Introduction to Distributed Systems

Ioan Raicu
Computer Science Department
Illinois Institute of Technology

CS554: Data-Intensive Computing
August 21st, 2013

Famous Quotes

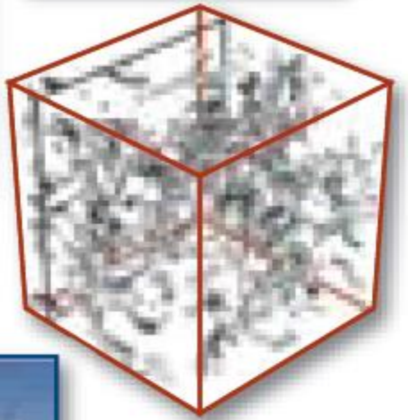
The advent of computation can be compared, in terms of the breadth and depth of its impact on research and scholarship, to the invention of writing and the development of modern mathematics.

Ian Foster, 2006

Science Paradigms

- Thousand years ago:
  science was **empirical**
  describing natural phenomena
- Last few hundred years:
  theoretical branch
  using models, generalizations

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{4\pi Gp}{3} - K\frac{c^2}{a^2}$$



- Last few decades:
  a **computational** branch
  simulating complex phenomena
- Today: **data exploration** (eScience)
  unify theory, experiment, and simulation
  - Data captured by instruments
    or generated by simulator
  - Processed by software
  - Information/knowledge stored in computer
  - Scientist analyzes database/files
    using data management and statistics

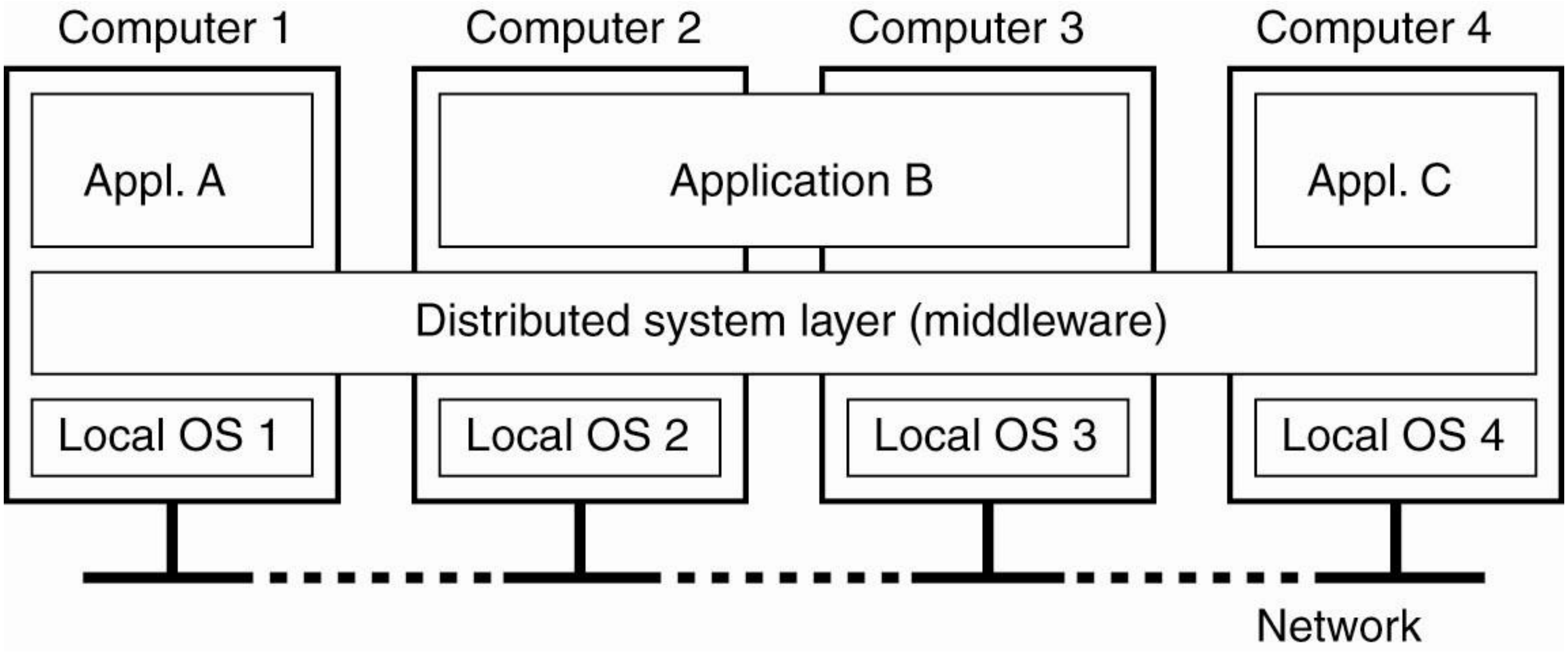
Distributed Systems

- What is a distributed system?

“A collection of independent computers that appears to its users as a single coherent system”

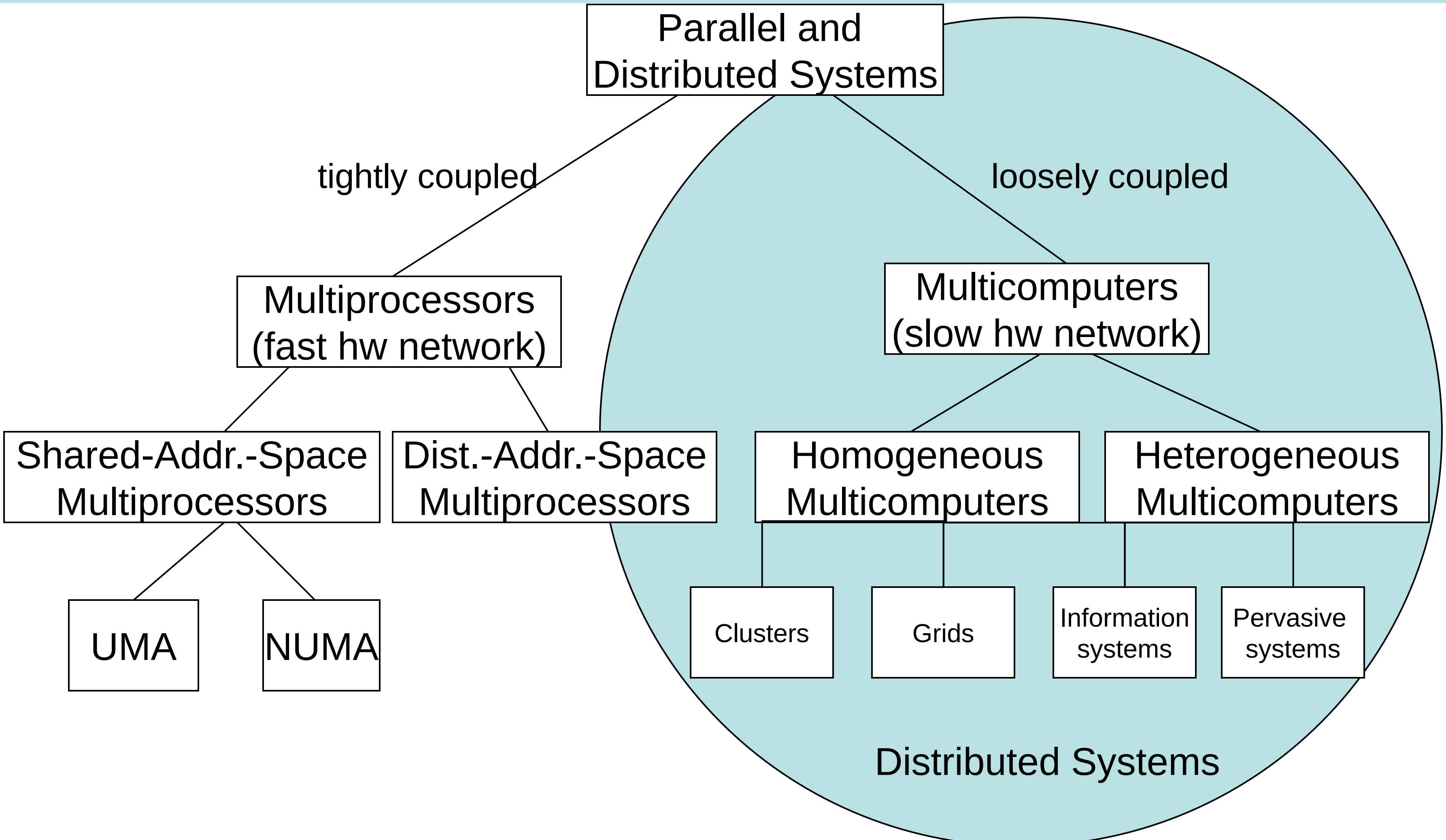
-A. Tanenbaum

Distributed Systems

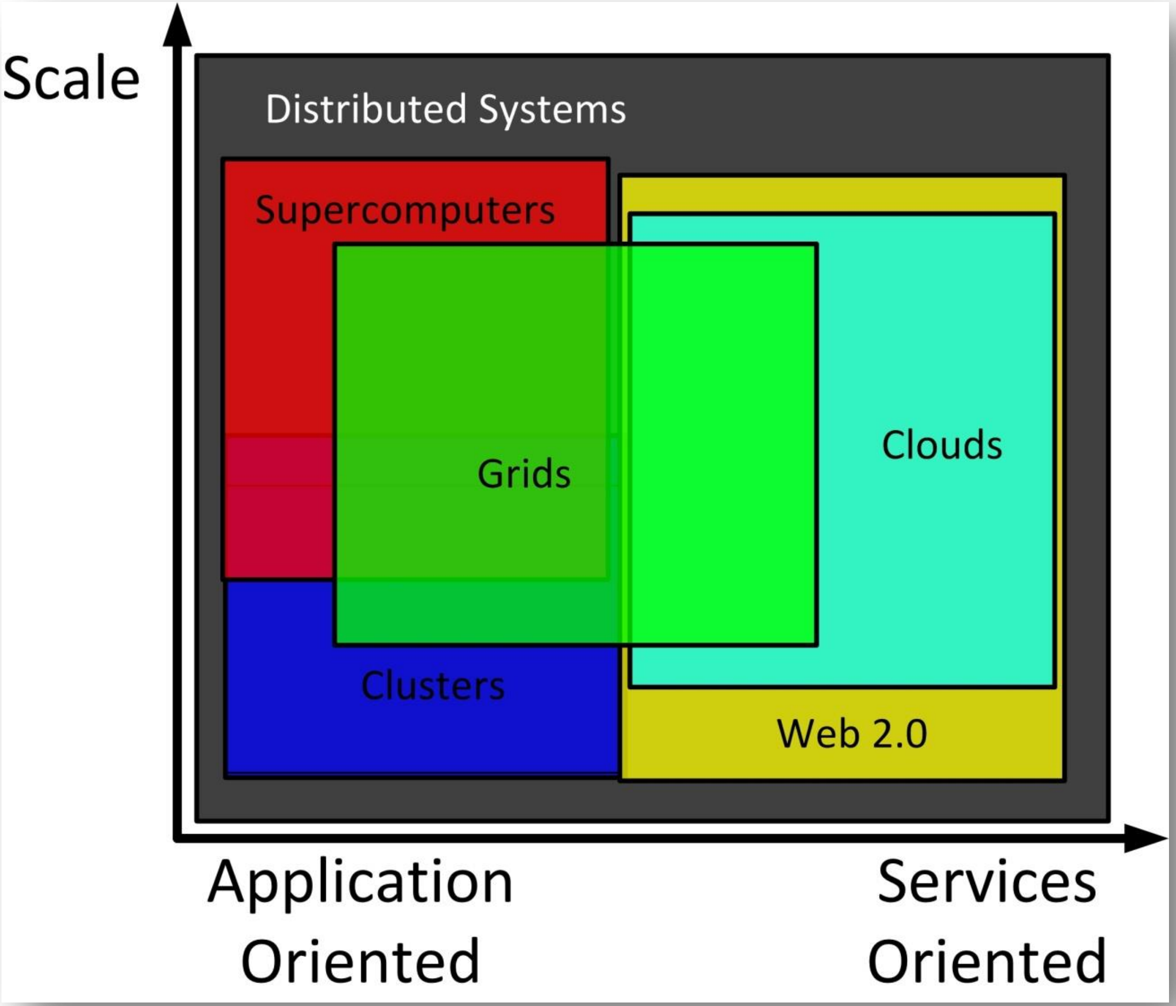


A distributed system organized as middleware. The middleware layer extends over multiple machines, and offers each application the same interface.

Distributed Systems

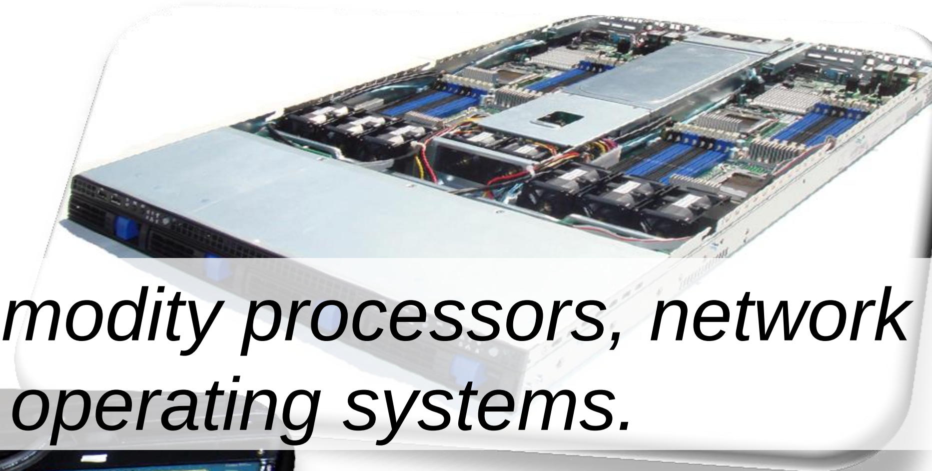


Distributed Systems:
Clusters, Grids, Clouds, and Supercomputers



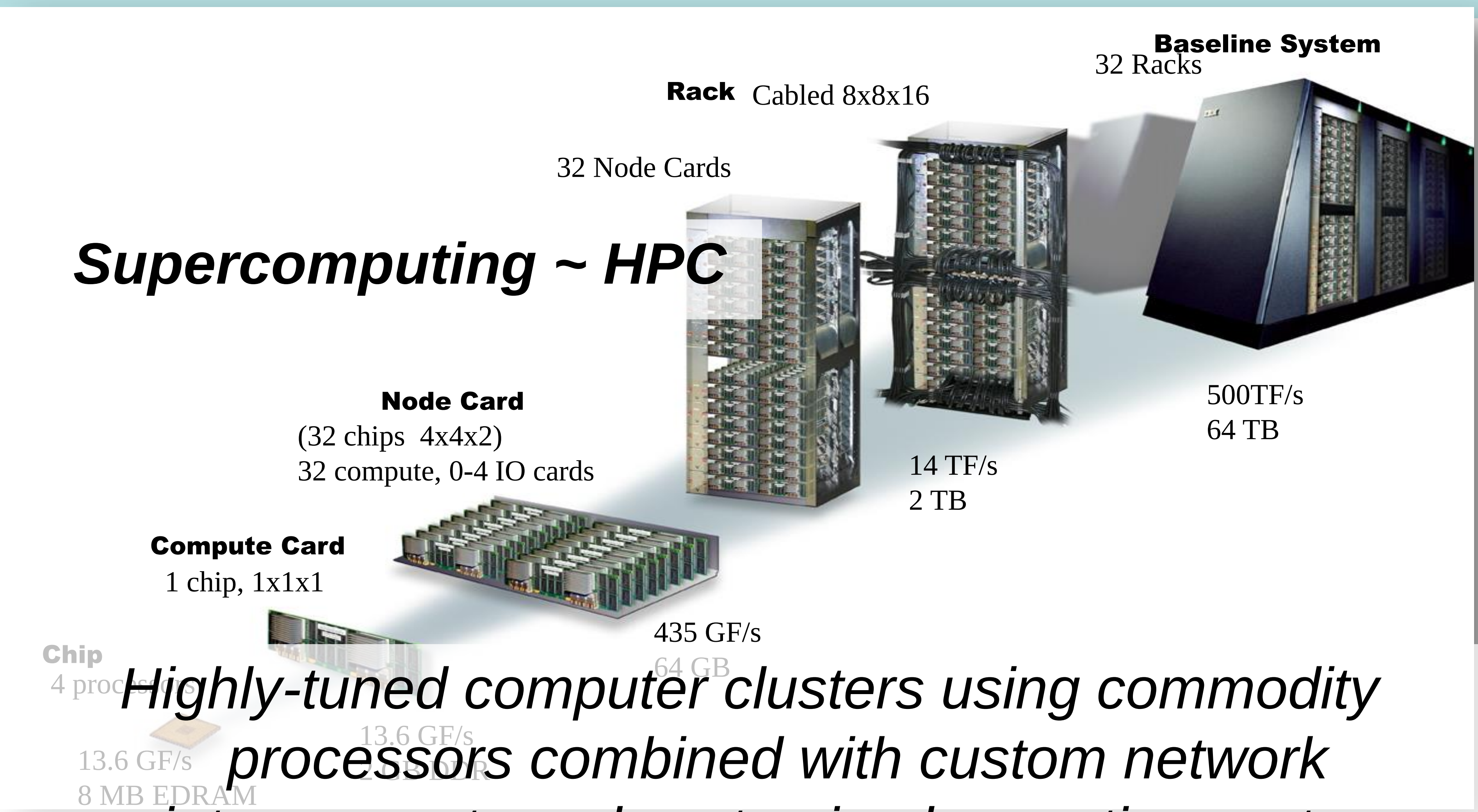
[GCE08] “Cloud Computing and Grid Computing 360-Degree Compared”

Cluster Computing



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

Supercomputing



Supercomputing ~ HPC

Highly-tuned computer clusters using commodity processors combined with custom network interconnects and customized operating system

Top Supercomputers from Top500

- Cray XT4 & XT5
  - Cray #1
  - Cielo #18
  - Hopper #19
- IBM BlueGene/L/P/Q
  - Sequia #2
  - Mira #4
  - Juqueen #5
  - Fermi #9
- GPU based
  - Titan #1
  - Tianhe-1A #8
  - Nebulae #12
- SGI Altix ICE
  - Plaiedas #14
- SPARC64 VIIIfx
  - K #3

Top500 List - November 2012

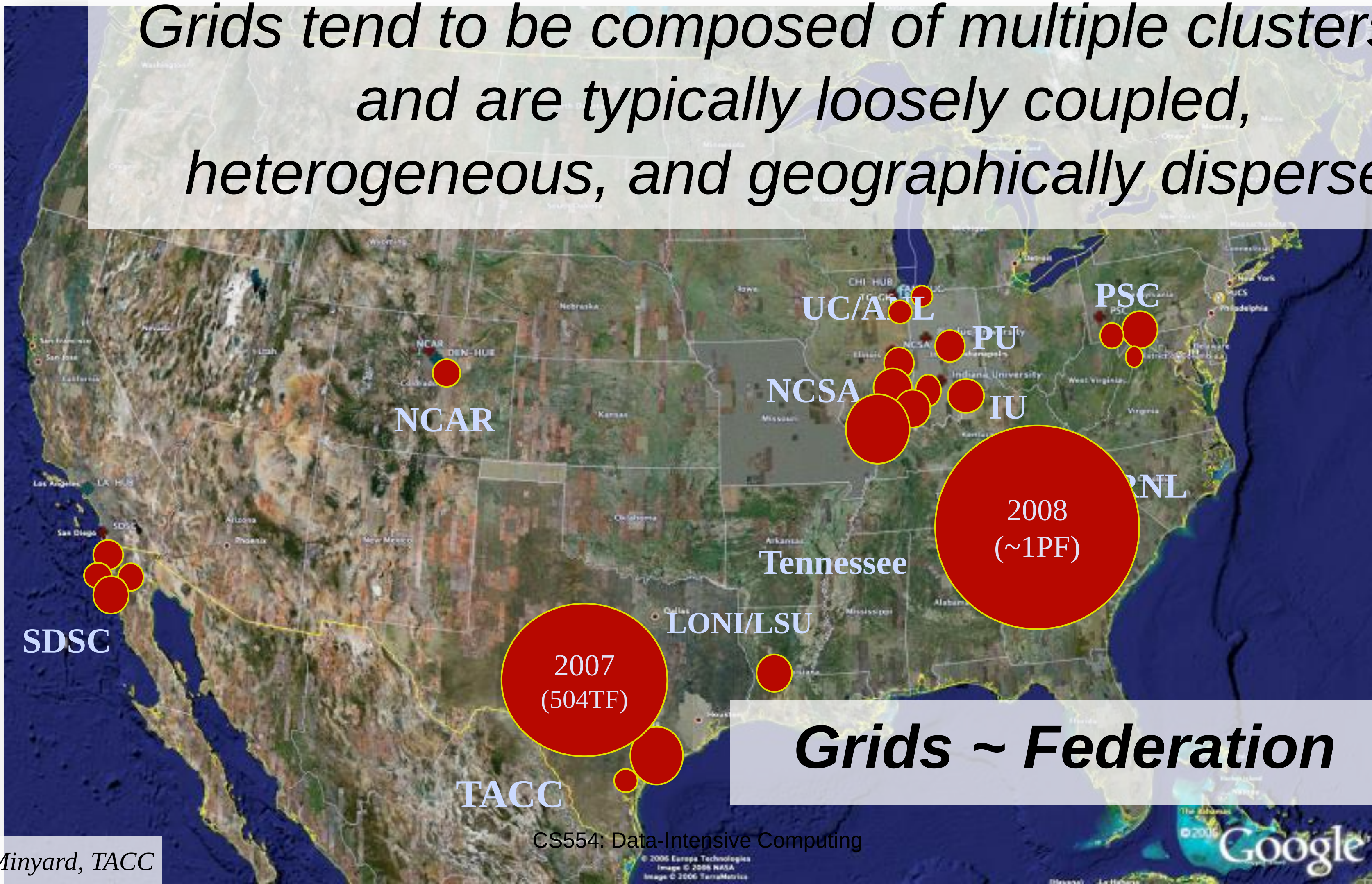
R_{max} and R_{peak} values are in TFlops. For more details about other fields, check the TOP500 description.

previous 1 2 3 4 5 next

Rank	Site	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	DOE/SC/Oak Ridge National Laboratory United States	**Titan** - Cray XK7 , Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x Cray Inc.	560640	17590.0	27112.5	8209
2	DOE/NNSA/LLNL United States	**Sequoia** - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom IBM	1572864	16324.8	20132.7	7890
3	RIKEN Advanced Institute for Computational Science (AICS) Japan	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect Fujitsu	705024	10510.0	11280.4	12660
4	DOE/SC/Argonne National Laboratory United States	**Mira** - BlueGene/Q, Power BQC 16C 1.60GHz, Custom IBM	786432	8162.4	10066.3	3945
5	Forschungszentrum Juelich (FZJ) Germany	**JUQUEEN** - BlueGene/Q, Power BQC 16C 1.600GHz, Custom Interconnect IBM	393216	4141.2	5033.2	1970

Grid Computing

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



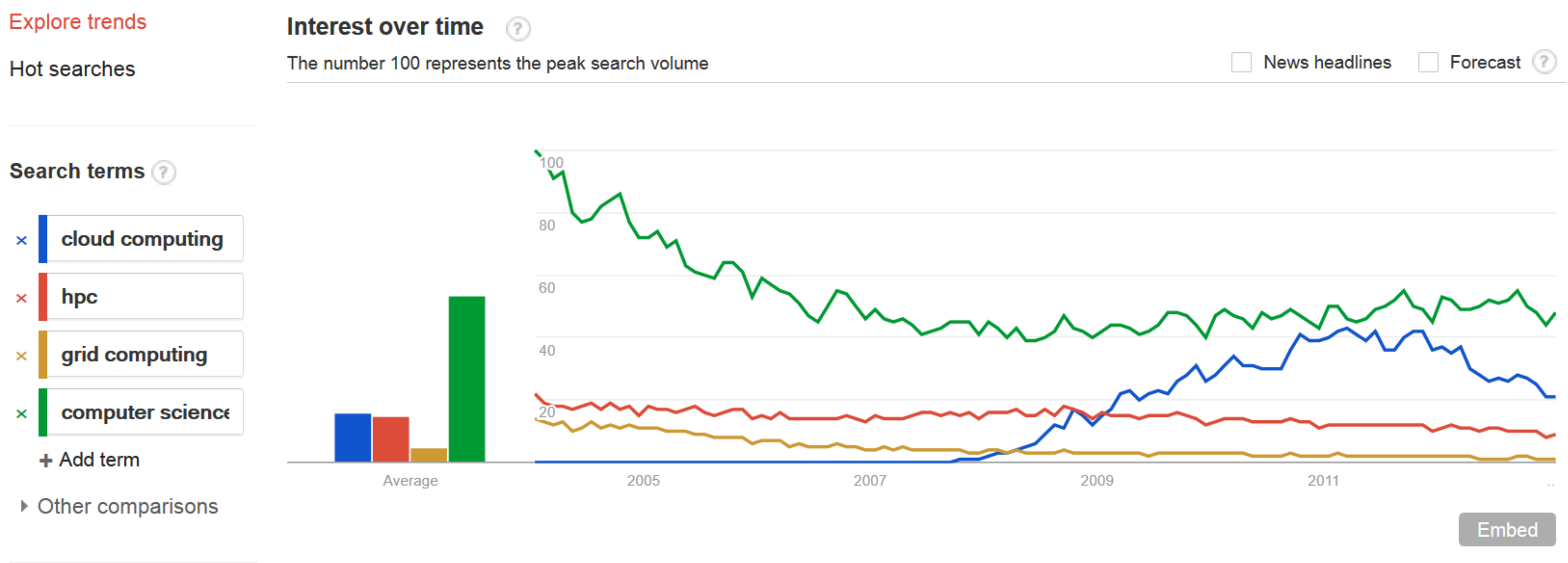
Grids ~ Federation

Tommy Minyard, TACC

Major Grids

- XSEDE (Formerly TeraGrid)
  - 200K-cores across 11 institutions and 22 systems over the US
- Open Science Grid (OSG)
  - 43K-cores across 80 institutions over the US
- Enabling Grids for E-sciencE (EGEE)
- LHC Computing Grid from CERN
- Middleware
  - Globus Toolkit
  - Unicore

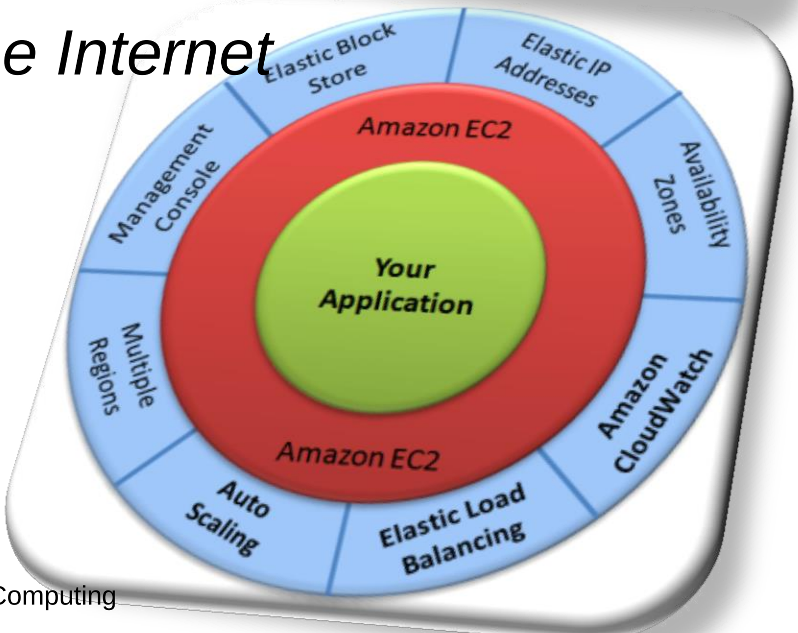
Cloud Computing: A Mature Paradigm



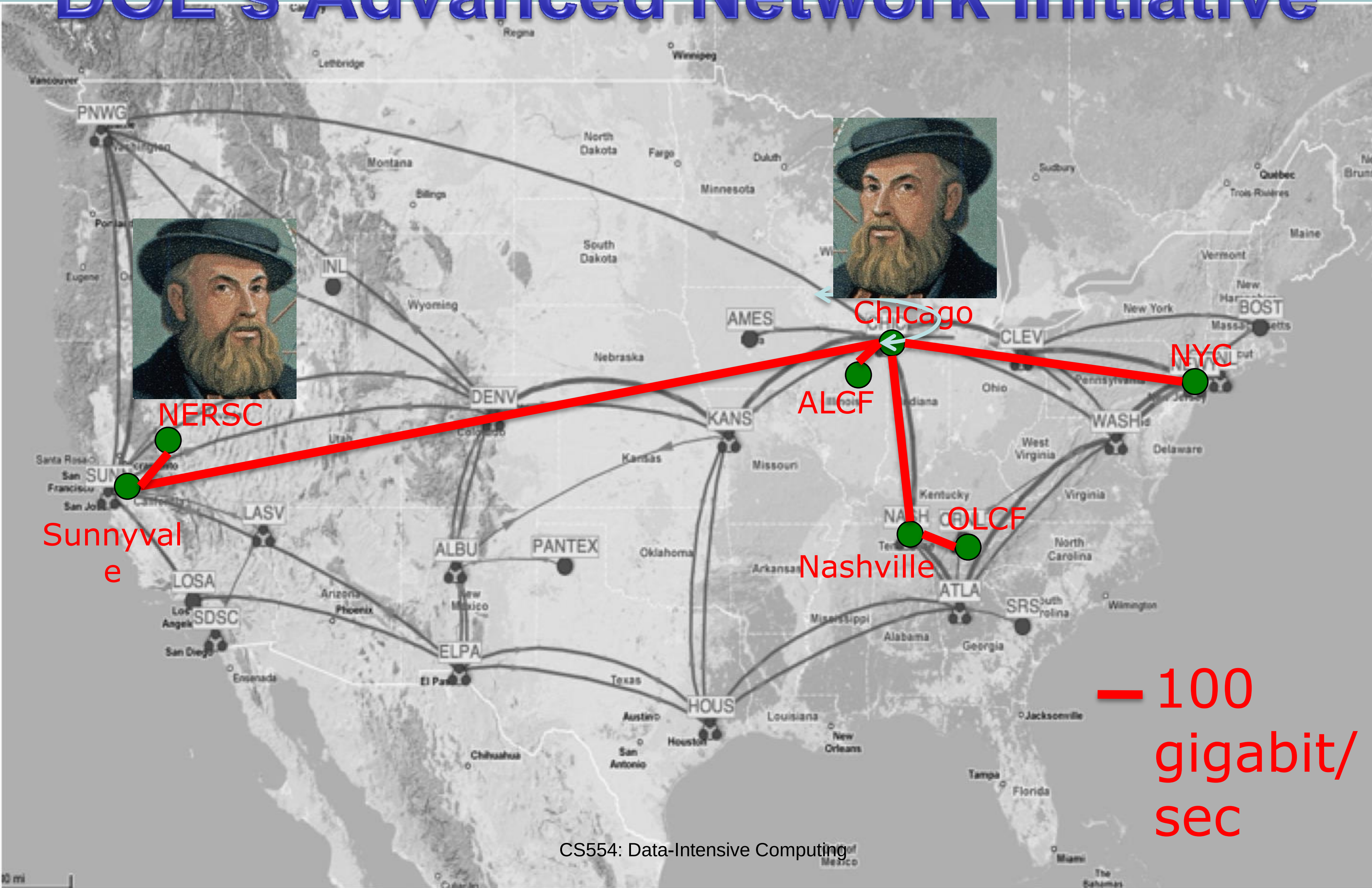
Cloud Computing

- *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



Magellan + DOE's Advanced Network Initiative



Major Clouds

- Industry
  - Google App Engine
  - Amazon
  - Windows Azure
  - Salesforce
- Academia/Government
  - Magellan
  - FutureGrid
- Opensource middleware
  - Nimbus
  - Eucalyptus
  - OpenNebula
  - OpenStack

Distributed vs. Single Systems

- Data sharing
  - Multiple users can access common database, data files,…
- Device/resource sharing
  - Printers, servers, CPUs,….
- Communication
  - Communication with other machines…
- Flexibility
  - Spread workload to different & most appropriate machines
- Extensibility
  - Add resources and software as needed

Distributed vs. Centralized Systems

- Economics
  - Microprocessors have better price/performance than mainframes
- Speed
  - Collective power of large number of systems
- Geographic and responsibility distribution
- Reliability
  - One machine's failure need not bring down the system
- Extensibility
  - Computers and software can be added incrementally

Disadvantages of Distributed Systems

- Software
  - Little software exists compared to PCs
- Networking
  - Still slow and can cause other problems (e.g. when disconnected)
- Security
  - Data may be accessed by unauthorized users

Key Characteristics of Distributed Systems

- Support for resource sharing
- Openness
- Concurrency
- Scalability
- Fault tolerance (reliability)
- Transparence

Resource Sharing

- Share hardware,software,data and information
- Hardware devices
  - Printers,disks,memory,…
- Software sharing
  - Compilers,libraries,toolkits,…
- Data
  - Databases,files,…

Openness

- Definition?
- Hardware extensions
  - Adding peripherals,memory,communication interfaces…
- Software extensions
  - Operating systems features
  - Communication protocols

Concurrency

- In a single system several processes are interleaved
- In distributed systems: there are many systems with one or more processors
  - Many users simultaneously invoke commands or applications
  - Many servers processes run concurrently, each responding to different client request

Scalability

- Scale of system
  - Few PCs servers ->dept level systems->local area networks->internetworked systems->wide are network…
  - Ideally, system and application software should not change as systems scales
- Scalability depends on all aspects
  - Hardware
  - Software
  - networks

Fault Tolerance

- Definition?
- Two approaches:
  - Hardware redundancy
  - Software recovery
- In distributed systems:
  - Servers can be replicated
  - Databases may be replicated
  - Software recovery involves the design so that state of permanent data can be recovered

Transparency in a Distributed System

Transparency	Description
Access	Hide differences in data representation and how a resource is accessed
Location	Hide where a resource is located
Migration	Hide that a resource may move to another location
Relocation	Hide that a resource may be moved to another location while in use
Replication	Means that users do not know whether a replica or a master provides a service.
Concurrency	Hide that a resource may be shared by several competitive users
Failure	Hide the failure and recovery of a resource
Persistence	Hide whether a (software) resource is in memory or on disk

Pitfalls When Developing Distributed Systems

- False assumptions made by first time developer:
  - The network is reliable.
  - The network is secure.
  - The network is homogeneous.
  - The topology does not change.
  - Latency is zero.
  - Bandwidth is infinite.
  - Transport cost is zero.
  - There is one administrator.

Questions

?