Computer Science 425
Distributed Systems

CS 425 / CSE 424 / ECE 428

Fall 2012

Indranil Gupta (Indy)

August 30, 2012

Lecture 2

What is Cloud Computing?

© 2012, I. Gupta

What are Clouds?

- **Oracle has a Cloud Computing Center.**
- **And yet…**
- **Larry Ellison – “Clouds are Water Vapor”**

The Hype!

- Gartner in 2009 - Cloud computing revenue will soar faster than expected and will exceed $150 billion by 2013. It will represent 19% of IT spending by 2015.
- IDC in 2009: "Spending on IT cloud services will triple in the next 5 years, reaching $42 billion."
- Forrester in 2010 – Cloud computing will go from $40.7 billion in 2010 to $241 billion in 2020.
- Companies and even Federal/state governments using cloud computing now: fedbizopps.gov

Sources: http://www.infosysblogs.com/cloudcomputing/2009/08/the_cloud_computing_quotes.htm and http://www.mytestbox.com

$$$

- Dave Power, Associate Information Consultant at Eli Lilly and Company**: "With AWS, Powers said, a new server can be up and running in three minutes (it used to take Eli Lilly seven and a half weeks to deploy a server internally) and a 64-node Linux cluster can be online in five minutes (compared with three months internally). … It's just shy of instantaneous."**
- Ingo Elfering, Vice President of Information Technology Strategy, GlaxoSmithKline**: "With Online Services, we are able to reduce our IT operational costs by roughly 30% of what we're spending"**
- Jim Swartz, CIO, Sybase**: "At Sybase, a private cloud of virtual servers inside its data centre has saved nearly $US2 million annually since 2006, Swartz says, because the company can share computing power and storage resources across servers."**
- **100s of startups in Silicon Valley can harness large computing resources without buying their own machines**

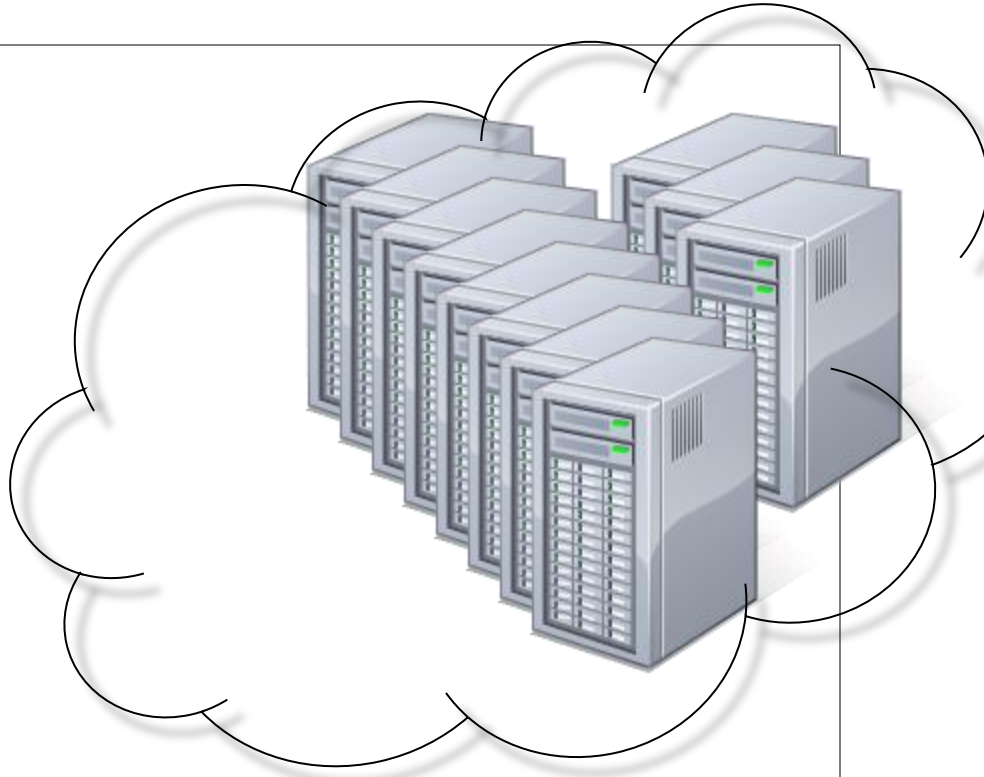
Lecture 2-4

Sources: http://www.infosysblogs.com/cloudcomputing/2009/08/the_cloud_computing_quotes.htm and http://www.mytestbox.com

What is a Cloud?

- **It's a cluster!**
- **It's a supercomputer!**
- **It's a datastore!**
- **It's superman!**

- **None of the above**
- **All of the above**



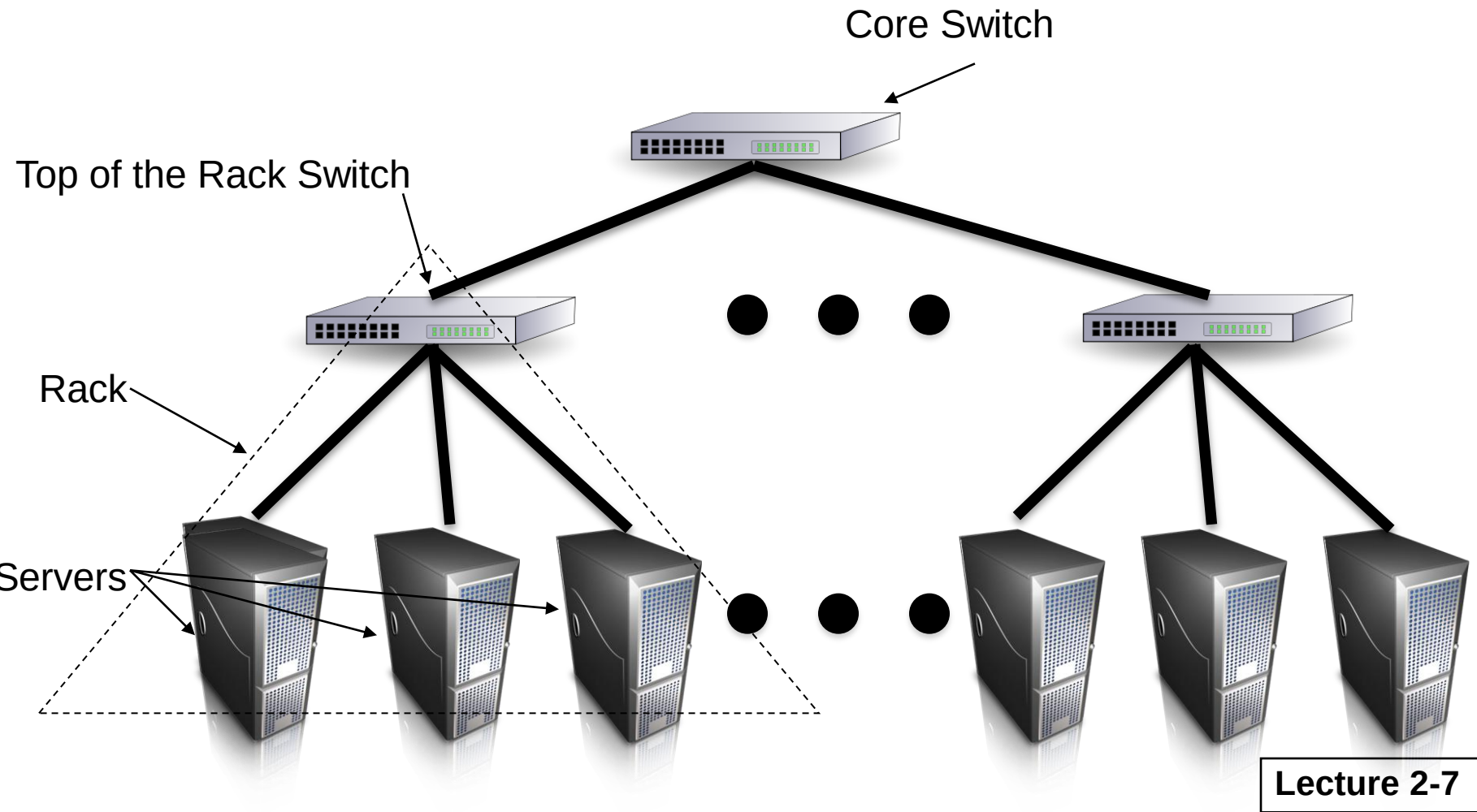
- Cloud = Lots of storage + compute cycles nearby

What is a Cloud?

- A single-site cloud (aka "Datacenter") consists of
  - Compute nodes (grouped into ***racks***)
  - Switches, connecting the racks
  - A network topology, e.g., hierarchical
  - Storage (backend) nodes connected to the network
  - Front-end for submitting jobs
  - Software Services
- A geographically distributed cloud consists of
  - Multiple such sites
  - Each site perhaps with a different structure and services

A Sample Cloud Topology

So then, what is a cluster?



What(’s new) in Today’s Clouds?

Four major features:

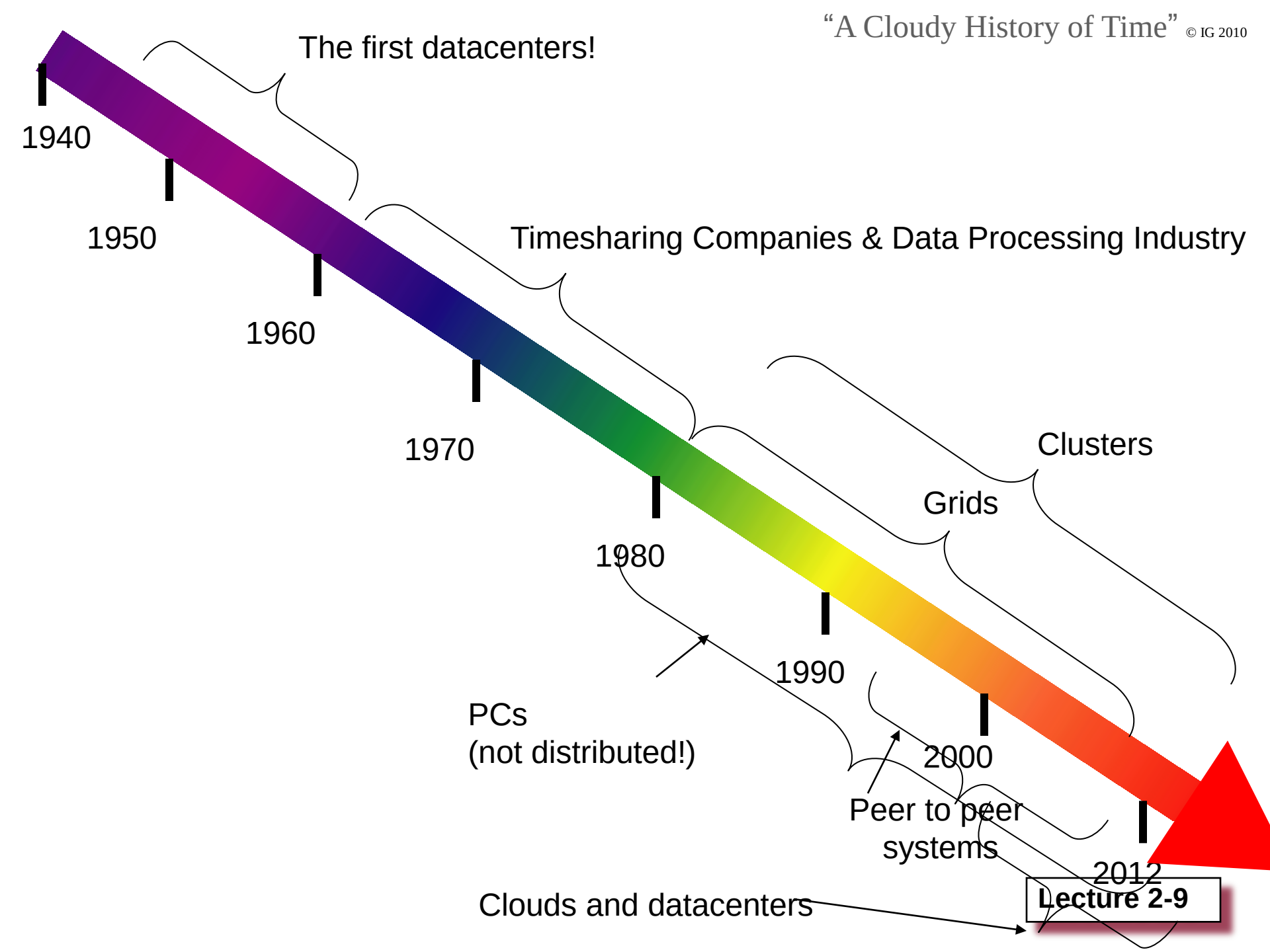
I. **Massive scale.**

II. **On-demand access: Pay-as-you-go, no upfront commitment.**
- Anyone can access it

III. **Data-intensive Nature: What was MBs has now become TBs, PBs and XBs.**
- Daily logs, forensics, Web data, etc.
- Do you know the size of Wikipedia dump?

IV. **New Cloud Programming Paradigms: MapReduce/Hadoop, NoSQL/Cassandra/MongoDB and many others.**
- High in accessibility and ease of programmability
- Lots of open-source

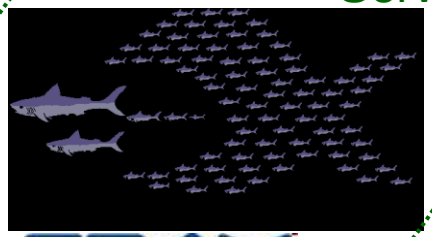
Combination of one or more of these gives rise to novel and unsolved distributed computing problems in cloud computing.


"A Cloudy History of Time" © IG 2010
The first datacenters!
1940
1950
1960
1970
1980
1990
2000
2012
Timesharing Companies & Data Processing Industry
Clusters
Grids
PCs
(not distributed!)
Peer to peer
systems
Clouds and datacenters
Lecture 2-9

First large datacenters: ENIAC, ORDVAC, ILLIAC
Many used vacuum tubes and mechanical relays

IBM

Berkeley NOW Project
Supercomputers
Server Farms (e.g., Oceano)



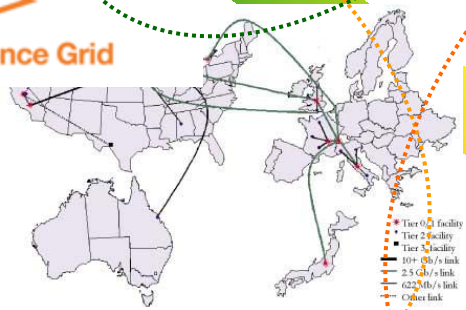
Honeywell

xerox

CRAY
THE SUPERCOMPUTER COMPANY

P2P Systems (90s-00s)
•Many Millions of users
•Many GB per day

Open Science Grid



Lime wire

BitTorrent

KaZaA

Data Processing Industry
- 1968: $70 M. 1978: $3.15 Billion.

Timesharing Industry (1975):
•Market Share: Honeywell 34%, IBM 15%,
•Xerox 10%, CDC 10%, DEC 10%, UNIVAC 10%
•Honeywell 6000 & 635, IBM 370/168,
Xerox 940 & Sigma 9, DEC PDP-10, UNIVAC 1108

Grids (1980s-2000s):
•GriPhyN (1970s-80s)
•Open Science Grid and Lambda Rail (2000s)
•Globus & other standards (1990s-2000s)

Clouds

Trends: Technology

- **Doubling Periods – storage: 12 mos, bandwidth: 9 mos, and (what law is this?) cpu compute capacity: 18 mos**
- **Then and Now**

  Bandwidth

  - **1985: mostly 56Kbps links nationwide**
  - **2012: Tbps links widespread**

  Disk capacity

  - **Today's PCs have TBs, far more than a 1990 supercomputer**

Trends: Users

- **Then and Now**

  Biologists:

  - **1990: were running small single-molecule simulations**
  - **2012: CERN's Large Hadron Collider producing many PB/year**

Prophecies

In 1965, MIT's Fernando Corbató and the other designers of the Multics operating system envisioned a computer facility operating "like a power company or water company".

Plug your thin client into the computing Utility

and Play your favorite Intensive Compute & Communicate Application

- **[Have today's clouds brought us closer to this reality?]**

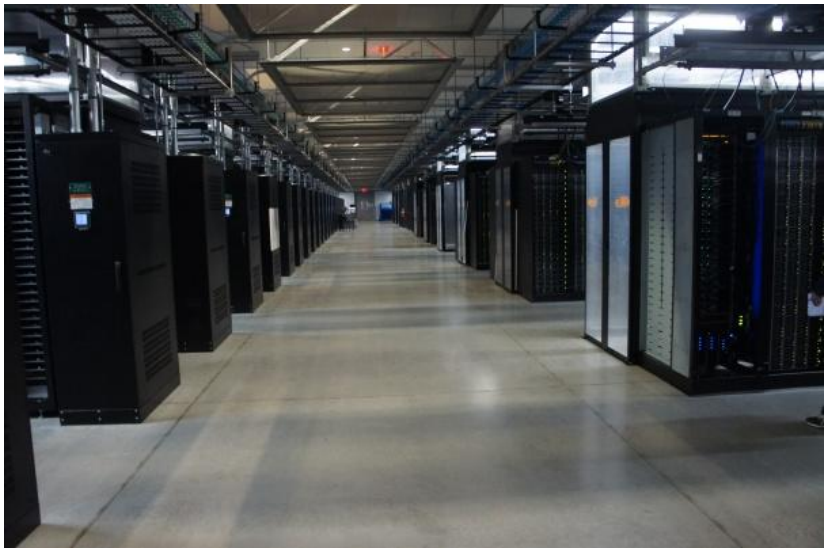
I. Massive Scale

- **Facebook [GigaOm, 2012]**
  - **30K in 2009 -> 60K in 2010 -> 180K in 2012**
- **Microsoft [NYTimes, 2008]**
  - **150Kmachines**
  - **Growth rate of 10K per month**
  - **80K total running Bing**
  - **300K in Chicago DC**
- **Yahoo! [2009]:**
  - **100K**
  - **Split into clusters of 4000**
- **AWS EC2 [Randy Bias, 2009]**
  - **40,000 machines**
  - **8 cores/machine**
- **eBay [2012]: 50K machines**
- **HP [2012]: 380K in 180 DCs**
- **Google: A lot**

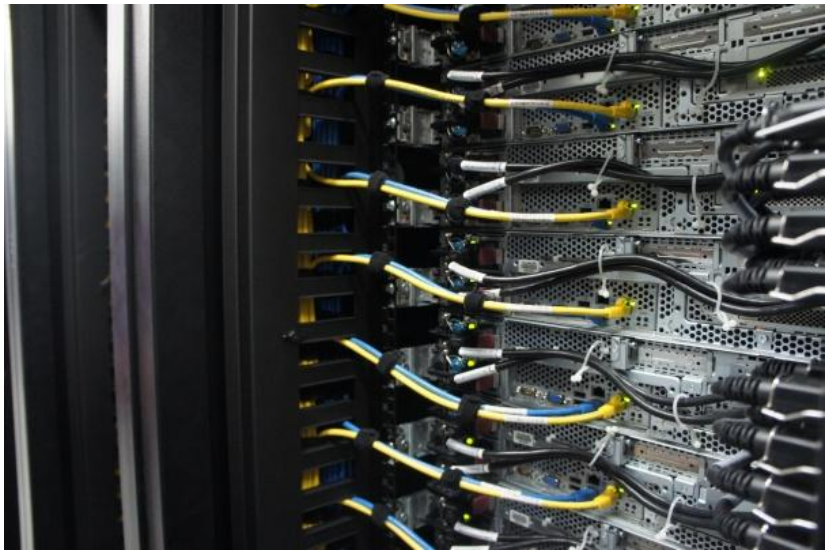
What does a datacenter look like from inside?

- **A virtual walk through Facebook's Datacenter in Prineville, Oregon (Facebook OpenCompute)**
- **Source: Gigaom article from 2012 - http://gigaom.com/cleantech/a-rare-look-inside-facebooks-oregon-data-center-photos-video/**

Servers

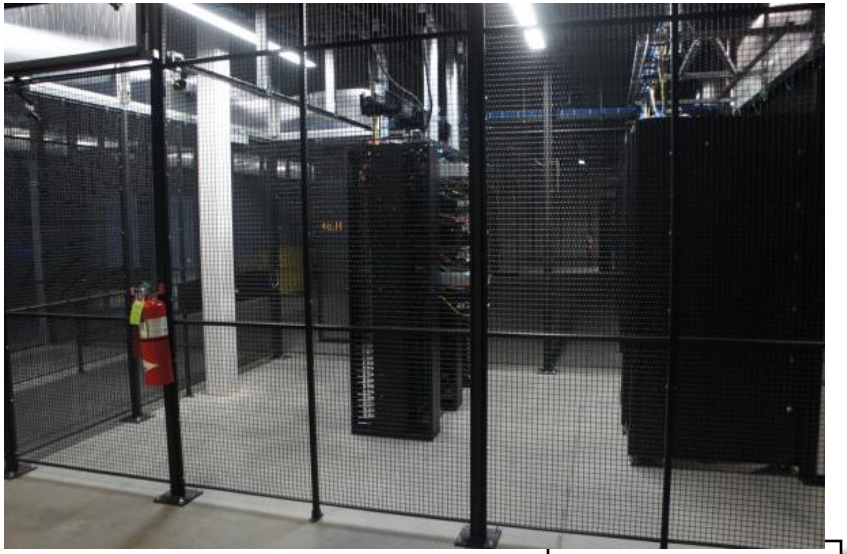


Front



Back

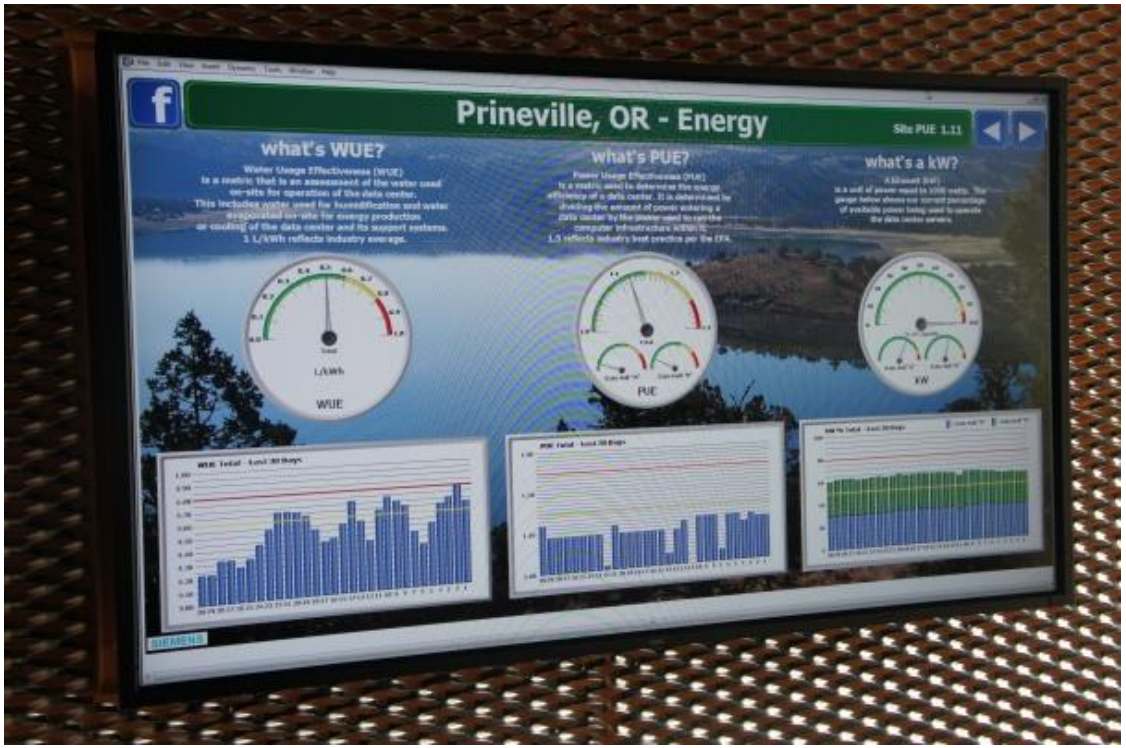
In



Some highly secure (e.g., financial info)

Power

Off-site



On-site

- WUE = Annual Water Usage / IT Equipment Energy (L/kWh)
- PUE = IT Equipment Power / Total facility Power

Cooling

Air sucked in from top

Water purified

Water sprayed into air

15 motors per server bank

Extra - Fun Videos to Watch

- **Microsoft GFS Datacenter Tour (Youtube)**
- **Timelapse of a Datacenter Construction on the Inside (Fortune 500 company)**

II. On-demand access: *aaS

Classification

On-demand: renting a cab vs (previously) renting a car, or buying one. E.g.:

- **AWS Elastic Compute Cloud (EC2): $0.080-$3.58 per CPU hour**
- **AWS Simple Storage Service (S3): $0.05-$0.12 per GB-month**

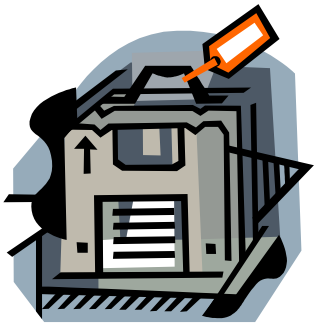
- **HaaS: Hardware as a Service**
  - **You get access to barebones hardware machines, do whatever you want with them, Ex: Your own cluster**
  - **Not always a good idea (why?)**
- **IaaS: Infrastructure as a Service**
  - **You get access to flexible computing and storage infrastructure. Virtualization is one way of achieving this (what's another way, e.g., using Linux). Often said to subsume HaaS.**
  - **Ex: Amazon Web Services (AWS: EC2 and S3), Eucalyptus, Rightscale, Microsoft Azure.**

I. On-demand access: *aaS

Classification

- **PaaS: Platform as a Service**
  - **You get access to flexible computing and storage infrastructure, coupled with a software platform (often tightly)**
  - **Ex: Google's AppEngine (Python, Java, Go)**
- **SaaS: Software as a Service**
  - **You get access to software services, when you need them. Often said to subsume SOA (Service Oriented Architectures).**
  - **Ex: Google docs, MS Office on demand**

III. Data-intensive Computing



- **Computation-Intensive Computing**
  - Example areas: MPI-based, High-performance computing, Grids
  - Typically run on supercomputers (e.g., NCSA Blue Waters)
- **Data-Intensive**
  - Typically store data at datacenters
  - Use compute nodes nearby
  - Compute nodes run computation services
- **In data-intensive computing, the focus shifts from computation to the data: CPU utilization no longer the most important resource metric**

IV. New Cloud Programming Paradigms

Easy to write and run highly parallel programs in new cloud programming paradigms:

- **Google: MapReduce and Sawzall**
- **Amazon: Elastic MapReduce service (pay-as-you-go)**
- **Google (MapReduce)**
  - **Indexing: a chain of 24 MapReduce jobs**
  - **~200K jobs processing 50PB/month (in 2006)**
- **Yahoo! (Hadoop + Pig)**
  - **WebMap: a chain of 100 MapReduce jobs**
  - **280 TB of data, 2500 nodes, 73 hours**
- **Facebook (Hadoop + Hive)**
  - **~300TB total, adding 2TB/day (in 2008)**
  - **3K jobs processing 55TB/day**
- **Similar numbers from other companies, e.g., Yieldex, eharmony.com, etc.**
- **NoSQL: MySQL is an industry standard, but Cassandra is 2400 times faster!**

Two Categories of Clouds

- **Industrial Clouds**
  - Can be either a (i) public cloud, or (ii) private cloud
  - Private clouds are accessible only to company employees
  - Public clouds provide service to any paying customer:
    - » Amazon S3 (Simple Storage Service): store arbitrary datasets, pay per GB-month stored
    - » Amazon EC2 (Elastic Compute Cloud): upload and run arbitrary images, pay per CPU hour used
    - » Google AppEngine: develop applications within their appengine framework, upload data that will be imported into their format, and run
- **Academic Clouds**
  - Allow researchers to innovate, deploy, and experiment
  - Google-IBM Cloud (U. Washington): run apps programmed atop Hadoop
  - Cloud Computing Testbed (CCT @ UIUC): first cloud testbed to support systems research. Runs: (i) apps programmed atop Hadoop and Pig, (ii) systems-level research on this first generation of cloud computing models (~HaaS), and (iii) (coming soon) OpenStack (~AWS EC2). http://cloud.cs.illinois.edu
    - » On the 4th floor of Siebel Center (if you care to look)
  - OpenCirrus: first federated cloud testbed. http://opencirrus.org

Single site Cloud: to Outsource or Own?

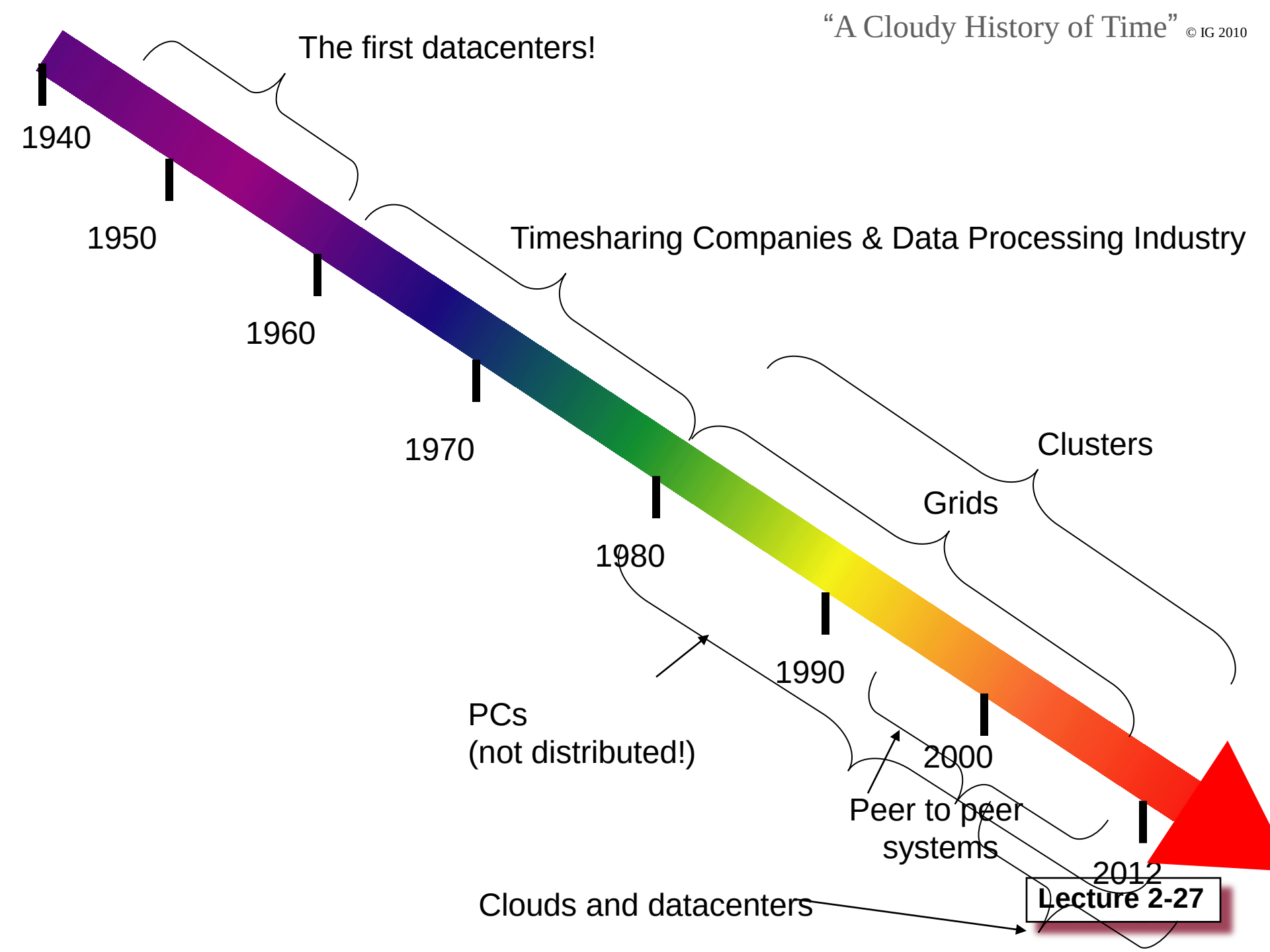
- **Medium-sized organization: wishes to run a service for *M* months**
  - **Service requires 128 servers (1024 cores) and 524 TB**
  - **Same as UIUC CCT cloud site**
- **Outsource (e.g., via AWS): monthly cost**
  - **S3 costs: $0.12 per GB month. EC2 costs: $0.10 per Cpu hour**
  - **Storage = $ 0.12 X 524 X 1000 ~ $62 K**
  - **Total = Storage + CPUs = $62 K + $0.10 X 1024 X 24 X 30 ~ $136 K**
- **Own: monthly cost**
  - **Storage ~ $349 K / *M***
  - **Total ~ $ 1555 K / *M* + 7.5 K (includes 1 sysadmin / 100 nodes)**
    - » **using 0.45:0.4:0.15 split for hardware:power:network and 3 year lifetime of hardware**

Single site Cloud: to Outsource or Own?

- **Breakeven analysis:** more preferable to own if**:**
  - **$349 K / *M* < $62 K (storage)**
  - **$ 1555 K / *M* + 7.5 K < $136 K (overall)**

  Breakeven points

  - ***M* > 5.55 months (storage)**
  - ***M* > 12 months (overall)**
    - **Startups use clouds a lot**
    - **Cloud providers benefit monetarily most from storage**


The first datacenters!
1940
1950
Timesharing Companies & Data Processing Industry
1960
1970
Clusters
Grids
1980
1990
PCs
(not distributed!)
2000
Peer to peer
systems
2012
Clouds and datacenters

But were there clouds before this?

- **Yes!**

emulab
total network testbed

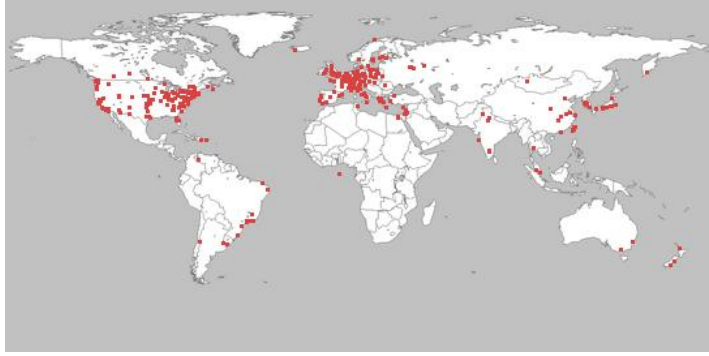
- A community resource open to researchers in academia and industry
- https://www.emulab.net/
- A cluster, with currently ~500 servers
- Founded and owned by University of Utah (led by Late Prof. Jay Lepreau)

- As a user, you can:
  - Grab a set of machines for your experiment
  - You get root-level (sudo) access to these machines
  - You can specify a network topology for your cluster
  - You can emulate any topology

All images © Emulab

PLANETLAB

An open platform for developing, deploying, and accessing planetary-scale services



- A community resource open to researchers in academia and industry
- http://www.planet-lab.org/
- Currently, ~ 1077 nodes at ~500 sites across the world
- Founded at Princeton University (led by Prof. Larry Peterson), but owned in a federated manner by the sites

- Node: Dedicated server that runs components of PlanetLab services.
- Site: A location, e.g., UIUC, that hosts a number of nodes.
- Sliver: Virtual division of each node. Currently, uses VMs, but it could also other technology. Needed for timesharing across users.
- Slice: A spatial cut-up of the PL nodes. Per user. A slice is a way of giving each user (Unix-shell like) access to a subset of PL machines, selected by the user. A slice consists of multiple slivers, one at each component node.

- Thus, PlanetLab allows you to run real world-wide experiments.
- Many services have been deployed atop it, used by millions (not just researchers): Application-level DNS services, Monitoring services, CoralCDN, etc.

All images © PlanetLab

Next Week

- **Tuesday**
  - **More cloud computing: MapReduce! (and other cool stuff)**
  - **MP1 and HW1 released**
- **Thursday**
  - **Failure detection**
  - **Readings: Section 15.1, parts of Section 2.4.2**