Network Programming
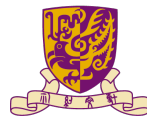## Introduction & Course Logistics

Minchen Yu

SDS@CUHK-SZ

Spring 2026

香港中文大學（深圳）
The Chinese University of Hong Kong, Shenzhen
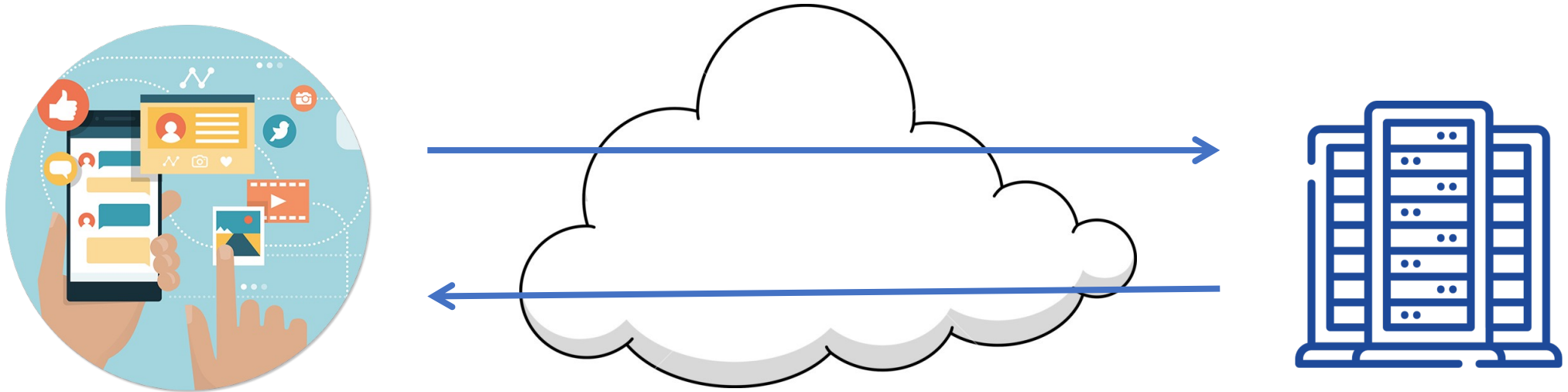
SCHOOL OF DATA SCIENCE
數據科學學院

About me

- Minchen Yu, Assistant Professor, SDS
  - Email: yuminchen@cuhk.edu.cn
  - Office: DaoYuan 420c
  - Office hour: Wednesday 2-3 PM

- Research interests
  - Cloud computing, distributed systems, and machine learning systems
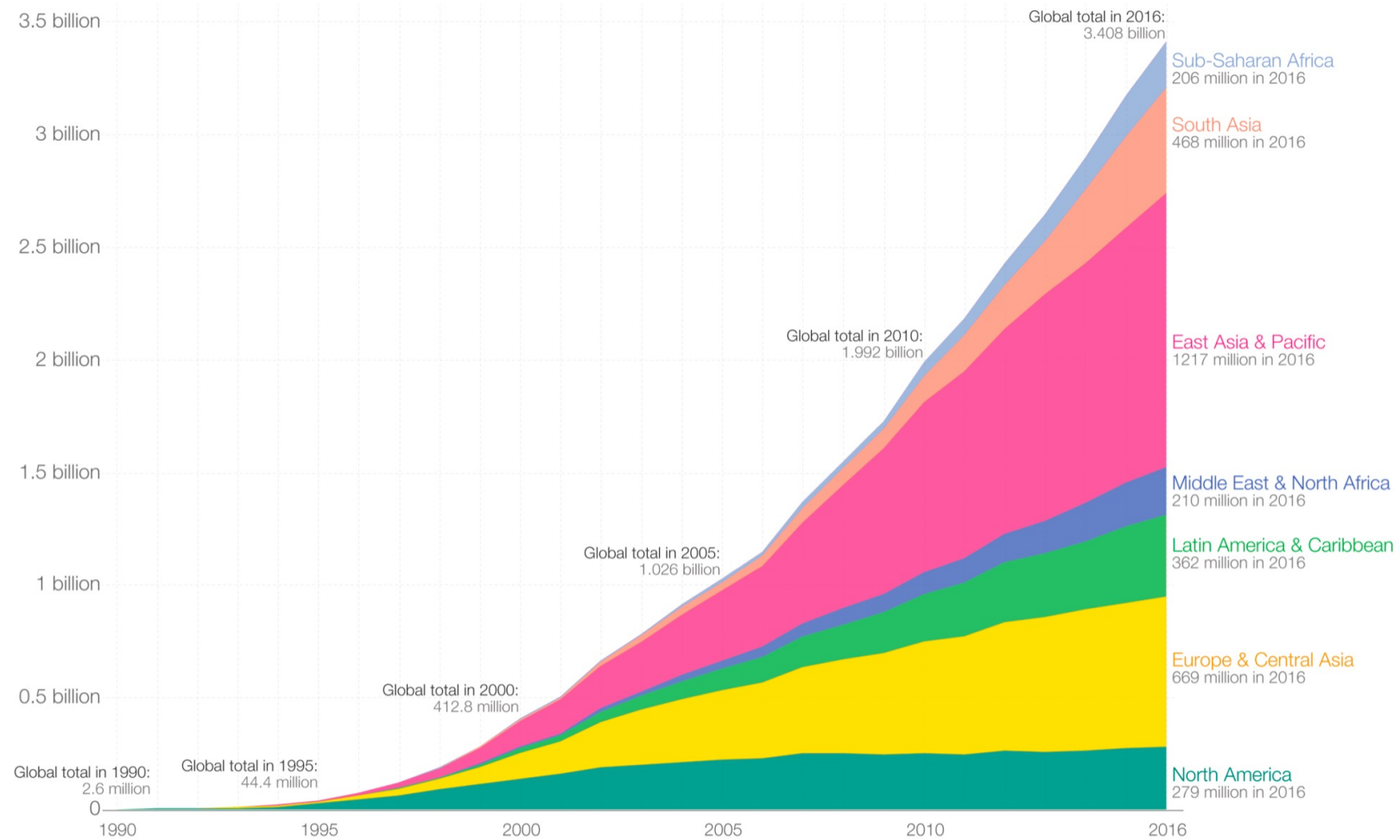
Internet applications today

- Social media
- Online shopping
- Search engine
- Video platforms
…

Access services over Internet/network

Internet user growth

Global total in 2016:
3.408 billion

Sub-Saharan Africa
206 million in 2016

South Asia
468 million in 2016

East Asia & Pacific
1217 million in 2016

Global total in 2010:
1.992 billion

Middle East & North Africa
210 million in 2016

Latin America & Caribbean
362 million in 2016

Global total in 2005:
1.026 billion

Europe & Central Asia
669 million in 2016

Global total in 2000:
412.8 million

North America
279 million in 2016

Global total in 1990:
2.6 million

Global total in 1995:
44.4 million

A view of service providers

- Establish network connections with users

- Handle many user requests

- Store and process a (large) volume of data

…

A view of service providers

Google

Processes 20 PB a day (2008)
Crawls 20B web pages a day (2012)
Search index is 100+ PB (5/2014)
Bigtable serves 2+ EB, 600M QPS (5/2014)

300 PB data in Hive +
600 TB/day (4/2014)

facebook.

amazon
web services™

S3: 2T objects, 1.1M
request/second (4/2013)
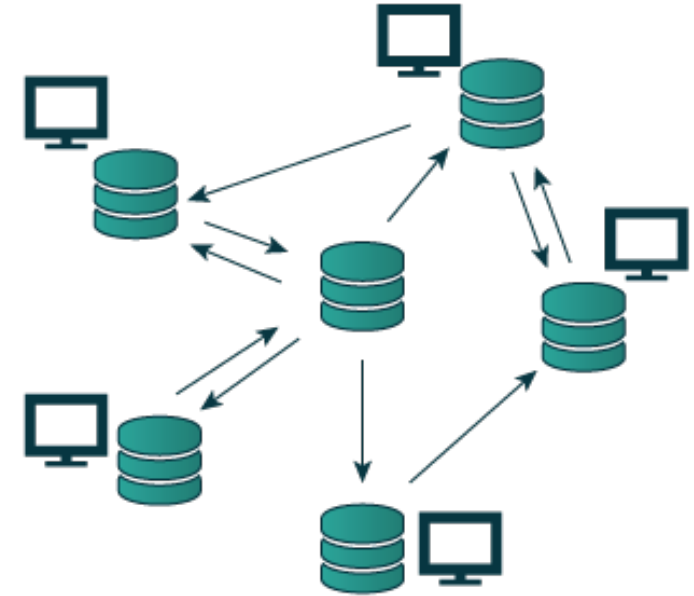
Datacenter

Copywrite: Google

Datacenter

- >100k servers
- Costs in billions of dollars
- Geographically distributed

Distributed systems

- Store big data across multiple servers
- Communications between servers
- Recover from failures if happened

…

Example of distributed systems

- Distributed storage systems
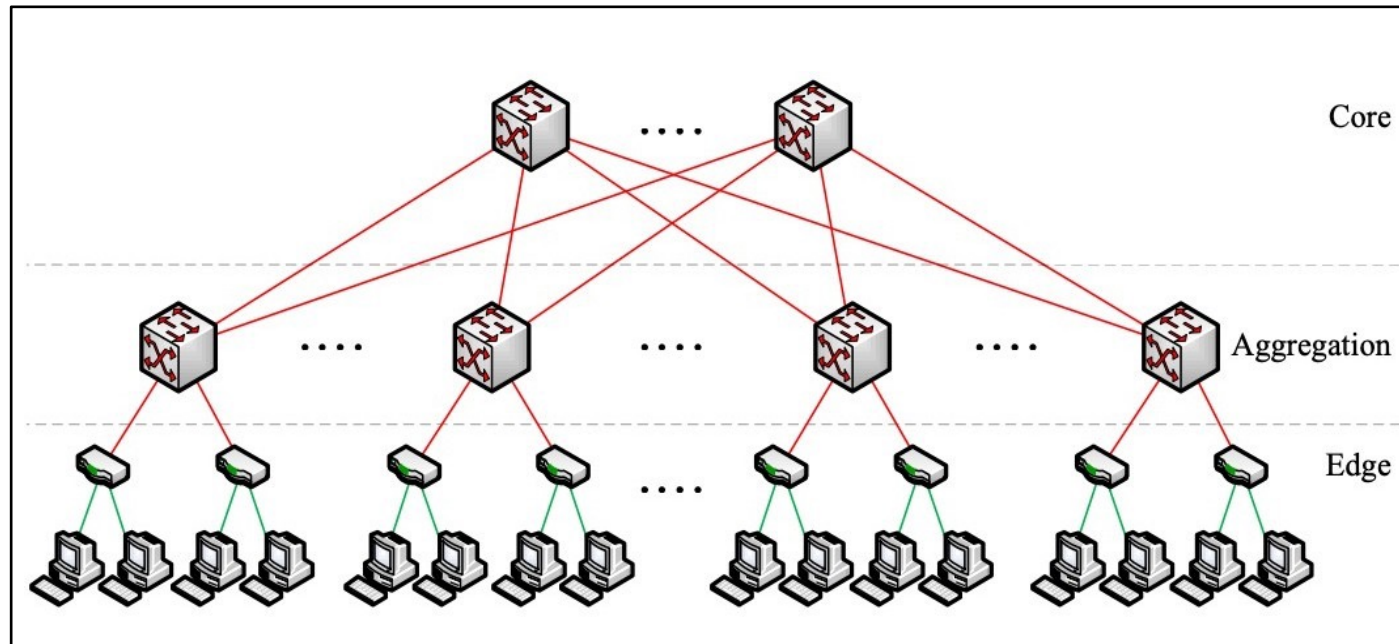 - Store a large amount of data in a distributed environment
  - GFS, HDFS

- Communication frameworks
  - Provide collective communication models
  - MPI

Data center networking

- Thousands of routers
- Hundreds of thousands of servers

About the course

- Website: see Blackboard
- Announcements and course materials are posted online on a regular basis
- TA: Ye Wang <yewang2@link.cuhk.edu.cn>
- USTF: Zihao Fang <zihaofang1@link.cuhk.edu.cn>

Prerequisites

- Be comfortable with Unix/Linux
- Be comfortable with Python Programming
- Be comfortable with C/C++ Programming

Accept an email invitation to get enrolled into an AWS Academy Learner Lab

A note about EC2

- Amazon EC2 credit
  - $50 USD credit from AWS Academy Classroom upon enrollment
  - AWS free tier: 750 hours per month for the **first 12 months**
- Shut down your virtual instances whenever you are done
- Extra usage at your own cost. We cannot financially help you in any means!

References

- References
  - Computer Networks: A Systems Approach
  - Unix Network Programming, Volume 1: The Sockets Networking API
  - Other online materials (check Blackboard)

- Learn things online

Assessment

- Programming assignments
  - 60%

- Final exam
  - 40%

Objectives

- Learn and understand the following aspects
  - Network layered design
  - Programming techniques for network communications
  - Distributed storage systems
  - Collective communication
  - Data center networking

Tips

- Attend the lecture and tutorial

- Get your hands dirty

- Learn things online

- Do all assignments by yourself

- Do well in the final

Learn programming by trying it yourself!

Academic honesty

- In short, don't cheat!

- Don't copy code or solutions from your classmates or third-party sources, and don't let others copy yours. Both cases are plagiarism and penalized in the same way

Protocol for Plagiarism

- TAs will detect possible plagiarism in your code/reports.

- Suspicious cases will be directly reported to the school.

- Minimum penalty: zero mark for the assignment.

Tentative lecture schedule

Week	Topic
1	Logistics, Computer Network Overview
2	Transport Layer I
3	Transport Layer II
4	Socket Programming
5	Network Layer
6	P2P
7	Application Layer
8	RPC
9	Distributed Storage System I
10	Distributed Storage System II
11	Collective Communication
12	Data Center Networking
13	Review
14	Final Exam

Subject to change.
Check back frequently!

Tentative tutorial schedule

- Monday 7-7:50 PM
- Teaching D, 307
- The first tutorial starts next week (1.12)
- Tutorial may be skipped with notification in advance

Any questions?