中山大學
SUN YAT-SEN UNIVERSITY

Lecture 1
Introduction to Cloud Computing

Cloud Application Development
(SE808, School of Software, Sun Yat-Sen University)

Yabo (Arber) Xu

Outline

- What is cloud computing?

- How it evolves?

- What are the underlying key technologies?
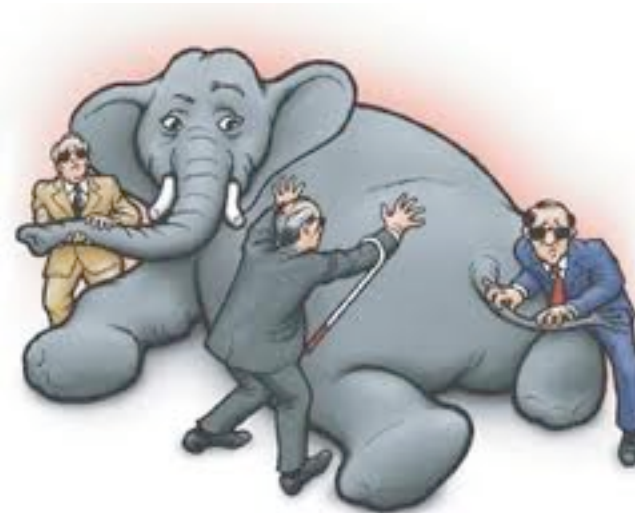
Question #1

What is Cloud Computing?

What is Cloud Computing?

5 people might have 8 answers ! ☹

Widely distributed,
network based,
storage,
computation,
……
utility computing,
HaaS,
PaaS,
SaaS.

A Customer-Oriented Definition

Anytime,

Anywhere,

With any device,

Accessing any services

How many of them you still store on your local computer?

- Email
- Calendars and contacts
- Photo/video sharing
- Document sharing, or
- Anything?

A Business-Oriented Definition

- Key Characteristic
  - Universal Access
  - Scalable Services
    - Infrastructure managing the scaling, not applications
    - Elasticity: Expenses only incurred when they are needed
  - New Application Service Models
    - XaaS = X as a Service
    - Pay-as-you-go

Public Cloud #1: Amazon

Amazon EC2

- Elastic Cloud Computing
- virtual servers for rent, called Amazon Machine Images (AMIs)
- based on Xen
- priced on per hour <u>from $0.085 to $1</u>

Amazon S3

- Simple Storage Service
- up to $0.18 per GB storage
- from $0.10 per GB transfer
- via
  - REST
  - SOAP
 - BitTorrent

Public Cloud #2: Google

- A web application development framework and hosting solution rolled into one

- That uses the infrastructure available at Google
  - so their servers + storage: BigTable

- Charges
  - 500MB of storage
  - up to 5 million page views a month
  - 10 applications per developer account
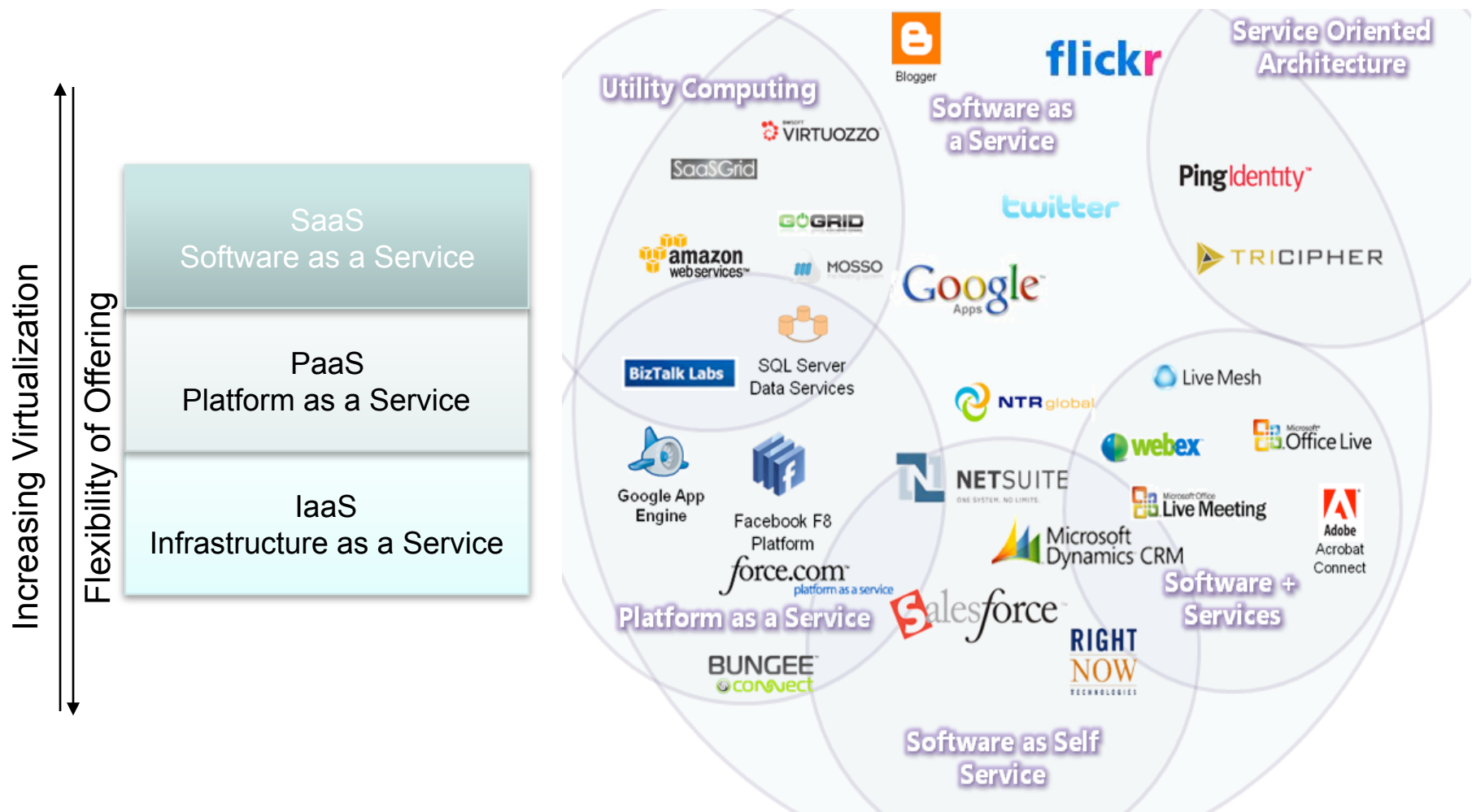 - pay for an extension

- Python/JAVA and GAE SDK

Public cloud #3: Microsoft Azure Services

- Released in Feb, 2010
- A cloud service operating system that supports the service development/ hosting/management environment
    - o .Net Services
    - o SQL Services
    - o Live Services

- Pricing
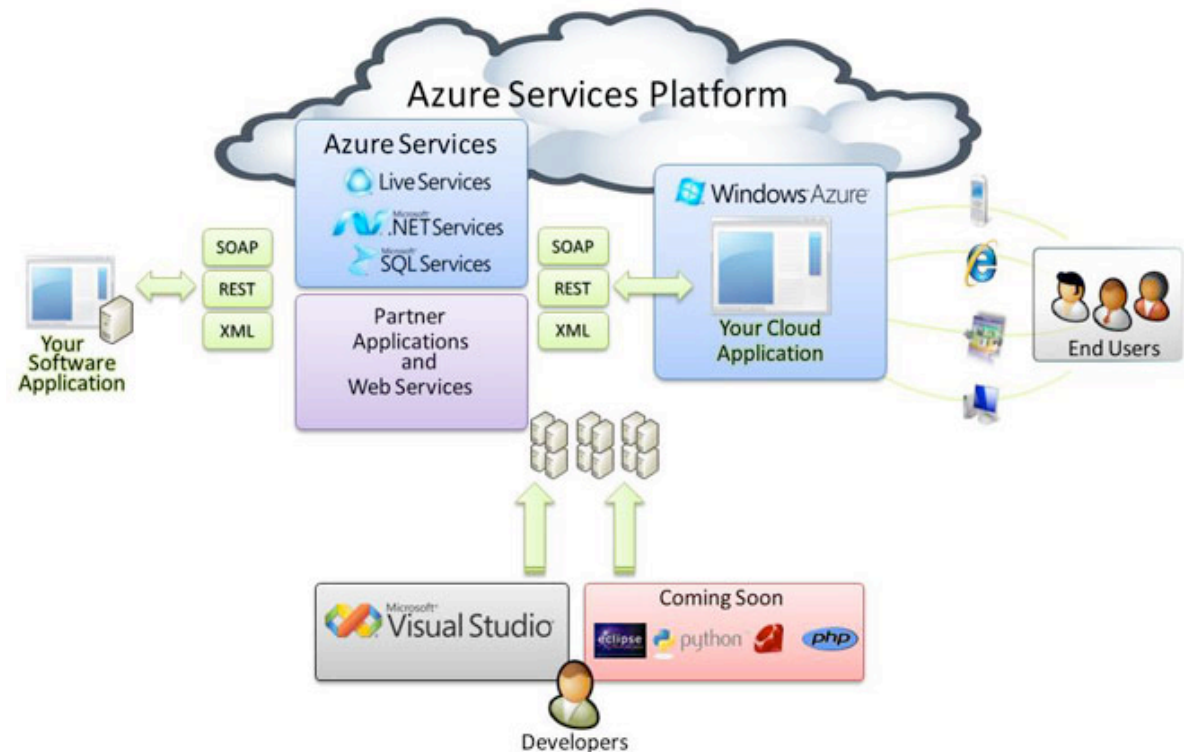
Cloud Services Taxonomy

Increasing Virtualization

Flexibility of Offering

| SaaS
Software as a Service |
| PaaS
Platform as a Service |
| IaaS
Infrastructure as a Service |

Everything as a Service

- AaaS   Architecture as a Service
- BaaS   Business as a Service
- CaaS   Computing as a Service
- DaaS   Data as a Service
- DBaaS   Database as a Service
- EaaS   Ethernet as a Service
- FaaS   Frameworks as a Service
- GaaS   Globalization or Governance as a Service
- HaaS   Hardware as a Service
- IMaaS   Information as a Service

- IaaS   Infrastructure as a Service

- IDaaS   Identity as a Service
- LaaS   Lending as a Service
- MaaS   Mashups as a Service
- OaaS   Organization or Operations as a Service

- SaaS   Software as a Service

- PaaS   Platform as a Service

- TaaS   Technology or Testing as a Service
- VaaS   Voice as a Service

How it evolves?

1. Web-scale problems

2. Large data centers

3. Different models of computing

Web-scale Problem

- Characteristics:
  - Definitely data-intensive
  - May also be processing intensive

- Examples:
  - Crawling, indexing, searching, mining the Web
  - "Post-genomics" life sciences research
  - Other scientific data (physics, astronomers, etc.)
  - Sensor networks
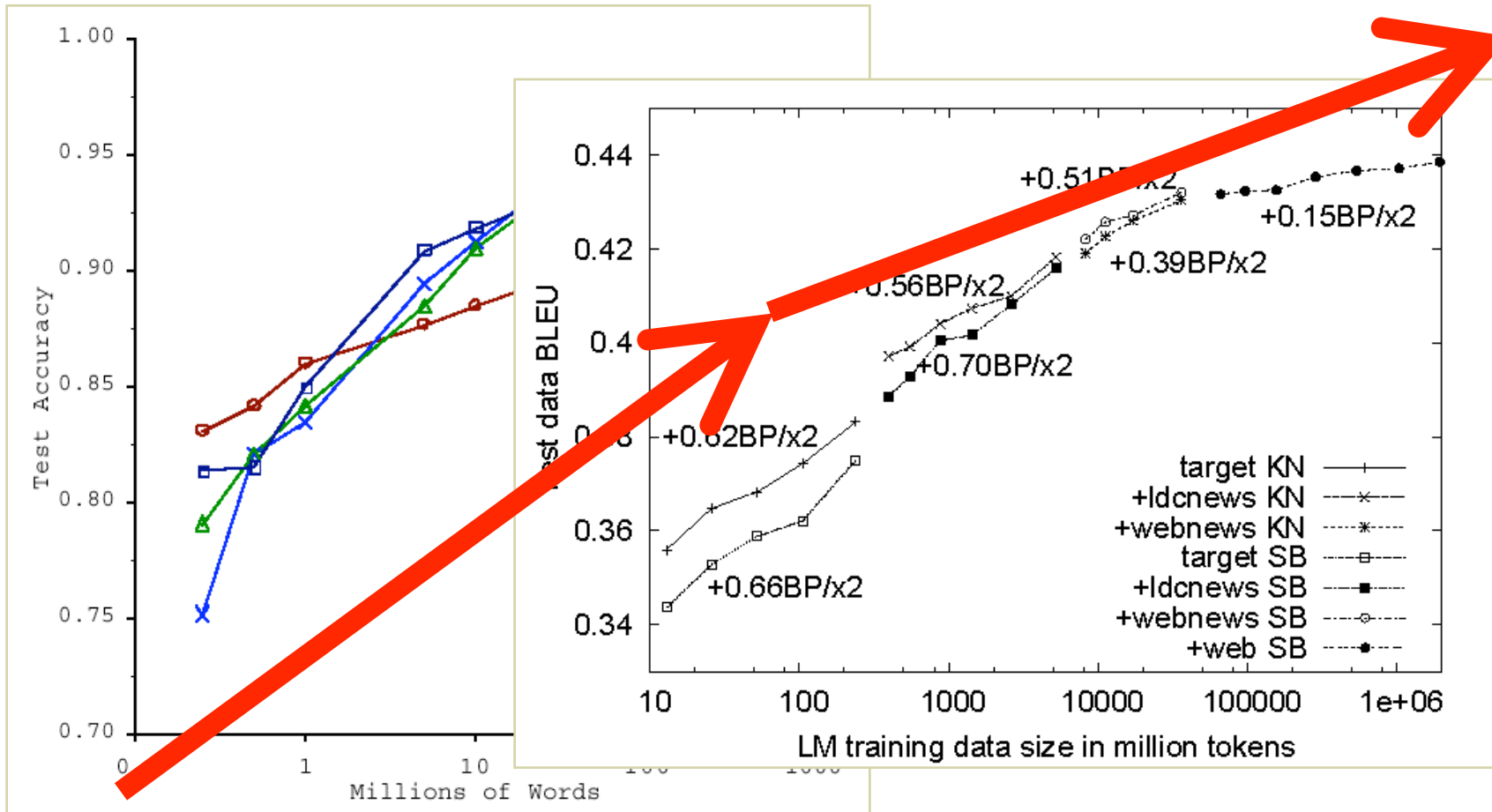 - Web 2.0 applications
  - …

How much data?

- Wayback Machine has 2 PB + 20 TB/month (2006)
- Google processes 20 PB a day (2008)
- "all words ever spoken by human beings" ~ 5 EB(1K PB)
- NOAA has ~1 PB climate data (2007)
- CERN's LHC will generate 15 PB a year (2008)

640K ought to be enough for anybody.

Data → Inspiration

(Banko and Brill, ACL 2001)
(Brants et al., EMNLP 2007)
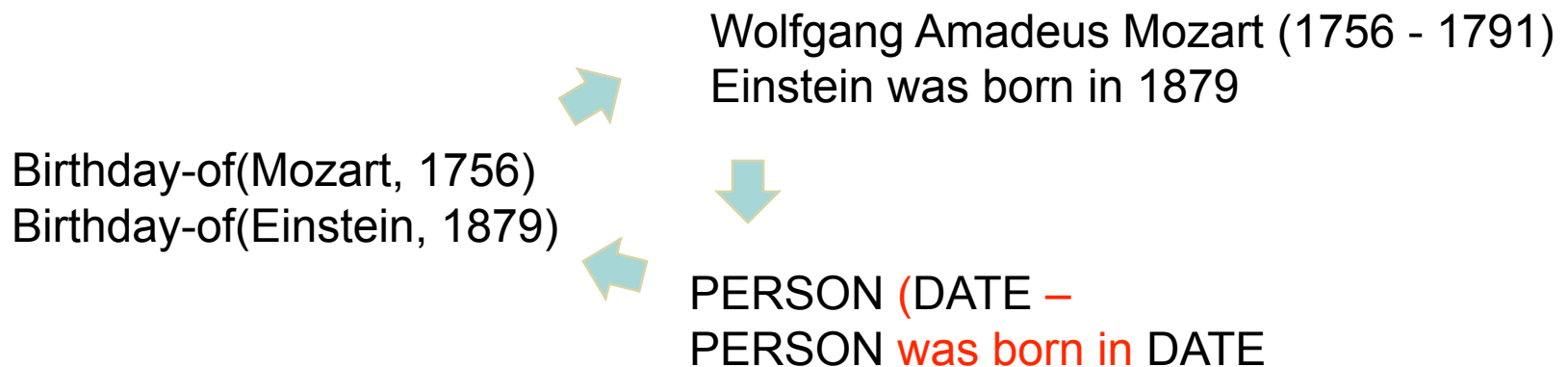
What to do with more data?

- Answering factoid questions
  - Pattern matching on the Web
  - Works amazingly well

    Who shot Abraham Lincoln? → X shot Abraham Lincoln

- Learning relations
  - Start with seed instances
  - Search for patterns on the Web
  - Using patterns to find more instances

Wolfgang Amadeus Mozart (1756 - 1791)
Einstein was born in 1879

Birthday-of(Mozart, 1756)
Birthday-of(Einstein, 1879)

PERSON (DATE –
PERSON was born in DATE

(Brill et al., TREC 2001; Lin, ACM TOIS 2007)
(Agichtein and Gravano, DL 2000; Ravichandran and Hovy, ACL 2002; …)

Large Data Centers

- Web-scale problems? Throw more machines at it!

- Clear trend: centralization of computing resources in large data centers
  - Necessary ingredients: commodity, network, juice, and space
  - Analogy to the power station

- Important Issues:
  - Redundancy
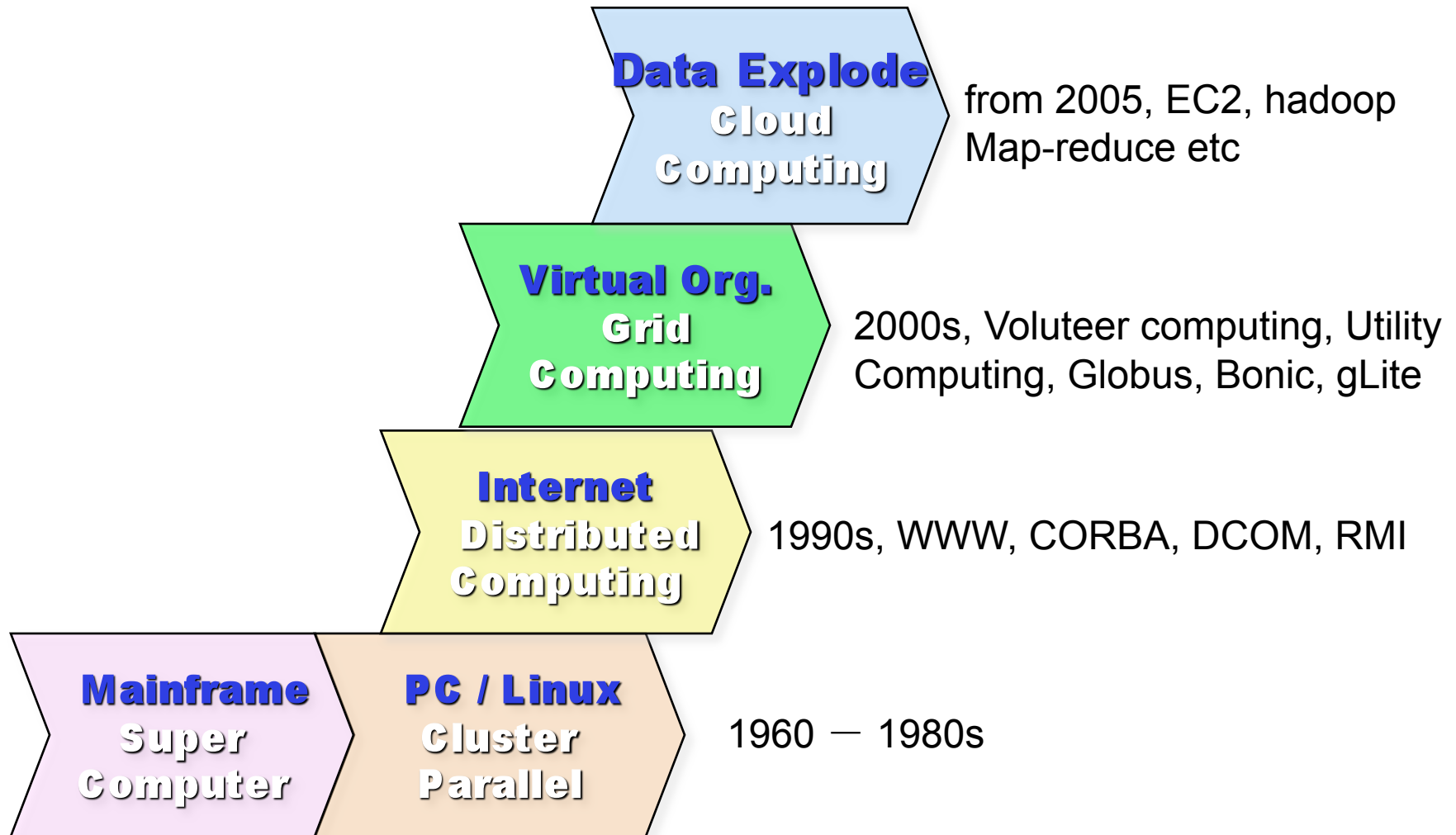 - Efficiency
  - Utilization
  - Management

A Sample Cloud Topology

Core Switch

Top of the Rack Switch

Rack

Servers
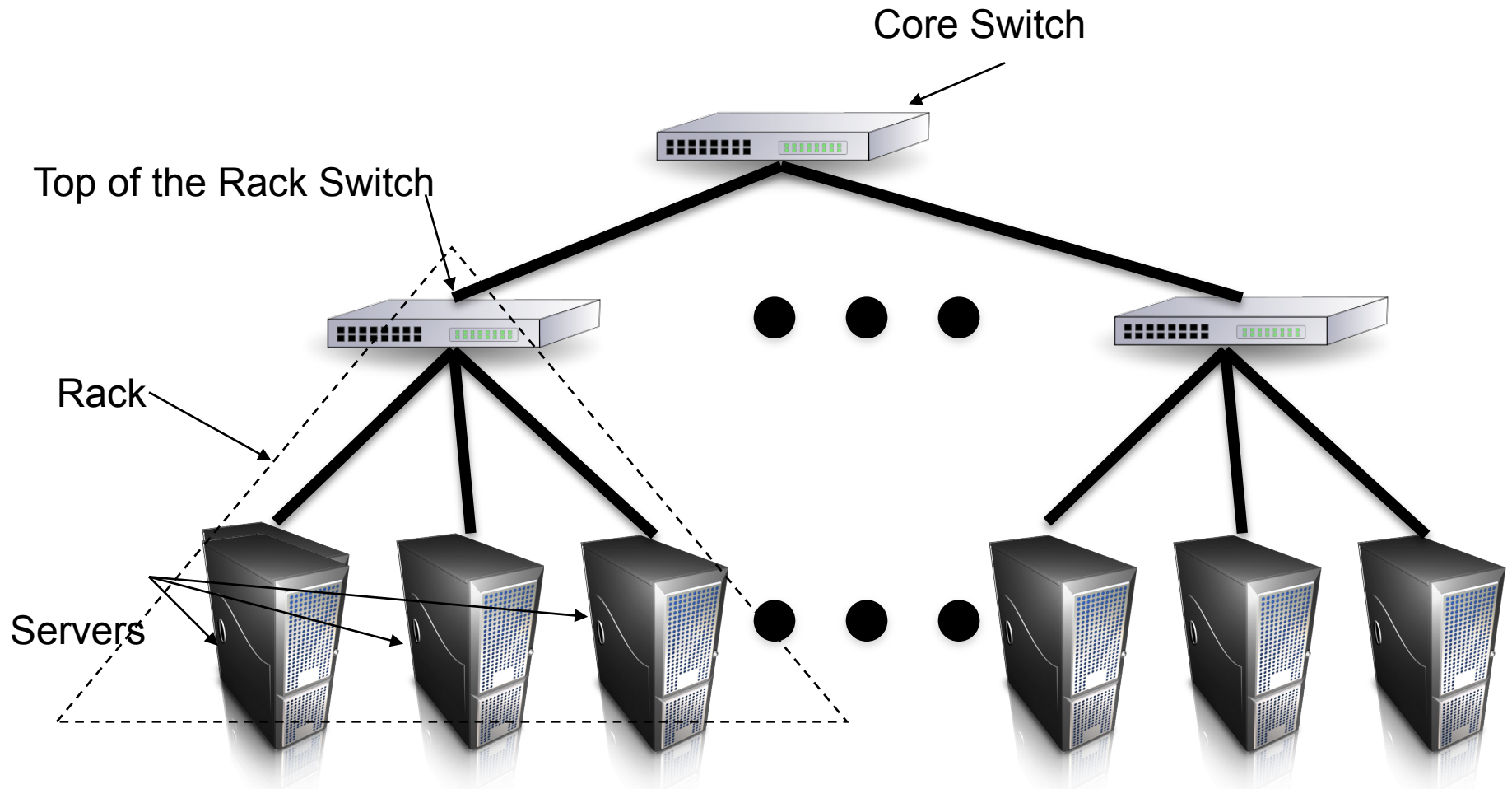
Scale of Industry Datacenters

- Microsoft [NYTimes, 2008]
  - 150,000 machines
  - Growth rate of 10,000 per month
  - Largest datacenter: 48,000 machines
  - 80,000 total running Bing

- Yahoo! [Hadoop Summit, 2009]
  - 25,000 machines
  - Split into clusters of 4000

- AWS EC2 (Oct 2009)
  - 40,000 machines
  - 8 cores/machine

- Google
  - (Rumored) several hundreds of thousands of machines

Different Models of Computing

Data Explode
Cloud Computing

from 2005, EC2, hadoop Map-reduce etc

Virtual Org.
Grid Computing

2000s, Voluteer computing, Utility Computing, Globus, Bonic, gLite

Internet
Distributed Computing

1990s, WWW, CORBA, DCOM, RMI

Mainframe
Super Computer

PC / Linux
Cluster Parallel

1960 － 1980s

Grid Computing Vs Cloud Computing

	Grid	Cloud
Underlying concept	Utility Computing	Utility Computing
Main benefit	Solve computationally complex problems	Provide a scalable standard environment for network-centric application development, testing and deployment
Resource distribution / allocation	Negotiate and manage resource sharing; schedulers	Simple user <-> provider model; pay-per-use
Domains	Multiple domains	Single domain
Character / history	Non-commercial, publicly funded	Commercial

Reference:http://www.slideshare.net/DSPIP/cloud-computing-introduction-2978287

4/17/11

What('s new) in Today's Clouds?

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

Data-intensive Nature: What was MBs has now become TBs.
- Daily logs, forensics, Web data, etc.

New Cloud Programming Paradigms: MapReduce/Hadoop, Pig Latin, DryadLinq, Swift, and many others.
- High in accessibility and ease of programmability

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

Outline

- What is cloud computing?

- Who is in this game?

- How it evolves?

- **What are the underlying key technologies?**

Service Model and Key Technologies

○ Infrastructure as a Service (IaaS)

- Why buy machines when you can rent cycles?
- Key Technology: Virtualization

○ Platform as a Service (PaaS)

- Give me nice API and take care of the implementation
- Key Technology: New cloud programming paradigm, i.e. MapReduce, PIG, HIVE etc

○ Software as a Service (SaaS)

- Just run it for me!
- Key Technology: Everybody has their own secret sauces, but Ajax is de-facto front-end.

Virtualization at a Glance
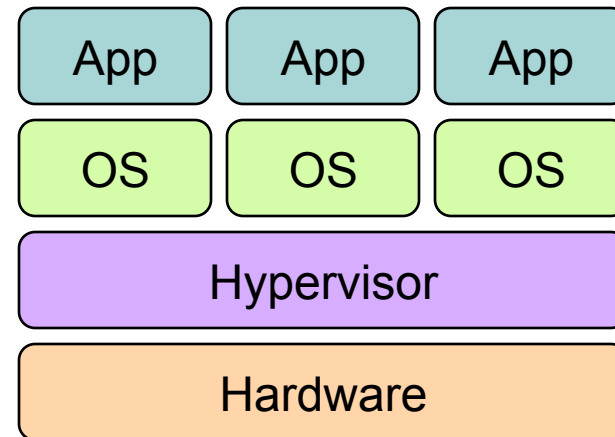
Run multiple virtual computers on one physical box

○ Lots of way to do it

  ○ Xen

  ○ VMWare

  ○ Parallels

  ○ Amazon AMI

  ○ Microsoft Hype V

Virtualization – Benefit

App	App	App

Operating System

Hardware
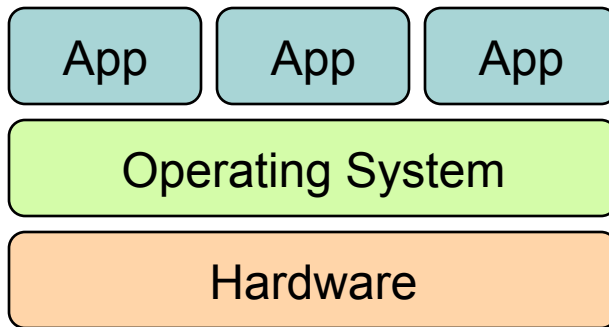
Traditional Stack

5 to 15 % utilization only

App	App	App

OS	OS	OS

Hypervisor

Hardware

Virtualized Stack

High utilization and standardization

New Cloud Programming Paradigms

○ A typical cloud-era task: Given 100 computers, how do you compute the frequency of words in 1T text files?

To utilize the underlying computing power, you need

a new paradigm for

storing and processing

large scale of data

New Cloud Programming Paradigms

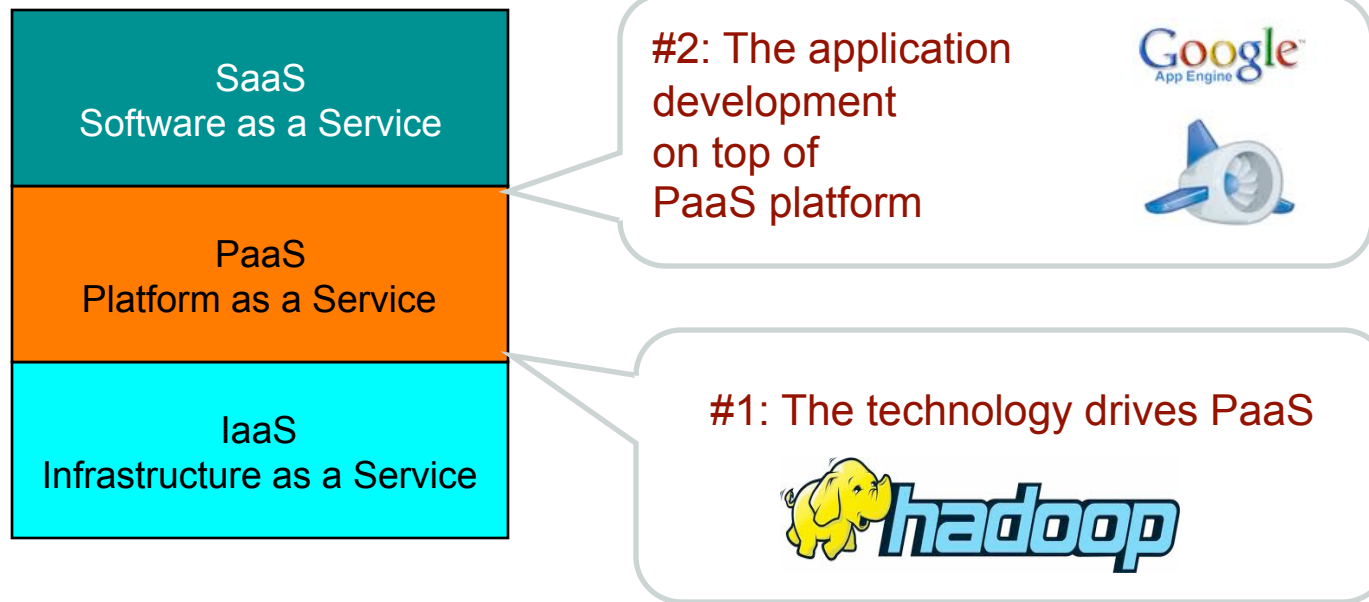|	**Google**	**Hadoop (Open Source)**	**Microsoft**
Dist. File System	GFS	HDFS	
No-SQL DB	BigTable	HBase	
Programming Framework	MapReduce	Hadoop MapReduce	Dryad (Generalized MR)
High-level Language	Sawzall	PIG (Yahoo) / Hive(FaceBook)	DryadLING

A Brief History in this Area

○ 2003, First MapReduce Lib developed in Google

○ 2003, 2004, and 2006, Google published papers on GFS/MapReduce/BigTable.

• 2005- Now, Hadoop project (open source version of GFS/BigTable/MapReduce), initiated by Doug Cutting, sponrsed by Yahoo

• 2008/2009, Yahoo/Facebook contributed PIG/Hive on top of hadoop.

Cloud Computing Usages

- Google (MapReduce)
  - Indexing: a chain of 24 MapReduce jobs
  - 450K nodes, ~200K jobs processing 50PB/month (in 2006)
- Yahoo! (Hadoop + Pig)
  - WebMap: a chain of 100 MapReduce jobs
  - 2500 nodes, 280 TB of data,
- Facebook (Hadoop + Hive)
  - 2250 nodes, adding 80-90TB/day (in 2010)
  - 25K jobs/day

- Taobao (Hadoop + TFS + Hbase)
  - 1300 nodes, 9.3PB (2010)
  - 1800 hadoop jobs per day
- Baidu
  - Their own implementation of hadoop in C++
  - 4000 nodes (2010)

Guess what is the main-stream configuration for each node?

What We Cover in this Course?

SaaS Software as a Service
PaaS Platform as a Service
IaaS Infrastructure as a Service

#2: The application development on top of PaaS platform

Google App Engine

#1: The technology drives PaaS

hadoop

Skills you would learn...

Only if you put enough efforts on it .

- Skill 1st: Know how to process Terabytes of data.
  - A basic skill for anyone who stays in IT industry in this digital-era.

- Skill 2nd:Know how to put up the platform if you are given the chance & resources.
  - Critical for anyone who want to become an excellent engineering in a big corporation

- Skill 3rd:Know how to quickly implement your ideas on top of cloud service, and make it big.
  - Critical for anyone who want to run your own startups and dreaming to be a billionaire. ☺

Thank you &
Welcome to Cloud Era!

SUN YAT-SEN UNIVERSITY