytics in Next-Generation Applications	102
<i>Oscar H. Mondragon, Patrick G. Bridges, Scott Levy, Kurt B. Ferreira, and Patrick Widener</i>	

Applications

CVSS: A Cost-Efficient and QoS-Aware Video Streaming Using Cloud Services	106
<i>Xiangbo Li, Mohsen Amini Salehi, Magdy Bayoumi, and Rajkumar Buyya</i>	
AMRZone: A Runtime AMR Data Sharing Framework for Scientific Applications	116
<i>Wenzhao Zhang, Houjun Tang, Steve Harenberg, Surendra Byna, Xiaocheng Zou, Dharshi Devendran, Daniel F. Martin, Kesheng Wu, Bin Dong, Scott Klasky, and Nagiza F. Samatova</i>	
Generalized GPU Acceleration for Applications Employing Finite-Volume Methods	126
<i>Jingheng Xu, Haohuan Fu, Lin Gan, Chao Yang, Wei Xue, Shizhen Xu, Wenlai Zhao, Xinliang Wang, Bingwei Chen, and Guangwen Yang</i>	
Optimizing Massively Parallel Simulations of Infection Spread Through Air-Travel for Policy Analysis	136
<i>Ashok Srinivasan, C. D. Sudheer, and Sirish Namilae</i>	

Tigres Workflow Library: Supporting Scientific Pipelines on HPC Systems	146
<i>Valerie Hendrix, James Fox, Devarshi Ghoshal, and Lavanya Ramakrishnan</i>	
Evaluation of In-Situ Analysis Strategies at Scale for Power Efficiency and Scalability	156
<i>Ivan Rodero, Manish Parashar, Aaditya G. Landge, Sidharth Kumar, Valerio Pascucci, and Peer-Timo Bremer</i>	
cuART: Fine-Grained Algebraic Reconstruction Technique for Computed Tomography Images on GPUs	165
<i>Xiaodong Yu, Hao Wang, Wu-Chun Feng, Hao Gong, and Guohua Cao</i>	
A Distributed System for Storing and Processing Data from Earth-Observing Satellites: System Design and Performance Evaluation of the Visualisation Tool	169
<i>Marek Szuba, Parinaz Ameri, Udo Grabowski, Jörg Meyer, and Achim Streit</i>	
Towards Fast Overlapping Community Detection	175
<i>Ismail El-Helw, Rutger Hofman, and Henri E. Bal</i>	
Infrastructure Cost Comparison of Running Web Applications in the Cloud Using AWS Lambda and Monolithic and Microservice Architectures	179
<i>Mario Villamizar, Oscar Garcés, Lina Ochoa, Harold Castro, Lorena Salamanca, Mauricio Verano, Rubby Casallas, Santiago Gil, Carlos Valencia, Angee Zambrano, and Mery Lang</i>	
Elastic Virtual Machine Scheduling for Continuous Air Traffic Optimization	183
<i>Shigeru Imai, Stacy Patterson, and Carlos A. Varela</i>	
Exploiting Sample Diversity in Distributed Machine Learning Systems	187
<i>Zhiqiang Liu, Xuanhua Shi, and Hai Jin</i>	
 Performance Modeling and Evaluation	
OptEx: A Deadline-Aware Cost Optimization Model for Spark	193
<i>Subhajit Sidhanta, Wojciech Golab, and Supratik Mukhopadhyay</i>	
Design and Experimental Evaluation of Distributed Heterogeneous Graph-Processing Systems	203
<i>Yong Guo, Ana Lucia Varbanescu, Dick Epema, and Alexandru Iosup</i>	
DocLite: A Docker-Based Lightweight Cloud Benchmarking Tool	213
<i>Blesson Varghese, Lawan Thamsuhang Subba, Long Thai, and Adam Barker</i>	
An Automated Tool Profiling Service for the Cloud	223
<i>Ryan Chard, Kyle Chard, Bryan Ng, Kris Bubendorfer, Alex Rodriguez, Ravi Madduri, and Ian Foster</i>	
Hubbub-Scale: Towards Reliable Elastic Scaling under Multi-tenancy	233
<i>Navaneeth Rameshan, Ying Liu, Leandro Navarro, and Vladimir Vlassov</i>	

Online Power Estimation of Graphics Processing Units	245
<i>Vignesh Adhinarayanan, Balaji Subramaniam, and Wu-Chun Feng</i>	
A Hybrid Simulation Model for Data Grids	255
<i>Martin Barisits, Eva Kühn, and Mario Lassnig</i>	
Impact and Limitations of Point-to-Point Performance on Collective Algorithms	261
<i>Shweta Jha and Edgar Gabriel</i>	
Distem: Evaluation of Fault Tolerance and Load Balancing Strategies in Real HPC Runtimes through Emulation	267
<i>Cristian Ruiz, Joseph Emeras, Emmanuel Jeanvoine, and Lucas Nussbaum</i>	
Diagnosing Performance Bottlenecks in Massive Data Parallel Programs	273
<i>Vinicius Dias, Ruens Moreira, Wagner Meira Jr., and Dorgival Guedes</i>	
Time Provisioning Evaluation of KVM, Docker and Unikernels in a Cloud Platform	277
<i>Bruno Xavier, Tiago Ferreto, and Luis Jersak</i>	

Programming Models and Runtime Systems

GoDB: From Batch Processing to Distributed Querying over Property Graphs	281
<i>Nitin Jamadagni and Yogesh Simmhan</i>	
OS-Based NUMA Optimization: Tackling the Case of Truly Multi-thread Applications with Non-partitioned Virtual Page Accesses	291
<i>Ilaria Di Gennaro, Alessandro Pellegrini, and Francesco Quaglia</i>	
iGiraph: A Cost-Efficient Framework for Processing Large-Scale Graphs on Public Clouds	301
<i>Safiollah Heidari, Rodrigo N. Calheiros, and Rajkumar Buyya</i>	
Faster: A Low Overhead Framework for Massive Data Analysis	311
<i>Matheus C. Santos, Wagner Meira Jr., Dorgival Guedes, and Virgílio F. Almeida</i>	
Flexible Data-Aware Scheduling for Workflows over an In-memory Object Store	321
<i>Francisco Rodrigo Duro, Javier Garcia Blas, Florin Isaila, Justin M. Wozniak, Jesus Carretero, and Rob Ross</i>	
Low Latency and Resource-Aware Program Composition for Large-Scale Data Analysis	325
<i>Masahiro Tanaka, Kenjiro Taura, and Kentaro Torisawa</i>	
Checkpointing to Minimize Completion Time for Inter-Dependent Parallel Processes on Volunteer Grids	331
<i>Mohammad Tanvir Rahman, Hien Nguyen, Jaspal Subhlok, and Gopal Pandurangan</i>	

mCluster: A Software Framework for Portable Device-Based Volunteer Computing	336
<i>Dimitris Theodoropoulos, Grigorios Chrysos, Iosif Koidis, George Charitopoulos, Emmanouil Pissadakis, Antonis Varikos, Dionisios Pnevmatikatos, Georgios Smaragdos, Christos Strydis, and Nikos Zervos</i>	
HPC-Reuse: Efficient Process Creation for Running MPI and Hadoop MapReduce on Supercomputers	342
<i>Thanh-Chung Dao and Shigeru Chiba</i>	

Storage and I/O

CLARISSE: A Middleware for Data-Staging Coordination and Control on Large-Scale HPC Platforms	346
<i>Florin Isaila, Jesus Carretero, and Rob Ross</i>	
Usage Pattern-Driven Dynamic Data Layout Reorganization	356
<i>Houjun Tang, Suren Byna, Steve Harenberg, Xiaocheng Zou, Wenzhao Zhang, Kesheng Wu, Bin Dong, Oliver Rubel, Kristofer Bouchard, Scott Klasky, and Nagiza F. Samatova</i>	
File System Scalability with Highly Decentralized Metadata on Independent Storage Devices	366
<i>Paul Hermann Lensing, Toni Cortes, Jim Hughes, and André Brinkmann</i>	
DTStorage: Dynamic Tape-Based Storage for Cost-Effective and Highly-Available Streaming Service	376
<i>Jaewon Lee, Jaehyung Ahn, Choongul Park, and Jangwoo Kim</i>	
OptCon: An Adaptable SLA-Aware Consistency Tuning Framework for Quorum-Based Stores	388
<i>Subhajit Sidhanta, Wojciech Golab, Supratik Mukhopadhyay, and Saikat Basu</i>	
Indexing Blocks to Reduce Space and Time Requirements for Searching Large Data Files	398
<i>Tzuhsien Wu, Hao Shyng, Jerry Chou, Bin Dong, and Kesheng Wu</i>	
On Efficient Hierarchical Storage for Big Data Processing	403
<i>K. R. Krish, Bharti Wadhwa, M. Safdar Iqbal, M. Mustafa Rafique, and Ali R. Butt</i>	
Towards Memory-Optimized Data Shuffling Patterns for Big Data Analytics	409
<i>Bogdan Nicolae, Carlos Costa, Claudia Misale, Kostas Katrinis, and Yoonho Park</i>	

Security, Privacy and Reliability

Spatial Support Vector Regression to Detect Silent Errors in the Exascale Era	413
<i>Omer Subasi, Sheng Di, Leonardo Bautista-Gomez, Prasanna Balaprakash, Osman Unsal, Jesus Labarta, Adrian Cristal, and Franck Cappello</i>	
Proof of Violation for Trust and Accountability of Cloud Database Systems	425
<i>Gwan-Hwan Hwang and Shih-Kai Fu</i>	
CACPPA: A Cloud-Assisted Conditional Privacy Preserving Authentication Protocol for VANET	434
<i>Ubaidullah Rajput, Fizza Abbas, Jian Wang, Hasoo Eun, and Heekuck Oh</i>	
Medusa: An Efficient Cloud Fault-Tolerant MapReduce	443
<i>Pedro A. R. S. Costa, Xiao Bai, Fernando M. V. Ramos, and Miguel Correia</i>	
Say Hello Again: Privacy Preserving Matchmaking Using Cloud in Encounter Based Mobile Social Networks	453
<i>Fizza Abbas, Ubaidullah Rajput, Jian Wang, Hasoo Eun, and Heekuck Oh</i>	
A Formal Approach for Service Composition in a Cloud Resources Sharing Context	458
<i>Kais Klai and Hanen Ochi</i>	

Mobile and Hybrid Clouds

CloudSwap: A Cloud-Assisted Swap Mechanism for Mobile Devices	462
<i>Dongju Chae, Joonsung Kim, Youngsok Kim, Jangwoo Kim, Kyung-Ah Chang, Sang-Bum Suh, and Hyogun Lee</i>	
KVLight: A Lightweight Key-Value Store for Distributed Access in Cloud	473
<i>Jiaan Zeng and Beth Plale</i>	
Efficient Heuristics for Placing Large-Scale Distributed Applications on Multiple Clouds	483
<i>Pedro Silva, Christian Perez, and Frédéric Desprez</i>	
Billing System CPU Time on Individual VM	493
<i>Boris Teabe, Alain Tchana, and Daniel Hagimont</i>	
COMPSs-Mobile: Parallel Programming for Mobile-Cloud Computing	497
<i>Francesc Lordan and Rosa M. Badia</i>	

Datacenters and Cyberinfrastructure

Dynamic Adaptation of Policies Using Machine Learning	501
<i>Alejandro Pelaez, Andres Quiroz, and Manish Parashar</i>	
De-Fragmenting the Cloud	511
<i>Mayank Mishra and Umesh Bellur</i>	

Towards Understanding Job Heterogeneity in HPC: A NERSC Case Study	521
<i>Gonzalo Pedro Rodrigo Alvarez, Per-Olov Östberg, Erik Elmroth, Katie Antypas, Richard Gerber, and Lavanya Ramakrishnan</i>	
An Improved Model for Live Migration in Data Centre Simulators	527
<i>Vincenzo De Maio, Gabor Kecskemeti, and Radu Prodan</i>	
Fogbow: A Middleware for the Federation of IaaS Clouds	531
<i>Francisco Brasileiro, Giovanni Silva, Francisco Araújo, Marcos Nóbrega Jr., Igor Silva, and Gustavo Rocha</i>	
Seeking for the Optimal Energy Modelisation Accuracy to Allow Efficient Datacenter Optimizations	535
<i>Edouard Outin, Jean-Emile Dartois, Olivier Barais, and Jean-Louis Pazat</i>	

Architecture and Networking

Creating Soft Heterogeneity in Clusters Through Firmware Re-configuration	540
<i>Xin Zhan, Mohammed Shoaib, and Sherief Reda</i>	
RMA-MT: A Benchmark Suite for Assessing MPI Multi-threaded RMA Performance	550
<i>Matthew G. F. Dosanjh, Taylor Groves, Ryan E. Grant, Ron Brightwell, and Patrick G. Bridges</i>	
sAXI: A High-Efficient Hardware Inter-Node Link in ARM Server for Remote Memory Access	560
<i>Ke Zhang, Yisong Chang, Lixin Zhang, Mingyu Chen, Lei Yu, and Zhiwei Xu</i>	
HPC Job Mapping over Reconfigurable Wireless Links	570
<i>Yao Hu, Ikki Fujiwara, and Michihiro Koibuchi</i>	

SCALE Challenge

Large Scale GPU Accelerated PPMLR-MHD Simulations for Space Weather Forecast	576
<i>Xiangyu Guo, Binbin Tang, Jian Tao, Zhaohui Huang, and Zhihui Du</i>	
Exploring Scalability in Pattern Finding in Galactic Structure Using MapReduce	582
<i>Anca Vulpe and Marc Frincu</i>	

Doctoral Symposium

Service Level Agreement Assurance between Cloud Services Providers and Cloud Customers	588
<i>Abdallah Ali Zainelabden Abdallah Ibrahim, Dzmityr Kliazovich, and Pascal Bouvry</i>	
Towards a Resource Manager for Scheduling Frameworks	592
<i>Aleksandra Kuzmanovska, Rudolf H. Mak, and Dick Epema</i>	

Multiobjective Workflow Scheduling in a Federation of Heterogeneous Green-Powered Data Centers	596
<i>Santiago Iturriaga, Sergio Nesmachnow, Andrei Tchernykh, and Bernabé Dorronsoro</i>	
High Performance On-demand Video Transcoding Using Cloud Services	600
<i>Xiangbo Li, Mohsen Amini Salehi, and Magdy Bayoumi</i>	
Cost-Efficient Elastic Stream Processing Using Application-Agnostic Performance Prediction	604
<i>Shigeru Imai, Stacy Patterson, and Carlos A. Varela</i>	

Poster Papers

EDISON: A Web-Based HPC Simulation Execution Framework for Large-Scale Scientific Computing Software	608
<i>Young-Kyoon Suh, Hoon Ryu, Hangi Kim, and Kum Won Cho</i>	
Graph-Oriented Code Transformation Approach for Register-Limited Stencils on GPUs	613
<i>Mengyao Jin, Haohuan Fu, Zihong Lv, and Guangwen Yang</i>	
Machine Learning Approach for Cloud NoSQL Databases Performance Modeling	617
<i>Victor A. E. Farias, Flávio R. C. Sousa, José G. R. Maia, João P. P. Gomes, and Javam C. Machado</i>	
Reusing Resource Coalitions for Efficient Scheduling on the Intercloud	621
<i>Teodora Selea, Adrian Spataru, and Marc Frincu</i>	

Workshop Papers

First IEEE/ACM International Workshop on Distributed Big Data Management (DBDM 2016)

Big Data Analytics Integrating a Parallel Columnar DBMS and the R. Language	627
<i>Yiqun Zhang, Carlos Ordonez, and Wellington Cabrera</i>	
A Quality-Driven Approach for Building Heterogeneous Distributed Databases: The Case of Data Warehouses	631
<i>Sabrina Abdellaoui, Ladjel Bellatreche, and Fahima Nader</i>	
Management of Distributed Big Data for Social Networks	639
<i>Carson K. Leung and Hao Zhang</i>	
Managing Big Data Analytics Workflows with a Database System	649
<i>Carlos Ordonez and Javier García-García</i>	

I-HASTREAM: Density-Based Hierarchical Clustering of Big Data Streams and Its Application to Big Graph Analytics Tools	656
<i>Marwan Hassani, Pascal Spaus, Alfredo Cuzzocrea, and Thomas Seidl</i>	

Environmental Computing Applications — Theory and Practice (ECATP)

Sensor Data Air Pollution Prediction by Kernel Models	666
<i>Petra Vidnerová and Roman Neruda</i>	

First Workshop on Bioinformatics Applications for Clusters and Clouds (WbioCC-2016)

Software Provisioning Inside a Secure Environment as Docker Containers Using Stroll File-System	674
<i>Abdulrahman Azab and Diana Domanska</i>	
Leveraging High Performance Computing for Bioinformatics: A Methodology that Enables a Reliable Decision-Making	684
<i>Mariza Ferro, Marisa F. Nicolas, Quadalupe Del Rosario Q. Saji, Antonio R. Mury, and Bruno Schulze</i>	

Fostering LATAM

Facilitating the Execution of HPC Workloads in Colombia through the Integration of a Private IaaS and a Scientific PaaS/SaaS Marketplace	693
<i>Harold Castro, Mario Villamizar, Oscar Garcés, Jose Pérez, Rodolfo Caliz, and Pedro F. Pérez Arteaga</i>	
Fostering Collaboration in Energy Research and Technological Developments Applying New Exascale HPC Techniques	701
<i>José María Cella, Philippe O. A. Navaux, Alvaro L. G. A. Coutinho, and Rafael Mayo-García</i>	
The Latin American Giant Observatory: A Successful Collaboration in Latin America Based on Cosmic Rays and Computer Science Domains	707
<i>H. Asorey, L.A. Núñez, M. Suárez-Durán, L.A. Torres-Niño, M. Rodríguez-Pascual, A.J. Rubio-Montero, and R. Mayo-García</i>	
Federated Campus Cloud Colombian Initiative	712
<i>César O. Díaz, Carlos E. Gómez, Harold E. Castro, Carlos J. Barrios, and Holman D. Bolívar</i>	
enerGyPU and enerGyPhi Monitor for Power Consumption and Performance Evaluation on Nvidia Tesla GPU and Intel Xeon Phi	718
<i>John A. García H., Esteban Hernandez B., Carlos E. Montenegro, Philippe O. A. Navaux, and Carlos J. Barrios H.</i>	

Additional Papers:

CUDA Kernel based Collective Reduction Operations on Large-scale GPU Clusters.....	726
<i>Ching-Hsiang Chu, Khaled Hamidouche, Akshay Venkatesh, Ammar Ahmad Awan and Dhabaleswar K. (DK) Panda</i>	
Fast Parallel Stream Compaction for IA-based Multi/Many-core Processors.....	736
<i>Qiao Sun, ChaoYang, Changmao Wu, Leisheng Li, Fangfang Liu</i>	
Author Index	746