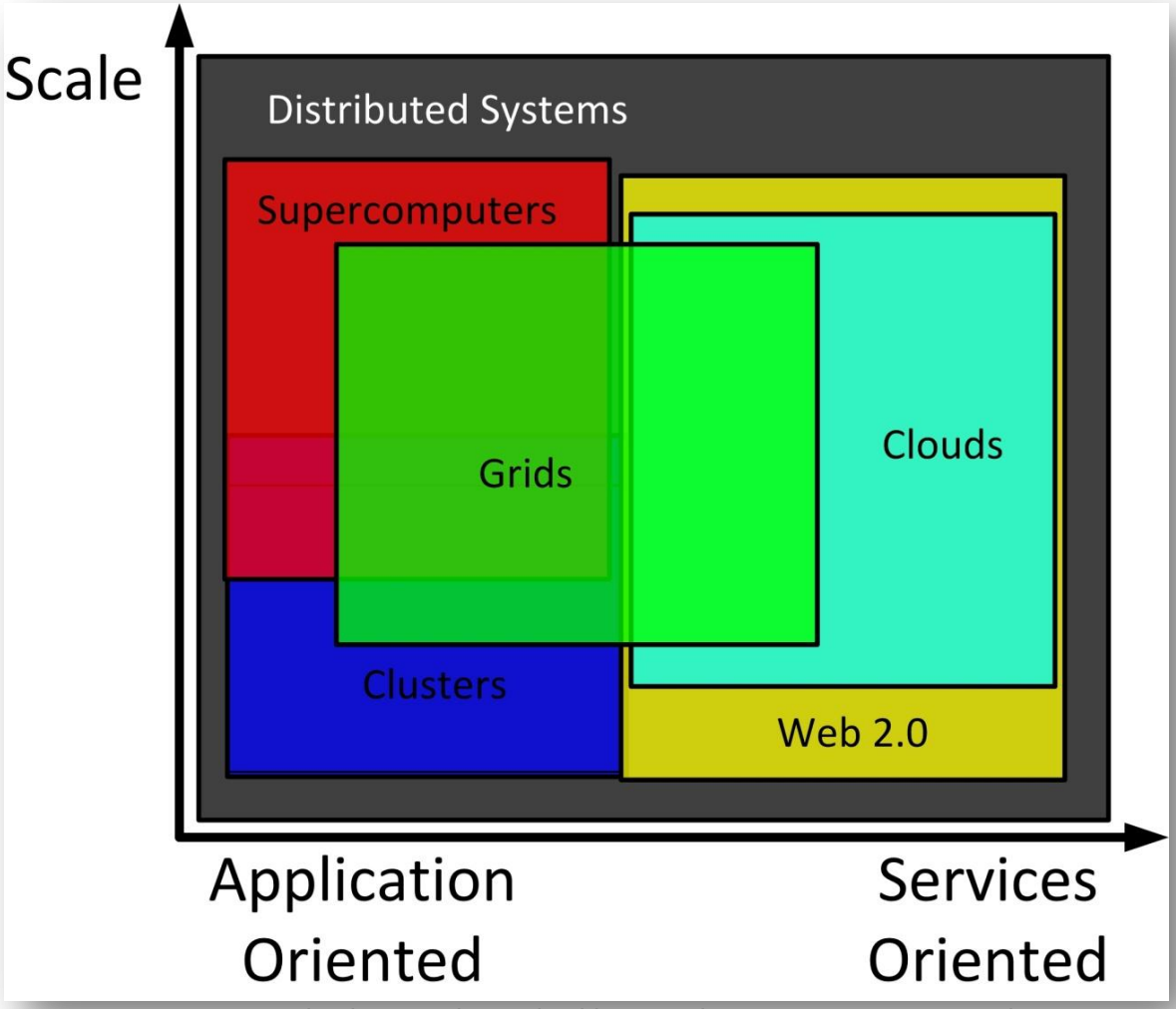
Cloud Computing and Grid Computing 360-Degree Compared

Ioan Raicu
Computer Science Department
Illinois Institute of Technology

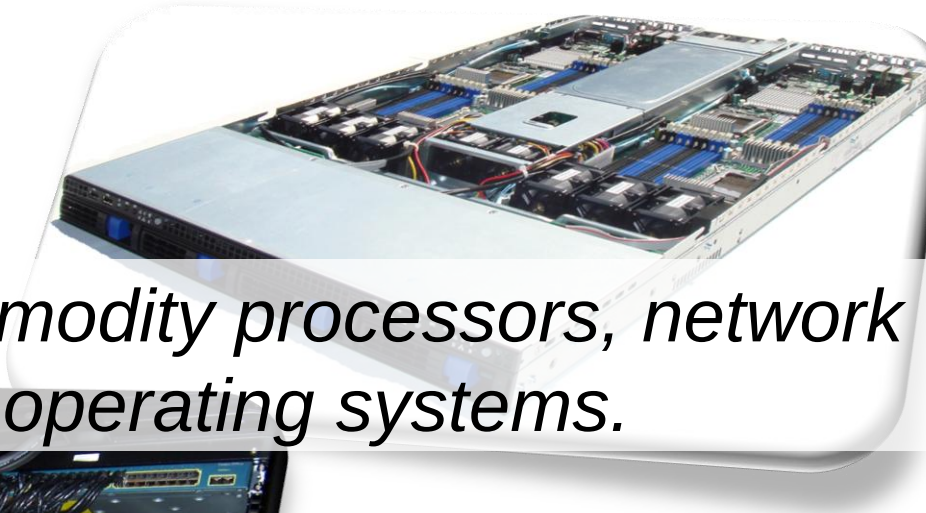
CS 595
Hot Topics in Distributed Systems: Data-Intensive Computing
September 15th, 2010

Clusters, Grids, Clouds, and Supercomputers

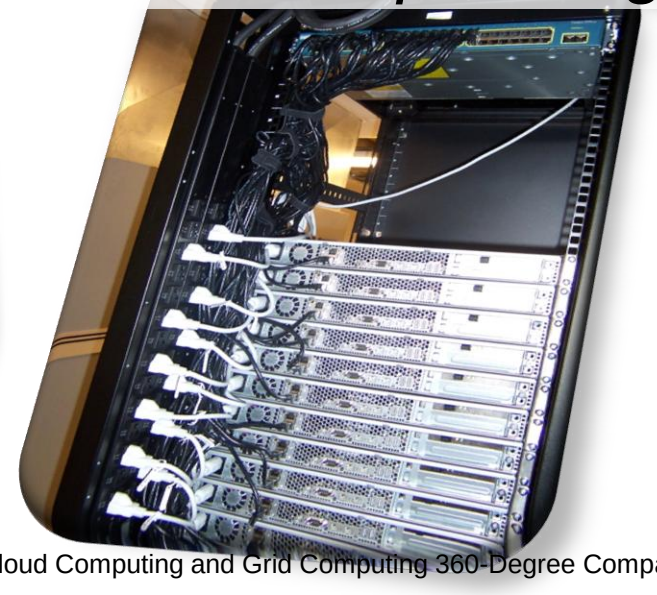


[GCE08] "Cloud Computing and Grid Computing 360-Degree Compared"

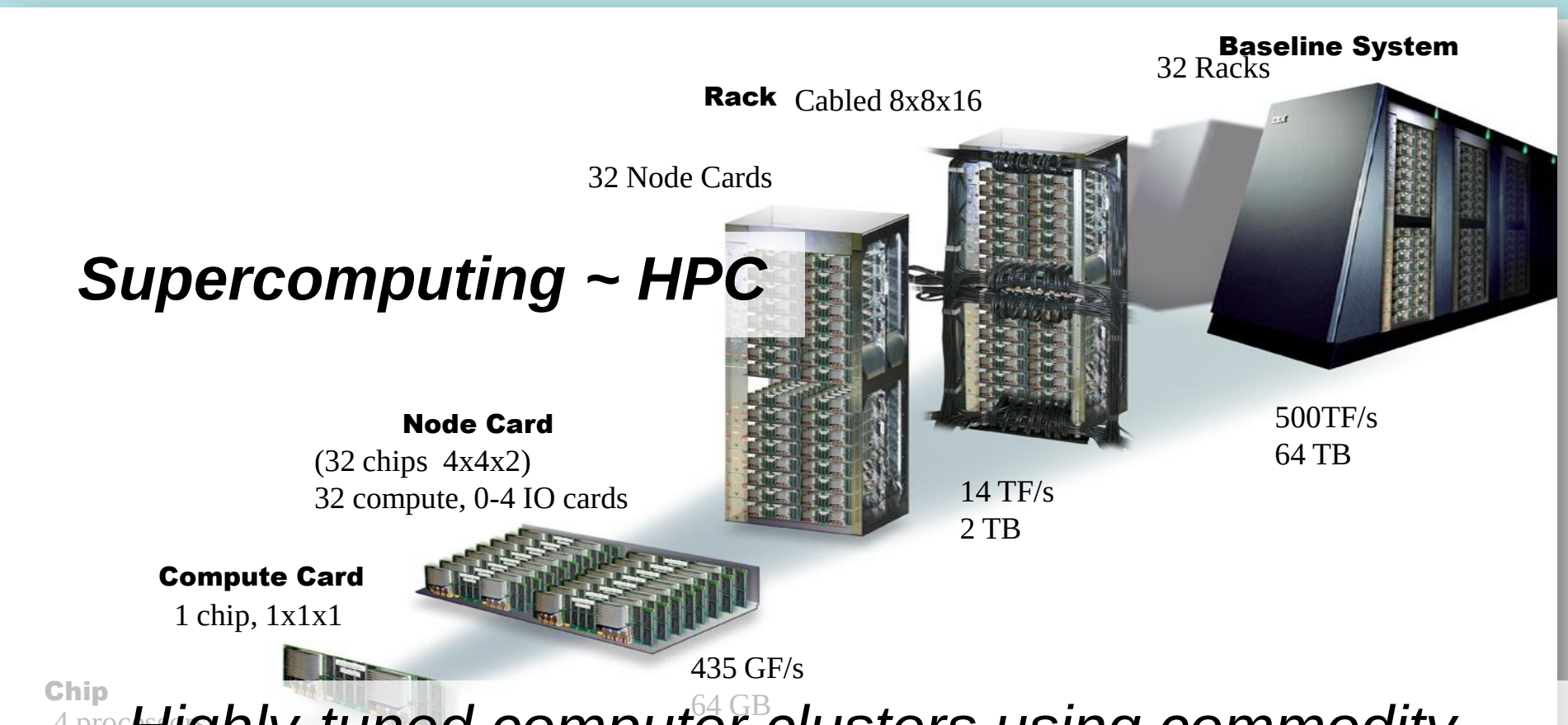
Cluster Computing



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



Supercomputing



Supercomputing ~ HPC

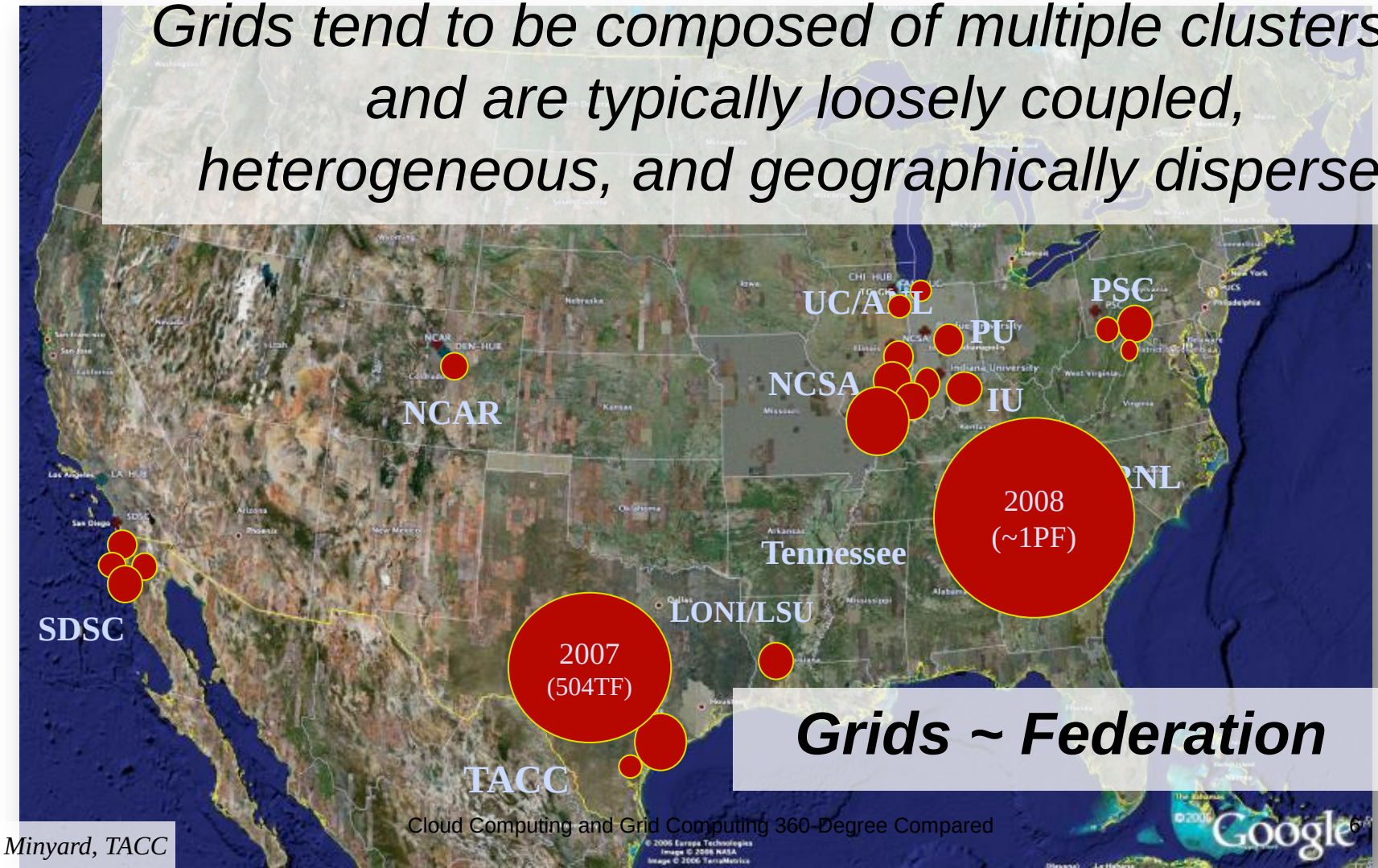
Highly-tuned computer clusters using commodity processors combined with custom network interconnects and customized operating system

Top 10 Supercomputers from Top500

- Cray XT4 & XT5
  - Jaguar #1
  - Kraken #3
- IBM BladeCenter Hybrid
  - Roadrunner #2
- IBM BlueGene/L & BlueGene/P
  - Jugene #4
  - Intrepid #8
  - BG/L #7
- NUDT (GPU based)
  - Tianhe-1 #5
- SGI Altix ICE
  - Plaiedas #6
- Sun Constellation
  - Ranger #9
  - Red Sky #10

Grid Computing

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



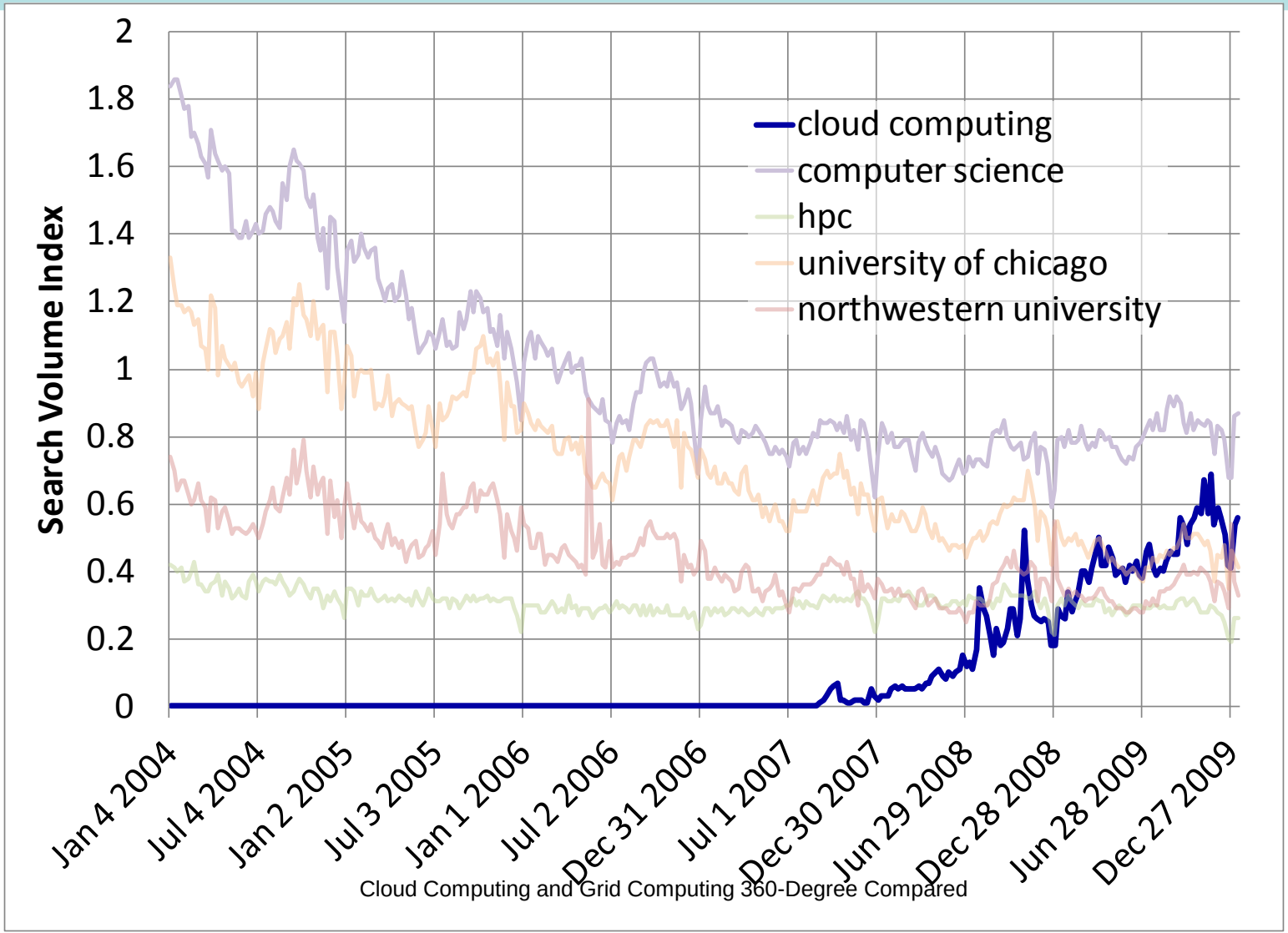
Grids ~ Federation

Tommy Minyard, TACC

Major Grids

- TeraGrid (TG)
  - 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

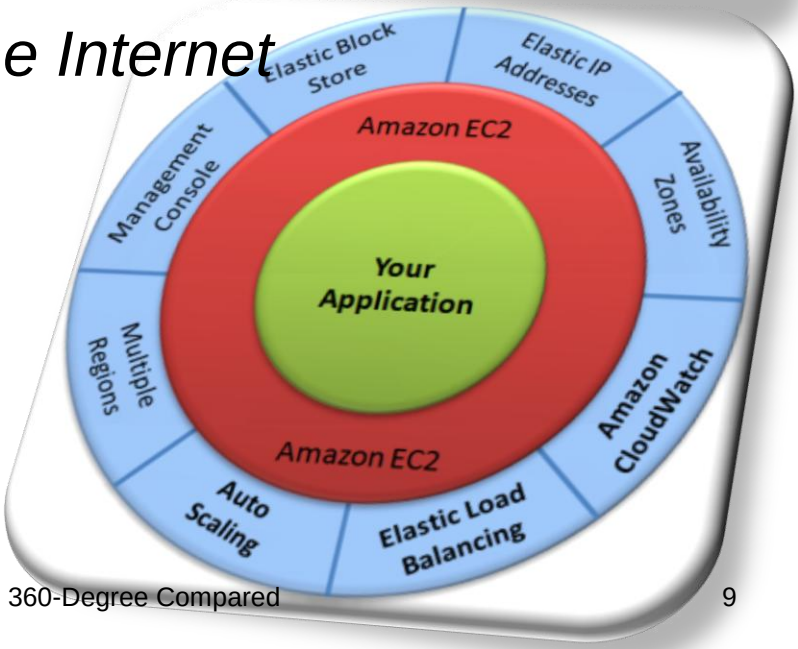
Cloud Computing: An Emerging Paradigm



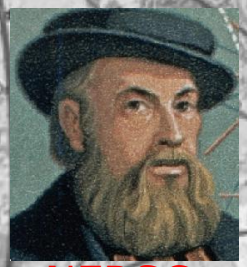
Cloud Computing

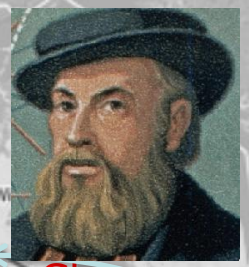
- *A large-scale distributed computing paradigm driven by:*
  1. *economies of scale*
  2. *virtualization*
  3. *dynamically-scalable resources*
  4. *delivered on demand over the Internet*

Clouds ~ hosting



Magellan + DOE's Advanced Network Initiative





Chicago

NERSC

ALCF

NYC

Sunnyvale

OLCF

Nashville

PNWG
INL
AMES
CLEV
BOST
DENV
KANS
WASH
SUN
LASV
ALBU
PANTEX
NASH
LOSA
SDSC
ATLA
SRS
ELPA
HOUS

— 100 gigabit/sec

Major Clouds

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

So is "Cloud Computing" just a new name for Grid?

- IT reinvents itself every five years
- The answer is complicated…

- **YES**: the vision is the same
  - to reduce the cost of computing
  - increase reliability
  - increase flexibility by transitioning from self operation to third party

So is "Cloud Computing" just a new name for Grid?

- **NO**: things are different than they were 10 years ago
  - New needs to analyze massive data, increased demand for computing
  - Commodity clusters are expensive to operate
  - We have low-cost virtualization
  - Billions of dollars being spent by Amazon, Google, and Microsoft to create real commercial large-scale systems with hundreds of thousands of computers
  - The prospect of needing only a credit card to get on-demand access to *infinite computers is exciting; *infinite<O(1000)

So is "Cloud Computing" just a new name for Grid?

- **YES:** the problems are mostly the same
  - How to manage large facilities
  - Define methods to discover, request, and use resources
  - How to implement and execute parallel computations
  - Details differ, but issues are similar

Questions