Panel: Many-Task Computing meets Big Data

Chairs
Ioan Raicu, Illinois Institute of Technology & Argonne National Laboratory
Justin Wozniak, Argonne National Laboratory
Ian Foster, University of Chicago & Argonne National Laboratory
Yong Zhao, University of Electronic Science and Technology of China, China

ACM MTAGS 2013
November 17th, 2013

Panelists

- **Dr. Robert Grossman**
  - Professor and Director, Division of Biological Sci… …nputation Institute, University of Chicago
- **Dr. Xian-He Sun**
  - Chair and Professor, Computer Science, Illinois I… …e
- **Dr. Judy Qiu**
  - Assistant Professor, Computer Science and Infor… …r University
- **Dr. Alexandru Iosup**
  - Assistant Professor, Faculty of Engineering, Mathematics and Science, Delft University of Technology, the Netherlands

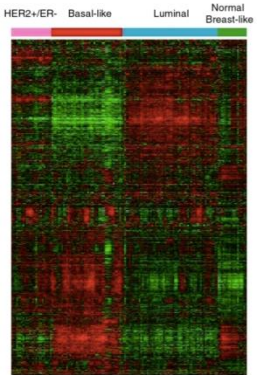
Robert Grossman

- We want to compute genomic variants.
- How can this be done as a distributed computation over science clouds?
- What are the APIs?
- What are the key common services?
- What is the governance structure?
- What is the sustainability model?

1,000,000 patients
1,000 PB

100,000 patients
100 PB

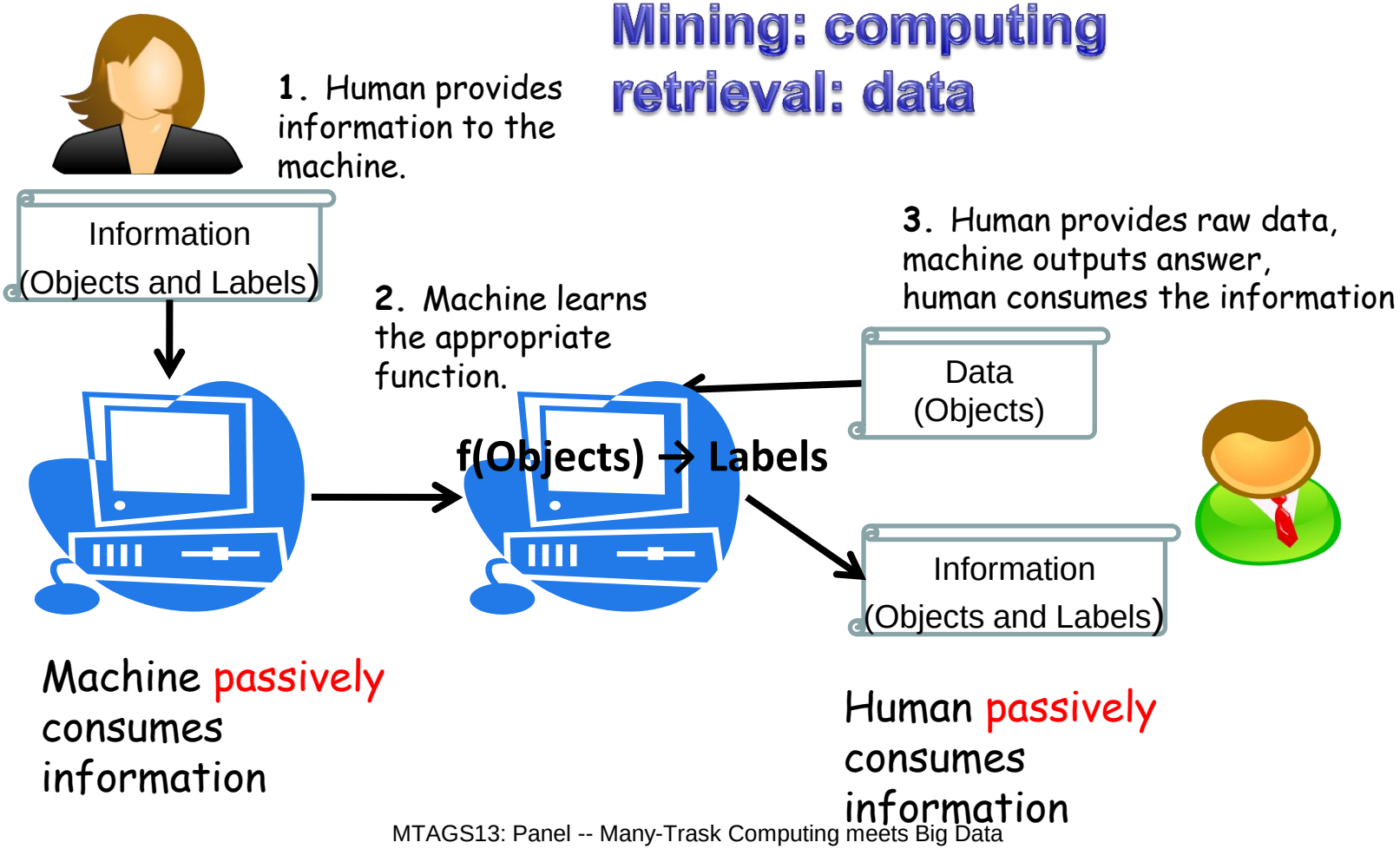
The Cancer Genome Atlas



10,000 patients
10 PB

1000 patients
1PB

Big Data require both HPC and HTC, that is MTC, and is mixed compute-intensive and data-intensive components



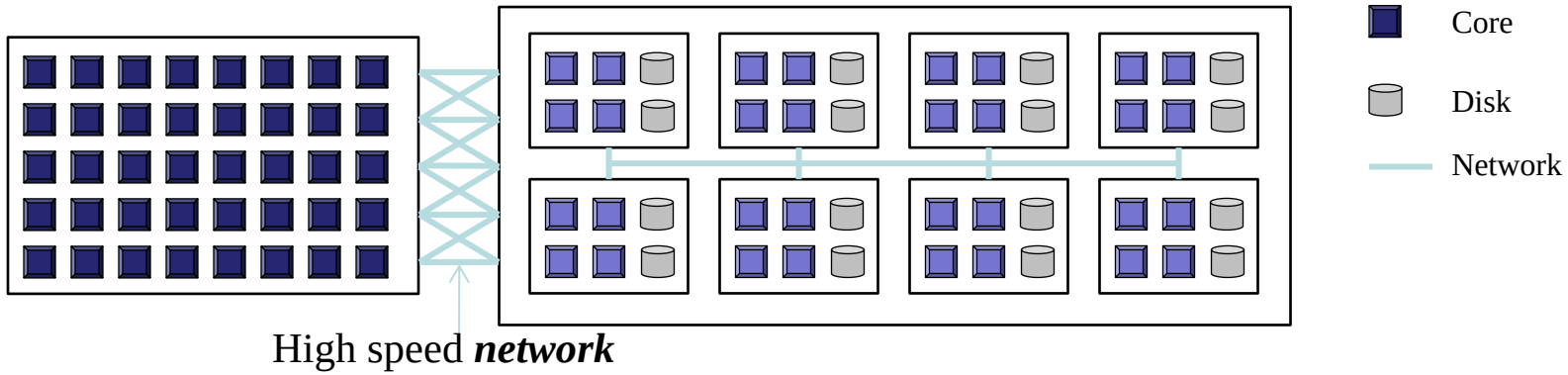
Support Both HPC and HTC: Ours Solution

Decoupled-Execution Paradigm:

- Handle computation- and data- intensive phases separatel
- One interface-Two systems, transparent to users
- Integration, scheduling, optimization

Supercomputer or
many-core computing system
for execution of computing
intensive part of an application

Data cloud or ***storage cluster***
for execution of data
intensive part of an application

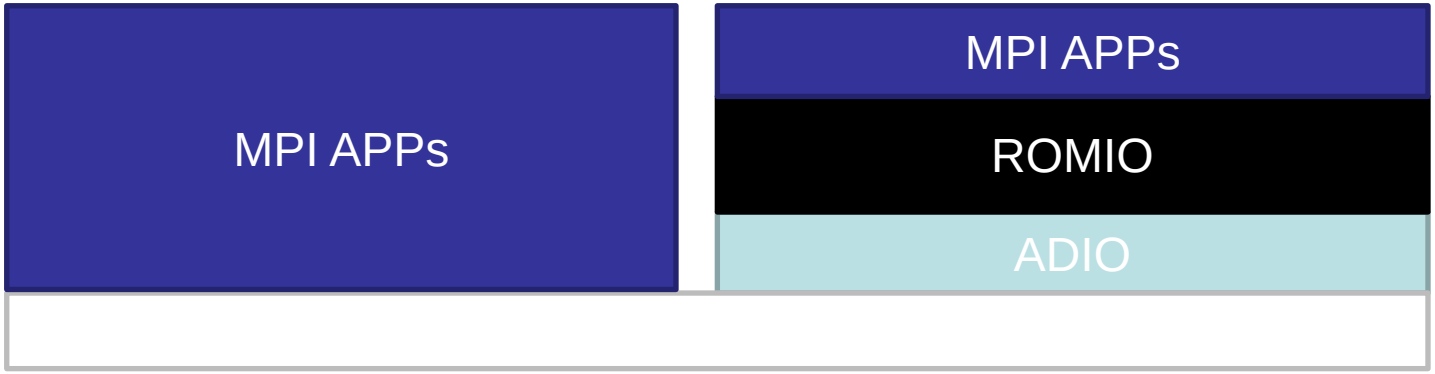


Y. Chen, C. Chen, X.-H. Sun, W. D. Gropp, and R. Thakur, "A Decoupled Execution Paradigm for Data-Intensive High-End Computing," IEEE Cluster'12, Sept, 2012

Important of data locality (consistence, non-consistence)

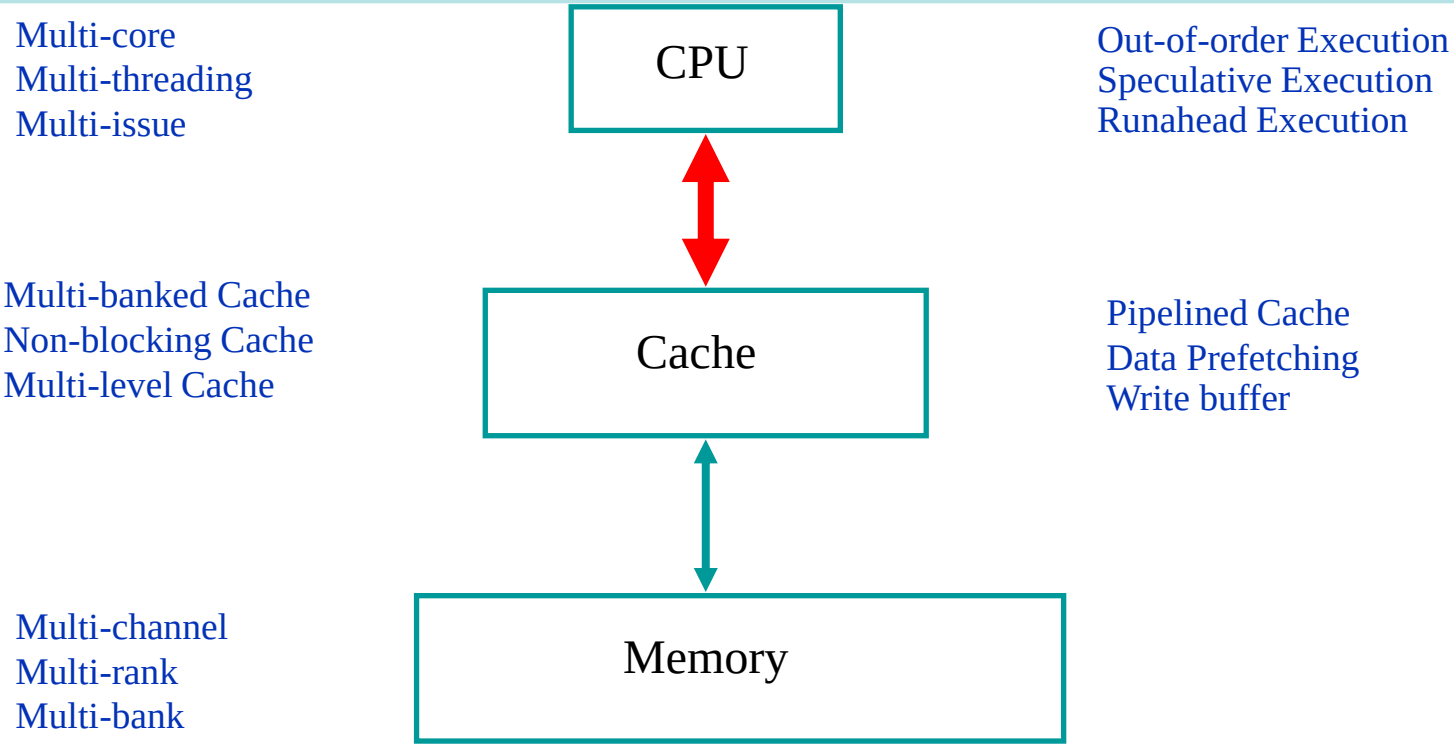
Interoperability between different file systems

- Enable MPI Apps to access data-intensive file systems
- HPC-Cloud, Data-Cloud



H. Jin, X.-H. Sun, et. al, "CHAIO: Enabling HPC Applications on Data-Intensive File Systems", **ICPP2012**.

Scalability of MTC: Memory-Parallelisms



Concurrent AMAT for Memory-Parallelism

- The traditional AMAT :

  HitCycle + MR×AMP.

- MR is the miss rate of cache accesses; and AMP is the average miss penalty
- The **Concurrent AMAT** :

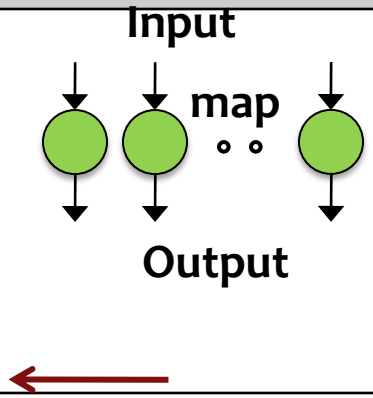
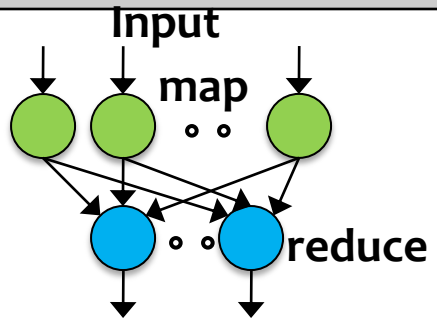
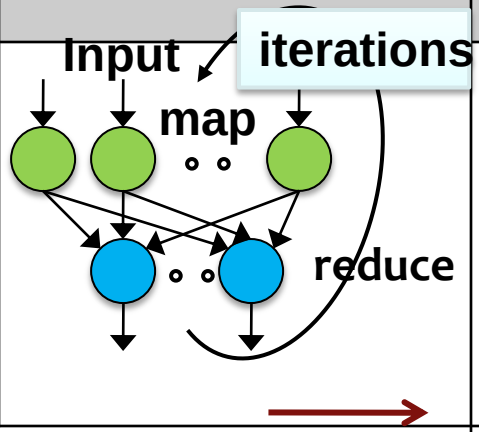
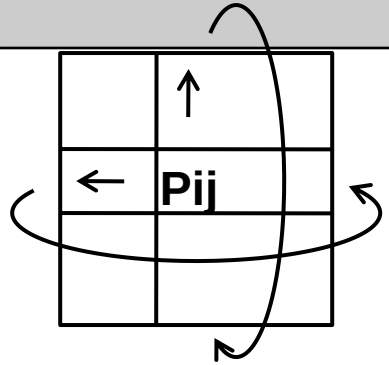
  $HitCycle/C_H + MR \times AMP/C_M$

- C_H is the hit concurrency; C_M is the pure miss concurrency
- Hit is always good, miss may not be necessary bad
- Design Choice of memory systems

X.-H. Sun and D. Wang, "Concurrent Average Memory Access Time", accepted to appear in ***IEEE Computers***, 2013.(IIT Technical Report, IIT/CS-SCS-2012-05)

Applications & Different Interconnection Patterns

Judy Qiu
Indiana University

(a) Map Only (Pleasingly Parallel)	(b) Classic MapReduce	(c) Iterative MapReduce	(d) Loosely Synchronous
			
Domain of MapReduce and Iterative Extensions			MPI
No Communication	Collective Communication		

Collective Patterns

MapReduce
- Wordcount, Grep

MapReduce-MergeBroadcast
- KMeansClustering, PageRank

Map-AllGather
- MDS-BCCalc
- Matrix Mult

Map-AllReduce
- KMeansClustering, MDS-StressCalc

Map-ReduceScatter
- PageRank, Belief Propagation

Alex Iosup
Thanks from the PDS Group at TU Delft.

VENI

Alexandru Iosup

Grids/Clouds
P2P systems
Big Data
Online gaming
Gamification

Dick Epema

Grids/Clouds
P2P systems
Video-on-demand
e-Science

VENI

Ana Lucia Varbanescu

HPC systems
Multi-cores
Big Data
e-Science

Henk Sips

HPC systems
Multi-cores
P2P systems

VENI

Johan Pouwelse

P2P systems
File-sharing
Video-on-demand

Home page

- www.pds.ewi.tudelft.nl

Publications

- see PDS publication database at publications.st.ewi.tudelft.nl

Alex Iosup

What's in a Name?

Applications from two worlds

- E-Science (incl. Big Data)
- Massively Multiplayer/Social Online Gaming (incl. Big Data)

10-years research in distributed systems

- System design, development, and evaluation
- Grid->Cloud computing, P2P->? Computing
- Performance measurements, evaluation, modeling, b'marking
- Grenchmark, Koala, Tribler, The Archives, [OpenTTD@large]

10 *operational* years research in comp. sci.

A. Iosup and D. Epema, On the Gamification of a Graduate Course on Cloud Computing, SC|13 Education Poster.

A. Iosup and D. Epema, An Experience Report on Using Gamification in Technical Higher Education, SIGCSE 2014. http://goo.gl/V97zSW

http://www.pds.ewi.tudelft.nl/~iosup/

Alex Iosup
Current work

1. In the future, will Small-and-Medium Enterprises use *elastic* infrastructure running multiple frameworks?
   - Many-Task Big-Data Processing on Clouds—GPUs
2. In the future, should we risk working on scheduling policies?
   - Portfolio Scheduling
3. In the future, what is the role of *job* throughput, next to task throughput and peak performance (HPC)?
4. In the future, will *social awareness* be at the core of our shared distributed systems?
5. In the future, will it be possible to rate and rank distributed computing systems (benchmarking, also commercial issue)?

Panel Overview

1. How do you see MTC intersecting with MapReduce, HTC, and HPC?
2. Importance of data locality for Big Data ==> how important is data-aware scheduling for Many-Task Computing
3. Supercomputers are designed for HPC applications today; in the future, should they be designed to support both MTC and/or Big Data?
4. With the growing scale of systems, has a centralized MTC system become obsolete? Is distributed MTC management (both scheduling and storage) a necessary next step?

Panel Overview

1. How do you see MTC intersecting with MapReduce, HTC, and HPC?
2. Importance of data locality for Big Data ==> how important is data-aware scheduling for Many-Task Computing
3. Supercomputers are designed for HPC applications today; in the future, should they be designed to support both MTC and/or Big Data?
4. With the growing scale of systems, has a centralized MTC system become obsolete? Is distributed MTC management (both scheduling and storage) a necessary next step?

Panel Overview

1. How do you see MTC intersecting with MapReduce, HTC, and HPC?
2. Importance of data locality for Big Data ==> how important is data-aware scheduling for Many-Task Computing
3. Supercomputers are designed for HPC applications today; in the future, should they be designed to support both MTC and/or Big Data?
4. With the growing scale of systems, has a centralized MTC system become obsolete? Is distributed MTC management (both scheduling and storage) a necessary next step?

Panel Overview

1. How do you see MTC intersecting with MapReduce, HTC, and HPC?
2. Importance of data locality for Big Data ==> how important is data-aware scheduling for Many-Task Computing
3. Supercomputers are designed for HPC applications today; in the future, should they be designed to support both MTC and/or Big Data?
4. With the growing scale of systems, has a centralized MTC system become obsolete? Is distributed MTC management (both scheduling and storage) a necessary next step?

Panel Overview

1. How do you see MTC intersecting with MapReduce, HTC, and HPC?
2. Importance of data locality for Big Data ==> how important is data-aware scheduling for Many-Task Computing
3. Supercomputers are designed for HPC applications today; in the future, should they be designed to support both MTC and/or Big Data?
4. With the growing scale of systems, has a centralized MTC system become obsolete? Is distributed MTC management (both scheduling and storage) a necessary next step?

More Information

- MTAGS 2013 Website:
  - http://datasys.cs.iit.edu/events/MTAGS13/
- Panel info:
  - http://datasys.cs.iit.edu/events/MTAGS13/panel.html
- Workshop program (7 exciting talks in the PM)
  - http://datasys.cs.iit.edu/events/MTAGS13/program.html
- Prize giveaway (win a Google Nexus 7):
  - http://datasys.cs.iit.edu/events/MTAGS13/prize.html
- Contact
  - iraicu@cs.iit.edu