

# **2012 IEEE International Conference on Cluster Computing (CLUSTER 2012)**

**Beijing, China  
24 – 28 September 2012**



**IEEE Catalog Number: CFP12235-PRT  
ISBN: 978-1-4673-2422-9**

# 2012 IEEE International Conference on Cluster Computing

## Cluster 2012

### Table of Contents

Greetings from the General Chairs.....	xii
Message from the Program Co-Chairs.....	xiii
Cluster 2012 Organization.....	xiv
Reviewers.....	xviii
Keynote Abstracts.....	xix

---

#### **Session 1: Applications and Algorithms for GPUs**

Adapting Irregular Computations to Large CPU-GPU Clusters in the MADNESS Framework .....	1
<i>Vlad Slavici, Raghu Varier, Gene Cooperman, and Robert J. Harrison</i>	
A GPU-accelerated Branch-and-Bound Algorithm for the Flow-Shop Scheduling Problem .....	10
<i>N. Melab, I. Chakroun, M. Mezmaz, and D. Tuyttens</i>	
A Node-based Parallel Game Tree Algorithm Using GPUs .....	18
<i>Liang Li, Hong Liu, Peiyu Liu, Taoying Liu, Wei Li, and Hao Wang</i>	
Using 1000+ GPUs and 10000+ CPUs for Sedimentary Basin Simulations .....	27
<i>Mei Wen, Huayou Su, Wenjie Wei, Nan Wu, Xing Cai, and Chunyuan Zhang</i>	

#### **Session 2: Conserving Energy**

Evaluating Power-Monitoring Capabilities on IBM Blue Gene/P and Blue Gene/Q .....	36
<i>Kazutomo Yoshii, Kamil Iskra, Rinku Gupta, Pete Beckman, Venkatram Vishwanath, Chenjie Yu, and Susan Coghlan</i>	
eco-IDC: Trade Delay for Energy Cost with Service Delay Guarantee for Internet Data Centers .....	45
<i>Jianying Luo, Lei Rao, and Xue Liu</i>	
Synergy: A Middleware for Energy Conservation in Mobile Devices .....	54
<i>Harshit Kharbanda, Manoj Krishnan, and Roy H. Campbell</i>	

### **Session 3: MapReduce and Services**

Affinity-aware Virtual Cluster Optimization for MapReduce Applications .....	63
<i>Cairong Yan, Ming Zhu, Xin Yang, Ze Yu, Min Li, Youqun Shi, and Xiaolin Li</i>	
Mastiff: A MapReduce-based System for Time-Based Big Data Analytics .....	72
<i>Sijie Guo, Jin Xiong, Weiping Wang, and Rubao Lee</i>	
Community Clustering for Distributed Publish/Subscribe Systems .....	81
<i>Wei Li, Songlin Hu, Jintao Li, and Hans-Arno Jacobsen</i>	
D2T: Doubly Distributed Transactions for High Performance and Distributed Computing .....	90
<i>Jay Lofstead, Jai Dayal, Karsten Schwan, and Ron Oldfield</i>	

### **Session 4: Distributed File Systems**

Cx: Concurrent Execution for the Cross-Server Operations in a Distributed File System .....	99
<i>Letian Yi, Jiwu Shu, Jiixin Ou, and Ying Zhao</i>	
Multicore-Enabled Smart Storage for Clusters .....	108
<i>Zhiyang Ding, Xunfei Jiang, Shu Yin, Xiao Qin, Kai-Hsiung Chang, Xiaojun Ruan, Mohammed I. Alghamdi, and Meikang Qiu</i>	
The Load Rebalancing Problem in Distributed File Systems .....	117
<i>Hsueh-Yi Chung, Che-Wei Chang, Hung-Chang Hsiao, and Yu-Chang Chao</i>	
Clover: A Distributed File System of Expandable Metadata Service Derived from HDFS .....	126
<i>Youwei Wang, Jiang Zhou, Can Ma, Weiping Wang, Dan Meng, and Jason Kei</i>	

### **Session 5: I/O Challenges**

sEBP: Event Based Polling for Efficient I/O Virtualization .....	135
<i>Kun Tian, Yaozu Dong, Xiang Mi, and Haibing Guan</i>	
The Power and Challenges of Transformative I/O .....	144
<i>Adam Manzanares, John Bent, Meghan Wingate, and Garth Gibson</i>	
Damaris: How to Efficiently Leverage Multicore Parallelism to Achieve Scalable, Jitter-free I/O .....	155
<i>Matthieu Dorier, Gabriel Antoniu, Franck Cappello, Marc Snir, and Leigh Orf</i>	

### **Session 6: Performance and QoS**

DOSAS: Mitigating the Resource Contention in Active Storage Systems .....	164
<i>Chao Chen, Yong Chen, and Philip C. Roth</i>	

High Performance and High Capacity Hybrid Shingled-Recording Disk System .....	173
<i>Jiguang Wan, Nannan Zhao, Yifeng Zhu, Jibin Wang, Yu Mao, Peng Chen, and Changsheng Xie</i>	

Replication Based QoS Framework for Flash Arrays .....	182
<i>Nihat Altiparmak and Ali Şaman Tosun</i>	

## **Session 7: Applications and Algorithms for Data and I/O**

Data Partitioning on Heterogeneous Multicore and Multi-GPU Systems Using Functional Performance Models of Data-Parallel Applications .....	191
<i>Ziming Zhong, Vladimir Rychkov, and Alexey Lastovetsky</i>	

A Decoupled Execution Paradigm for Data-Intensive High-End Computing .....	200
<i>Yong Chen, Chao Chen, Xian-He Sun, William D. Gropp, and Rajeev Thakur</i>	

Improving I/O Throughput with PRIMACY: Preconditioning ID-Mapper for Compressing Incompressibility .....	209
<i>Neil Shah, Eric R. Schendel, Sriram Lakshminarasimhan, Saurabh V. Pendse, Terry Rogers, and Nagiza F. Samatova</i>	

## **Session 8: Performance and Visualization**

Performance Analysis of Parallel Processing Systems with Horizontal Decomposition .....	220
<i>Hanwei Chen, Jianwei Yin, and Calton Pu</i>	

Characterization and Comparison of Cloud versus Grid Workloads .....	230
<i>Sheng Di, Derrick Kondo, and Walfredo Cirne</i>	

DisplayCluster: An Interactive Visualization Environment for Tiled Displays .....	239
<i>Gregory P. Johnson, Gregory D. Abram, Brandt Westing, Paul Navrátil, and Kelly Gaither</i>	

## **Session 9: Applications and Algorithms for Scientific Computing**

An Out-of-Core Eigensolver on SSD-equipped Clusters .....	248
<i>Zheng Zhou, Erik Saule, Hasan Metin Aktulga, Chao Yang, Esmond G. Ng, Pieter Maris, James P. Vary, and Umit V. Çatalyürek</i>	

On Optimal and Balanced Sparse Matrix Partitioning Problems .....	257
<i>Anael Grandjean, Johannes Langguth, and Bora Uçar</i>	

Autotuning Stencil-Based Computations on GPUs .....	266
<i>Azamat Mamejanov, Daniel Lowell, Ching-Chen Ma, and Boyana Norris</i>	

Accelerating Expectation-Maximization Algorithms with Frequent Updates .....	275
<i>Jiangtao Yin, Yanfeng Zhang, and Lixin Gao</i>	

## Session 10: Reliability and Consistency

SDM: A Stripe-Based Data Migration Scheme to Improve the Scalability of RAID-6 .....	284
<i>Chentao Wu, Xubin He, Jizhong Han, Huailiang Tan, and Changsheng Xie</i>	
Harmony: Towards Automated Self-Adaptive Consistency in Cloud Storage .....	293
<i>Housseem-Eddine Chihoub, Shadi Ibrahim, Gabriel Antoniu, and María S. Pérez</i>	
Accelerating Distributed Updates with Asynchronous Ordered Writes in a Parallel File System .....	302
<i>Youyou Lu, Jiwu Shu, Shuai Li, and Letian Yi</i>	
HerpRap: A Hybrid Array Architecture Providing Any Point-in-Time Data Tracking for Datacenter .....	311
<i>Lingfang Zeng, Dan Feng, Bo Mao, Jianxi Chen, Qingsong Wei, and Wenguo Liu</i>	

## Session 11: Communications

A New End-to-End Flow-Control Mechanism for High Performance Computing Clusters .....	320
<i>Javier Prades, Federico Silla, José Duato, Holger Fröning, and Mondrian Nüssle</i>	
Minimizing Network Contention in InfiniBand Clusters with a QoS-Aware Data-Staging Framework .....	329
<i>Ragunath Rajachandrasekar, Jai Jaswani, Hari Subramoni, and Dhabaleswar K. Panda</i>	
Adjustable Credit Scheduling for High Performance Network Virtualization .....	337
<i>Zhibo Chang, Jian Li, Ruhui Ma, Zhiqiang Huang, and Haibing Guan</i>	
Low-Latency Collectives for the Intel SCC .....	346
<i>Adan Kohler, Martin Radetzki, Philipp Gschwandtner, and Thomas Fahringer</i>	

## Session 12: Resilience and Load Balancing

Hierarchical Clustering Strategies for Fault Tolerance in Large Scale HPC Systems .....	355
<i>Leonardo Bautista Gomez, Thomas Ropars, Naoya Maruyama, Franck Cappello, and Satoshi Matsuoka</i>	
Hiding Checkpoint Overhead in HPC Applications with a Semi-Blocking Algorithm .....	364
<i>Xiang Ni, Esteban Meneses, and Laxmikant V. Kalé</i>	
Automated Load Balancing Invocation Based on Application Characteristics .....	373
<i>Harshitha Menon, Nikhil Jain, Gengbin Zheng, and Laxmikant Kalé</i>	
Overlay-Centric Load Balancing: Applications to UTS and B&B .....	382
<i>Trong-Tuan Vu, Bilel Derbel, Asim Ali, Ahcène Bendjoudi, and Nouredine Melab</i>	

## Session 13: Applications and Algorithms

PIC: Partitioned Iterative Convergence for Clusters .....	391
<i>Reza Farivar, Anand Raghunathan, Srimat Chakradhar, Harshit Kharbanda, and Roy H. Campbell</i>	
Improving Resource Utilization in MapReduce .....	402
<i>Zhenhua Guo, Geoffrey Fox, Mo Zhou, and Yang Ruan</i>	
Differentiating Your Friends for Scaling Online Social Networks .....	411
<i>Yewei Huang, Qianni Deng, and Yanmin Zhu</i>	
Generic Parallel Programming for Massive Remote Sensing Data Processing .....	420
<i>Yan Ma, Lizhe Wang, Dingsheng Liu, Peng Liu, Jun Wang, and Jie Tao</i>	

## Session 14: Memory and I/O

KNOWAC: I/O Prefetch via Accumulated Knowledge .....	429
<i>Jun He, Xian-He Sun, and Rajeev Thakur</i>	
Evaluation and Optimization of Breadth-First Search on NUMA Cluster .....	438
<i>Zehan Cui, Licheng Chen, Mingyu Chen, Yungang Bao, Yongbing Huang, and Huiwei Lv</i>	
GPGPU Memory Estimation and Optimization Targeting OpenCL Architecture .....	449
<i>Junfeng Zhu, Gang Chen, and Baifeng Wu</i>	
Adaptive and Scalable Optimizations for High Performance SR-IOV .....	459
<i>Zhiqiang Huang, Ruhui Ma, Jian Li, Zhibo Chang, and Haibing Guan</i>	

## Session 15: MPI

Enabling Fast, Noncontiguous GPU Data Movement in Hybrid MPI+GPU Environments .....	468
<i>John Jenkins, James Dinan, Pavan Balaji, Nagiza F. Samatova, and Rajeev Thakur</i>	
Designing an Offloaded Nonblocking MPI_Allgather Collective Using CORE-Direct .....	477
<i>Grigori Inozemtsev and Ahmad Afsahi</i>	
Optimizing Process-to-Core Mappings for Application Level Multi-dimensional MPI Communications .....	486
<i>Christer Karlsson, Teresa Davies, and Zizhong Chen</i>	
On the Effects of CPU Caches on MPI Point-to-Point Communications .....	495
<i>Simone Pellegrini, Torsten Hoefler, and Thomas Fahringer</i>	

## Session 16: Task and Job Scheduling

ANOLE: A Profiling-Driven Adaptive Lock Waiter Detection Scheme for Efficient MP-guest Scheduling .....	504
<i>Jian Zhang, Yaozu Dong, and Jiangang Duan</i>	
Backfilling under Two-tier Virtual Machines .....	514
<i>Xiaocheng Liu, Chen Wang, Xiaogang Qiu, Bing Bing Zhou, Bin Chen, and Albert Y. Zomaya</i>	
A Job Scheduling Design for Visualization Services Using GPU Clusters .....	523
<i>Wei-Hsien Hsu, Chun-Fu Wang, Kwan-Liu Ma, Hongfeng Yu, and Jacqueline H. Chen</i>	

## Poster Session

Cache Promotion Policy Using Re-reference Interval Prediction .....	534
<i>Gangyong Jia, Xi Li, Chao Wang, Xuehai Zhou, and Zongwei Zhu</i>	
Built-in Device Simulator for OS Performance Evaluation .....	538
<i>Junjie Mao, Yu Chen, and Yaozu Dong</i>	
Dynamic Network Forecasting Using SimGrid Simulations .....	542
<i>Matthieu Imbert and Eddy Caron</i>	
BWCC: A FS-Cache Based Cooperative Caching System for Network Storage System .....	546
<i>Liu Shi, Zhenjun Liu, and Lu Xu</i>	
Towards a Cost-Aware Data Migration Approach for Key-Value Stores .....	551
<i>Xiulei Qin, Wenbo Zhang, Wei Wang, Jun Wei, Xin Zhao, and Tao Huang</i>	
DFG-base Dynamic Operation Partitioning for Heterogeneous Multicluster VLIW DSP Processor .....	557
<i>Yangzhao Yang, Zeng Zhao, and Naijie Gu</i>	
Memvisor: Application Level Memory Mirroring via Binary Translation .....	562
<i>Haoliang Dong, Wei Sun, Bin Wang, Haiyang Sun, Zhengwei Qi, Haibing Guan, and Yaozu Dong</i>	
A Locality-based Performance Model for Load-and-Compute Style Computation .....	566
<i>Liang Yuan and Yunquan Zhang</i>	
Transactional Multi-row Access Guarantee in the Key-Value Store .....	572
<i>Yaoguang Wang, Weiming Lu, and Baogang Wei</i>	
Building a TaaS Platform for Web Service Load Testing .....	576
<i>Minzhi Yan, Hailong Sun, Xu Wang, and Xudong Liu</i>	
Design of Hardware-Based Communication Performance Measurement Tool .....	580
<i>Zhan Wang, Zheng Cao, Xiaoli Liu, Yong Su, Feilong Liu, and Xuejun An</i>	
Phase Detection for Loop-Based Programs on Multicore Architectures .....	584
<i>Chao Wang, Xi Li, Dong Dai, Gangyong Jia, and Xuehai Zhou</i>	

SOME: Selective Offloading for a Mobile Computing Environment .....	588
<i>Sehoon Park, Youngil Choi, Qichen Chen, and Heon Y. Yeom</i>	
clone_n(): Parallel Thread Creation for Upcoming Many-Core Architectures .....	592
<i>Balazs Gerofi, Atsushi Hori, and Yutaka Ishikawa</i>	
D4D: Inter-datacenter Bulk Transfers with ISP Friendliness .....	597
<i>Yangyang Li, Hongbo Wang, Peng Zhang, Jiankang Dong, and Shiduan Cheng</i>	
Cloud Based Short Read Mapping Service .....	601
<i>Dong Dai, Xi Li, Chao Wang, and Xuehai Zhou</i>	
Memory Affinity: Balancing Performance, Power, Thermal and Fairness for Multi-core Systems .....	605
<i>Gangyong Jia, Xi Li, Chao Wang, Xuehai Zhou, and Zongwei Zhu</i>	
ME2: Efficient Live Migration of Virtual Machine with Memory Exploration and Encoding .....	610
<i>Yanqing Ma, Hongbo Wang, Jiankang Dong, Yangyang Li, and Shiduan Cheng</i>	
A Reliable and High-Performance Distributed Storage System for P2P-VoD Service .....	614
<i>Jing Zhao, Hongbo Wang, Jiankang Dong, and Shiduan Cheng</i>	
HAaaS: Towards Highly Available Distributed Systems .....	618
<i>Yaoguang Wang, Weiming Lu, Bin Yu, and Baogang Wei</i>	
Towards Fault-Tolerant Energy-Efficient High Performance Computing in the Cloud .....	622
<i>Kurt L. Keville, Rohan Garg, David J. Yates, Kapil Arya, and Gene Cooperman</i>	
<b>Author Index</b> .....	<b>627</b>