2011 IEEE International Conference on Cluster Computing

(CLUSTER 2011)

Austin, Texas, USA
26 – 30 September 2011

IEEE Catalog Number: CFP11235-PRT
ISBN: 978-1-4577-1355-2

2011 IEEE International Conference on Cluster Computing

CLUSTER 2011

Table of Contents

Message from the Program Co-chairs xiii

Organizing Committee xiv

Technical Program Committee xv

External Reviewers xvii

Advisory Committee xviii

PPAC 2011 Workshop Organizing Committee xix

IASDS 2011 Workshop Organizing Committee xx

Technical Papers: Case Studies

Multiphase LBM Distributed over Multiple GPUs 1
Carlos Rosales

Performance Emulation of Cell-Based AMR Cosmology Simulations 8
Jingjin Wu, Roberto E. González, Zhiling Lan, Nickolay Y. Gnedin, Andrey V. Kravtsov, Douglas H. Rudd, and Yongen Yu

BMF: Bitmapped Mass Fingerprinting for Fast Protein Identification 17
Weikuan Yu, K. John Wu, Wei-Shinn Ku, Cong Xu, and Juan Gao

Technical Papers: Virtualization

Optimizing Network I/O Virtualization with Efficient Interrupt Coalescing and Virtual Receive Side Scaling 26
Yaozu Dong, Dongxiao Xu, Yang Zhang, and Guangdeng Liao

RDMA Based Replication of Multiprocessor Virtual Machines over High-Performance Interconnects 35
Balazs Gerofi and Yutaka Ishikawa

ResourceExchange: Latency-Aware Scheduling in Virtualized Environments with High Performance Fabrics 45
Adit Ranadive, Ada Gavrilovska, and Karsten Schwan

Technical Papers: Large Scale Algorithms

Large-Scale Simulator for Global Data Infrastructure Optimization 54
Sergio Herrero-Lopez, John R. Williams, and Abel Sanchez

Achieving Scalable Parallelization for the Hessenberg Factorization 65
Anthony M. Castaldo and R. Clint Whaley

Design and Implementation of Broadcast Algorithms for Extreme-Scale Systems 74
Pavel Shamis, Richard Graham, Manjunath Gorentla Venkata, and Joshua Ladd

Model-Driven Simulation to Evaluate Performance Impact of Workload Features on Parallel Systems 84
Tran Ngoc Minh and Lex Wolters

Technical Papers: Storage

EDO: Improving Read Performance for Scientific Applications through Elastic Data Organization 93
Yuan Tian, Scott Klasky, Hasan Abbasi, Jay Lofstead, Ray Grout, Norbert Podhorszki, Qing Liu, Yandong Wang, and Weikuan Yu

PIDX: Efficient Parallel I/O for Multi-resolution Multi-dimensional Scientific Datasets 103
Sidharth Kumar, Venkatram Vishwanath, Philip Carns, Brian Summa, Giorgio Scorzelli, Valerio Pascucci, Robert Ross, Jacqueline Chen, Hemanth Kolla, and Ray Grout

AA-Dedupe: An Application-Aware Source Deduplication Approach for Cloud Backup Services in the Personal Computing Environment 112
Yinjin Fu, Hong Jiang, Nong Xiao, Lei Tian, and Fang Liu

Technical Papers: Node Architecture

Incorporating Network RAM and Flash into Fast Backing Store for Clusters 121
Tia Newhall and Douglas Woos

Design of HPC Node with Heterogeneous Processors 130
Zheng Cao, Hongwei Tang, Qiang Li, Bo Li, Fei Chen, Kai Wang, Xuejun An, and Ninghui Sun

Performance Analysis and Benchmarking of the Intel SCC 139
Philipp Gschwandtner, Thomas Fahringer, and Radu Prodan

Technical Papers: Resource Management

Supporting Computing Element Heterogeneity in P2P Grids 150
Jaehwan Lee, Pete Keleher, and Alan Sussman

DARE: Adaptive Data Replication for Efficient Cluster Scheduling 159
Cristina L. Abad, Yi Lu, and Roy H. Campbell

A Framework for Data-Intensive Computing with Cloud Bursting 169
Tekin Bicer, David Chiu, and Gagan Agrawal

Technical Papers: Message Passing

Automatic Hybrid OpenMP + MPI Program Generation for Dynamic Programming Problems 178
Denny R. Vandenberg and Quentin F. Stout

On Scalability for MPI Runtime Systems 187
George Bosilca, Thomas Herault, Ala Rezmerita, and Jack Dongarra

Process Distance-Aware Adaptive MPI Collective Communications 196
Teng Ma, Thomas Herault, George Bosilca, and Jack J. Dongarra

Technical Papers: Workload and Performance Characterization

Experience on Comparison of Operating Systems Scalability on the Multi-core Architecture 205
Yan Cui, Yingxin Wang, Yu Chen, and Yuanchun Shi

Automatic Computer System Characterization for a Parallelizing Compiler 216
Alan Sussman, Norman Lo, and Timothy Anderson

Energy Templates: Exploiting Application Information to Save Energy 225
Darren J. Kerbyson, Abhinav Vishnu, and Kevin J. Barker

Performance Characterization and Optimization of Atomic Operations on AMD GPUs 234
Marwa Elteir, Heshan Lin, and Wu-Chun Feng

Technical Papers: System Performance

Analyzing the Performance Bottlenecks of the POWER7-IH Network 244
Darren J. Kerbyson and Kevin J. Barker

Play It Again, SimMR! 253
Abhishek Verma, Ludmila Cherkasova, and Roy H. Campbell

An ISO-Energy-Efficient Approach to Scalable System Power-Performance Optimization 262
Shuaiwen Song, Matthew Grove, and Kirk W. Cameron

Technical Papers: Fault Tolerance

High Performance Dense Linear System Solver with Soft Error Resilience 272
Peng Du, Piotr Luszczek, and Jack Dongarra

Dynamic Load Balance for Optimized Message Logging in Fault Tolerant HPC Applications 281
Esteban Meneses, Laxmikant V. Kalé, and Greg Bronevetsky

Accelerating Galois Field Arithmetic for Reed-Solomon Erasure Codes in Storage Applications 290
Sebastian Kalcher and Volker Lindenstruth

Technical Papers: Communication

A Sampling-Based Approach for Communication Libraries Auto-Tuning 299
Élisabeth Brunet, François Trahay, Alexandre Denis, and Raymond Namyst

Optimized Non-contiguous MPI Datatype Communication for GPU Clusters: Design, Implementation and Evaluation with MVAPICH2 308
Hao Wang, Sreeram Potluri, Miao Luo, Ashish Kumar Singh, Xiangyong Ouyang, Sayantan Sur, and Dhabaleswar K. Panda

Design and Evaluation of Network Topology-/Speed- Aware Broadcast Algorithms for InfiniBand Clusters 317
H. Subramoni, K. Kandalla, J. Vienne, S. Sur, B. Barth, K. Tomko, R. Mclay, K. Schulz, and D.K. Panda

Technical Papers: Scheduling

An RMS for Non-predictably Evolving Applications 326
Cristian Klein and Christian Pérez

Automatic Task Re-organization in MapReduce 335
Zhenhua Guo, Marlon Pierce, Geoffrey Fox, and Mo Zhou

Evolutionary Scheduling of Parallel Tasks Graphs onto Homogeneous Clusters 344
Sascha Hunold and Joachim Lepping

Symphony: A Scheduler for Client-Server Applications on Coprocessor-Based Heterogeneous Clusters 353
M. Mustafa Rafique, Srihari Cadambi, Kunal Rao, Ali R. Butt, and Srimat Chakradhar

PPAC 2011 Workshop

Multicore/GPGPU Portable Computational Kernels via Multidimensional Arrays 363
H. Carter Edwards, Daniel Sunderland, Chris Amsler, and Sam Mish

Implementation of Multigrid on QPACE 371
Matthias Bolten, Daniel Brinkers, Ulrich Rüde, and Markus Stürmer

Heterogeneous Cloud Computing 378
Steve Crago, Kyle Dunn, Patrick Eads, Lorin Hochstein, Dong-In Kang, Mikyung Kang, Devendra Modium, Karandeep Singh, Jinwoo Suh, and John Paul Walters

Exploring Fine-Grained Task-Based Execution on Multi-GPU Systems 386
Long Chen, Oreste Villa, and Guang R. Gao

Performance Portability of a GPU Enabled Factorization with the DAGuE Framework 395
George Bosilca, Aurelien Bouteiller, Thomas Herault, Pierre Lemarinier, Narapat Ohm Saengpatsa, Stanimire Tomov, and Jack J. Dongarra

CULZSS: LZSS Lossless Data Compression on CUDA 403
Adnan Ozsoy and Martin Swany

Quartile and Outlier Detection on Heterogeneous Clusters Using Distributed Radix Sort 412
Kyle L. Spafford, Jeremy S. Meredith, and Jeffrey S. Vetter

MPI Alltoall Personalized Exchange on GPGPU Clusters: Design Alternatives and Benefit 420
Ashish Kumar Singh, Sreeram Potluri, Hao Wang, Krishna Kandalla, Sayantan Sur, and Dhabaleswar K. Panda

IASDS 2011 Workshop

Automatically Selecting the Number of Aggregators for Collective I/O Operations 428
Mohamad Chaarawi and Edgar Gabriel

Improving I/O Forwarding Throughput with Data Compression 438
Benjamin Welton, Dries Kimpe, Jason Cope, Christina M. Patrick, Kamil Iskra, and Robert Ross

Application I/O and Data Management 446
William W. Dai

FastQuery: A Parallel Indexing System for Scientific Data 455
Jerry Chou, Kesheng Wu, and Prabhat

Parallel I/O Performance for Application-Level Checkpointing on the Blue Gene/P System 465
Jing Fu, Misun Min, Robert Latham, and Christopher D. Carothers

Methodology for Performance Evaluation of the Input/Output System on Computer Clusters 474
Sandra Méndez, Dolores Rexachs, and Emilio Luque

Can a Decentralized Metadata Service Layer Benefit Parallel Filesystems? 484
Vilobh Meshram, Xavier Besseron, Xiangyong Ouyang, Raghunath Rajachandrasekar, Ravi Prakash Darbha, and Dhabaleswar K. Panda

Asynchronous Collective Output with Non-dedicated Cores 494
Phil Miller, Shen Li, and Chao Mei

Posters: Hardware

Improving PCM Endurance with Randomized Address Remapping in Hybrid Memory System 503
Gang Wu, Jian Gao, Huxing Zhang, and Yaozu Dong

Posters: Resource Scheduling and Management

HEaRS: A Hierarchical Energy-Aware Resource Scheduler for Virtualized Data Centers 508
Hui Chen, Meina Song, Junde Song, Ada Gavrilovska, and Karsten Schwan

Parallel Greedy Genetic Algorithm for Job Scheduling in Cluster Enviornments 513
Gholamali Rahnavard, Jharrod Lafon, and Hadi Sharifi

Scheduling Workflows in Opportunistic Environments 517
Maria Del Mar Lopez, Elisa Heymann, and Miquel Angel Senar

TDP-Shell: A Generic Framework to Improve Interoperability between Batch Queue Systems and Monitoring Tools 522
Vicente J. Ivars, Miquel A. Senar, and Elisa Heymann

Locality-Aware Parallel Process Mapping for Multi-core HPC Systems 527
Joshua Hursey, Jeffrey M. Squyres, and Terry Dontje

Evaluating Performance Impacts of Delayed Failure Repairing on Large-Scale Systems 532
Zhou Zhou, Wei Tang, Ziming Zheng, Zhiling Lan, and Narayan Desai

Reservation-Based Overbooking for HPC Clusters 537
Georg Birkenheuer and André Brinkmann

Posters: Communications

Investigating Scenario-Conscious Asynchronous Rendezvous over RDMA 542
Judicael A. Zounmevo and Ahmad Afsahi

Implementing High Performance Remote Method Invocation in CCA 547
Jian Yin, Khushbu Agarwal, Manoj Krishnan, Daniel Chavarría-Miranda, Ian Gorton, and Tom Epperly

Predictive and Distributed Routing Balancing for High Speed Interconnection Networks 552
Carlos Núñez Castillo, Diego Lugones, Daniel Franco, and Emilio Luque

Posters: Applications, Models and Performance

Improving MapReduce Performance via Heterogeneity-Load-Aware Partition Function 557
Huifeng Sun, Junliang Chen, Chuanchang Liu, Zibin Zheng, Nan Yu, and Zhi Yang

Scalability of Semi-implicit Time Integrators for Nonhydrostatic Galerkin-Based Atmospheric Models on Large Scale Clusters 561
James F. Kelly, Frank X. Giraldo, and Gabriele Jost

Performance Behavior Prediction Scheme for Shared-Memory Parallel Applications 566
John Corredor, Juan Carlos Moure, Dolores Rexachs, Daniel Franco, and Emilio Luque

Performance Optimization of Data Structures Using Memory Access Characterization 570
Ashay Rane and James Browne

Experimental and Numerical Study of the Effect of Geometric Parameters on Liquid Single-Phase Pressure Drop in Micro-scale Pin-Fin Arrays 575
Valerie Pezzullo and Steven Voinier

Posters: Data Centric and Cloud Computing

Data Partitioning on Heterogeneous Multicore Platforms 580
Ziming Zhong, Vladimir Rychkov, and Alexey Lastovetsky

Frequent Itemset Mining on Large-Scale Shared Memory Machines 585
Yan Zhang, Fan Zhang, and Jason Bakos

GPApriori: GPU-Accelerated Frequent Itemset Mining 590
Fan Zhang, Yan Zhang, and Jason Bakos

An Energy-Efficient Scheme for Cloud Resource Provisioning Based on CloudSim 595
Yuxiang Shi, Xiaohong Jiang, and Kejiang Ye

Performance of a Virtual Cluster in a General-Purpose Teaching Laboratory 600
Eric Johnson, Patrick Garrity, Timothy Yates, and Richard A. Brown

Posters: I/O and Storage

Datamation: A Quarter of a Century and Four Orders of Magnitude Later 605
Paolo Bertasi, Michele Bonazza, Marco Bressan, and Enoch Peserico

Author Index 610