22nd International ACM Symposium on High-Performance Parallel and Distributed Computing

HPDC'13

New York City

17-21 June 2013

Program

General Co-chairs:
Manish Parashar, Rutgers University, USA
Jon Weissman, University of Minnesota, USA

Program Co-chairs:
Dick Epema, Delft University of Technology and Eindhoven University of Technology, the Netherlands
Renato Figueiredo, University of Florida, USA and Vrije Universiteit, the Netherlands

Posters Chair:
Ivan Rodero, Rutgers University, USA

Workshops Chair:
Abhishek Chandra, University of Minnesota, USA

HPDC 2013 Workshops Overview

Monday June 17	
Full day **Room - Sutton**	**EEHPC - Workshop on Energy Efficient High Performance Parallel and Distributed Computing**
Full day **Room - Grammercy**	**ScienceCloud - 4th Workshop on Scientific Cloud Computing**
Half day **Room - Herald**	**OrmaCloud - Workshop on Optimization Techniques for Resources Management in Clouds**

Tuesday June 18	
Full day **Room - Sutton**	**FTXS - 3rd Workshop on Fault-Tolerance for HPC at Extreme Scale**
Full day **Room - Grammercy**	**VTDC - 7th International Workshop on Virtualization Technologies in Distributed Computing**
Half day **Room - Herald**	**HPPN – Workshop on High Performance and Programmable Networking**
Half day **Room - Herald**	**CLHS - Workshop on Changing Landscapes in HPC Security**

- The **workshop rooms** are located on the **third floor**
- **Main conference** is in **Crystal Ballroom, 2nd floor**

- **All lunches, banquet and reception** are in the **Grand Ballroom, 2nd floor**
- **Breakfast** on the 19, 20, 21 is at 7:45am, **Grand Ballroom, 2nd floor**

HPDC 2013 Conference Program Overview

Wednesday June 19	
07:45 – 08:45	Breakfast
08:45 – 09:00	Conference Opening
09:00 - 10:00	Session 1: Keynote 1
10:00 – 10:30	Break
10:30 – 12:10	Session 2: I/O- and Data-Intensive Computing
12:10 – 13:30	Lunch
13:30 – 14:45	Session 3: Networks
14:45 – 15:10	Break
15:10 – 16:00	Session 4: Communication
16:00 – 17:00	Session 5: Poster Presentations
17:00 – 18:30	Session 6: Posters + Reception

Thursday June 20	
07:45 – 09:00	Breakfast
09:00 – 10:00	Session 7: Keynote 2
10:00 – 10:30	Break
10:30 – 12:10	Session 8: Checkpointing, Bugs and Errors
12:10 – 13:30	Lunch
13:30 – 15:00	Session 9: Industry session
15:00 – 15:30	Break
15:30 – 17:10	Session 10: Multicore and GPUs
17:10 – 19:00	Social Event
19:00 – 21:00	Banquet

Friday June 21	
07:45 – 09:00	Breakfast
09:00 – 10:00	Session 11: Keynote 3
10:00 – 10:20	Break
10:20 – 11:35	Session 12: Virtualization and Clouds
11:35 – 11:45	Short break
11:45 – 12:45	Session 13: Panel
12:45	Conference Closing (with Best Paper Award)

Wednesday, June 19

08:45-09:00   **Conference Opening**

09:00-10:00   **Session 1: Keynote 1**
Chair: Dick Epema, Delft University of Technology and Eindhoven University of Technology

Keynote speaker: Garth Gibson, Carnegie Mellon University

10:00-10:30   **Break**

10:30-12:10   **Session 2: I/O- and Data-Intensive Computing**
Chair: *Doug Thain, University of Notre Dame*

Scalable In Situ Scientific Data Encoding for Analytical Query Processing
Sriram Lakshminarasimhan, David A. Boyuka II, Saurabh V. Pendse, Xiaocheng Zou, John Jenkins (North Carolina State University, Oak Ridge National Laboratory), Venkatram Vishwanath (Argonne National Laboratory), Michael E. Papka (Argonne National Laboratory, Northern Illinois University), Nagiza F. Samatova (North Carolina State University, Oak Ridge National Laboratory)

Taming Massive Distributed Datasets: Data Sampling Using Bitmap Indices
Yu Su, Gagan Agrawal (Ohio State University), Jonathan Woodring, Kary Myers, Joanne Wendelberger, James Ahrens (Los Alamos National Laboratory)

I/O Acceleration with Pattern Detection
Jun He (Illinois Institute of Technology), John Bent (EMC), Aaron Torres, Gary Grider (Los Alamos National Laboratory), Garth Gibson (Carnegie Mellon University), Carlos Maltzahn (University of California, Santa Cruz), Xian-He Sun (Illinois Institute of Technology)

MTC Envelope: Defining the Capability of Large Scale Computers in the Context of Parallel Scripting Applications
Zhao Zhang, Daniel S. Katz, Michael Wilde (University of Chicago), Justin Wozniak (Argonne National Laboratory), Ian Foster (University of Chicago, Argonne National Laboratory)

12:10-13:30   **Lunch - Grand Ballroom, 2nd floor**

13:30-14:45   **Session 3: Networks**
Chair: *Kartik Gopalan, Binghamton University*

Virtual TCP Offload: Optimizing Ethernet Overlay Performance on Advanced Interconnects
Zheng Cui, Patrick G. Bridges (University of New Mexico), John R. Lange (University of Pittsburgh), Peter A. Dinda (Northwestern University)

Scalanytics: A Declarative Multi-core Platform for Scalable Composable Traffic Analytics
Harjot Gill, Dong Lin, Xianglong Han, Cam Nguyen, Tanveer Gill, Boon Thau Loo (University of Pennsylvania)

CamCubeOS: A Key-based Network Stack for 3D Torus Cluster Topologies
Paolo Costa, Austin Donnelly, Greg O'Shea, Antony Rowstron (Microsoft Research Cambridge)

14:45-15:10   **Break**

15:10-16:00   **Session 4: Communication**
Chair: *Matthew Wolf, Georgia Tech*

NUMA-Aware Shared Memory Collective Communication for MPI
Shigang Li (University of Illinois at Urbana-Champaign), Torsten Hoefler (ETH Zurich), Marc Snir (University of Illinois at Urbana-Champaign)

Modeling Communication in Cache-Coherent SMP Systems - A Case-Study with Xeon Phi
Sabela Ramos Garea (Universidade da Coruña), Torsten Hoefler (ETH Zurich)

16:00-17:00   **Session 5: Poster Presentations**
Chair: *Ivan Rodero, Rutgers University*

17:00-18:30   **Session 6: Posters + Conference Reception - Grand Ballroom, 2nd floor**

Thursday, June 20

09:00-10:00   **Session 7: Keynote 2**
Chair: *Andrew Chien, University of Chicago and Argonne National Lab*

Anton: A Special-Purpose Machine That Achieves a Hundred-Fold Speedup in Biomolecular Simulations
David E. Shaw, DE Shaw Research

10:00-10:30   **Break**

10:30-12:10   **Session 8: Checkpointing, Bugs and Errors**
Chair: *Naoya Maruyama, RIKEN Advanced Institute for Computational Science*

WuKong: Automatically Detecting and Localizing Bugs that Manifest at Large System Scales
Bowen Zhou, Jonathan Too, Milind Kulkarni, Saurabh Bagchi (Purdue University)

A 1 PB/s File System to Checkpoint Three Million MPI Tasks
Raghunath Rajachandrasekar (The Ohio State University), Adam Moody, Kathryn Mohror (Lawrence Livermore National Laboratory), Dhabaleswar K. Panda (The Ohio State University)

AI-Ckpt: Leveraging Memory Access Patterns for Adaptive Asynchronous Incremental Checkpointing
Bogdan Nicolae (IBM Research), Franck Cappello (INRIA Saclay)

Correcting Soft Errors Online in LU Factorization
Teresa Davies (Colorado School of Mines), Zizhong Chen (University of California, Riverside)

12:10-13:30   **Lunch - Grand Ballroom, 2nd floor**

13:30-15:00   **Session 9: Industry Session**
Chair: *Manish Parashar, Rutgers University*

From Academia to Industry: Perspectives on research directions in large scale computation geared to industrial impact
Oscar Boykin (Twitter)

Exploring the Capabilities of a Massively Scalable, Compute-in-Storage Architecture by Close Integration of Solid State Storage (Flash) into the IBM Blue Gene/Q System
Blake G. Fitch (IBM Research)

Bridging the Gap Between Applications and Networks in Data Centers
Paolo Costa (Microsoft Research Cambridge)

15:00-15:30   **Break**

15:30-17:10   **Session 10: Multicore and GPUs**
Chair: *Henri Bal, Vrije Universiteit*

A Preemption-based Runtime to Efficiently Schedule Multi-process Applications on Heterogeneous Clusters with GPUs
Kittisak Sajjapongse, Xiang Wang, Michela Becchi (University of Missouri)

On the Efficacy of GPU-Integrated MPI for Scientific Applications

Ashwin M. Aji, Lokendra S. Panwar (Virginia Tech), Feng Ji (North Carolina State University), Milind Chabbi, Karthik Murthy (Rice University), Pavan Balaji (Argonne National Laboratory), Keith R. Bisset (Virginia Bioinformatics Institute), James Dinan (Argonne National Laboratory), Wu-chun Feng (Virginia Tech), John Mellor-Crummey (Rice University), Xiaosong Ma (North Carolina State University), Rajeev Thakur (Argonne National Laboratory)

VGRIS: Virtualized GPU Resource Isolation and Scheduling in Cloud Gaming
Miao Yu, Chao Zhang, Zhengwei Qi, Jianguo Yao (Shanghai Jiao Tong University)

COSMIC: Middleware for High Performance and Reliable Multiprocessing on Intel Manycore Coprocessors
Srihari Cadambi, Giuseppe Coviello, Cheng-Hong Li, Rajat Phull, Kunal Rao, Murugan Sankaradass, Srimat Chakradhar (NEC Laboratories America, Inc.)

17:10-19:00   Social Event - Guided walking tour of NYC

19:00-21:00   Banquet - Grand Ballroom, 2nd floor

Friday, June 21

09:00-10:00   Session 11: Keynote 3
Chair: *Jon Weissman, University of Minnesota*

From Principles to Capabilities - the Birth and Evolution of High Throughput Computing
Miron Livny, University of Wisconsin-Madison

10:00-10:20   Break

10:20-11:35   Session 12: Virtualization and Clouds
Chair: *Peter Dinda, Northwestern University*

Interference and Locality-Aware Task Scheduling for MapReduce Applications in Virtual Clusters
Xiangping Bu (Wayne State University), Jia Rao (University of Colorado at Colorado Springs), Cheng-Zhong Xu (Wayne State University)

A Comparative Study of High-Performance Computing on the Cloud
Aniruddha Marathe, Rachel Harris, David K. Lowenthal (University of Arizona), Bronis R. de Supinski, Barry Rountree, Martin Schulz (Lawrence Livermore National Laboratory), Xin Yuan (Florida State University)

kMemvisor: Flexible System Wide Memory Mirroring in Virtual Environments
Bin Wang, Wei Sun, Haoliang Dong, Zhengwei Qi (Shanghai Jiao Tong University)

11:35-11:45   Short break

11:45-12:45   Session 13: Panel - How big is your 'Big Data', and how can HPDC help?
Moderator: *Tevfik Kosar, University at Buffalo (SUNY)*

Panelists:
Gagan Agrawal, Ohio State University
Peter Dinda, Northwestern University
Dan Katz, University of Chicago and Argonne National Laboratory
Carlos Maltzahn, University of California - Santa Cruz
Douglas Thain, University of Notre Dame

12:45-13:00   Conference Closing (with Best Paper Award)

Accepted Posters (based on full submitted papers)

IBIS: Interposed Big-data I/O Scheduler
Yiqi Xu, Adrian Suarez, Ming Zhao (Florida International University)

ACIC: Automatic Cloud I/O Configurator for Parallel Applications
Mingliang Liu (Tsinghua University), Ye Jin (North Carolina State University), Jidong Zhai (Tsinghua University), Yan Zhai (University of Wisconsin-Madison), Qianqian Shi (Tsinghua University), Xiaosong Ma (North Carolina State University & Oak Ridge National Laboratory), Wenguang Chen (Tsinghua University)

ElastMan: Autonomic Elasticity Manager for Cloud-Based Key-Value Stores
Ahmad Al-Shishtawy, Vladimir Vlassov (KTH Royal Institute of Technology)

Supporting Parallel Soft Real-Time Applications in Virtualized Environment
Like Zhou, Song Wu, Huahua Sun, Hai Jin, Xuanhua Shi (Huazhong University of Science and Technology)

Building and Scaling Virtual Clusters with Residual Resources from Interactive Clouds
R. Benjamin Clay, Zhiming Shen (North Carolina State University), Xiaosong Ma (North Carolina State University & Oak Ridge National Laboratory)

SCDA: SLA-aware Cloud Datacenter Architecture for Efficient Content Storage and Retrieval
Debessay Fesehaye, Klara Nahrstedt (University of Illinois at Urbana-Champaign)

Load Balancing in Large-scale Epidemiological Simulations
Tariq Kamal, Keith R. Bisset, Ali R. Butt, Youngyun Chungbaek, Madhav Marathe (Virginia Tech)

Efficient Analytics on Ordered Datasets Using MapReduce
Jiangtao Yin (University of Massachusetts, Amherst), Yong Liao, Mario Baldi (Narus Inc.), Lixin Gao (University of Massachusetts, Amherst), Antonio Nucci (Narus Inc.)

A Framework for Auto-Tuning HDF5 Applications
Babak Behzad (University of Illinois at Urbana-Champaign), Joseph Huchette (Rice University, Huong Vu Thanh Luu (University of Illinois at Urbana-Champaign), Ruth Aydt (The HDF Group), Surendra Byna, Yushu Yao (Lawrence Berkeley National Laboratory), Quincey Koziol (The HDF Group), Prabhat (Lawrence Berkeley National Laboratory)

Accepted Posters (based on the call for posters)

Monitoring Large-Scale Cloud Systems with Layered Gossip Protocols
Jonathan Ward, Adam Barker (University of St Andrews)

A Performance Study on Virtual Machine Provisioning and Backup for Storage Design of the HPC Cloud
Yusuke Tanimura, Ryousei Takano, Takahiro Hamanishi, Hidemoto Nakada, Yoshio Tanaka (National Institute of Advanced Industrial Science and Technology)

Hybrid parallelization of a multi-material fluid flow solver on heterogeneous architectures
Mathieu Peybernes, Renaud Motte, Jean-Philippe Braeunig (CEA), Jean-Michel Ghidaglia (ENS de Cachan), Joris Costes (Eurobios)

Taking the Elephant to the Market: Improving Hadoop Market Awareness for Auction-based Clouds
Moussa Taifi (Temple University)

P-HGRMS: A Parallel Hypergraph Based Root Mean Square Algorithm for Image Denoising
Tejaswi Agarwal, Saurabh Jha, Rajesh Kanna Baskaran (Vellore Institute of Technology - Chennai)

Parallel simulation of large population dynamics
Cristina Montañola-Sales, Josep Casanovas-Garcia, Jose Maria Cela-Espín (Universitat Politècnica de Catalunya - Barcelona Supercomputing Center), Bhakti S. S. Onggo (Lancaster University), Adriana Kaplan-Marcusán (Universitat Autònoma de Barcelona)

Security-Aware Models for Clouds
Arnaud Lefray (INRIA-ENS-ENSIB), Eddy Caron (University of Lyon-ENS-INRIA), Jonathan Rouzaud-Cornabas, Huaxi Yulin Zhang (INRIA-ENS), Aline Bousquet, Jérémy Briffaut, Christian Toinard (ENSIB)

V-BOINC: The Virtualization of BOINC
Gary McGilvary (The University of Edinburgh), Adam Barker (University of St Andrews), Ashley Lloyd, Malcolm Atkinson (The University of Edinburgh)

Prototype of computer cluster monitoring data analysis system on the basis of support vector method
Sergey Konyuhov (RSC Technologies), Ekaterina Tyutlyaeva (Russian Academy of Sciences), Alexander Moskovsky (RSC Technologies)

ConPaaS: an Integrated Runtime Environment for Elastic Cloud Applications
Guillaume Pierre (IRISA/Universite de Rennes 1), Thilo Kielmann, Emanuele Rocca, Kaveh Razavi, Bert Ijff, Hector Fernandez, Renato Figueiredo, Alexandru Uta, Alexandra Vintila (Vrije Universiteit), Ana-Maria Oprescu (University of Amsterdam), Thorsten Schuett, Michael Berlin (ZIB), Matej Artac, Ales Cernivec (XLAB)

An Instruction Level GPU Power Modeling Method
Qi Zhao, Hailong Yang, Zhongzhi Luan, Depei Qian (Beihang University)

Uncovering the Perfect Place: Optimising Workflow Engine Deployment in the Cloud
Michael Luckeneder, Adam Barker (University of St Andrews)

A Distributed Data Component for the Open Modeling Interface
Tom Bulatewicz, Daniel Andresen, Sarah Auvenshine, Jeffrey Peterson, David Steward (Kansas State University)

Towards a multi-gigapixel deconvolution tool for the next generation of astronomical imaging
Michael Sherry, Andy Shearer (NUI Galway)

A Generic Capability Model for Analyzing Modification Effects in HPC Infrastructures
Christian Straube, Dieter Kranzlmüller (Ludwig-Maximilians-Universitat Munchen)

D2T: Doubly Distributed Transactions for High Performance and Distributed Computing
Jai Dayal (Georgia Institute of Technology), Jay Lofstead (Sandia National Labs), Karsten Schwan (Georgia Institute of Technology), Ron Oldfield (Sandia National Labs)