Data-Intensive Computing at the Intersection of Cloud Computing and Supercomputing

Ioan Raicu
Computer Science Department
Illinois Institute of Technology

Who We Are, What We Are Seminar at Illinois Institute of Technology
October 4th, 2012

- **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**
  - 1 Faculty Member
  - 5 PhD Students
  - 4 MS Students
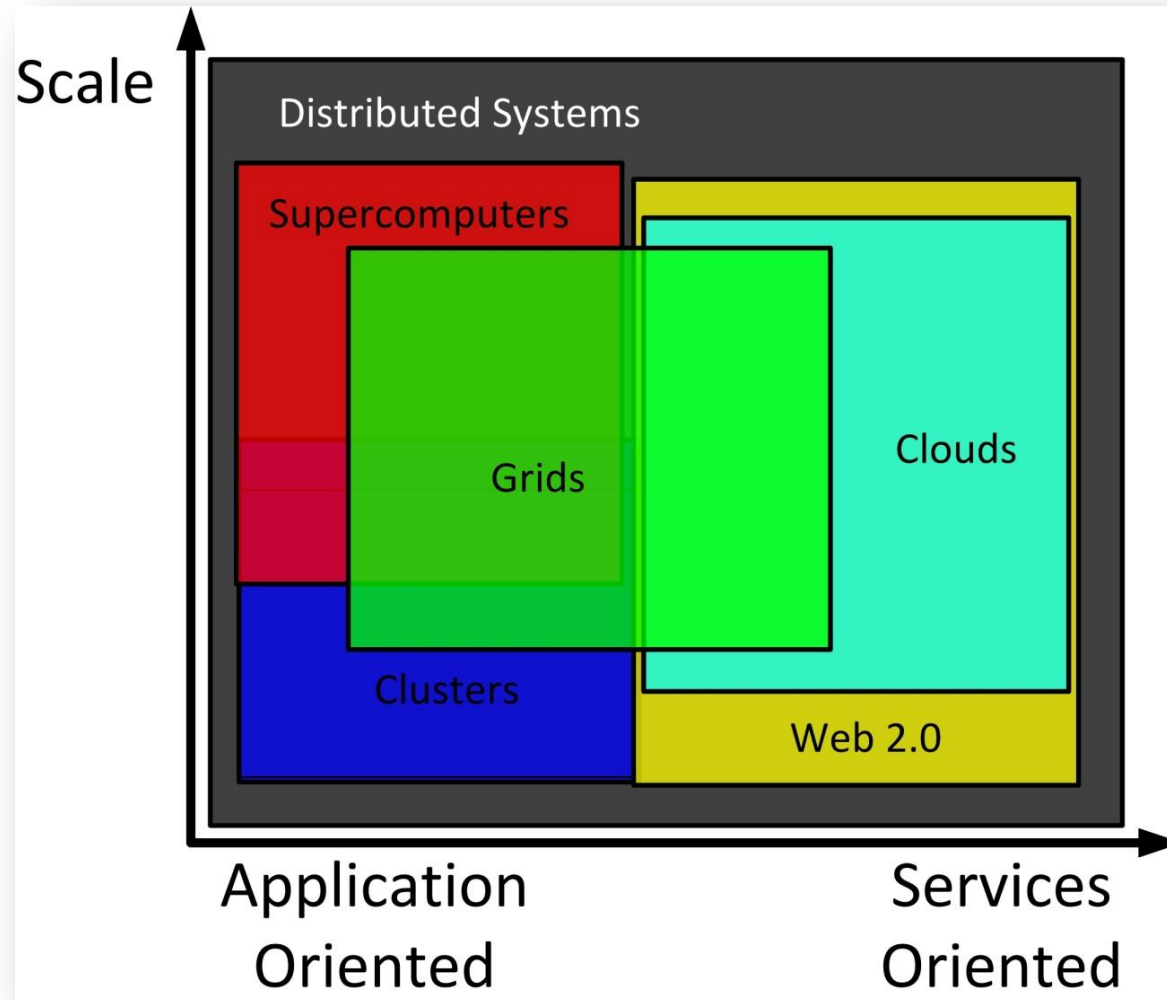 - 2 UG Students

- **More information**
  - http://datasys.cs.iit.edu/
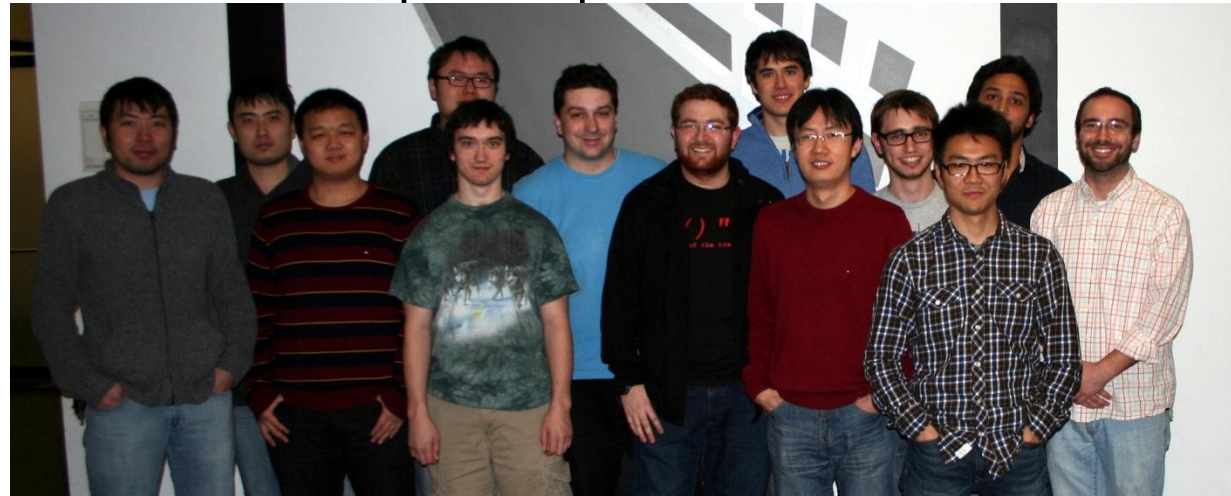
Distributed Systems

- What is a distributed system?

  "A collection of independent computers that appears to its users as a single coherent system"
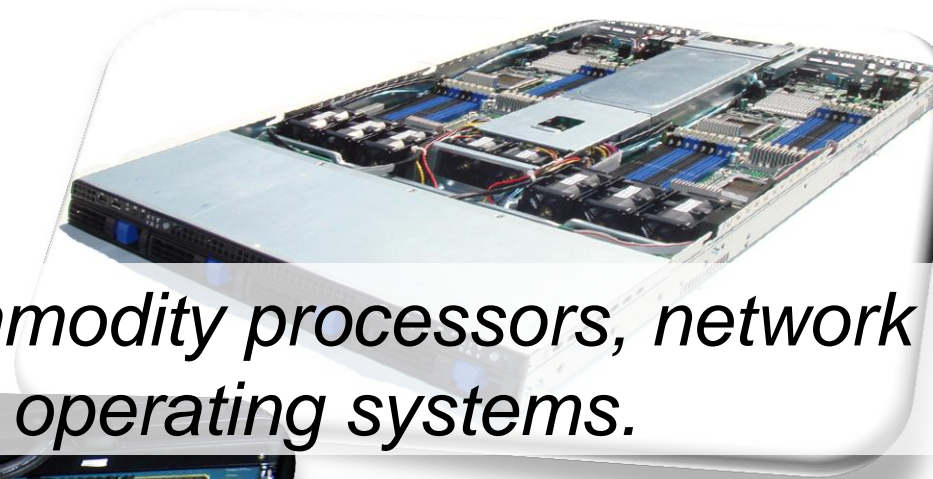
  -A. Tanenbaum

Distributed Systems:
Clusters, Grids, Clouds, and Supercomputers

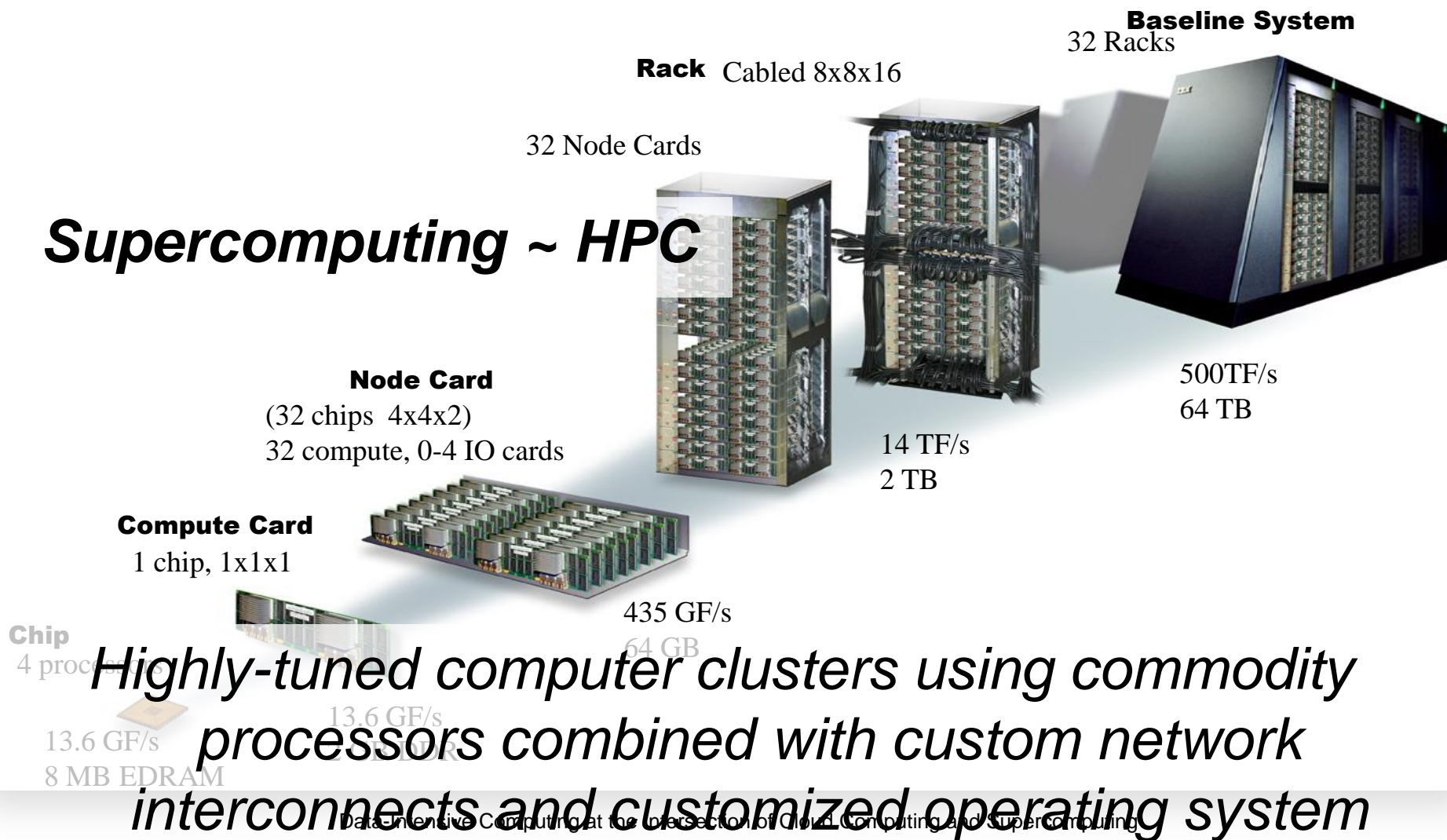Data-Intensive Computing at the Intersection of Cloud Computing and Supercomputing

[GCE08] "Cloud Computing and Grid Computing 360-Degree Compared"

Cluster Computing

Computer clusters using commodity processors, network interconnects, and operating systems.

Supercomputing

Baseline System
32 Racks

Rack Cabled 8x8x16

32 Node Cards

Supercomputing ~ HPC

500TF/s
64 TB

Node Card
(32 chips 4x4x2)
32 compute, 0-4 IO cards

14 TF/s
2 TB

Compute Card
1 chip, 1x1x1

435 GF/s
64 GB

Chip
4 processors

13.6 GF/s
8 MB EDRAM

13.6 GF/s
2.0 GB DDR

Highly-tuned computer clusters using commodity processors combined with custom network interconnects and customized operating system
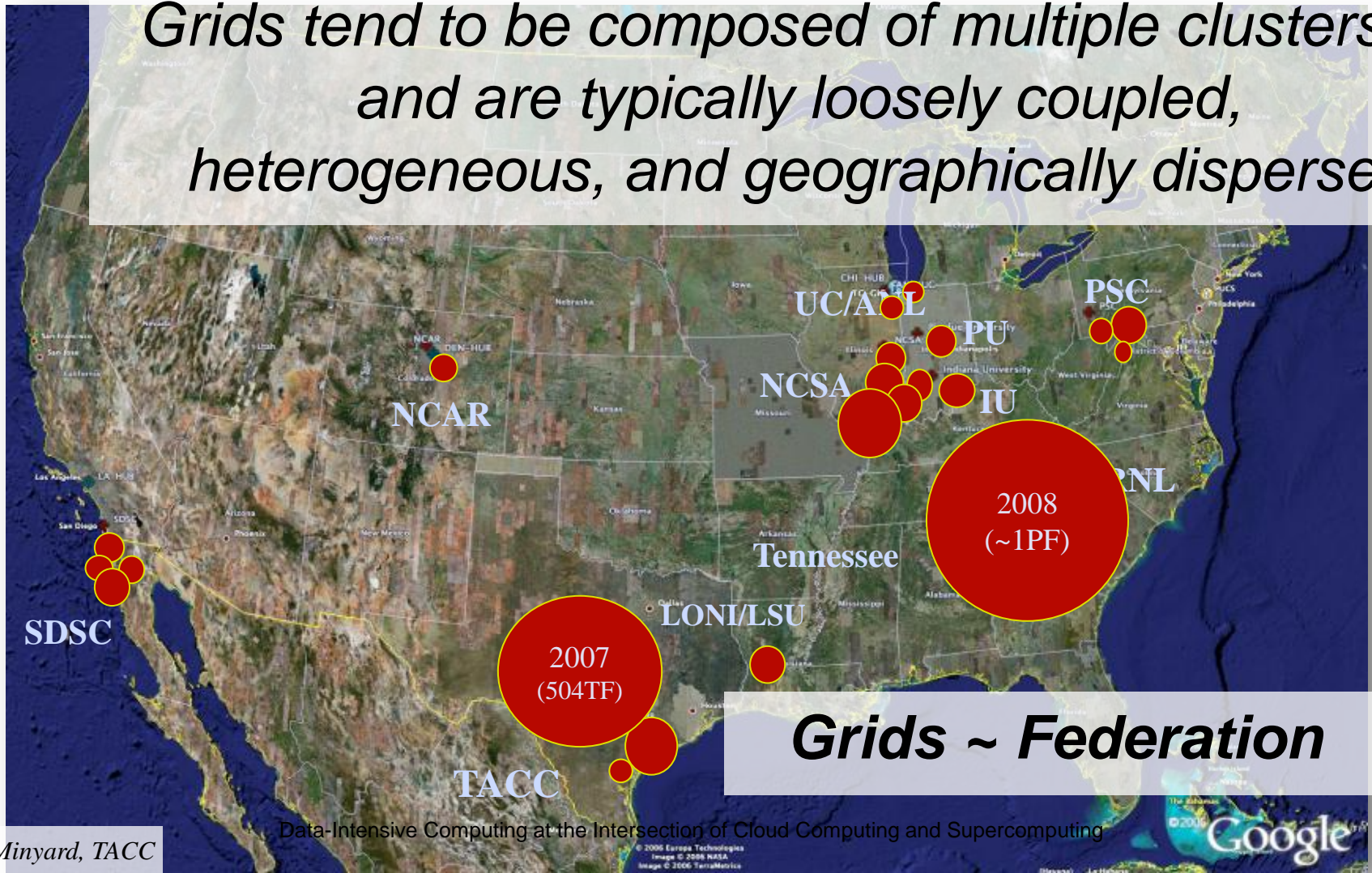
Data-Intensive Computing at the Intersection of Cloud Computing and Supercomputing

Grid Computing

Grids tend to be composed of multiple clusters, and are typically loosely coupled, heterogeneous, and geographically dispersed

UC/ANL

PSC

PU

NCSA

IU

NCAR

PNL

2008
(~1PF)

Tennessee

LONI/LSU

SDSC

2007
(504TF)

Grids ~ Federation

TACC

Data-Intensive Computing at the Intersection of Cloud Computing and Supercomputing
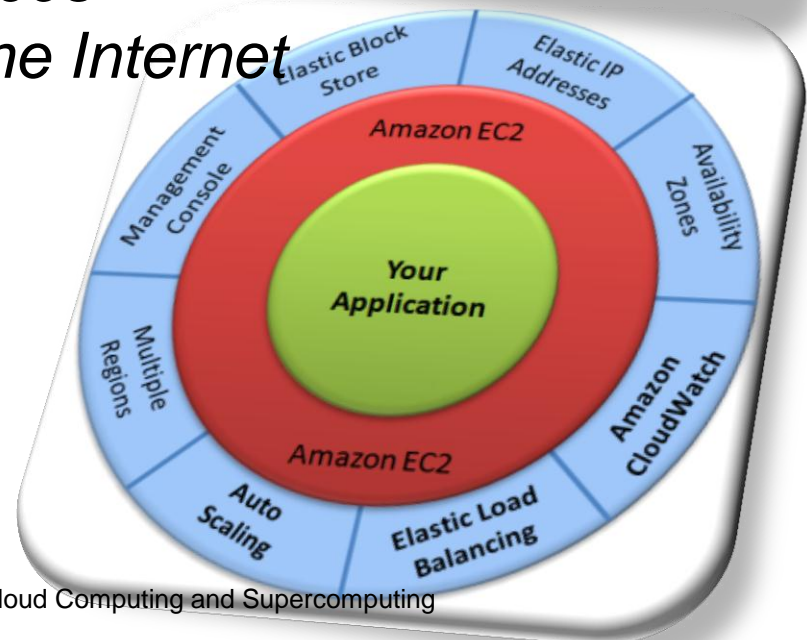
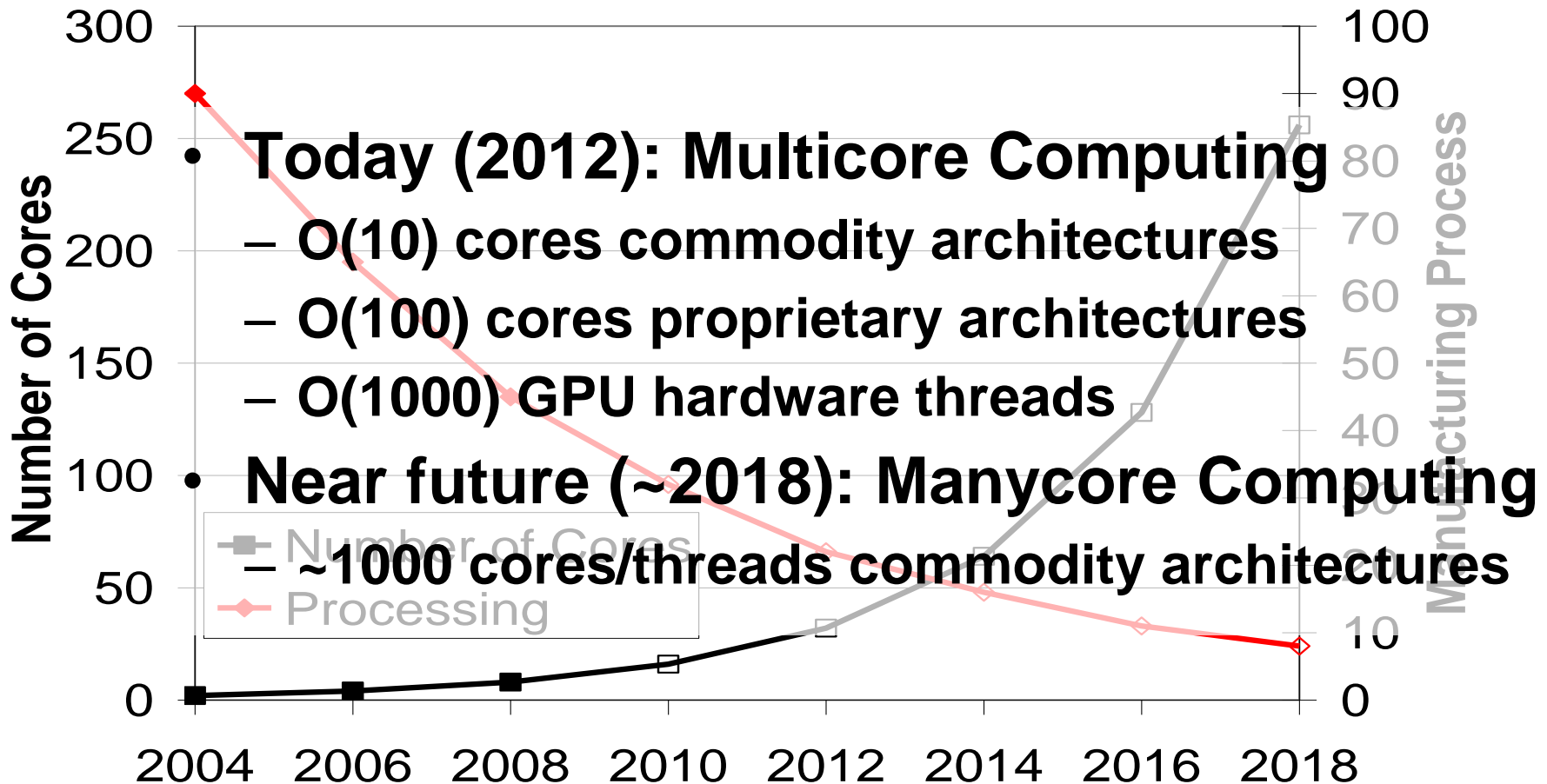*Tommy Minyard, TACC*

Cloud Computing

- *A large-scale distributed computing paradigm driven by:*
  1. *economies of scale*
  2. *virtualization*
  3. *dynamically-scalable resources*
  4. *delivered on demand over the Internet*

Clouds ~ hosting

Manycore Computing

- **Today (2012): Multicore Computing**
  - **O(10) cores commodity architectures**
  - **O(100) cores proprietary architectures**
  - **O(1000) GPU hardware threads**
- **Near future (~2018): Manycore Computing**
  - **~1000 cores/threads commodity architectures**

Chart legend: Number of Cores, Processing

Y-axis (left): Number of Cores — 0, 50, 100, 150, 200, 250, 300
Y-axis (right): Manufacturing Process — 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100
X-axis: 2004 2006 2008 2010 2012 2014 2016 2018

Exascale Computing

- ## Today (2012): Petascale Computing
  - ### O(100K) nodes
  - ### O(1M) cores
- ## Near future (~2018): Exascale Computing
  - ### ~1M nodes (10X)
  - ### ~1B processor-cores/threads (1000X)
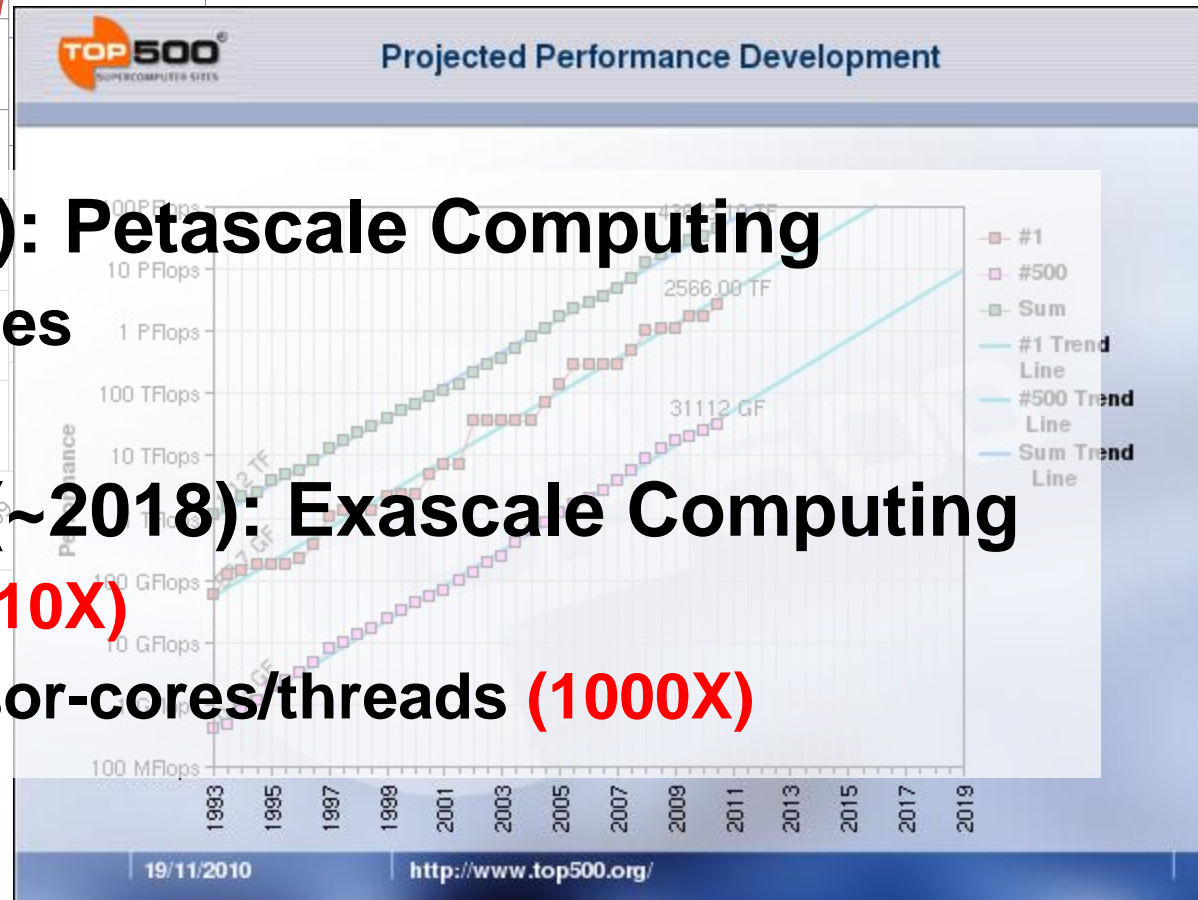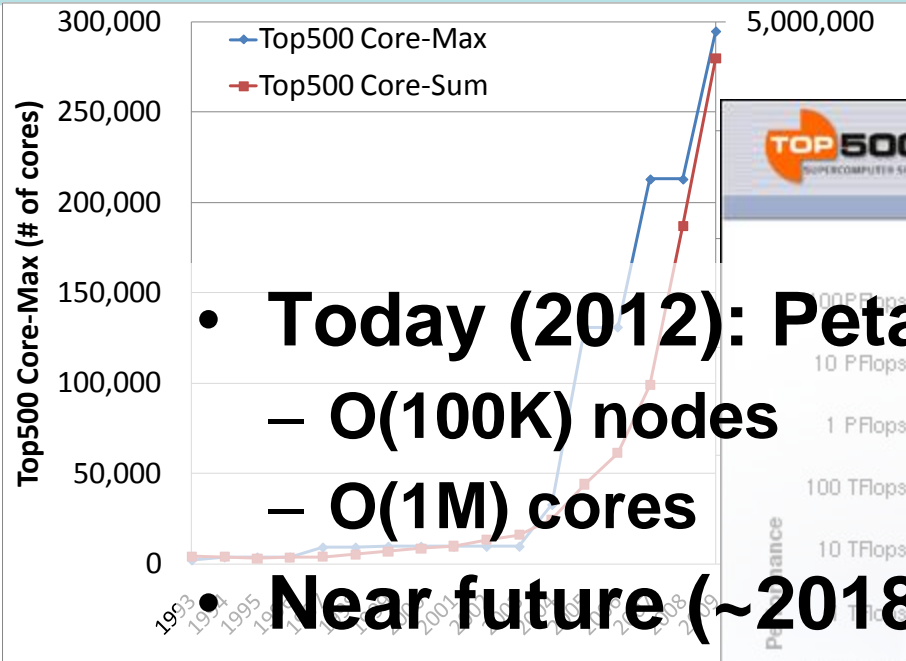
Top500 Projected Development,
http://www.top500.org/lists/2010/11/performance_development

10

Cloud Computing

- Relatively new paradigm… 3~4 years old
- Amazon in 2009
  - 40K servers split over 6 zones
    - 320K-cores, 320K disks
    - $100M costs + $12M/year in energy costs
    - Revenues about $250M/year
    - http://www.siliconvalleywatcher.com/mt/archives/2009/10/meausuring_amaz.php
- Amazon in 2018
  - Will likely look similar to exascale computing
    - 100K~1M nodes, ~1B-cores, ~1M disks
    - $100M~$200M costs + $10M~$20M/year in energy
    - Revenues 100X~1000X of what they are today

Common Challenges

- ## Power efficiency
  - Will limit the number of cores on a chip (Manycore)
  - Will limit the number of nodes in cluster (Exascale and Cloud)
  - Will dictate a significant part of the cost of ownership

- ## Programming models/languages
  - Automatic parallelization
  - Threads, MPI, workflow systems, etc
  - Functional, imperative
  - Languages vs. Middleware

Common Challenges

- Bottlenecks in scarce resources
  - Storage (Exascale and Clouds)
  - Memory (Manycore)
- Reliability
  - How to keep systems operational in face of failures
  - Checkpointing (Exascale)
  - Node-level replication enabled by virtualization (Exascale and Clouds)
  - Hardware redundancy and hardware error correction (Manycore)

Research Directions

- ***Decentralization is critical***
  - Computational resource management (e.g. LRMs)
  - Storage systems (e.g. parallel file systems)
- ***Data locality must be maximized, while preserving I/O interfaces***
  - POSIX I/O on shared/parallel file systems ignore locality
  - Data-aware scheduling coupled with distributed file systems that expose locality is the key to scalability over the next decade

Main Message

- ***Preserving locality is critical!***

- *Segregating storage from compute resources is* **BAD**

- *Parallel file systems + distributed file systems + distributed hash tables + nonvolatile memory*
  ➔ new storage architecture for extreme-scale HEC

- *Co-locating storage and compute is* **GOOD**

  - *Leverage the abundance of processing power, bisection bandwidth, and local I/O*

Projects
Many-Task Computing

- [SimMatrix: Simulator for MAny-Task computing execution fabRIc at eXascales](#)

- [MATRIX: MAny-Task computing execution fabRIc at eXascales](#)

- [Falkon: Fast and Light-weight tasK executiON framework](#)

- [Swift: Fast, Reliable, Loosely Coupled Parallel Computation](#)

Projects
Storage

- [FusionFS: Fusion distributed File System](#) ([PDF](#))
- PAFS: Provenance-Aware Distributed File System ([PDF](#))
- [HyCache: A Hybrid User-Level File System with SSD Caching](#) ([PDF](#))
- [ZHT: Zero-Hop Distributed Hash Table](#) ([PDF](#))
- IDAStore: Information Dispersal Algorithms for Distributed Storage Systems
- NoVoHT: Non-Volatile Hash Table (PDF)

Projects
HPC/Cloud/Many-Core

- **High-Performance Computing**
  - SimHEC: Simulator for High-End Computing Systems
- **Cloud Computing**
  - CloudStorage: Understanding the Cost of Cloud Storage
- **Many-Core Computing**
  - ManyCoreSim: Scheduling Direct Acyclic Graphs on Massively Parallel Processors
  - GeMTC: GPU Enabled Many-Task Computing
- **Mobile Computing**
  - CiteSearcher: a Google Scholar frontend for iOS and Android mobile devices

Data-Intensive Computing at the Intersection of Cloud Computing and Supercomputing

Coursework

- CS495: Introduction to Distributed Computing
- CS 546 Parallel and Distributed Processing
- CS 550 Advanced Operating Systems
- CS 552 Distributed Real-Time Systems
- **CS 553 Cloud Computing**
- CS 570 Advanced Computer Architecture
- CS 595 Data-Intensive Distributed Computing

Specializations

- Undergraduate Level
  - **Distributed and Cloud Computing**
  - **Data Science**
- Master Level
  - **Distributed and Cloud Computing**
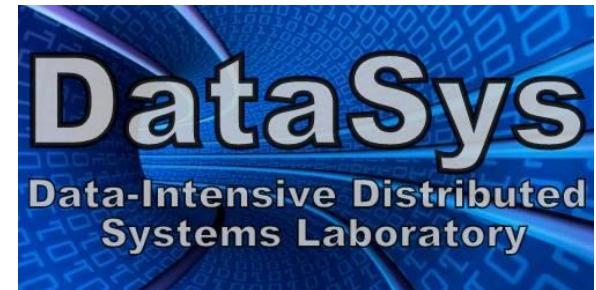 - **Cyber-Physical Systems**
  - **Data Analytics**

Faculty

- Xian-He Sun

- Zhiling Lan

- Shangping Ren

- Ioan Raicu

Scalable Computing Software Laboratory

DataSys — Data-Intensive Distributed Systems Laboratory

Employment Opportunities

- Google
- Yahoo
- Microsoft
- Amazon
- IBM
- Apple
- VMWare
- Netflix
- Cray
- Intel

- NVIDIA
- Facebook
- LinkedIn
- Salesforce.com
- Rackspace
- Red Hat
- Cleversafe
- UnivaUD
- Greenplum
- AsterData

- Proprietary Trading Companies
- Department of Energy Laboratories
- NASA
- Academic supercomputer centers
- Many more…

More Information

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