

Introduction to Distributed Systems

Ioan Raicu
Computer Science Department
Illinois Institute of Technology

CS554: Data-Intensive Computing
January 12th, 2023

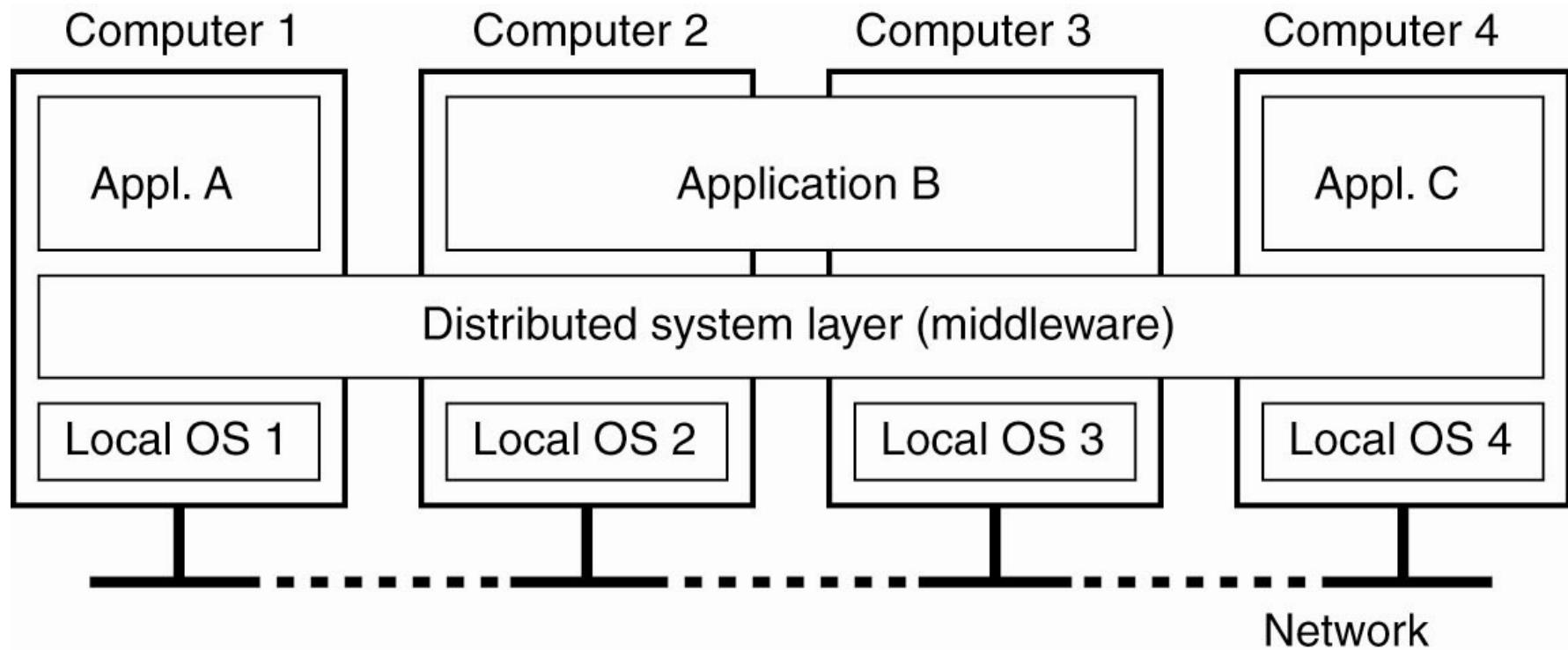
Distributed Systems

- What is a distributed system?

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

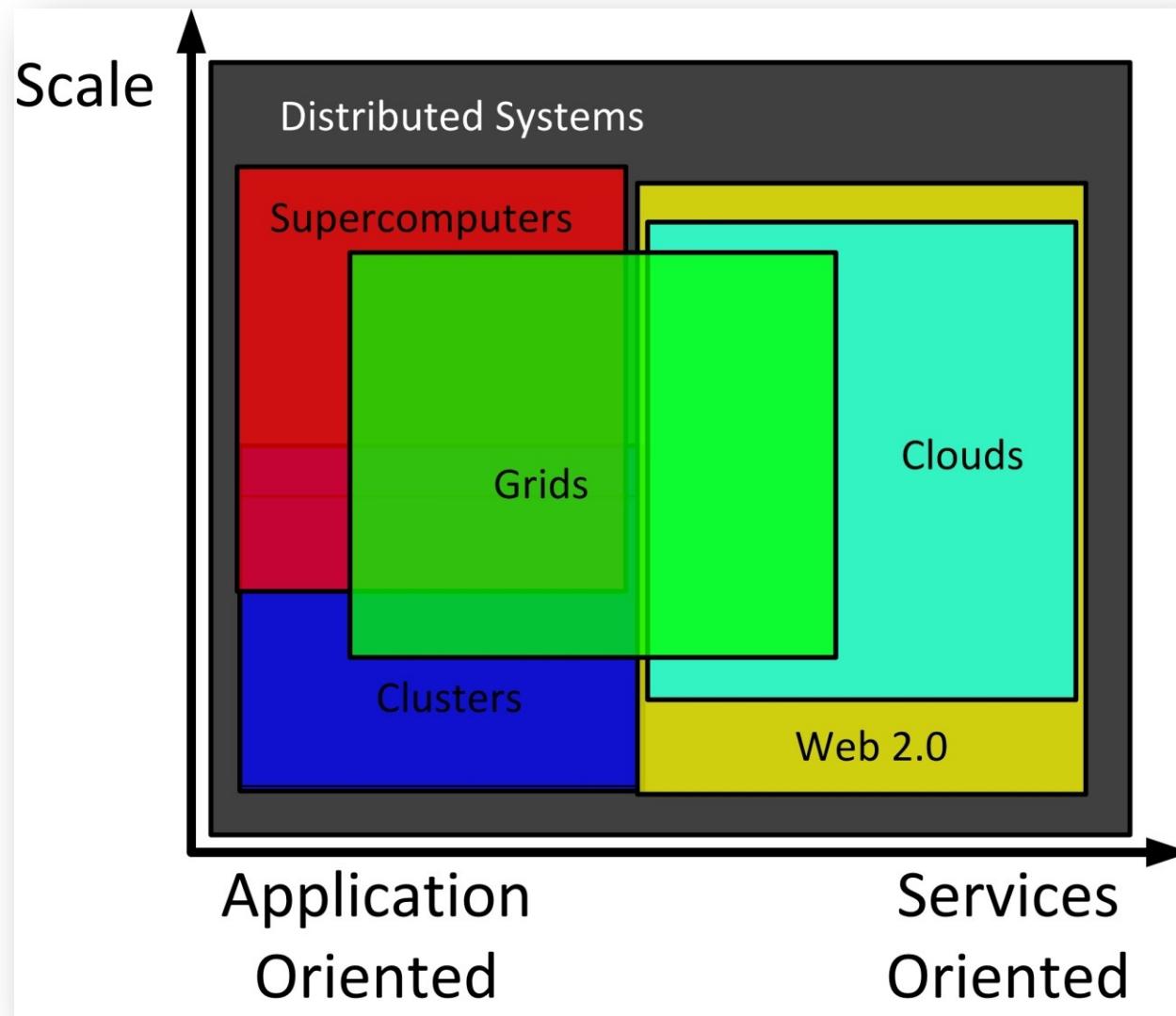
-A. Tanenbaum

Distributed Systems

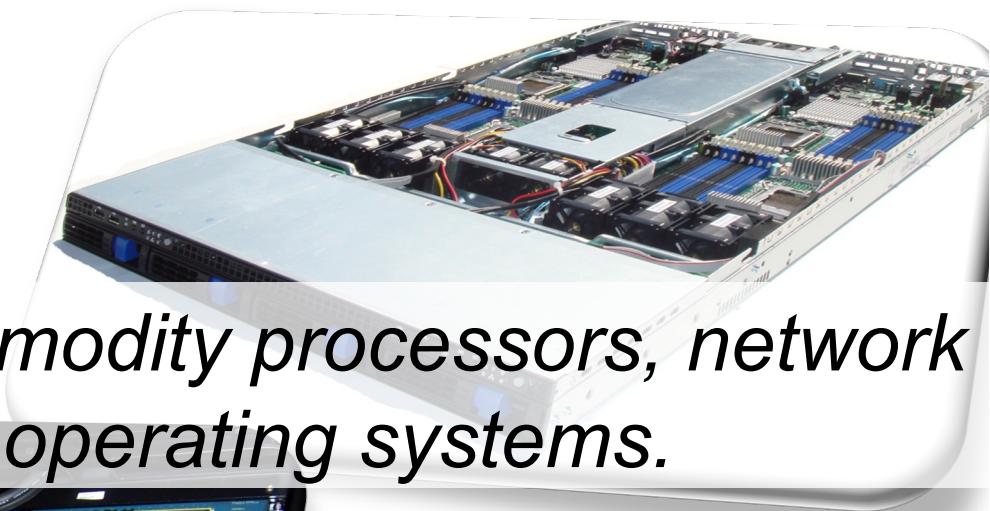
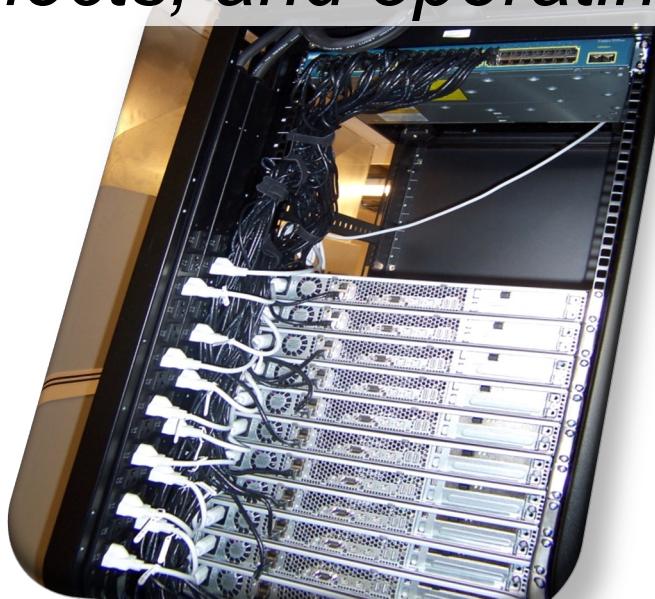


A distributed system organized as middleware. The middleware layer extends over multiple machines, and offers each application the same interface.

Distributed Systems: Clusters, Grids, Clouds, and Supercomputers



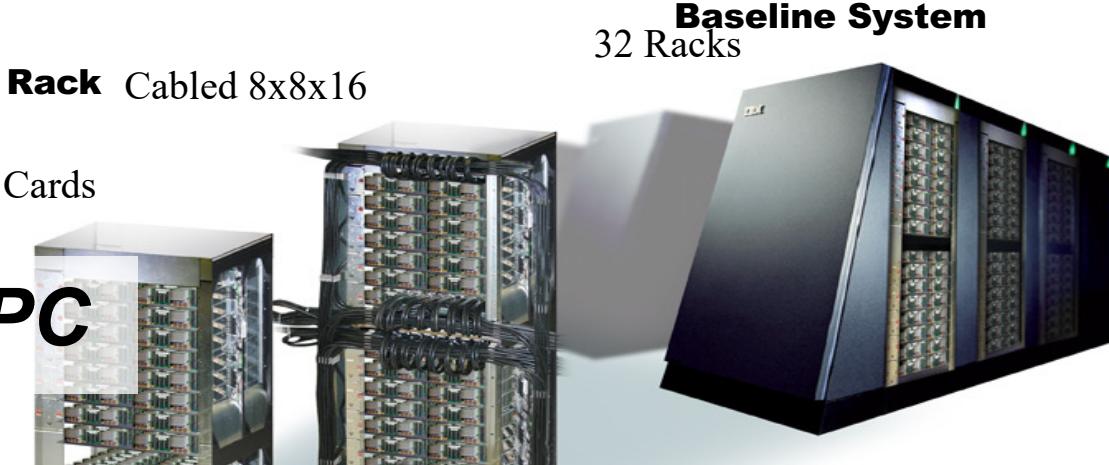
Cluster Computing



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

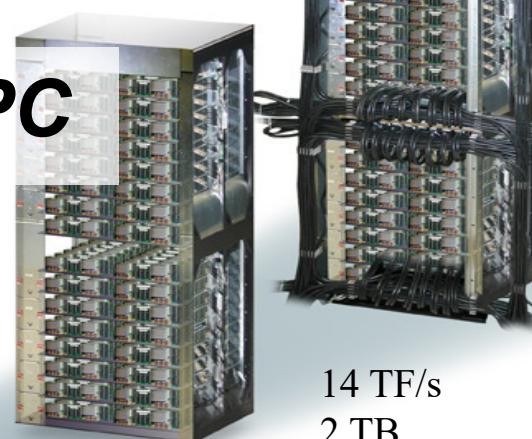
Supercomputing

Supercomputing ~ HPC



Rack Cabled 8x8x16

32 Node Cards



Node Card

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



435 GF/s

64 GB

Compute Card

1 chip, 1x1x1

Chip

4 processors



13.6 GF/s

8 MB EDRAM

13.6 GF/s

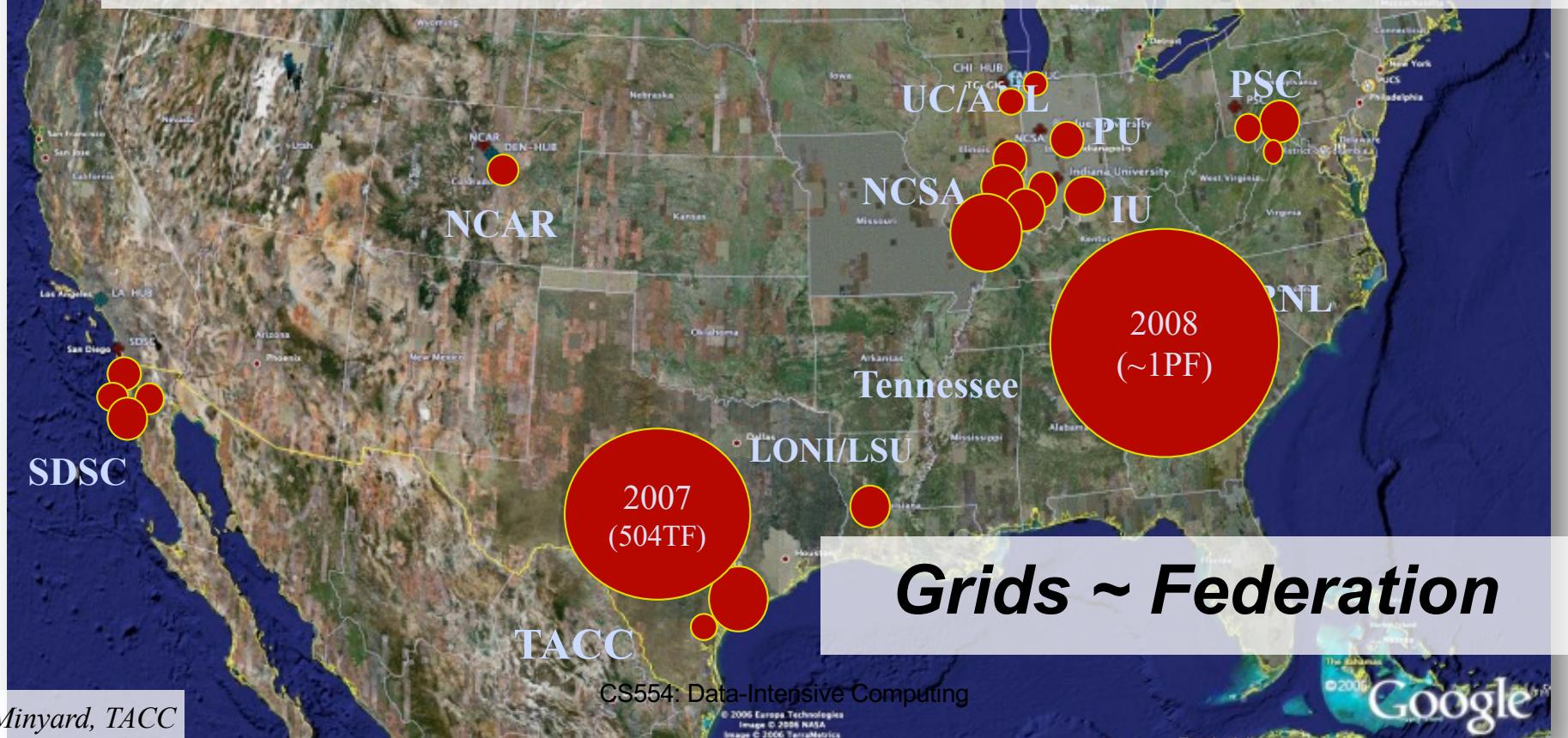
256 MB EDRAM

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

CS554 Data Intensive Computing

Grid Computing

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

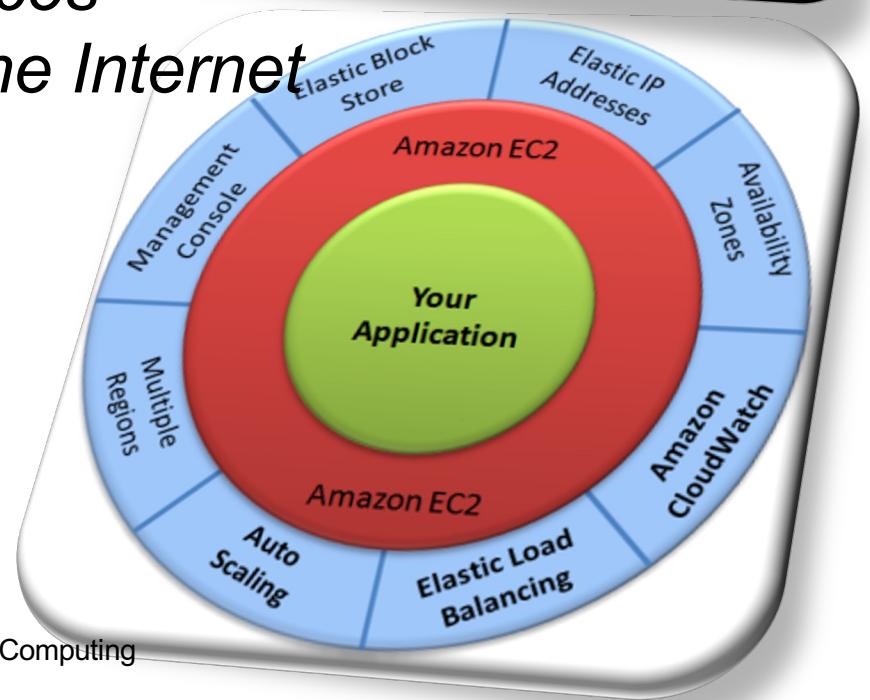


Cloud Computing

- 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



Reading Material

- Maarten van Steen, Andrew S. Tanenbaum. [A brief introduction to distributed systems.](#)
- Ian Foster, Yong Zhao, Ioan Raicu, Shiyong Lu. "[Cloud Computing and Grid Computing 360-Degree Compared](#)", IEEE Grid Computing Environments (GCE08) 2008, co-located with IEEE/ACM Supercomputing 2008.
- Satoshi Nakamoto. [Bitcoin: A Peer-to-Peer Electronic Cash System](#)

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

Key Characteristics of Distributed Systems

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

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

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

