

IEEE DataCloud 2012 Concluding Remarks

Chairs

Ioan Raicu, Illinois Institute of Technology & Argonne National Laboratory
Tevfik Kosar, University of Buffalo (SUNY)
Roger Barga, Microsoft Research

IEEE DataCloud 2012
November 11th, 2012

IEEE DataCloud 2012

3rd International Workshop on Data Intensive Computing in the Clouds 2012

<http://www.cse.buffalo.edu/faculty/tkosar/datacloud2012/>

Keynotes



Daniel A. Reed, Ph.D.
VP for Research and Economic Development
Chair in Computational Science and Bioinformatics
University of Iowa

*"Clouds:
From Both Sides, Now"*



Xian-He Sun, Ph.D.
Chair and Professor of Computer Science
Illinois Institute of Technology

*"Advancing Computing
Systems from the
Data-Cloud Point-of-View"*

| | |
|----------------------|---|
| 9:00 am | Opening Remarks Tevfik Kosar, <i>University at Buffalo, USA</i> Ioan Raicu, <i>Illinois Institute of Technology, USA</i> Roger Barga, <i>Microsoft Research, USA</i> |
| 9:10 am | Keynote Address - I: "Clouds: From Both Sides, Now" Daniel A. Reed, <i>University of Iowa, USA</i> |
| 10:10 am | Coffee Break |
| SESSION - I | |
| 10:30am | The Design and Role of a Community Science Cloud: The Open Science Data Cloud Perspective Heidi Alvarez, <i>Florida International University, USA</i> Matthew Greenway, <i>University of Chicago, USA</i> Robert Grossman, <i>University of Chicago, USA</i> Allison Heath, <i>University of Chicago, USA</i> Joel Mambretti, <i>Northwestern University, USA</i> Ray Powell, <i>University of Chicago, USA</i> Rafael Suarez, <i>University of Chicago, USA</i> Walt Wells, <i>Open Cloud Consortium, USA</i> |
| 11:00am | A Social Content Delivery Network for Scientific Cooperation: Vision, Design, and Architecture Kyle Chard, <i>University of Chicago, USA</i> Simon Caton, <i>Karlsruher Institut für Technologie, Germany</i> Omer Rana, <i>Cardiff University, UK</i> Dan Katz, <i>University of Chicago, USA</i> |
| 11:30am | Your Data, Your Way: The IPlant Foundation API Data Services Rion Dooley, <i>Texas Advanced Computing Center, USA</i> Matthew Vaughn, <i>Texas Advanced Computing Center, USA</i> Edwin Skidmore, <i>University of Arizona, USA</i> Steven Terry, <i>Texas Advanced Computing Center, USA</i> Nirav Merchant, <i>University of Arizona, USA</i> Dan Stanzione, <i>Texas Advanced Computing Center, USA</i> |
| 12:00 pm | LUNCH BREAK |
| 1:00 pm | Keynote Address - II: "Advancing Computing Systems from the Data-Cloud Point-of-View" Xian-He Sun, <i>Illinois Institute of Technology, USA</i> |
| SESSION - II | |
| 2:00 pm | Supporting Bulk Synchronous Parallelism in Map-Reduce Queries Leonidas Fegaras, <i>Univ. of Texas at Arlington, USA</i> |
| 2:30 pm | Incremental and Parallel Analytics on Astrophysical Data Streams Dmitry Mishin, <i>Johns Hopkins University, USA</i> Tamas Budavari, <i>Johns Hopkins University, USA</i> Alexander Szalay, <i>Johns Hopkins University, USA</i> Yanil Ahmad, <i>Johns Hopkins University, USA</i> |
| 3:00 pm | Coffee Break |
| SESSION - III | |
| 3:30 pm | Deploying Bioinformatics Workflows on Clouds with Galaxy and Globus Provision Bo Liu, <i>University of Chicago, USA</i> Borja Sotomayor, <i>University of Chicago, USA</i> Ravi Madduri, <i>Argonne National Laboratory, USA</i> Kyle Chard, <i>University of Chicago, USA</i> Ian Foster, <i>University of Chicago, USA</i> |
| 4:00 pm | FRIEDA: Flexible Robust Intelligent Elastic Data Management in Cloud Environments Devarshi Ghoshal, <i>Indiana University, USA</i> Lavanya Ramakrishnan, <i>Lawrence Berkeley National Laboratory, USA</i> |
| 4:30 pm | An Approach to Protect the Privacy of Cloud Data from Data Mining Based Attacks Ilmeel Dew, <i>Bangladesh Univ. of Engineering and Technology, Bangladesh</i> Tannoy Sen, <i>Bangladesh Univ. of Engineering and Technology, Bangladesh</i> Madhusudan Basak, <i>Bangladesh Univ. of Engineering and Technology, Bangladesh</i> Mohammed Eunus Ali, <i>Bangladesh Univ. of Engineering and Technology, Bangladesh</i> |
| 5:00 pm | Closing Remarks Tevfik Kosar, <i>University at Buffalo, USA</i> Ioan Raicu, <i>Illinois Institute of Technology, USA</i> Roger Barga, <i>Microsoft Research, USA</i> |

Chairs



Ioan Raicu, Ph.D.
Data-Intensive Distributed Systems Lab.
Computer Science Department
Illinois Institute of Technology
Math and Computer Science Div.
Argonne National Laboratory



Roger Barga, Ph.D.
Cloud Computing Futures (CCF)
eXtreme Computing Group (XCG)
Microsoft Research



Tevfik Kosar, Ph.D.
Data Intensive Distributed Computing Lab.
Computer Science & Engineering
University at Buffalo (SUNY)



Sponsorship

COMPUTER SCIENCE

IIT College of
Science and Letters



DataSys
Data-Intensive Distributed
Systems Laboratory



GRADUATE COLLEGE

ILLINOIS INSTITUTE OF TECHNOLOGY

Argonne
NATIONAL LABORATORY



Microsoft®

Research



The State University
of New York

IEEE MTAGS 2012

5th ACM Workshop on Many-Task Computing on Grids and Supercomputers

<http://datasys.cs.iit.edu/events/MTAGS12/>

Keynote



Dr. Laxmikant (Sanjay) Kale
Professor
Parallel Programming Laboratory
Department of Computer Science
University of Illinois, Urbana Champaign

Invited Speakers

Petascale Challenge Award

Zhao Zhang
PhD Candidate
Distributed Systems Laboratory
Computer Science Department
Computation Institute
University of Chicago



Biggest Impact Award

| Time | Title | Authors | Affiliation |
|--|--|--|----------------------------------|
| 9:00AM | Opening Statements | Ioan Raicu Yong Zhao Ian Foster Justin Wozniak | IIT/ANL UESTC UChicago/ANL |
| 9:10AM | Keynote -- Adaptive Runtime Systems meet needs of many task computing | Laxmikant V. Kale | UIUC |
| 10:10AM | Break | | |
| Session I -- Middleware (Chair Reagan Moore, UNC) | | | |
| 10:30AM | Invited Talk -- Petascale Challenge Award Data Management for Parallel Scripting | Zhao Zhang | UChicago |
| 11:00AM | A Scalable Master-Worker Architecture for PaaS Clouds | Vibhor Aggarwal Shubhashis Sengupta Vibhu Saujanya Sharma Aravindan Santharam | Accenture |
| 11:30AM | HOG:Distributed Hadoop MapReduce on the Grid | Chen He Derek Weitzel David Swanson Ying Lu | UNL |
| 12:00PM | Lunch | | |
| Session II -- MapReduce (Chair Daniel S. Katz, NSF) | | | |
| 1:30PM | Invited Talk -- Biggest Impact Award IaaS Cloud Benchmarking: Approaches, Challenges, and Experience | Alexandru Iosup | TU Delft |
| 2:00PM | Resource Management for Dynamic MapReduce Clusters in Multicluster Systems | Bogdan Ghit Nezih Yigitbasi Dick Epema | TU Delft |
| 2:30PM | A Hybrid Scheduling Approach for Scalable Heterogeneous Hadoop Systems | Aysan Rasooli Douglas Down | McMaster |
| 3:00PM | Break | | |
| Session III -- Applications (Chair Yong Chen, TTU) | | | |
| 3:30PM | Invited Talk -- Cloud Challenge Award Portable Data Mining on Azure and HPC Platforms | Judy Qiu | IndianaU |
| | | Dan Gunter Shreyas Cholia | 4 |

Call for Papers



<http://www.pds.ewi.tudelft.nl/ccgrid2013/>

GENERAL CHAIR

Dick Epema, Delft University of Technology, the Netherlands

PROGRAM CHAIR

Thomas Fahringer, University of Innsbruck, Austria

PROGRAM VICE-CHAIRS

Rosa Badia, Barcelona Supercomputing Center, Spain
Henri Bal, Vrije Universiteit, the Netherlands
Marios Dikaiakos, University of Cyprus, Cyprus
Kirk Cameron, VirginiaTech, USA
Daniel Katz, University of Chicago & Argonne Nat Lab, USA
Kate Keahey, Argonne National Laboratory, USA
Martin Schulz, Lawrence Livermore National Laboratory, USA
Douglas Thain, University of Notre Dame, USA
Cheng-Zhong Xu, Shenzhen Inst. of Advanced Techn, China

WORKSHOPS CHAIR

Shantenu Jha, Rutgers and Louisiana State University, USA

DOCTORAL SYMPOSIUM CHAIRS

Yogesh Simmhan, University of Southern California, USA
Ana Varbanescu, Delft Univ of Technology, the Netherlands

SUBMISSIONS AND PROCEEDINGS CHAIR

Pavan Balaji, Argonne National Laboratory, USA

FINANCE AND REGISTRATION CHAIR

Alexandru Iosup, Delft Univ of Technology, the Netherlands

PUBLICITY CHAIRS

Nazareno Andrade, Univ Federal de Campina Grande, Brazil
Gabriel Antoniu, INRIA, France

Rapid advances in architectures, networks, and systems and middleware technologies are leading to new concepts in and platforms for computing, ranging from Clusters and Grids to Clouds and Datacenters. CCGrid is a series of very successful conferences, sponsored by the IEEE Computer Society Technical Committee on Scalable Computing (TCSC) and the ACM, with the overarching goal of bringing together international researchers, developers, and users to provide an international forum to present leading research activities and results on a broad range of topics related to these concepts and platforms, and their applications. The conference features keynotes, technical presentations, workshops, tutorials, and posters, as well as the SCALE challenge featuring live demonstrations.

In 2013, CCGrid will come to **the Netherlands** for the first time, and will be held in **Delft, a historical, picturesque city that is less than one hour away from Amsterdam-Schiphol airport**. The main conference will be held on May 14-16 (Tuesday to Thursday), with tutorials and affiliated workshops taking place on May 13 (Monday).

TOPICS OF INTEREST

CCGrid 2013 will have a focus on important and immediate issues that are significantly influencing all aspects of cluster, cloud and grid computing. Topics of interest include, but are not limited to:

- **Applications and Experiences:** Applications to real and complex problems in science, engineering, business, and society; User studies; Experiences with large-scale deployments, systems, or applications
- **Architecture:** System architectures, design and deployment; Power and cooling; Security and reliability; High availability solutions
- **Autonomic Computing and Cyberinfrastructure:** Self-managed behavior, models and technologies; Autonomic paradigms and systems (control-based, bio-inspired, emergent, etc.); Bio-inspired optimizations and computing
- **Cloud Computing:** Cloud architectures; Software tools and techniques for clouds

GENERAL CO-CHAIRS

Manish Parashar, Rutgers University
Jon Weissman, University of Minnesota

PROGRAM CO-CHAIRS

Dick Epema, Delft University of Technology
Renato Figueiredo, University of Florida

WORKSHOPS CHAIR

Abhishek Chandra, University of Minnesota

LOCAL ARRANGEMENTS CHAIR

Daniele Scarpazza, DEShaw Research

SPONSORSHIP CHAIR

Dean Hildebrand, IBM Almaden

PUBLICITY CO-CHAIRS

Alexandru Iosup, Delft University of Technology
Ioan Raicu, Illinois Institute of Technology
Kenjiro Taura, University of Tokyo
Bruno Schulze, LNCC

PROGRAM COMMITTEE

David Abramson, Monash University
Kento Aida, National Institute of Informatics
Gabriel Antoniu, INRIA
Henri Bal, Vrije Universiteit
Adam Barker, University of St Andrews
Michela Becchi, University of Missouri - Columbia
John Bent, EMC
Ali Butt, Virginia Tech
Kirk Cameron, Virginia Tech
Franck Cappello, INRIA & Univ. of Illinois Urbana-Champaign
Henri Casanova, University of Hawaii
Abhishek Chandra, University of Minnesota
Andrew Chien, Univ. of Chicago & Argonne National Lab
Paolo Costa, Imperial College London
Peter Dinda, Northwestern University
Gilles Fedak, INRIA
Ian Foster, Univ. of Chicago & Argonne National Lab
Clemens Grelck, University of Amsterdam
Dean Hildebrand, IBM Research
Fabrice Huet, INRIA-University of Nice
Adriana Iamnitchi, University of South Florida
Alexandru Iosup, Delft University of Technology
Kate Keahey, Argonne National Laboratory
Thilo Kielmann, Vrije Universiteit
Charles Kilian, Purdue University
Zhiling Lan, Illinois Institute of Technology
John Lange, University of Pittsburgh
Barney Maccabe, Oak Ridge National Laboratory
Carlos Maltzahn, University of California, Santa Cruz
Naoya Maruyama, RIKEN
Satoshi Matsuoka, Tokyo Institute of Technology

Call for
Papers



<http://www.hpdc.org/2013/>

The ACM International Symposium on [High-Performance Parallel and Distributed Computing](http://www.hpdc.org/2013/) (HPDC) is the premier annual conference for presenting the latest research on the design, implementation, evaluation, and the use of parallel and distributed systems for high-end computing.

The 22nd HPDC will take place June 17-21, 2013 in the heart of iconic New York City at the New Yorker Hotel.

SCOPE AND TOPICS

Submissions are welcomed on high-performance parallel and distributed computing topics including but not limited to: clusters, clouds, grids, data-intensive computing, massively multicore, and global-scale computing systems. New scholarly research showing empirical and reproducible results in architectures, systems, and networks is strongly encouraged, as are experience reports of operational deployments that can provide insights for future research on HPDC applications and systems. All papers will be evaluated for their originality, technical depth and correctness, potential impact, relevance to the conference, and quality of presentation. Research papers must clearly demonstrate research contributions and novelty, while experience reports must clearly describe lessons learned and demonstrate impact.

In the context of high-performance parallel and distributed computing, the topics of interest include, but are not limited to:

- Systems, networks, and architectures for high-end computing
- Massively multicore systems
- Resource virtualization
- Programming languages and environments
- I/O, storage systems, and data management
- Resource management and scheduling, including energy-aware techniques
- Performance modeling and analysis
- Fault tolerance, reliability, and availability
- Data-intensive computing
- Applications of parallel and distributed computing

ACM ScienceCloud 2013



4th Workshop on Scientific Cloud Computing (ScienceCloud) 2013 - Mozilla Firefox

File Edit View History Bookmarks Tools Help

4th Workshop on Scienti... +

datasys.cs.iit.edu/events/ScienceCloud2013/index.html

Google

Google Google Desktop Google Scholar Microsoft Academic Sear... Ioan Raicu's Web Site

Home

News

Projects

People

Publications

Theses & Dissertations

Events

CS Seminar

DataSys Seminar

Conference

[CFP](#) (PDF, TXT) | [News](#) | [Topics](#) | [Dates](#) | [Submission](#) | [Organization](#) | [Program](#) | [Sponsors](#)

4th Workshop on Scientific Cloud Computing (ScienceCloud) 2013

Co-located with [ACM HPDC 2013](#)
New York City, NY, USA -- June 17th, 2013

[Overview](#)

Computational and Data-Driven Sciences have become the third and fourth pillar of scientific discovery in addition to experimental and theoretical sciences. Scientific Computing has already begun to change how science is done, enabling scientific breakthroughs through new kinds of experiments that would have been impossible only a decade ago. Today's "Data" science is generating datasets that are increasing exponentially in both complexity and volume, making their analysis, archival, and sharing one of the grand challenges of the 21st century. The support for data intensive computing is critical to advance modern science as storage systems have exposed a widening gap between their capacity and their bandwidth by more than 10-fold over the last decade. There is a growing need for advanced techniques to manipulate, visualize and interpret large datasets. Scientific Computing is the key to solving "grand challenges" in many domains and providing breakthroughs in knowledge, and it comes in many shapes and forms: high-performance computing (HPC) which is heavily focused on computing intensive applications; high-throughput computing (HTC) which focuses on using many computing resources over long periods of time to accomplish its computational tasks; many-task computing (MTC) which aims to bridge the gap between HPC and HTC by focusing on using many resources over short periods of time; and data-intensive computing which is heavily focused on data distribution, data-parallel execution, and harnessing data locality by scheduling of computations close to the data.

IEEE CCGrid 2014

- **What: Host CCGrid**
- **When: 2014**
- **Where: Chicago**
- **Who:**



- **Institutions:**

- Illinois Institute of Technology (IIT)
- Argonne National Laboratory (ANL)
- University of Chicago (UChicago)

ILLINOIS INSTITUTE
OF TECHNOLOGY

Argonne
NATIONAL LABORATORY



THE UNIVERSITY OF
CHICAGO

- **General Chairs:**

- Xian-He Sun (IIT)
- Ian T. Foster (UChicago/ANL)





Data-Intensive Distributed Systems Laboratory

[Illinois Institute of Technology](#)
[Department of Computer Science](#)

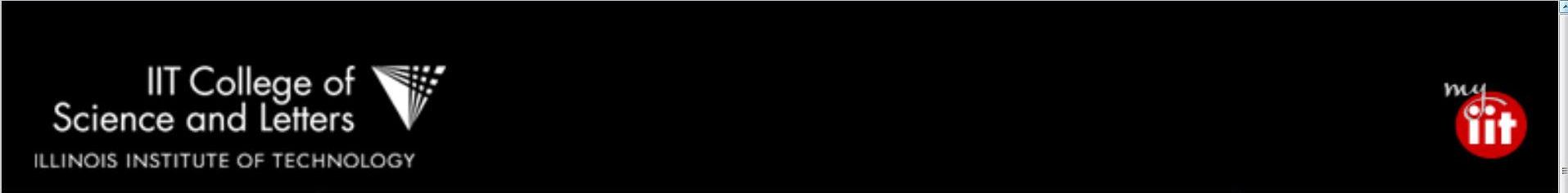
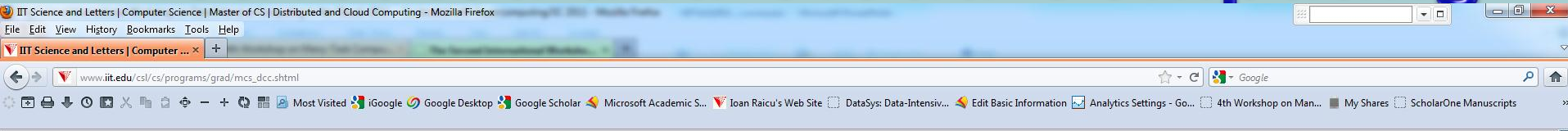
- [Home](#)
- [News](#)
- [Projects](#)
- [People](#)
- [Publications](#)
- [Theses & Dissertations](#)
- [Events](#)
- [CS Seminar](#)
- [DataSys Seminar](#)

The DataSys Lab is a research lab in the Department of Computer Science at Illinois Institute of Technology, being lead by Dr. Ioan Raicu. The DataSys Lab conducts research in various areas of distributed systems with an emphasis on designing, implementing, and evaluating systems, protocols, and middleware with the goal of supporting data-intensive applications at extreme scales. The lab's mission is to investigate challenging, high-impact research projects to support data-intensive distributed computing on a variety of systems, from many-core systems, clusters, grids, clouds, and supercomputers.

[People](#)



New MS Specialization in Distributed and Cloud Computing



- CSL Home
- Graduate Admission
- Undergrad Admission
- Financial Aid
- CSL Programs
- Student Life



Quick Links

Search Content

[Contact CSL](#)

- COMPUTER SCIENCE HOME**
- ABOUT PEOPLE**
- ACADEMIC PROGRAMS**
- Undergraduate
 - Degrees Offered
 - Course Descriptions
- Graduate**
- Degrees Offered
- Course Descriptions
- Accelerated Courses

Academic Programs / Graduate

Computer Science Graduate Programs

Master Of Computer Science With a Specialization in Distributed and Cloud Computing

30 credit hours

The Master of Computer Science With a Specialization in Distributed and Cloud Computing is intended for students who are interested to learn about distributed systems and how they are applied to real world problems, as well as how emerging cloud computing technologies can be used to implement some of the worlds most popular services and applications.

Tenure Track Faculty Positions

IIT Science and Letters | Computer Science | Tenure-track Opening for Fall 2013 - Mozilla Firefox

File Edit View History Bookmarks Tools Help

IIT Science and Letters | ...

www.iit.edu/csl/cs/announcements/2012/2012_asst_prof_posit Google

Google Google Desktop Google Scholar

IIT College of Science and Letters ILLINOIS INSTITUTE OF TECHNOLOGY

my iit

IIT.EDU HOME PROSPECTIVE STUDENTS ALUMNI BUSINESS & INDUSTRY VISITORS Apply Now >>

CSL Home Graduate Admission Undergrad Admission Financial Aid CSL Programs Student Life

COMPUTER SCIENCE HOME ABOUT PEOPLE ACADEMIC PROGRAMS RESEARCH RESOURCES ANNOUNCEMENTS News News Archive

Tenure Track Faculty Positions

- Security
- Information Retrieval

Announcements

Tenure-Track Assistant Professor Position Available for Fall 2013

Illinois Institute of Technology
Department of Computer Science

The Department of Computer Science at Illinois Institute of Technology seeks applications for one or two tenure-track positions at the rank of Assistant Professor, starting Fall 2013. Applicants must have a Ph.D. in computer science or a closely-related field, demonstrated success in research, significant potential for attracting external research funding, and a strong commitment to excellence in teaching. Candidates in information retrieval,

More Information

- DataCloud 2012 Website:
 - <http://www.cse.buffalo.edu/faculty/tkosar/datacloud2012/>
- Prize giveaway (win an Amazon Kindle Fire HD):
 - <https://docs.google.com/spreadsheet/viewform?formkey=dDVUUGVxeU1iQ25UWlJONWU1dVBtT3c6MA#gid=0>
- Contact:
 - Ioan Raicu: iraicu@cs.iit.edu
 - Tevfik Kosar: tkosar@buffalo.edu
 - Roger Barga: barga@microsoft.com