

Supporting Data-Intensive Computing at Extreme Scales

Ioan Raicu

Computer Science Department, Illinois Institute of Technology
Math and Computer Science Division, Argonne National Laboratory

July 12th, 2013

Google – Wisconsin Madison



DataSys: Data-Intensive Distributed Systems Laboratory

- **Research Focus**

- Emphasize designing, implementing, and evaluating systems, protocols, and middleware with the goal of supporting **data-intensive applications on extreme scale distributed systems**, from many-core systems, clusters, grids, clouds, and supercomputers

- **People**

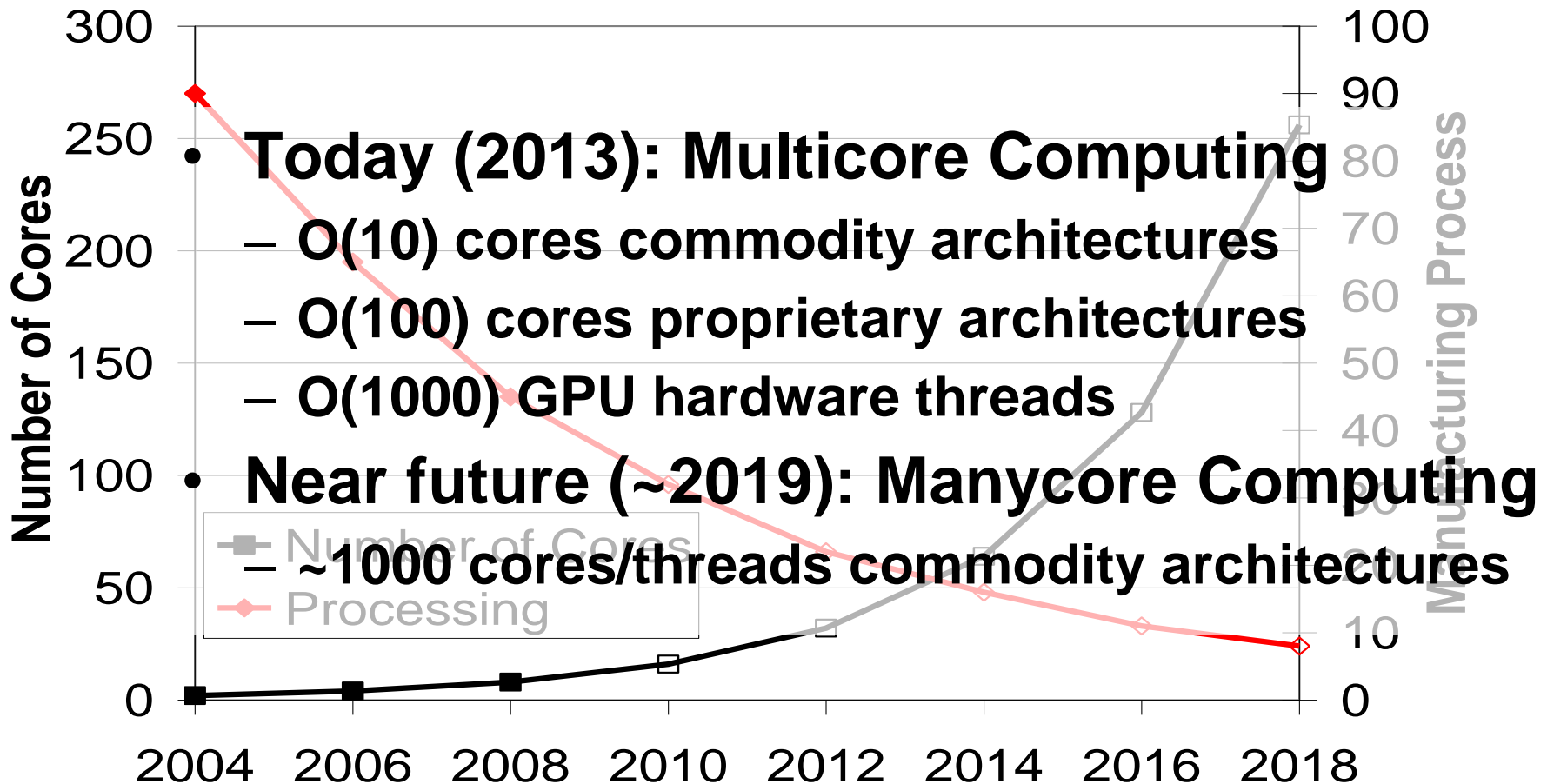
- Dr. Ioan Raicu (Director)
- 6 PhD Students
- 2 MS Students
- 4 UG Students

- **Contact**

- <http://datasys.cs.iit.edu/>
- iraicu@cs.iit.edu



Manycore Computing

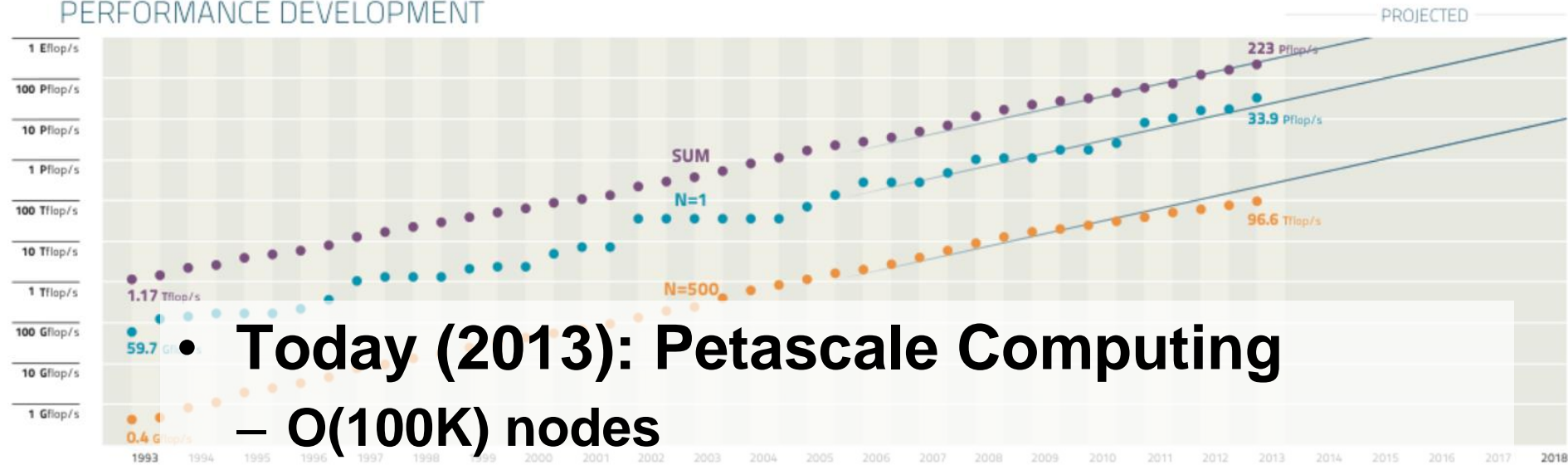


Pat Helland, Microsoft, The Irresistible Forces Meet the Movable

Objects, November 9th, 2007
Supporting Data-Intensive Computing at Extreme Scales

Exascale Computing

PERFORMANCE DEVELOPMENT



- Today (2013): Petascale Computing
 - O(100K) nodes
 - O(1M) cores
- <http://www.top500.org/> Near future (~2018): Exascale Computing
 - ~1M nodes (10X)
 - ~1B processor-cores/threads (1000X)

http://s.top500.org/static/lists/2013/06/TOP500_201306_Poster.png

Active Projects

- **Storage**

- [FusionFS: Fusion distributed File System](#)
 - [HyCache](#), [FusionProv](#), IStore, RXSim
- [ZHT: Zero-Hop Distributed Hash Table](#)
 - NoVoHT

- **Computing**

- **Many-Task Computing**

- [MATRIX: MAny-Task computing execution framework at eXascales](#)
 - [SimMatrix](#)
- [Falkon: Fast and Light-weight task execution framework](#)
 - FalkonCloud
- [Swift: Fast, Reliable, Loosely Coupled Parallel Computation](#)

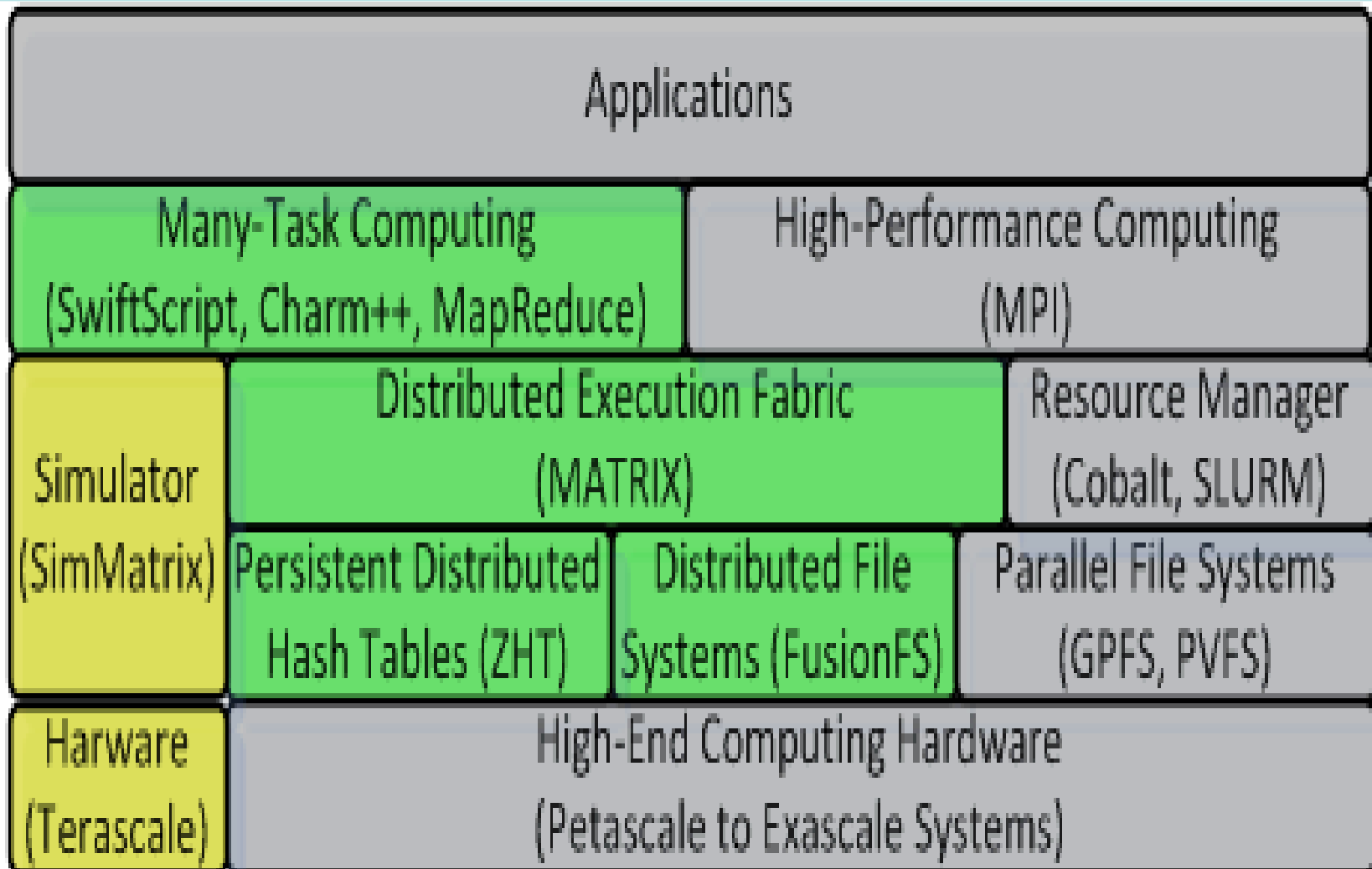
- **Many-Core Computing**

- [GeMTC: Virtualizing GPUs to Support MTC Applications](#)

- **Cloud Computing**

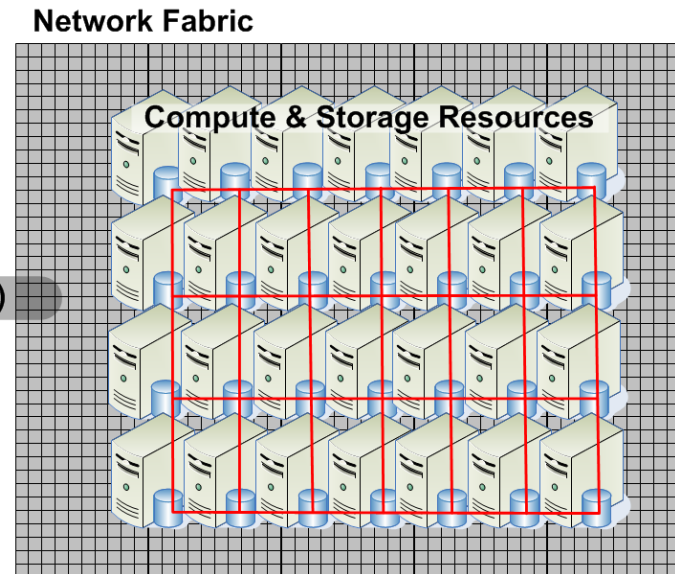
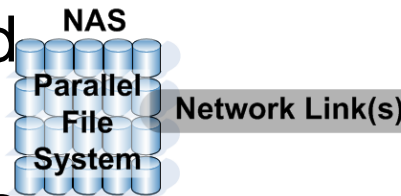
- CloudBench: Optimizing Cloud Infrastructure for Scientific Computing Applications

Proposed Software Stack in Large-Scale Distributed Systems

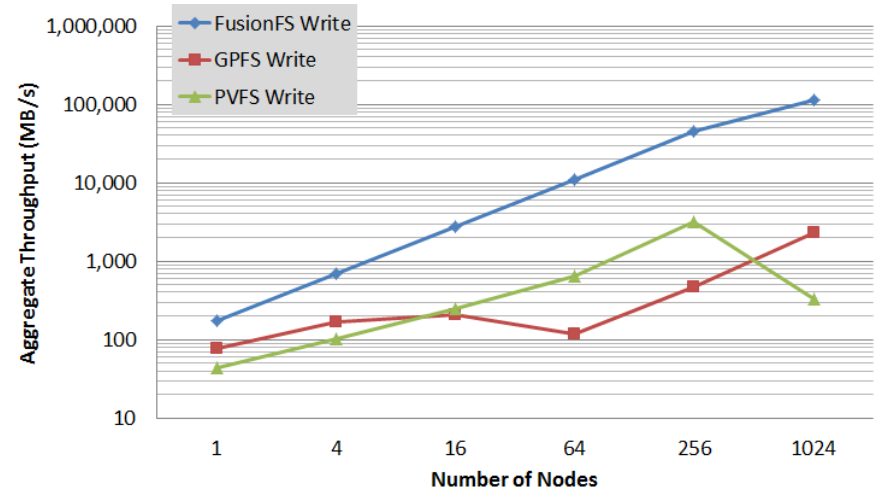
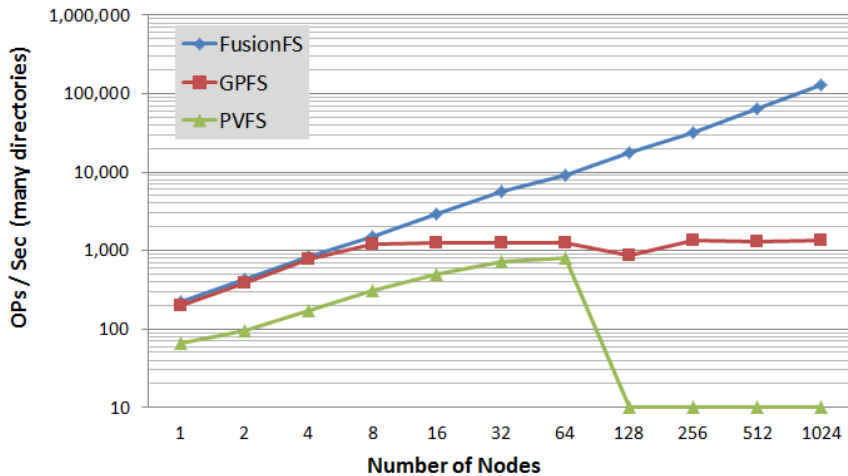


FusionFS Project

- A distributed file system co-locating storage and computations, while supporting POSIX
- Everything is decentralized and distributed
- Aims for millions of servers and clients scales
- Aims at orders of magnitude higher performance than current state of the art parallel file systems

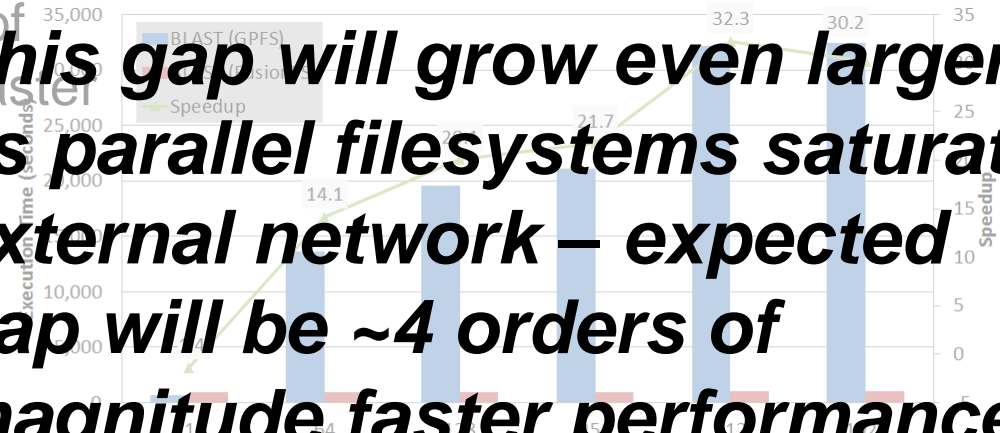


FusionFS Project



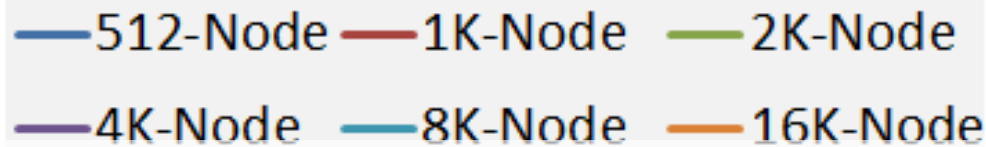
^ ~2 orders of magnitude faster metadata

This gap will grow even larger as parallel filesystems saturate external network – expected gap will be ~4 orders of magnitude faster performance



^ ~1.5 order of magnitude faster I/O
 < ~1.5 order of magnitude faster runtime for real application

FusionFS Project

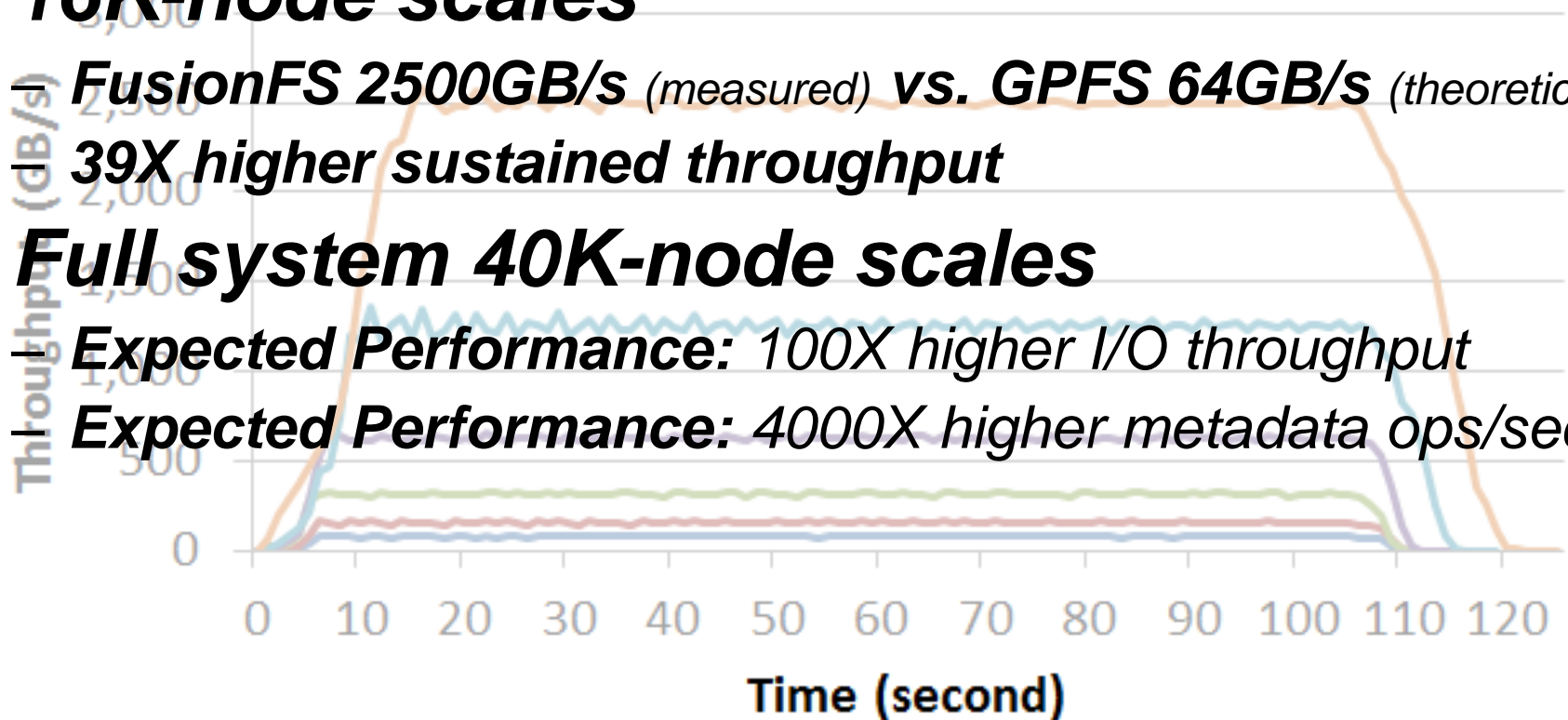


- **16K-node scales**

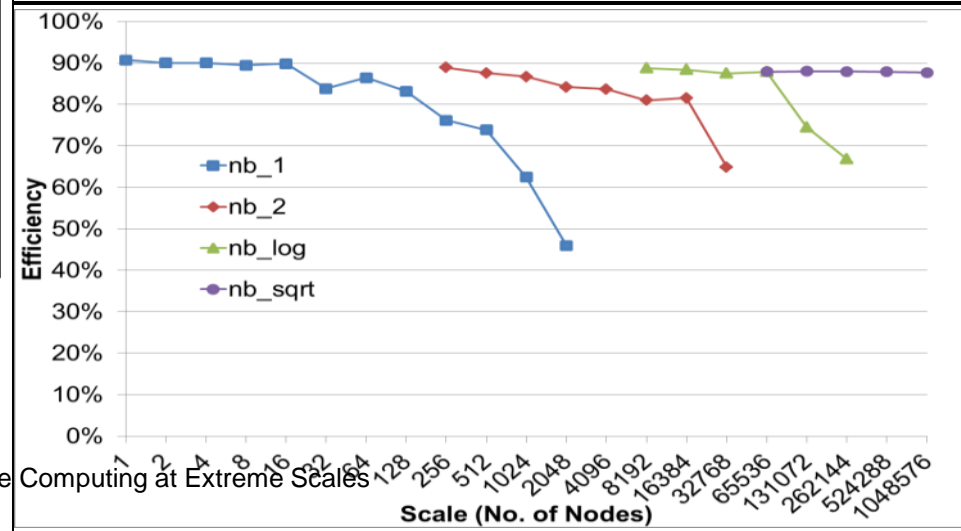
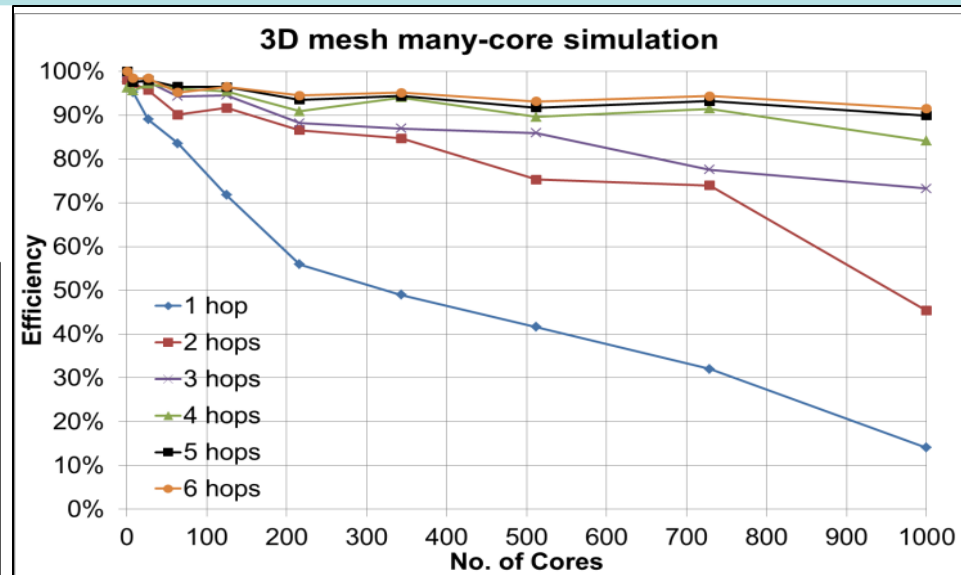
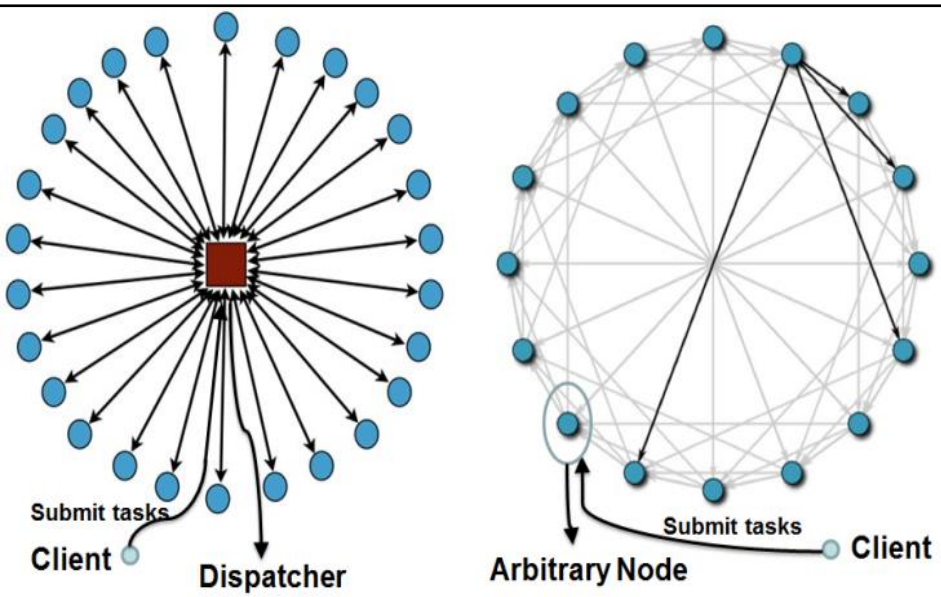
FusionFS 2500GB/s (measured) vs. GPFS 64GB/s (theoretical)
39X higher sustained throughput

- **Full system 40K-node scales**

Expected Performance: 100X higher I/O throughput
Expected Performance: 4000X higher metadata ops/sec



Many-Task Computing



Active Collaborations

National Labs and Industry

- **National Laboratories**

- **ANL:** Kamil Iskra, Rob Ross, Mike Wilde, Marc Snir, Pete Beckman, Justin Wozniak
- **FNAL:** Gabriele Garzoglio
- **LANL:** Mike Lang
- **ORNL:** Arthur Barney Maccabe
- **LBL:** Lavanya Ramakrishnan

- **Industry**

- **Cleversafe:** Chris Gladwin
- **EMC:** John Bent
- **Accenture Technology Laboratory:** Teresa Tung
- **Microsoft:** Roger Barga
- **SchedMD:** Morris Jette, Danny Auble
- **Oracle:** Hui Jin
- **INRIA:** Gabriel Antoniu
- **IBM:** Bogdan Nicolae

Active Collaborations

Academia

- **Academia**

- **IIT:** Xian-He Sun, Zhiling Lan, Shlomo Argamon
- **UChicago:** Ian Foster, Tanu Malik, Zhao Zhang, Kyle Chard
- **UEST China:** Yong Zhao
- **SUNY:** Tevfik Kosar
- **WSU:** Shiyong Lu
- **USC:** Yogesh Simmhan
- **Georgia Tech:** Jeffrey Vetter
- **Columbia:** Glen Hocky

Active Funding (\$)

- **NSF CAREER 2011 – 2015: \$486K**
 - “*Avoiding Achilles’ Heel in Exascale Computing with Distributed File Systems*”, NSF CAREER
- **DOE Fermi 2011 – 2013: \$84K**
 - “Networking and Distributed Systems in High-Energy Physics”, DOE FNAL
- **DOE LANL 2013: \$75K**
 - “Investigation of Distributed Systems for HPC System Services”, DOE LANL
- **IIT STARR 2013: \$15K**
 - “*Towards the Support for Many-Task Computing on Many-Core Computing Platforms*”, IIT STARR Fellowship
- **Amazon 2011 - 2013: \$18K**
 - “*Distributed Systems Research on the Amazon Cloud Infrastructure*”, Amazon
- **NVIDIA 2013 – 2014: \$12K**
 - “CUDA Teaching Center”, NVIDIA

Funding (Time)

- **DOE 2011 – 2013: 450K hours**
 - “*FusionFS: Distributed File Systems for Exascale Computing*”, DOE ANL ALCF; 450,000 hours on the IBM BlueGene/P
- **XSEDE 2013: 200K hours**
 - “*Many-Task Computing with Many-Core Accelerators on XSEDE*”, NSF XSEDE; 200K hours on XSEDE
- **GLCPC 2013: 6M hours**
 - “*Implicitly-parallel functional dataflow for productive hybrid programming on Blue Waters*”, Great Lakes Consortium for Petascale Computation (GLCPC); 6M hours on the Blue Waters Supercomputer
- **NICS 2013: 320K hours**
 - “*Many-Task Computing with Many-Core Accelerators on Beacon*”, National Institute for Computational Sciences (NICS); 320K hours on the Beacon system

Service Activities

- IEEE Transactions on Cloud Computing
 - Special Issue on Scientific Cloud Computing
- Springer's Journal of Cloud Computing: Advances, Systems and Applications
- IEEE/ACM MTAGS 2013 @ SC13
- IEEE/ACM DataCloud 2013 @ SC13
- ACM ScienceCloud 2014 @ HPDC14
- IEEE CCGrid 2014 in Chicago
- GCASR 2014 in Chicago
- Others:
 - IEEE/ACM SC 2013, ACM HPDC 2014, IEEE IPDPS 2014, IEEE ICDCS 2014, IEEE eScience 2014



IEEE TRANSACTIONS ON CLOUD COMPUTING

The *IEEE Transactions on Cloud Computing* will publish peer reviewed articles that provide innovative research ideas and applications results in all areas relating to cloud computing. Topics relating to novel theory, algorithms, performance analyses and applications of techniques relating to all areas of cloud computing will be considered for the transactions. The transactions will consider submissions specifically in the areas of cloud security, tradeoffs between privacy and utility of cloud, cloud standards, the architecture of cloud computing, cloud development tools, cloud software, cloud backup and recovery, cloud interoperability, cloud applications management, cloud data analytics, cloud communications protocols, mobile cloud, liability issues for data loss on clouds, data integration on clouds, big data on clouds, cloud education, cloud skill sets, cloud energy consumption, cloud applications in commerce, education and industry. This title will also consider submissions on Infrastructure as a Service (IaaS), Platform as a Service (PaaS), Software as a Service (SaaS), and Business Process as a Service (BPaaS).

ACM MTAGS 2013

6th Workshop on Many-Task Computing on Clouds, Grids, and Supercomputers

co-located with ACM/IEEE SC13 (International Conference for High Performance, Networking, Storage and Analysis), Denver, Colorado -- November 17th, 2013

Committee Members

Workshop Chairs

Ioan Raicu, Illinois Institute of Technology &
Argonne Nat. Lab

Ian Foster, Univ. of Chicago & Argonne Nat. Lab

Yong Zhao, University of Electronic Science and
Technology of China

Justin Wozniak, Argonne National Laboratory

Steering Committee

David Abramson, Monash University, Australia

Jack Dongarra, University of Tennessee, USA

Geoffrey Fox, Indiana University, USA

Manish Parashar, Rutgers University, USA

Marc Snir, Argonne National Laboratory & University of
Illinois at Urbana Champaign, USA

Xian-He Sun, Illinois Institute of Technology, USA

Weimin Zheng, Tsinghua University, China

Program Committee

Samer Al-Kiswani (University of British Columbia)

Mihai Budiu (Microsoft Research)

Kyle Chard (University of Chicago)

Yong Chen (Texas Tech University)

Evangelinou Constantinos (Massachusetts Institute of
Technology)

Catalin Dumitrescu (Fermi National Labs)

Alexandru Iosup (Delft University of Technology -
Netherlands)

The 6th workshop on Many-Task Computing on Grids and Supercomputers (MTAGS) will provide the scientific community a dedicated forum for presenting new research, development, and deployment efforts of large-scale many-task computing (MTC) applications on large scale clusters, Grids, Supercomputers, and Cloud Computing infrastructure. MTC, the theme of the workshop encompasses loosely coupled applications, which are generally composed of many tasks (both independent and dependent tasks) to achieve some larger application goal. This workshop will cover challenges that can hamper efficiency and utilization in running applications on large-scale systems, such as local resource manager scalability and granularity, efficient utilization of raw hardware, parallel file system contention and scalability, data management, I/O management, reliability at scale, and application scalability. We welcome paper submissions on all theoretical, simulations, and systems topics related to MTC, but we give special consideration to papers addressing petascale to exascale challenges. Papers will be peer-reviewed, and accepted papers will be published in the workshop proceedings as part of the ACM digital library (pending approval). The workshop will be co-located with the IEEE/ACM Supercomputing 2013 Conference in Denver Colorado on November 18th, 2013.

We invite the submission of original work that is related to the topics

ACM DataCloud 2013 @ SC13

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

The Fourth International Workshop on Data Intensive Computing in the Clouds (DataCloud) 2013

**Co-located with [Supercomputing/SC 2013](#)
Denver Colorado -- November 17th, 2013**

News

- [The Third International Workshop on Data Intensive Computing in the Clouds \(DataCloud 2012\)](#) attracted over 100 attendees. See workshop website for full program details.

Overview

Applications and experiments in all areas of science are becoming increasingly complex and more demanding in terms of computational and data requirements. Some applications generate data volumes reaching hundreds of terabytes and even petabytes. Analyzing, visualizing, and disseminating these large data sets has become a major challenge and data intensive computing is now considered as the "fourth paradigm" in scientific discovery after theoretical, experimental, and



IEEE/ACM CCGrid 2014

14th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing

May 26-29, 2014 -- Chicago, IL, USA

<http://datasys.cs.iit.edu/events/CCGrid2014/>

Rapid advances in architectures, networks, and systems and middleware technologies are leading to new concepts and platforms for computing, ranging from Clusters and Grids to Clouds and Datacenters. The 14th Annual IEEE/ACM International Symposium in Cluster, Cloud, and Grid Computing (CCGrid 2014) is a forum bringing together international researchers, developers, and practitioners 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 2014, CCGrid will return to the USA and be held in Chicago, the third largest city in the United States. The main conference will be held on May 27-29, 2014, with tutorials and affiliated workshops taking place on May 26, 2014.

IMPORTANT DATES

Papers Due:	11 November 2013 Anywhere on Earth
Author Notifications:	24 January 2014
Camera Ready Papers Due:	17 February 2014

TOPICS OF INTEREST

CCGrid 2014 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 and Accelerators:** Design and use of emergent system architectures including but not limited to accelerators; Multicores; 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:** Development, evaluation, and optimization of cloud architectures; Software tools and techniques for clouds
- **Modeling and Evaluation of Performance and Energy:** Prediction, modeling and analysis; Performance tools; Benchmarks and testbeds
- **Programming Models, Systems, and Fault-Tolerant Computing:** Programming models, systems software and environments for cluster, cloud, and grid computing; Fault-tolerant systems, programs and algorithms
- **Scheduling and Resource Management:** Techniques to schedule jobs and resources on cluster, cloud, and grid computing platforms; SLA

IEEE Transactions on Cloud Computing

Special Issue on Scientific Cloud Computing

<http://datasys.cs.iit.edu/events/ScienceCloud2014-TCC/>

Guest Editors

Kate Keahey Argonne National Laboratory
Ioan Raicu Illinois Institute of Technology
Kyle Chard University of Chicago
Bogdan Nicolae IBM Research



Topics:

- Scientific application cases studies on Clouds
- Performance evaluation of Cloud technologies
- Fault tolerance and reliability in cloud systems
- Data-intensive workloads and tools on Clouds
- Programming models such as Map-Reduce
- Storage cloud architectures
- I/O and Data management in the Cloud
- Workflow and resource management in the Cloud
- NoSQL databases for scientific applications
- Data streaming and dynamic applications on Clouds
- Dynamic resource provisioning
- Many-Task Computing in the Cloud

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. It is the key to solving “grand challenges” in many domains and providing breakthroughs in new knowledge, and it comes in many shapes and forms: high-performance computing (HPC) which is heavily focused on compute-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. Today’s “Big Data” trend 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. Not surprisingly, it becomes increasingly difficult to design and operate large scale systems capable of addressing these grand challenges.

This journal Special Issue on Scientific Cloud Computing in the IEEE Transactions on Cloud Computing will provide the scientific community a dedicated forum for discussing new research, development, and deployment efforts in running these kinds of scientific computing workloads on Cloud Computing infrastructures. This special issue will focus on the use of cloud-based technologies to meet new compute-intensive and data-intensive scientific challenges that are not well served by the current supercomputers, grids and HPC clusters. The special issue will aim to address questions such as:

ACM ScienceCloud 2013

4th Workshop on Scientific Cloud Computing (ScienceCloud) 2013

<http://datasys.cs.iit.edu/events/ScienceCloud2013/>

Keynote



Dr. Ian T. Foster

- Senior Scientist & Distinguished Fellow
- Math and Computer Science Division (MCS), Argonne National Laboratory
- Director
- Computation Institute (CI), University of Chicago
- Arthur Holly Compton Distinguished Service Professor
- Department of Computer Science (CS), University of Chicago
- Professor
- Physical Sciences, University of Chicago

The notion of science as a service was originally positioned in 2005 as a means of publishing and accessing scientific data and applications through internet accessible services. At that time, researchers were only just grasping the benefits of employing the same service oriented architectures commonly used in other domains. Since this time we have indeed seen a huge uptake in researchers leveraging services to disseminate and share data and applications in fields as diverse as genomics, climate science, and physical sciences. In addition, commercial software as a service (SaaS) products like Google Docs and Gmail are now used by many researchers in everyday activities. The major benefit of a SaaS approach is that researchers are able to invoke applications or access data remotely over the internet without needing to know the inner workings of the service. Our vision of science as a service worked well in a world when computing resources were scarce; when we needed to federate heterogeneous resources and make them ac-

Time	Description	Presenter	Institution
9:00AM	Opening Remarks		
9:05AM	Keynote -- Science as a Service: How On-Demand Computing Can Accelerate Discovery	Ian T. Foster	ANL UChicago
10:00AM	Break		
Session 1: Applications and Services			
10:30AM	High Performance Risk Aggregation: Addressing the Data Processing Challenge the Hadoop MapReduce Way	A. Rau-Chaplin B. Varghese Z. Yao	Dalhousie Univ.
11:00AM	Performance Evaluation of a MongoDB and Hadoop Platform for Scientific Data Analysis	E. Dede M. Govindaraju D. Gunter R. Canon L. Ramakrishnan	SUNY LBL
11:30AM	VIDAS: Object-based Virtualized Data Sharing for High Performance Storage I/O	Pablo Llopis Javier Garcia Blas Florin Isaila Jesus Carretero	Univ. Carlos III de Madrid
12:00PM	Lunch		
Session 2: Cloud Services			
1:30PM	Invited talk -- To Cloud My Big Data or Not To? Musings at the Intersection of Big Data, Intense Computing and Clouds	Radu Sion	Stony Brook
2:30PM	StorkCloud: Data Transfer Scheduling and Optimization as a Service (transfer service)	Tevfik Kosar Engin Arslan Brandon Ross Bing Zhang	SUNY
3:00PM	Break		
Session 3: Technology and Infrastructure			
3:30PM	Rebalancing in a Multi-Cloud Environment Supporting Data-Intensive Computing at Extreme Scales	Dmitry Duplyakin Ali Alzabarah Paul Marshall Kate Keahey Henry Tufo	Univ. of Colorado, Boulder ANL UChicago

Greater Chicago Area Systems Research Workshop



[GCASR13](#) | [News](#) | [Topics](#) | [Dates](#) | [Location](#) | [Submission](#) |
[Organization](#) | [Register](#) | [Program](#) | [Sponsors](#)

2nd Greater Chicago Area System Research Workshop (GCASR) 2013

Northwestern University, Evanston IL -- May 3rd, 2013

Workshop Program

We are pleased to announce that the final program is composed of [1 keynote presentation](#), [14 oral presentations](#), and 42 poster presentations (separated into [session 1](#) and [session 2](#)). We have 57 workshop presenters, representing 11 different institutions (ANL, DePaul, Google, IIT, Northwestern, Notre Dame, Purdue, UChicago, UIC, UIUC, and Loyola). The workshop will start at 8AM with breakfast, and will end at 6:30PM. For a detailed program which includes all presentation titles and abstracts, click [here](#).

In order for us to be prepared for the right number of people at this event, we ask everyone to please register ahead of time for the event. The registration deadline is Thursday April 25th, 2013. The event is free of charge, due to the [generous financial sponsorship of Northwestern University](#). To register, click [here](#).

Keynote Presentation

- [Pete Beckman](#), Argonne National Laboratory

Interested

- Collaborations with industry
- Access to industry logs, data sets and applications
 - Data access patterns, failure logs, applications, resources
- Compare and contrast technology stack
 - MapReduce vs. Swift/Falkon
 - GFS/Colosus vs. FusionFS
 - Redis/Scalaris vs. ZHT
 - LevelDB vs. NoVoHT
- Help with journals, conferences, and workshops
- Student internships (3 months, UG/MS/PhD)
- Deep immersion in Google projects (1 year for myself during sabbatical)

More Information

- More information:
 - <http://www.cs.iit.edu/~iraicu/>
 - <http://datasys.cs.iit.edu/>
- Contact:
 - iraicu@cs.iit.edu
- Questions?