Z6110X0035: Cloud Computing

Lecturer: Prof. Zichen Xu

Self-Introduction

- Dr. Zichen Xu（徐子晨）
- Professor in NCU
- Former Googler, Buckeye, Gator, and Bull

- Main field:
  - Green computing, Data storage,
  - Database management systems,
  - Data analysis, and AI.

Information About the Course (Cont.)

- We have a regular office hour
  - Open office time: Tuesday 14:00-15:00, Location: IEB A608-1
  - We can have other meeting time if you email to schedule

- Lab session, starting on 5^{th}/6^{th} week
  - Wednesday 9:45AM-12:10PM

- Course Link:
  - http://good.ncu.edu.cn/~xuz/cc/homepage.html

Learning Objectives

- What is Cloud Computing?
- Why Cloud Computing?
- *Infrastructure* for CC

Course Description

- We cannot argue that it is a matter of whether cloud computing will become ubiquitous.

- The economic forces drive cloud computing are inescapable.

- Else rather what can we do to assess enterprise governance, risk assessment, and development of strong internal controls, in the implementation and management of ever increasing cloud computing environments.

Course Description

- This course will begin with first establishing the definition of cloud computing, then describing the various service delivery models of a cloud computing architecture, and the ways in which clouds can be deployed as public, private, hybrid, and federated clouds, followed by a much deeper review of the security and privacy issues related to cloud computing environments.

Topics

- Understand various basic concepts related to cloud computing technologies
- Understand the architecture and concept of different cloud models: IaaS, PaaS, SaaS
- Understand the underlying principle of cloud virtualization, cloud storage, data management and data visualization
- Understand different cloud programming platforms and tools
- Link/network/transport layer protocols
- Be familiar with cloud programming using Google's 'Go' programming language
- Create application by utilizing cloud platforms
- Learn basic concepts of MapReduce programming models for big data analysis on cloud

Course Organization

- We plan to have 4 – 5 homework
  - Approximately once in two/three weeks
- We will NOT have mid-term exam
  - A class report will do
- We may not have final exam, a project maybe

Course Organization (Cont.)

- We do have a lot of reading assignment
  - Reading papers, articles, and books
  - Writing your after-read notes
  - Of course, in markdown or latex style

- We will recommend you many many books to read
  - Most books do have electronic copies, no worries
  - Some books are worthy to keep a paper version, I will mention it then
  - I welcome counter-reading suggestions

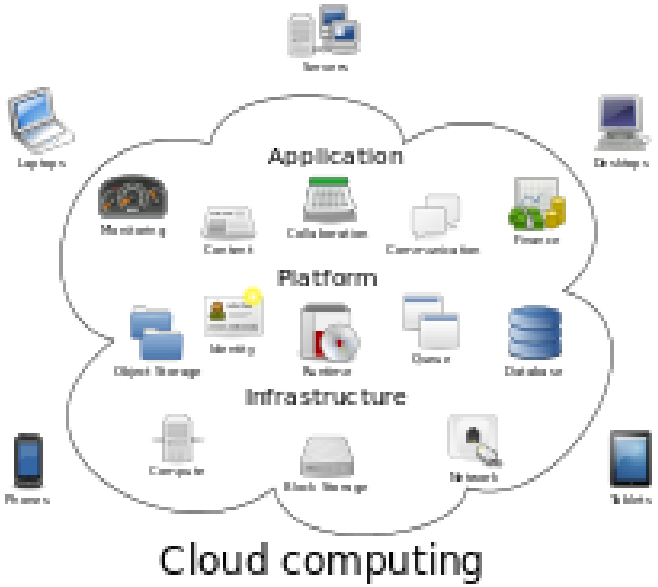
Lecture 1:
Cloud Computing 101

Interlude: Cloud Computing

Source: Wikipedia (Clouds)

We are in the Cloud

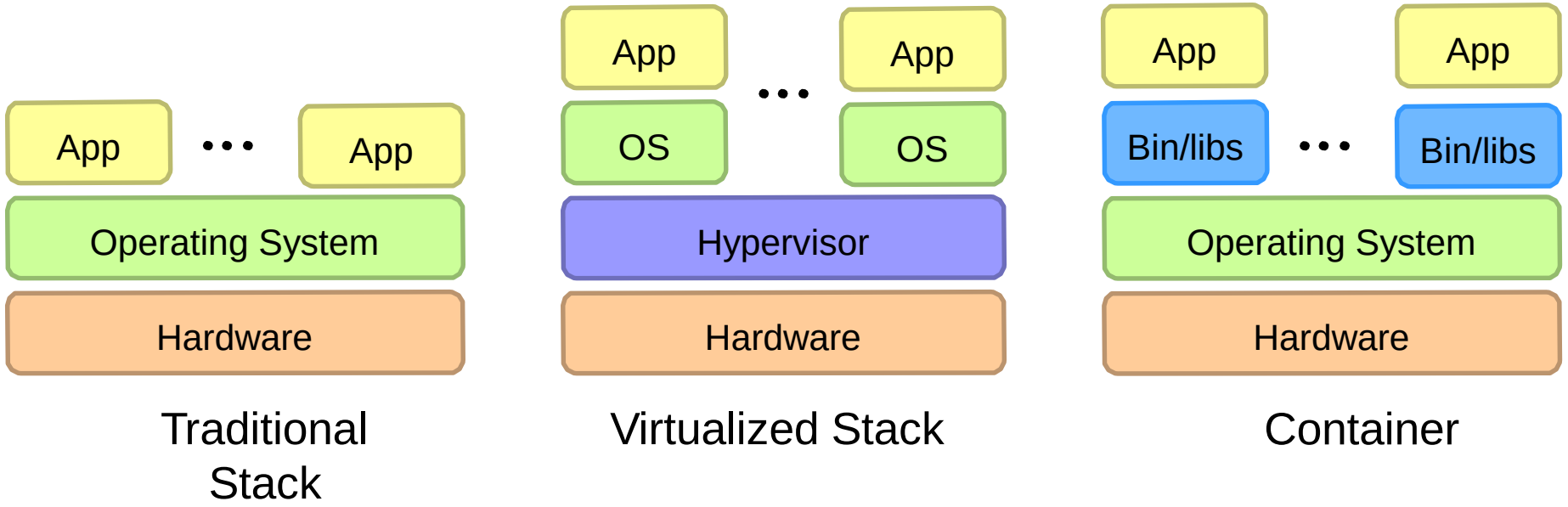
- Before clouds…
  - Grids
  - Connection machine
  - Vector supercomputers
  - …
- Cloud computing means many different things:
  - Big data
  - Rebranding of web 2.0
  - Utility computing
  - Everything as a service



Utility Computing

- What?
  - Computing resources as a metered service (“pay as you go”)
  - Ability to dynamically provision virtual machines
- Why?
  - Cost: capital vs. operating expenses
  - Scalability: “infinite” capacity
  - Elasticity: scale up or down on demand
- Does it make sense?
  - Benefits to cloud users
  - Business case for cloud providers

Enabling Technology: Virtualization and Container



Everything as a Service

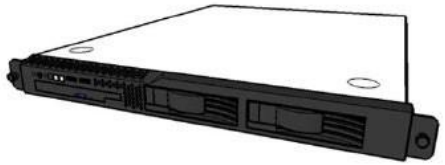
- Utility computing = Infrastructure as a Service (IaaS)
  - Why buy machines when you can rent cycles?
  - Examples: Amazon's EC2, Rackspace
- Platform as a Service (PaaS)
  - Give me nice API and take care of the maintenance, upgrades, …
  - Example: Google App Engine
- Software as a Service (SaaS)
  - Just run it for me!
  - Example: Gmail, Salesforce

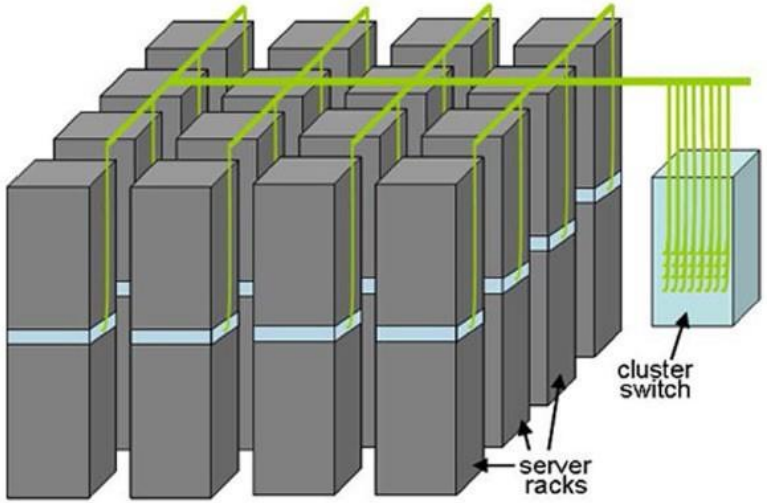
Who cares?

- A source of problems…
  - Cloud-based services *generate* big data
  - Clouds make it easier to start companies that *generate* big data
- As well as a solution…
  - Ability to provision analytics clusters on-demand in the cloud
  - Commoditization and democratization of big data capabilities

Source: Google

Building Blocks





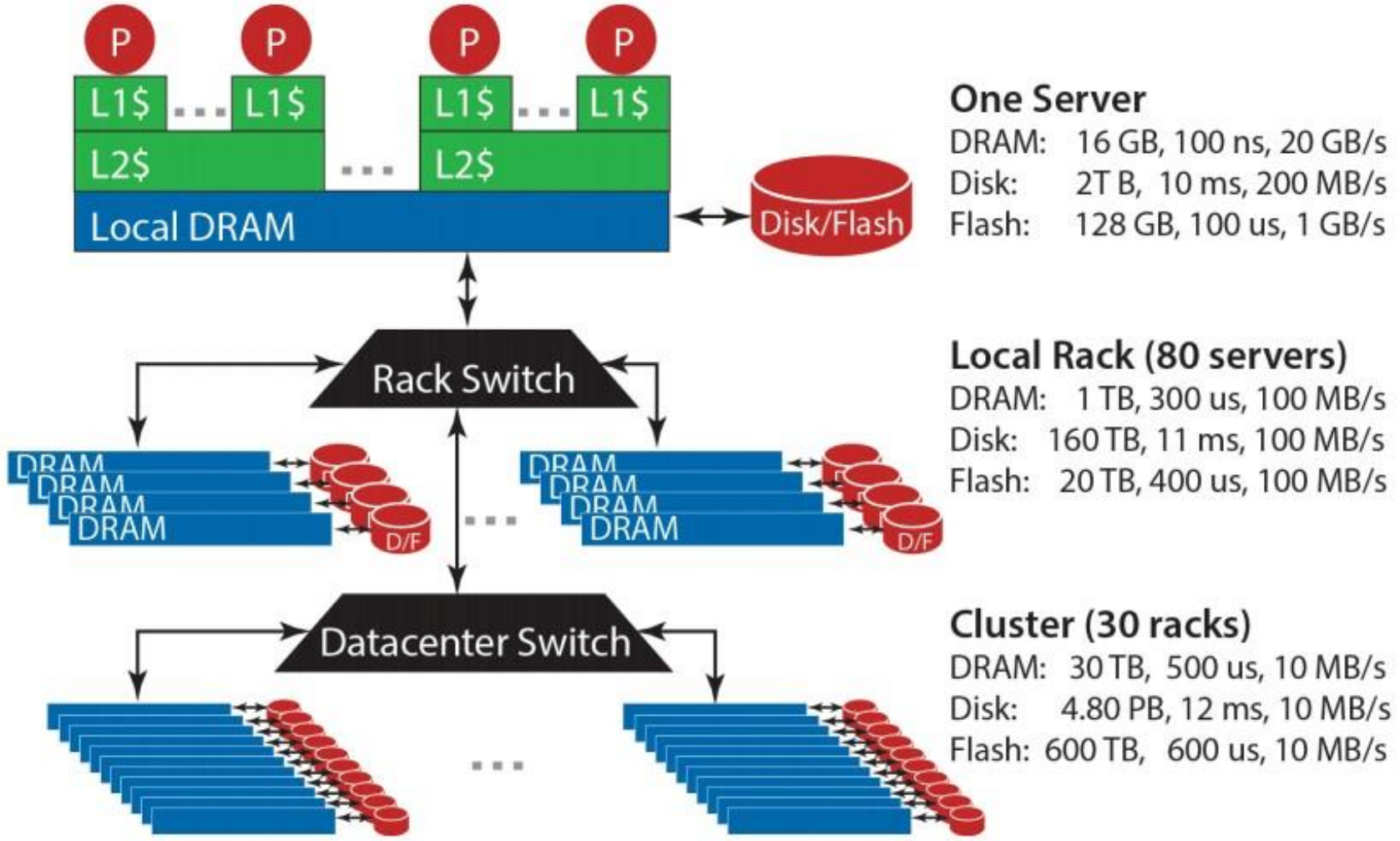
Source: Barroso and Urs Hölzle (2009)

Source: Google

Source: Google

Source: Facebook

Storage Hierarchy



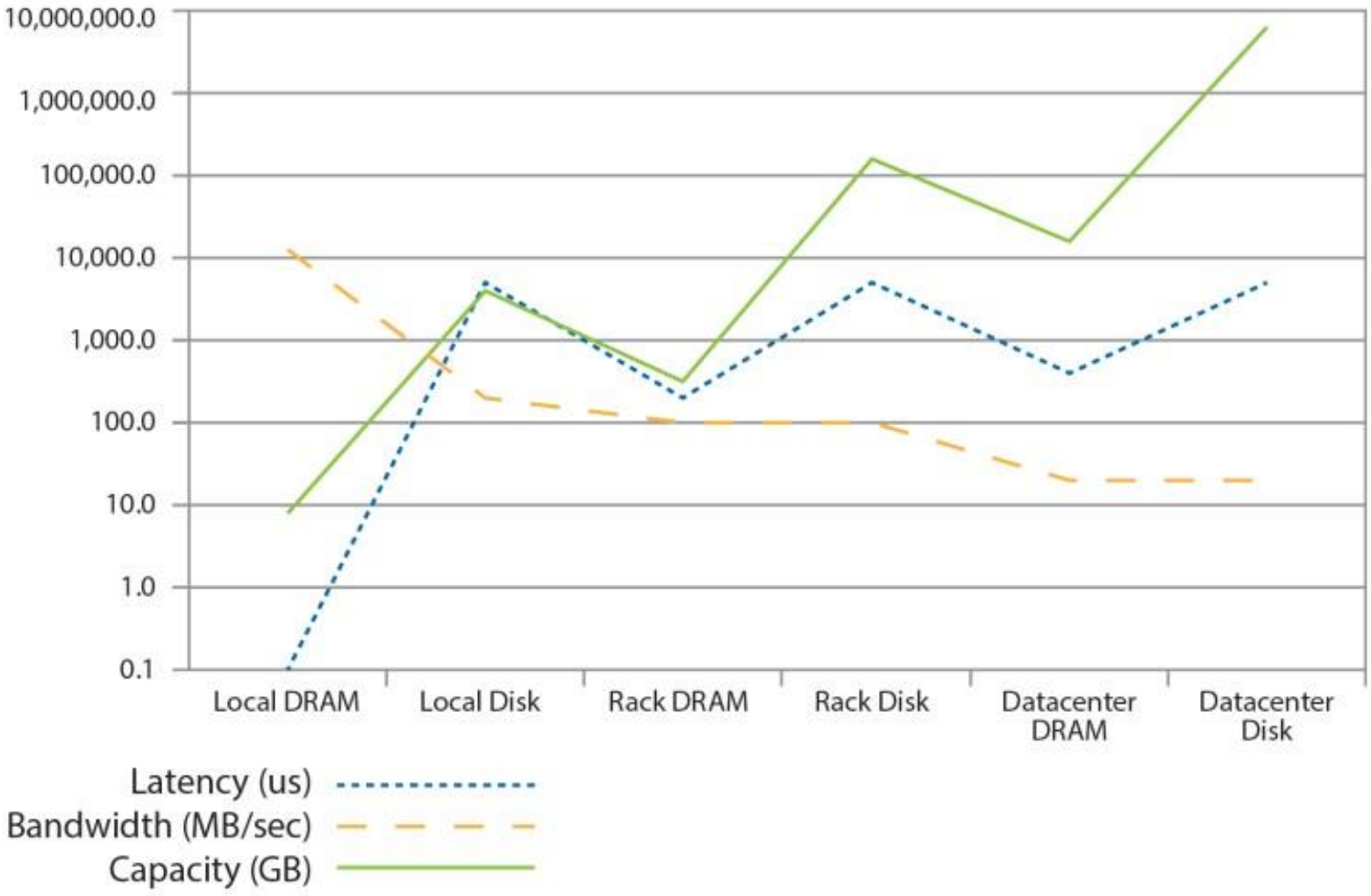
One Server
DRAM: 16 GB, 100 ns, 20 GB/s
Disk: 2T B, 10 ms, 200 MB/s
Flash: 128 GB, 100 us, 1 GB/s

Local Rack (80 servers)
DRAM: 1 TB, 300 us, 100 MB/s
Disk: 160 TB, 11 ms, 100 MB/s
Flash: 20 TB, 400 us, 100 MB/s

Cluster (30 racks)
DRAM: 30 TB, 500 us, 10 MB/s
Disk: 4.80 PB, 12 ms, 10 MB/s
Flash: 600 TB, 600 us, 10 MB/s

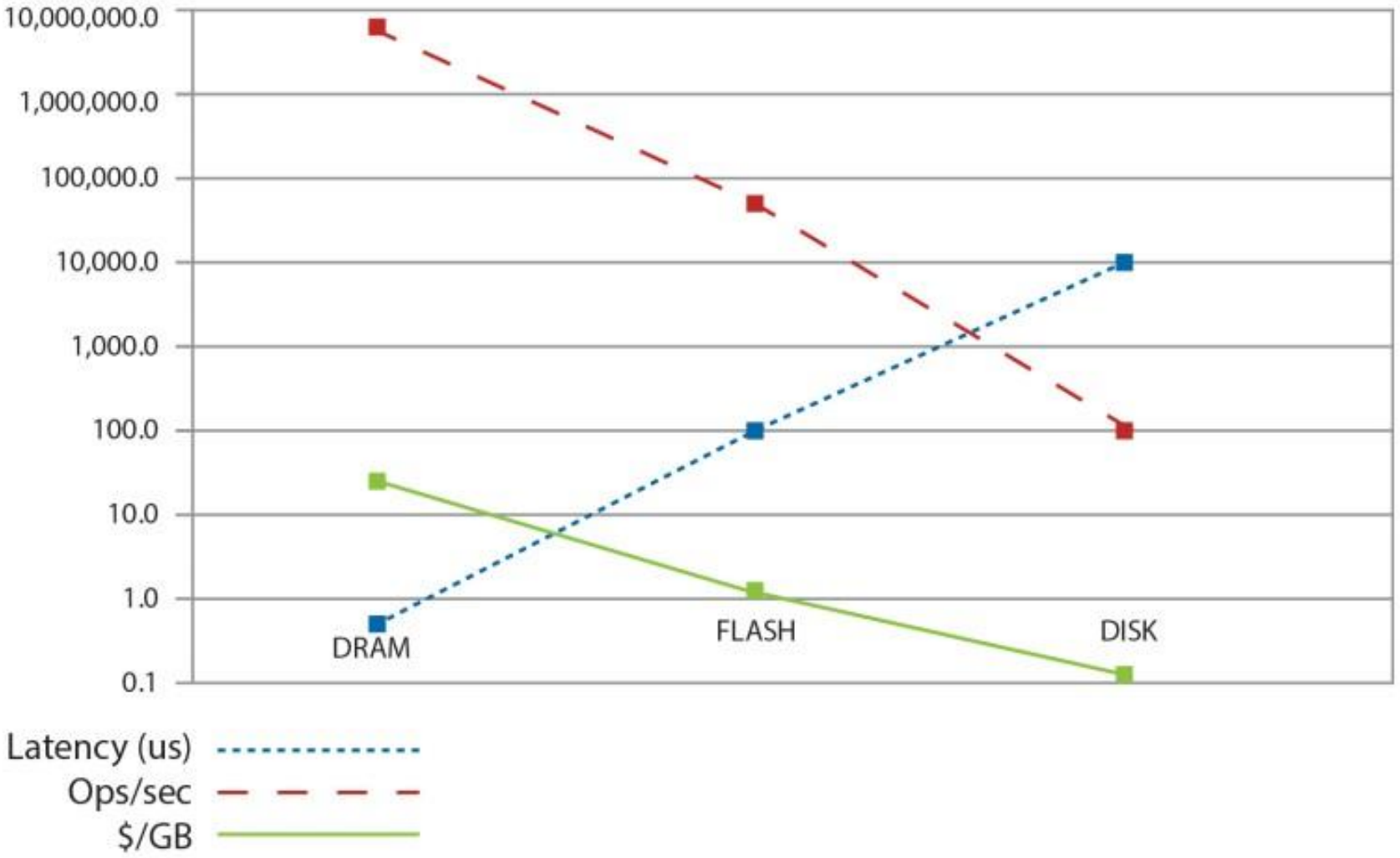
Source: Barroso and Urs Hölzle (2013)

The Cost of Moving Data Around Data Center



Source: Barroso and Urs Hölzle (2013)

The Cost of Moving Data Within a Server



Source: Barroso and Urs Hölzle (2013)

“Big Ideas” of Massive Data Processing in Data Centers

- Scale “out”, not “up”
  - Limits of SMP and large shared-memory machines
- Move processing to the data
  - Clusters have limited bandwidth
- Process data sequentially, avoid random access
  - Seeks are expensive, disk throughput is reasonable
- Seamless scalability
  - From the mythical man-month to the tradable machine-hour

A Case Study: Google Data Center

https://www.youtube.com/watch?v=XZmGGAbHqa0

Take-away Message

- Cloud computing and data centres are natural infrastructures for data science.
- Cloud computing distribute data contains value and knowledge
- Cloud computing creates infrastructure to build/support EaaS
- Further readings:

Chapter 1. Jimmy Lin and Chris Dyer. 2010. Data-Intensive Text Processing with Mapreduce. Morgan and Claypool Publishers. https://lintool.github.io/MapReduceAlgorithms/MapReduce-book-final.pdf

Questions?

How do you want from a Cloud?

Acknowledgement

- Some slides are adopted/revised from
  - Jimmy Lin, http://lintool.github.io/UMD-courses/bigdata-2015-Spring/
  - Jure Leskovec, Anand Rajaraman, and Jeffrey David Ullman. 2014. Mining of Massive Datasets (2nd ed.). Cambridge University Press. http://www.mmds.org/
  - Bingsheng He, School of Computing, National University of Singapore, CS4225/CS5425 Big Data Systems for Data Science