2010 IEEE Second International Conference on Cloud Computing Technology and Science

(CloudCom 2010)

Indianapolis, Indiana, USA
30 November – 3 December 2010

IEEE Catalog Number: CFP10CLU-PRT
ISBN: 978-1-4244-9405-7

2nd IEEE International Conference on Cloud Computing Technology and Science

CloudCom 2010

Table of Contents

Message from the General Chair xiii
Message from the Program Chairs xiv
Message from the Cloud Computing Association xv
Message from the Workshop Chairs xvi
Conference Organization xvii

Main Track

LEMO-MR: Low Overhead and Elastic MapReduce Implementation Optimized for Memory and CPU-Intensive Applications 1
Zacharia Fadika and Madhusudhan Govindaraju

Voronoi-Based Geospatial Query Processing with MapReduce 9
Afsin Akdogan, Ugur Demiryurek, Farnoush Banaei-Kashani, and Cyrus Shahabi

LEEN: Locality/Fairness-Aware Key Partitioning for MapReduce in the Cloud 17
Shadi Ibrahim, Hai Jin, Lu Lu, Song Wu, Bingsheng He, and Li Qi

Applying Twister to Scientific Applications 25
Bingjing Zhang, Yang Ruan, Tak-Lon Wu, Judy Qiu, Adam Hughes, and Geoffrey Fox

Evaluation of MapReduce for Gridding LIDAR Data 33
Sriram Krishnan, Chaitanya Baru, and Christopher Crosby

Correlation Based File Prefetching Approach for Hadoop 41
Bo Dong, Xiao Zhong, Qinghua Zheng, Lirong Jian, Jian Liu, Jie Qiu, and Ying Li

Reliability Support in Virtual Infrastructures 49
Guilherme Koslovski, Wai-Leong Yeow, Cedric Westphal, Tram Truong Huu, Johan Montagnat, and Pascale Vicat-Blanc

Self-Organizing Agents for Service Composition in Cloud Computing 59
J. Octavio Gutierrez-Garcia and Kwang-Mong Sim

elasticLM: A Novel Approach for Software Licensing in Distributed Computing Infrastructures 67
Claudio Cacciari, Francesco D'Andria, Miriam Gozalo, Björn Hagemeier, Daniel Mallmann, Josep Martrat, David García Peréz, Angela Rumpl, Wolfgang Ziegler, and Csilla Zsigri

A Mechanism of Flexible Memory Exchange in Cloud Computing Environments 75
Takeshi Okuda, Eiji Kawai, and Suguru Yamaguchi

A Novel Approach for Cooperative Overlay-Maintenance in Multi-overlay Environments 81
Chin-Jung Hsu, Wu-Chun Chung, Kuan-Chou Lai, Kuan-Ching Li, and Yeh-Ching Chung

Fine-Grained Data Access Control Systems with User Accountability in Cloud Computing 89
Jin Li, Gansen Zhao, Xiaofeng Chen, Dongqing Xie, Chunming Rong, Wenjun Li, Lianzhang Tang, and Yong Tang

Trusted Data Sharing over Untrusted Cloud Storage Providers 97
Gansen Zhao, Chunming Rong, Jin Li, Feng Zhang, and Yong Tang

A Token-Based Access Control System for RDF Data in the Clouds 104
Arindam Khaled, Mohammad Farhan Husain, Latifur Khan, Kevin W. Hamlen, and Bhavani Thuraisingham

Image Distribution Mechanisms in Large Scale Cloud Providers 112
Romain Wartel, Tony Cass, Belmiro Moreira, Ewan Roche, Manuel Guijarro, Sebastien Goasguen, and Ulrich Schwickerath

A Hybrid and Secure Mechanism to Execute Parameter Survey Applications on Local and Public Cloud Resources 118
Hao Sun and Kento Aida

Combinatorial Auction-Based Allocation of Virtual Machine Instances in Clouds 127
Sharrukh Zaman and Daniel Grosu

Cost-Optimal Outsourcing of Applications into the Clouds 135
Immanuel Trummer, Frank Leymann, Ralph Mietzner, and Walter Binder

Towards a Reference Architecture for Semantically Interoperable Clouds 143
Nikolaos Loutas, Vassilios Peristeras, Thanassis Bouras, Eleni Kamateri, Dimitrios Zeginis, and Konstantinos Tarabanis

CloudView: Describe and Maintain Resource View in Cloud 151
Dehui Zhou, Liang Zhong, Tianyu Wo, and Junbin Kang

Performance Analysis of High Performance Computing Applications on the Amazon Web Services Cloud 159
Keith R. Jackson, Lavanya Ramakrishnan, Krishna Muriki, Shane Canon, Shreyas Cholia, John Shalf, Harvey J. Wasserman, and Nicholas J. Wright

Cost-Effective HPC: The Community or the Cloud? 169
Adam G. Carlyle, Stephen L. Harrell, and Preston M. Smith

Combining Grid and Cloud Resources by Use of Middleware for SPMD Applications 177
Brian Amedro, Françoise Baude, Fabrice Huet, and Elton Mathias

Analyzing Electroencephalograms Using Cloud Computing Techniques 185
Kathleen Ericson, Shrideep Pallickara, and Charles W. Anderson

Rapid Processing of Synthetic Seismograms Using Windows Azure Cloud 193
Vedaprakash Subramanian, Liqiang Wang, En-Jui Lee, and Po Chen

Finding Tropical Cyclones on a Cloud Computing Cluster: Using Parallel Virtualization for Large-Scale Climate Simulation Analysis 201
D. Hasenkamp, A. Sim, M. Wehner, and K. Wu

Performing Large Science Experiments on Azure: Pitfalls and Solutions 209
Wei Lu, Jared Jackson, Jaliya Ekanayake, Roger S. Barga, and Nelson Araujo

Exploring Architecture Options for a Federated, Cloud-Based System Biology Knowledgebase 218
Ian Gorton, Yan Liu, and Jian Yin

Usage Patterns to Provision for Scientific Experimentation in Clouds 226
Eran Chinthaka Withana and Beth Plale

Semantics Centric Solutions for Application and Data Portability in Cloud Computing 234
Ajith Ranabahu and Amit Sheth

Affinity-Aware Dynamic Pinning Scheduling for Virtual Machines 242
Zhi Li, Yuebin Bai, Huiyong Zhang, and Yao Ma

Achieving High Throughput by Transparent Network Interface Virtualization on Multi-core Systems 250
Huiyong Zhang, Yuebin Bai, Zhi Li, Niandong Du, and Wentao Yang

Xenrelay: An Efficient Data Transmitting Approach for Tracing Guest Domain 258
Hai Jin, Wenzhi Cao, Pingpeng Yuan, and Xia Xie

Power-Saving in Large-Scale Storage Systems with Data Migration 266
Koji Hasebe, Tatsuya Niwa, Akiyoshi Sugiki, and Kazuhiko Kato

Evaluation and Analysis of GreenHDFS: A Self-Adaptive, Energy-Conserving Variant of the Hadoop Distributed File System 274
Rini T. Kaushik, Milind Bhandarkar, and Klara Nahrstedt

Data Replication and Power Consumption in Data Grids 288
Susan V. Vrbsky, Ming Lei, Karl Smith, and Jeff Byrd

Resource Provisioning for Enriched Services in Cloud Environment 296
Rosy Aoun, Elias A. Doumith, and Maurice Gagnaire

Using Global Behavior Modeling to Improve QoS in Cloud Data Storage Services 304
Jesús Montes, Bogdan Nicolae, Gabriel Antoniu, Alberto Sánchez, and María S. Pérez

Building a Distributed Block Storage System for Cloud Infrastructure 312
Xiaoming Gao, Yu Ma, Marlon Pierce, Mike Lowe, and Geoffrey Fox

REMEM: REmote MEMory as Checkpointing Storage 319
Hui Jin, Xian-He Sun, Yong Chen, and Tao Ke

Resource Allocation with a Budget Constraint for Computing Independent Tasks in the Cloud 327
Weiming Shi and Bo Hong

A Multi-agent Approach for Semantic Resource Allocation 335
Jorge Ejarque, Raül Sirvent, and Rosa M. Badia

Investigating Business-Driven Cloudburst Schedulers for E-Science Bag-of-Tasks Applications 343
David Candeia, Ricardo Araújo, Raquel Lopes, and Francisco Brasileiro

Bag-of-Tasks Scheduling under Budget Constraints 351
Ana-Maria Oprescu and Thilo Kielmann

A Novel Heuristic-Based Task Selection and Allocation Framework in Dynamic Collaborative Cloud Service Platform 360
Biao Song, M.M. Hassan, and Eui-nam Huh

CloudBATCH: A Batch Job Queuing System on Clouds with Hadoop and HBase 368
Chen Zhang and Hans De Sterck

A Novel Parallel Traffic Control Mechanism for Cloud Computing 376
Zheng Li, Nenghai Yu, and Zhuo Hao

Work in Progress/Short Papers

Exploring the Performance Fluctuations of HPC Workloads on Clouds 383
Yaakoub El-Khamra, Hyunjoo Kim, Shantenu Jha, and Manish Parashar

Scheduling Hadoop Jobs to Meet Deadlines 388
Kamal Kc and Kemafor Anyanwu

Petri Net Modeling of the Reconfigurable Protocol Stack for Cloud Computing Control Systems 393
Hui Chen, Chunjie Zhou, Yuanqing Qin, Art Vandenberg, Athanasios V. Vasilakos, and Naixue Xiong

Tree-Based Consistency Approach for Cloud Databases 401
Md. Ashfakul Islam and Susan V. Vrbsky

Application-Oriented Remote Verification Trust Model in Cloud Computing 405
Xiaofei Zhang, Hui Liu, Bin Li, Xing Wang, Haiqiang Chen, and Shizhong Wu

Recommendations for Virtualization Technologies in High Performance Computing 409
Nathan Regola and Jean-Christophe Ducom

A Comparison and Critique of Eucalyptus, OpenNebula and Nimbus 417
Peter Sempolinski and Douglas Thain

Exploratory Project: State of the Cloud, from University of Michigan and Beyond 427
Traci L. Ruthkoski

Self-Caring IT Systems: A Proof-of-Concept Implementation in Virtualized Environments 433
Selvi Kadirvel and José A.B. Fortes

Dynamic Resource Provisioning for Data Streaming Applications in a Cloud Environment 441
Smita Vijayakumar, Qian Zhu, and Gagan Agrawal

User Demand Prediction from Application Usage Pattern in Virtual Smartphone 449
Joon Heo, Kenji Terada, Masashi Toyama, Shunsuke Kurumatani, and Eric Y. Chen

Forecasting for Grid and Cloud Computing On-Demand Resources Based on Pattern Matching 456
Eddy Caron, Frédéric Desprez, and Adrian Muresan

Dynamic Request Allocation and Scheduling for Context Aware Applications Subject to a Percentile Response Time SLA in a Distributed Cloud 464
Keerthana Boloor, Rada Chirkova, Yannis Viniotis, and Tiia Salo

Initial Findings for Provisioning Variation in Cloud Computing 473
M. Suhail Rehman and Majd F. Sakr

VDBench: A Benchmarking Toolkit for Thin-Client Based Virtual Desktop Environments 480
Alex Berryman, Prasad Calyam, Matthew Honigford, and Albert M. Lai

Attaching Cloud Storage to a Campus Grid Using Parrot, Chirp, and Hadoop 488
Patrick Donnelly, Peter Bui, and Douglas Thain

Power of Clouds in Your Pocket: An Efficient Approach for Cloud Mobile Hybrid Application Development 496
Ashwin Manjunatha, Ajith Ranabahu, Amit Sheth, and Krishnaprasad Thirunarayan

CSAL: A Cloud Storage Abstraction Layer to Enable Portable Cloud Applications 504
Zach Hill and Marty Humphrey

Sustainable Network Resource Management System for Virtual Private Clouds 512
Takahiro Miyamoto, Michiaki Hayashi, and Kosuke Nishimura

SafeVanish: An Improved Data Self-Destruction for Protecting Data Privacy 521
Lingfang Zeng, Zhan Shi, Shengjie Xu, and Dan Feng

BetterLife 2.0: Large-Scale Social Intelligence Reasoning on Cloud 529
Dexter H. Hu, Yinfeng Wang, and Cho-Li Wang

Intercloud Security Considerations 537
David Bernstein and Deepak Vij

Performance Considerations of Data Acquisition in Hadoop System 545
Baodong Jia, Tomasz Wiktor Wlodarczyk, and Chunming Rong

Abstractions for Loosely-Coupled and Ensemble-Based Simulations on Azure 550
André Luckow and Shantenu Jha

Research Issues for Software Testing in the Cloud 557
Leah Muthoni Riungu, Ossi Taipale, and Kari Smolander

MapReduce in the Clouds for Science 565
Thilina Gunarathne, Tak-Lon Wu, Judy Qiu, and Geoffrey Fox

Efficient Metadata Generation to Enable Interactive Data Discovery over Large-Scale Scientific Data Collections 573
Sangmi Lee Pallickara, Shrideep Pallickara, Milija Zupanski, and Stephen Sullivan

Special Session: Cloud Computing, HCI, & Design--Sustainability and Social Impacts

Energy Use in the Media Cloud: Behaviour Change, or Technofix? 581
Chris Preist and Paul Shabajee

Enabling Sustainable Clouds via Environmentally Opportunistic Computing 587
Michał Witkowski, Paul Brenner, Ryan Jansen, David B. Go, and Eric Ward

Social Impact of Privacy in Cloud Computing 593
Rui Máximo Esteves and Chunming Rong

On the Sustainability Impacts of Cloud-Enabled Cyber Physical Space 597
Tomasz Wiktor Wlodarczyk and Chunming Rong

Framing the Issues of Cloud Computing & Sustainability: A Design Perspective 603
Yue Pan, Siddharth Maini, and Eli Blevis

An Interface Design for Future Cloud-Based Visualization Services 609
Yuzuru Tanahashi, Cheng-Kai Chen, Stéphane Marchesin, and Kwan-Liu Ma

The Ethics of Cloud Computing: A Conceptual Review 614
Job Timmermans, Bernd Carsten Stahl, Veikko Ikonen, and Engin Bozdag

User Experience and Security in the Cloud – An Empirical Study in the Finnish Cloud Consortium 621
Nilay Oza, Kaarina Karppinen, and Reijo Savola

Cloud Computing for Enhanced Mobile Health Applications 629
M.T. Nkosi and F. Mekuria

International Workshop on Cloud Privacy, Security, Risk, and Trust

OpenPMF SCaaS: Authorization as a Service for Cloud & SOA Applications 634
Ulrich Lang

Security Services Lifecycle Management in On-Demand Infrastructure Services Provisioning 644
Yuri Demchenko, Cees de Laat, Diego R. Lopez, and Joan A. García-Espín

Modeling the Runtime Integrity of Cloud Servers: A Scoped Invariant Perspective 651
Jinpeng Wei, Calton Pu, Carlos V. Rozas, Anand Rajan, and Feng Zhu

Inadequacies of Current Risk Controls for the Cloud 659
M. Auty, S. Creese, M. Goldsmith, and P. Hopkins

A Privacy Impact Assessment Tool for Cloud Computing 667
David Tancock, Siani Pearson, and Andrew Charlesworth

A Framework for Evaluating Clustering Algorithm 677
Joshua Ojo Nehinbe

Do You Get What You Pay For? Using Proof-of-Work Functions to Verify Performance Assertions in the Cloud 687
Falk Koeppe and Joerg Schneider

Privacy, Security and Trust Issues Arising from Cloud Computing 693
Siani Pearson and Azzedine Benameur

CloudSEC: A Cloud Architecture for Composing Collaborative Security Services 703
Jia Xu, Jia Yan, Liang He, Purui Su, and Dengguo Feng

Trust and Cloud Services - An Interview Study 712
Ilkka Uusitalo, Kaarina Karppinen, Arto Juhola, and Reijo Savola

International Workshop on Theory and Practice of MapReduce (MAPRED 2010)

HAMA: An Efficient Matrix Computation with the MapReduce Framework 721
Sangwon Seo, Edward J. Yoon, Jaehong Kim, Seongwook Jin, Jin-Soo Kim, and Seungryoul Maeng

The Two Quadrillionth Bit of Pi is 0! Distributed Computation of Pi with Apache Hadoop 727
Tsz-Wo Sze

Hybrid Map Task Scheduling for GPU-Based Heterogeneous Clusters 733
Koichi Shirahata, Hitoshi Sato, and Satoshi Matsuoka

Pepper: An Elastic Web Server Farm for Cloud Based on Hadoop 741
Subramaniam Krishnan and Jean Christophe Counio

Characterization of Hadoop Jobs Using Unsupervised Learning 748
Sonali Aggarwal, Shashank Phadke, and Milind Bhandarkar

SSS: An Implementation of Key-Value Store Based MapReduce Framework 754
Hirotaka Ogawa, Hidemoto Nakada, Ryousei Takano, and Tomohiro Kudoh

Implementation and Performance Evaluation of a Hybrid Distributed System for Storing and Processing Images from the Web 762
Murali Krishna, Balaji Kannan, Anand Ramani, and Sriram J. Sathish

Cogset vs. Hadoop: Measurements and Analysis 768
Steffen Viken Valvåg, Dag Johansen, and Åge Kvalnes

Howdah - A Flexible Pipeline Framework for Analyzing Genomic Data 776
Steven Lewis, Sheila Reynolds, Hector Rovera, Mike O'Leary, Sarah Killcoyne, Ilya Shmulevich, and John Boyle

Scaling Populations of a Genetic Algorithm for Job Shop Scheduling Problems Using MapReduce 780
Di-Wei Huang and Jimmy Lin

A Study in Hadoop Streaming with Matlab for NMR Data Processing 786
Kalpa Gunaratna, Paul Anderson, Ajith Ranabahu, and Amit Sheth

A MapReduce-Based Architecture for Rule Matching in Production System 790
Bin Cao, Jianwei Yin, Qi Zhang, and Yanming Ye

Author Index 796