

# **2009 9th IEEE International Symposium on Cluster Computing and the Grid**

**(CCGrid)**

**Shanghai, China  
18 – 21 May 2009**



**IEEE Catalog Number: CFP09276-PRT  
ISBN: 978-1-4244-3935-5**

# 9th IEEE/ACM International Symposium on Cluster Computing and the Grid

---

## *CCGrid 2009*

# Table of Contents

**Message from the General Chairs**  
**Message from the Program Co-Chairs**  
**Chairs and Committees**  
**Program Committee**  
**Workshop Organizers**  
**Reviewers**

---

### **Keynote Speaker Abstracts**

Market-Oriented Cloud Computing: Vision, Hype, and Reality of Delivering Computing as the 5th Utility .....	1
<i>Rajkumar Buyya</i>	
Challenges and Opportunities on Parallel/Distributed Programming for Large-scale: From Multi-core to Clouds .....	2
<i>Denis Caromel</i>	
Online Storage and Content Distribution System at a Large Scale: Peer-Assistance and Beyond .....	3
<i>Bo Li</i>	

### **Session: Scheduling in Grid 1**

Efficient Grid Task-Bundle Allocation Using Bargaining Based Self-Adaptive Auction .....	4
<i>Han Zhao and Xiaolin Li</i>	
Scheduling Strategies for Cycle Scavenging in Multicluster Grid Systems .....	12
<i>Ozan Sonmez, Bart Grundeken, Hashim Mohamed, Alexandru Iosup, and Dick Epema</i>	
Developing Scheduling Policies in gLite Middleware .....	20
<i>A. Kretsis, P. Kokkinos, and E. Varvarigos</i>	

Runtime Estimations, Reputation and Elections for Top Performing Distributed Query Scheduling .....	28
<i>Rogério Luís de Carvalho Costa and Pedro Furtado</i>	

### **Session: Peer-to-Peer 1**

Extending Pastry by an Alphanumerical Overlay .....	36
<i>Dominic Battré, André Höing, Martin Raack, Ulf Rerrer-Brusch, and Odej Kao</i>	
Self-Chord: A Bio-inspired Algorithm for Structured P2P Systems .....	44
<i>Agostino Forestiero, Carlo Mastroianni, and Michela Meo</i>	
BloomCast: Efficient Full-Text Retrieval over Unstructured P2Ps with Guaranteed Recall .....	52
<i>Hanhua Chen, Hai Jin, Xucheng Luo, Yunhao Liu, and Lionel M. Ni</i>	
Multicast Trees for Collaborative Applications .....	60
<i>Krzysztof Rzadca, Jackson Tan Teck Yong, and Anwitaman Datta</i>	

### **Session: Power Management**

Energy-Efficient Cluster Computing via Accurate Workload Characterization .....	68
<i>S. Huang and W. Feng</i>	
Markov Model Based Disk Power Management for Data Intensive Workloads .....	76
<i>Rajat Garg, Seung Woo Son, Mahmut Kandemir, Padma Raghavan, and Ramya Prabhakar</i>	
Optimal Power Management for Server Farm to Support Green Computing .....	84
<i>Dusit Niyato, Sivadon Chaisiri, and Lee Bu Sung</i>	
Minimizing Energy Consumption for Precedence-Constrained Applications Using Dynamic Voltage Scaling .....	92
<i>Young Choon Lee and Albert Y. Zomaya</i>	

### **Session: Scheduling in Grid 2**

The Grid Enablement and Sustainable Simulation of Multiscale Physics Applications .....	100
<i>Yingwen Song, Yoshio Tanaka, Hiroshi Takemiya, Aiichiro Nakano, Shuji Ogata, and Satoshi Sekiguchi</i>	
Reliability-Oriented Genetic Algorithm for Workflow Applications Using Max-Min Strategy .....	108
<i>Xiaofeng Wang, Rajkumar Buyya, and Jinshu Su</i>	
Hybrid Re-scheduling Mechanisms for Workflow Applications on Multi-cluster Grid .....	116
<i>Yang Zhang, Charles Koelbel, and Keith Cooper</i>	

## Session: Cloud Computing

The Eucalyptus Open-Source Cloud-Computing System .....	124
<i>Daniel Nurmi, Rich Wolski, Chris Grzegorzczak, Graziano Obertelli, Sunil Soman, Lamia Youseff, and Dmitrii Zagorodnov</i>	
GenLM: License Management for Grid and Cloud Computing Environments .....	132
<i>Mathias Dalheimer and Franz-Josef Pfreundt</i>	
On-Demand Resource Provisioning for BPEL Workflows Using Amazon's Elastic Compute Cloud .....	140
<i>Tim Dörnemann, Ernst Juhnke, and Bernd Freisleben</i>	
Multi-Tiered On-Demand Resource Scheduling for VM-Based Data Center .....	148
<i>Ying Song, Hui Wang, Yaqiong Li, Binqun Feng, and Yuzhong Sun</i>	

## Session: Peer-to-Peer 2

Reliable P2P Feed Delivery .....	156
<i>Anwitaman Datta and Liu Xin</i>	
PAIS: A Proximity-Aware Interest-Clustered P2P File Sharing System .....	164
<i>Haiying Shen</i>	
Adaptive Resource Indexing Technique for Unstructured Peer-to-Peer Networks .....	172
<i>Sumeth Lerthirunwong, Naoya Maruyama, and Satoshi Matsuoka</i>	
Range Query Using Learning-Aware RPS in DHT-Based Peer-to-Peer Networks .....	180
<i>Ze Deng, Dan Feng, Ke Zhou, Zhan Shi, and Chao Luo</i>	

## Session: I/O and File Systems

GMount: An Ad Hoc and Locality-Aware Distributed File System by Using SSH and FUSE .....	188
<i>Nan Dun, Kenjiro Taura, and Akinori Yonezawa</i>	
Improving Parallel Write by Node-Level Request Scheduling .....	196
<i>Kazuki Ohta, Hiroya Matsuba, and Yutaka Ishikawa</i>	
File Clustering Based Replication Algorithm in a Grid Environment .....	204
<i>Hitoshi Sato, Satoshi Matsuoka, and Toshio Endo</i>	
Latency Hiding File I/O for Blue Gene Systems .....	212
<i>Florin Isailă, Javier Garcia Blas, Jesus Carretero, Robert Latham, Samuel Lang, and Robert Ross</i>	

## Session: Workflow

Utility Driven Adaptive Workflow Execution .....	220
<i>Kevin Lee, Norman W. Paton, Rizos Sakellariou, and Alvaro A.A. Fernandes</i>	
Hierarchical Caches for Grid Workflows .....	228
<i>David Chiu and Gagan Agrawal</i>	

Performance under Failures of DAG-based Parallel Computing .....	236
<i>Hui Jin, Xian-He Sun, Ziming Zheng, Zhiling Lan, and Bing Xie</i>	
Combined Fault Tolerance and Scheduling Techniques for Workflow Applications on Computational Grids .....	244
<i>Yang Zhang, Anirban Mandal, Charles Koelbel, and Keith Cooper</i>	
<b>Session: Resource Management</b>	
Harvesting Large-Scale Grids for Software Resources .....	252
<i>Asterios Katsifodimos, George Pallis, and Marios D. Dikaiakos</i>	
Resource Allocation Using Virtual Clusters .....	260
<i>Mark Stillwell, David Schanzenbach, Frédéric Vivien, and Henri Casanova</i>	
Resource Information Aggregation in Hierarchical Grid Networks .....	268
<i>P. Kokkinos and E.A. Varvarigos</i>	
Measuring Fragmentation of Two-Dimensional Resources Applied to Advance Reservation Grid Scheduling .....	276
<i>Julius Gehr and Jörg Schneider</i>	
<b>Session: Data Management</b>	
BLAST Application with Data-Aware Desktop Grid Middleware .....	284
<i>Haiwu He, Gilles Fedak, Bing Tang, and Franck Cappello</i>	
AVSS: An Adaptable Virtual Storage System .....	292
<i>Jian Ke, Xu-dong Zhu, Wen-wu Na, and Lu Xu</i>	
Memory-Mapped File Approach for On-Demand Data Co-allocation on Grids .....	300
<i>Po-Cheng Chen, Jyh-Biau Chang, Yi-Chang Zhuang, Ce-Kuen Shieh, and Tyng-Yeu Liang</i>	
Performance Issues in Parallelizing Data-Intensive Applications on a Multi-core Cluster .....	308
<i>Vignesh T. Ravi and Gagan Agrawal</i>	
<b>Session: Performance Modeling</b>	
Using Templates to Predict Execution Time of Scientific Workflow Applications in the Grid .....	316
<i>Farrukh Nadeem and Thomas Fahringer</i>	
Modeling Job Arrival Process with Long Range Dependence and Burstiness Characteristics .....	324
<i>Tran Ngoc Minh and Lex Wolters</i>	
Modeling Job Lifespan Delays in Volunteer Computing Projects .....	331
<i>Trilce Estrada, Michela Taufer, and Kevin Reed</i>	
A Hybrid Intelligent Method for Performance Modeling and Prediction of Workflow Activities in Grids .....	339
<i>Rubing Duan, Farrukh Nadeem, Jie Wang, Yun Zhang, Radu Prodan, and Thomas Fahringer</i>	

## Session: Virtualization

A Scalable Security Model for Enabling Dynamic Virtual Private Execution Infrastructures on the Internet .....	348
<i>Pascale Vicat-Blanc Primet, Jean-Patrick Gelas, Olivier Mornard, Guilherme Koslovski, Vincent Roca, Lionel Giraud, Johan Montagnat, and Tram Truong Huu</i>	
Self-Tuning Virtual Machines for Predictable eScience .....	356
<i>Sang-Min Park and Marty Humphrey</i>	
Dynamic Provisioning of Virtual Organization Clusters .....	364
<i>Michael A. Murphy, Brandon Kagey, Michael Fenn, and Sebastien Goasguen</i>	
Failure-Aware Construction and Reconfiguration of Distributed Virtual Machines for High Availability Computing .....	372
<i>Song Fu</i>	

## Session: High-Performance Communications and Fault Tolerance

Natively Supporting True One-Sided Communication in MPI on Multi-core Systems with InfiniBand .....	380
<i>G. Santhanaraman, P. Balaji, K. Gopalakrishnan, R. Thakur, W. Gropp, and D.K. Panda</i>	
MPISec I/O: Providing Data Confidentiality in MPI-I/O .....	388
<i>Ramya Prabhakar, Christina Patrick, and Mahmut Kandemir</i>	
Dynamic and Distributed Multipath Routing Policy for High-Speed Cluster Networks .....	396
<i>D. Lugones, D. Franco, and E. Luque</i>	
Handling Persistent States in Process Checkpoint/Restart Mechanisms for HPC Systems .....	404
<i>Pierre Riteau, Adrien Lèbre, and Christine Morin</i>	

## Session: Monitoring and Visualization

Collusion Detection for Grid Computing .....	412
<i>Eugen Staab and Thomas Engel</i>	
Multi-scale Real-Time Grid Monitoring with Job Stream Mining .....	420
<i>Xiangliang Zhang, Michèle Sebag, and Cécile Germain-Renaud</i>	
Towards Visualization Scalability through Time Intervals and Hierarchical Organization of Monitoring Data .....	428
<i>Lucas Mello Schnorr, Guillaume Huard, and Philippe Olivier Alexandre Navaux</i>	

## Session: Matching and Adaptation

A Skyline Approach to the Matchmaking Web Service .....	436
<i>Hyuck Han, Hyungsoo Jung, Shingyu Kim, and Heon Y. Yeom</i>	
Flexible and Efficient In-Vivo Enhancement for Grid Applications .....	444
<i>Dong Kwan Kim, Yang Jiao, and Eli Tilevich</i>	

Evaluating the Divisible Load Assumption in the Context of Economic Grid Scheduling with Deadline-Based QoS guarantees .....	452
<i>Wim Depoorter, Ruben Van den Bossche, Kurt Vanmechelen, and Jan Broeckhove</i>	

### **International Workshop on Cloud Computing (Cloud 2009)**

A Live Storage Migration Mechanism over WAN for Relocatable Virtual Machine Services on Clouds .....	460
<i>Takahiro Hirofuchi, Hirotaka Ogawa, Hidemoto Nakada, Satoshi Itoh, and Satoshi Sekiguchi</i>	
A Model-Based Algorithm for Optimizing I/O Intensive Applications in Clouds Using VM-Based Migration .....	466
<i>Kento Sato, Hitoshi Sato, and Satoshi Matsuoka</i>	
C-Meter: A Framework for Performance Analysis of Computing Clouds .....	472
<i>Nezih Yigitbasi, Alexandru Iosup, Dick Epema, and Simon Ostermann</i>	
Programming Abstractions for Data Intensive Computing on Clouds and Grids .....	478
<i>Chris Miceli, Michael Miceli, Shantenu Jha, Hartmut Kaiser, and Andre Merzky</i>	
Online Risk Analytics on the Cloud .....	484
<i>Hyunjoo Kim, Shivangi Chaudhari, Manish Parashar, and Christopher Marty</i>	
The “Any-Schedulability” Criterion for Providing QoS Guarantees through Advance Reservation Requests .....	490
<i>Shikharesh Majumdar</i>	

### **Workshop on Challenges in the Application of Grids in Healthcare**

Towards a Virtual Research Environment for Paediatric Endocrinology across Europe .....	496
<i>Jipu Jiang, Richard Sinnott, Anthony Stell, John Watt, and Faisal Ahmed</i>	
Grid-Based Sleep Research: Analysis of Polysomnographies Using a Grid Infrastructure .....	502
<i>Dagmar Krefting, Sebastian Canisius, Andreas Hoheisel, Thomas Tolxdorff, and Thomas Penzel</i>	
A Clinical Grid Infrastructure Supporting Adverse Hypotensive Event Prediction .....	508
<i>Anthony Stell, Richard Sinnott, and Jipu Jiang</i>	
A Novel Collaborative Grid Framework for Distributed Healthcare .....	514
<i>Hoang M. Phung, Doan B. Hoang, and Elaine Lawrence</i>	
An E-infrastructure to Support Collaborative Embryo Research .....	520
<i>Adam Barker, Jano I. van Hemert, Richard A. Baldock, and Malcolm P. Atkinson</i>	

## **Workshop on Service-Oriented P2P Networks and Grid Systems (ServP2P 2009)**

Risk Informed Computer Economics .....	526
<i>Bin Li and Lee Gillam</i>	
Towards a Cluster Based Incentive Mechanism for P2P Networks .....	532
<i>Kan Zhang and Nick Antonopoulos</i>	
Block-Based Concurrent and Storage-Aware Data Streaming for Grid Applications with Lots of Small Files .....	538
<i>Wen Zhang, Junwei Cao, Yisheng Zhong, Lianchen Liu, and Cheng Wu</i>	
Framework for Decentralizing Legacy Applications .....	544
<i>Jani Hautakorpi, Gonzalo Camarillo, and David López</i>	
Distributed Indexing for Resource Discovery in P2P Networks .....	550
<i>Marco Hentschel, Maozhen Li, Mahesh Ponraj, and Man Qi</i>	
Transient Analysis of an Overlay Multicast Protocol .....	556
<i>Xiaobing Hou and Shuju Wu</i>	

## **Workshop on Workflow Systems in E-science (WSES09)**

Automating Malware Scanning Using Workflows .....	562
<i>David Stirling, Ian Welch, Peter Komisarzcuk, and Christian Seifert</i>	
Experiment and Workflow Management Using Cyberaide Shell .....	568
<i>Gregor von Laszewski, Andrew Younge, Xi He, Kumar Mahinthakumar, and Lizhe Wang</i>	
An Agent-Based Approach for Autonomic Web Service Workflow Model .....	574
<i>Hongxia Tong, Jian Cao, and Shensheng Zhang</i>	
Cesar-FD: An Effective Stateful Fault Detection Mechanism in Drug Discovery Grid .....	580
<i>Yongjian Wang, Yinan Ren, Ting Chen, Yuanqiang Huang, Zhongzhi Luan, Zhongxin Wu, and Depei Qian</i>	

## **Workshop on E-science/E-research Visualization**

Supporting Distributed Visualization Services for High Performance Science and Engineering Applications—A Service Provider Perspective .....	586
<i>Lakshmi Sastry, Ronald Fowler, Srikanth Nagella, and Jonathan Churchill</i>	
Domain-Specific Groupware Environment for E-research on Chemistry .....	591
<i>Dongmei Yue, Ruisheng Zhang, Chen Zhao, Ruipeng Wei, and Lian Li</i>	
Web-Based Visualization of Atmospheric Nucleation Processes Using Java3D .....	597
<i>Rui Ding, Jinzhu Gao, Bin Chen, J. Ilja Siepmann, and Yi Liu</i>	
EnVision: A Web-Based Tool for Scientific Visualization .....	603
<i>Gregory P. Johnson, Stephen A. Mock, Brandt M. Westing, and Gregory S. Johnson</i>	
Distributed Collaborative Visualization Using Light Field Rendering .....	609
<i>A. Al-Saidi, N.J. Avis, I.J. Grimstead, and O.F. Rana</i>	
WSRF-Based Distributed Visualization .....	615
<i>Yi Liu and Shu Gao</i>	

## **Author Index**