INFORMATION CODING
Linköping University

TSIN02 Internetworking

Lecture 7 – Data center networks

Literature

- Networked life: 20 Questions and Answers, Mung Chiang

Chapter 16: What's inside the cloud of iCloud?

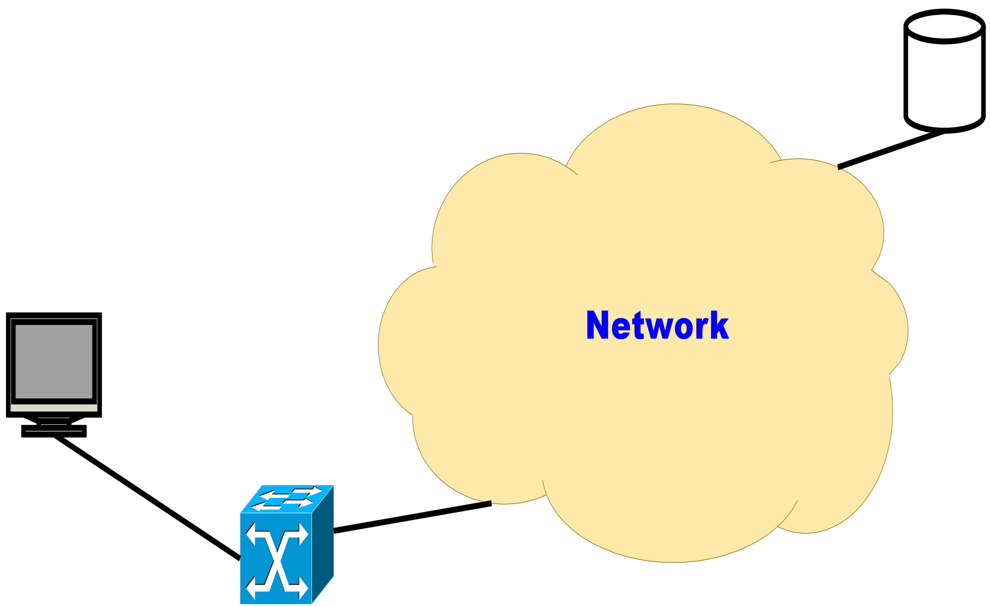
INFORMATION CODING
Linköping University

Outline

- Cloud computing
- Data center networks
- Ingredients, challenges, and advantages of cloud services
- Scalability problem in intra-datacenter networks
- Interconnection networks (Clos network)
- Inter-datacenter networks

What is Cloud?

- Cloud has often been used as a metaphor for Internet in the network diagram.
- Cloud computing is a new IT delivery model accessed over the network (Internet or intranet).



Computer network

The term "cloud computing" is coined by University of Texas professor Ramnath Chellappa in a talk on a "new computing paradigm" in 1997.

What is Cloud?

- "Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction." defined by the National Institute of Standards and Technology (NIST).

- **Example of cloud-based services**
  - Apple iCloud.
  - Web-based email systems: Gmail etc.
  - Web-based software systems: Microsoft Office 365.
  - Google Docs and Dropbox.

Characteristics of Cloud Computing

- On-demand self-service: provisioning (or de-provisioning) computing resources as needed in an automated fashion without human intervention.

- Ubiquitous network access: computing facilities can be accessed from anywhere over the network using any sort of thin or thick clients (e.g., smartphones, tablets, laptops, personal computers).

- Resource pooling: computing resources are pooled to meet the demand of the consumers so that resources can be dynamically assigned, reassigned or de-allocated as per the requirement.

Characteristics of Cloud Computing

- Rapid elasticity: resources can be elastically provisioned or released according to demand. For example, cloud-based online services should be able to handle a sudden peak in traffic demand by expanding the resources elastically. When the peak subsides, unnecessary resources can be released automatically.

- Measured service: consumers only pay for the computing resources they have used. This concept is similar to utilities like water or electricity.

History of Cloud Computing

- It was a gradual evolution that started in the 1950s with mainframe computing.
- Multiple users were capable of accessing a central computer through dumb terminals, whose only function was to provide access to the mainframe.
- Around 1970, the concept of virtual machines (VMs) was created.
- Using virtualization software like VMware, it became possible to execute one or more operating systems simultaneously in an isolated environment.
- The VM operating system took the 1950s' shared access mainframe to the next level, permitting multiple distinct computing environments to reside on one physical environment.

History of Cloud Computing

- In the 1990s, telecommunications companies started offering virtualized private network connections.
- The newly offered virtualized private network connections had the same service quality as their dedicated services at a reduced cost.
- The concept of delivering enterprise applications via a simple website started in 1999 by Salesforce.
- The real physical implementation of cloud computing came when Amazon announced Elastic Computing 2 (EC2) on August 25, 2006.

Elastic Computing 2 (EC2) Folklore

"When Amazon's CEO visited the company computing facility, he was amazed by the number of computers. Since computing facility is designed to avoid crashes when overloaded, the normal utilization of systems is low. The Amazon CEO, therefore, asked to figure out a way to manage the hardware in a programmatic manner where all the management could be done easily remotely using application programming interfaces (APIs). This allowed them to rent out the unused capacity; so began the computer rental business we now call cloud computing."

Source: "Network virtualization and software defined networking for cloud computing: A survey", by Raj Jain, Washington University.

Data Centers

- Data centers are facilities hosting many servers and connecting them via many switches.
- Size can be over 300,000 square feet, house half a million servers, and cost hundreds of millions of dollars to build.
- Can be dedicated facilities, i.e. serving a particular company; For example Google services are run by their dedicated data centers.
- Can be rented and shared facilities; **For example** Target (second largest retail store chain in the US) runs its web and inventory control in a rented cloud.
- Three major shared data center providers: Amazon's EC2 (Elastic Computing 2), Microsoft's Azure, Google's AppEngine.

Facebook Luleå Data Center
The Backbone of Facebook in Europe

90, 000 square meters – 11 soccer fields server hall

Inside Facebook Luleå Data Center

Racks of Servers

Inside Facebook Luleå Data Center

Racks of Servers

INFORMATION CODING
Linköping University

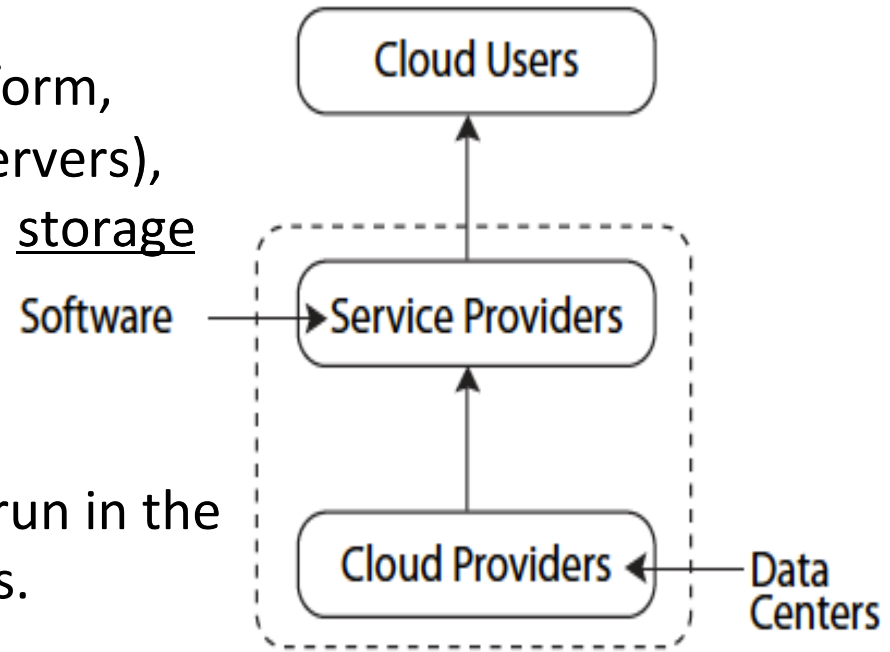
Inside Facebook Luleå Data Center

Cooling System

Cloud Service Industry

Cloud Providers

- Build and manage the hardware platform, consisting of computing resources (servers), networking resources (switches), and storage resources (memory device).



Service Providers

- Offer software and applications that run in the data centers, and interface with users.

Cloud Users

- You and me.
- Big enterprises.

Cloud Service Models

<u>Software as a Service (SaaS)</u>: the capability provided to the cloud consumer is to use the provider's applications running on a cloud infrastructure. The applications are accessible from various client devices through a thin client interface such as a web browser. <u>For example Microsoft Office 365, Google Docs</u>.

The <u>consumer does not manage or control the underlying cloud infrastructure</u> including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings.

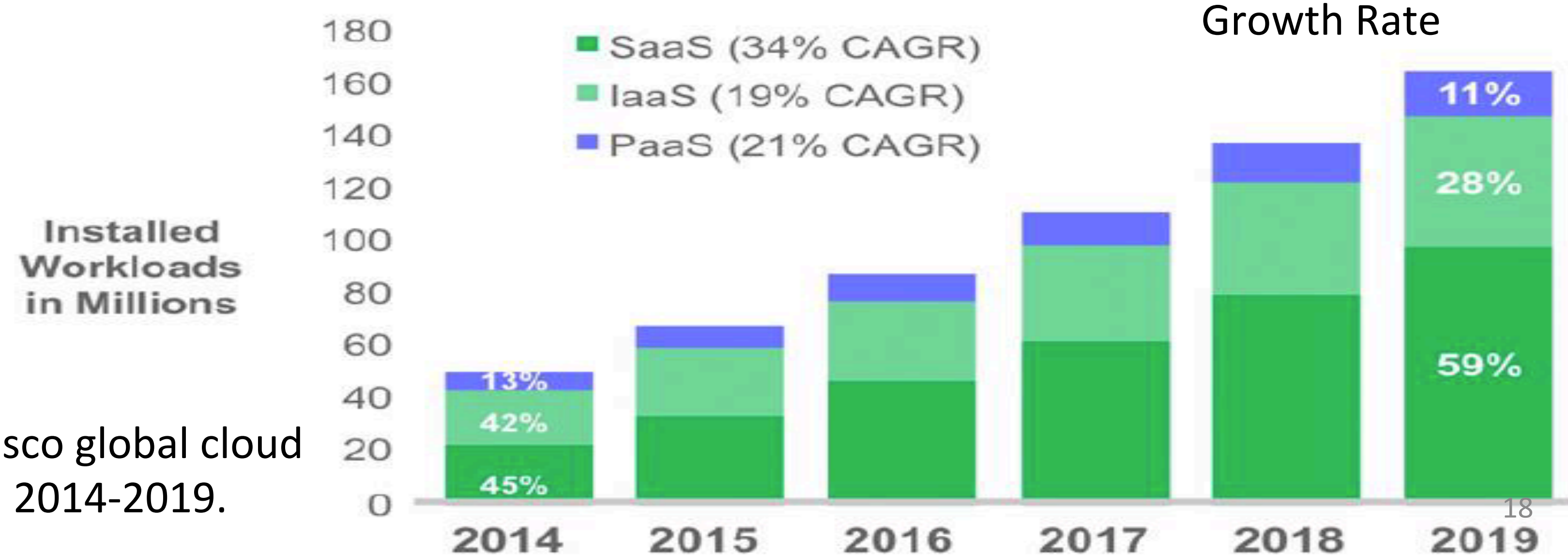
<u>Platform as a Service (PaaS)</u>: the capability provided to the cloud consumer is to <u>deploy onto the cloud infrastructure</u> consumer-created or acquired <u>applications created using programming languages</u> and tools supported by the provider. The consumer does not manage or control the underlying cloud infrastructure, but has control over the deployed applications.

INFORMATION CODING
Linköping University

Cloud Service Models

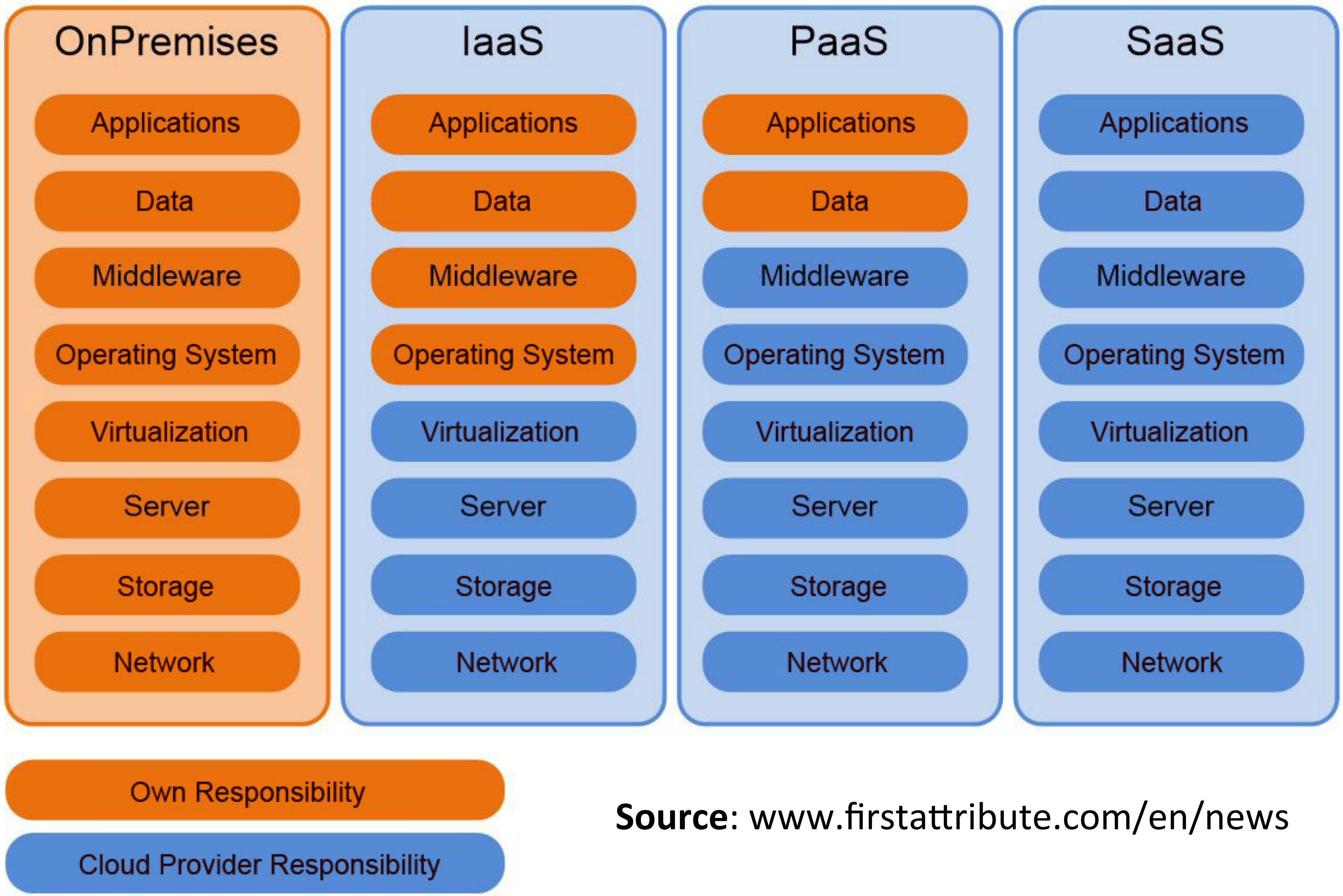
Infrastructure as a Service: the capability provided to the cloud consumer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems, storage, deployed applications, and possibly limited control of select networking components (e.g., host firewalls).

CAGR:Compound Annual Growth Rate



Source: Cisco global cloud index, 2014-2019.

Cloud Service Models: Level of Control



Source: www.firstattribute.com/en/news

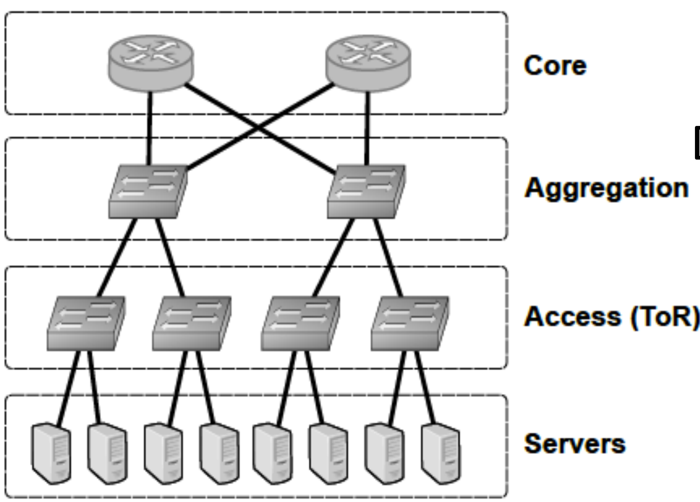
OnPremises: Server and applications are in your own computer center.

Cloud Services Ingredients

Large-scale computing and storage systems.

Software: graphic user interface (GUI), digital rights management (DRM), security and privacy, billing and charging, etc.

Networking: connecting storage and server devices within a data center through switches (intra-datacenter network).



Intra-datacenter network

Data center networking cost amounts to about **15%** of the total cost of the data center.

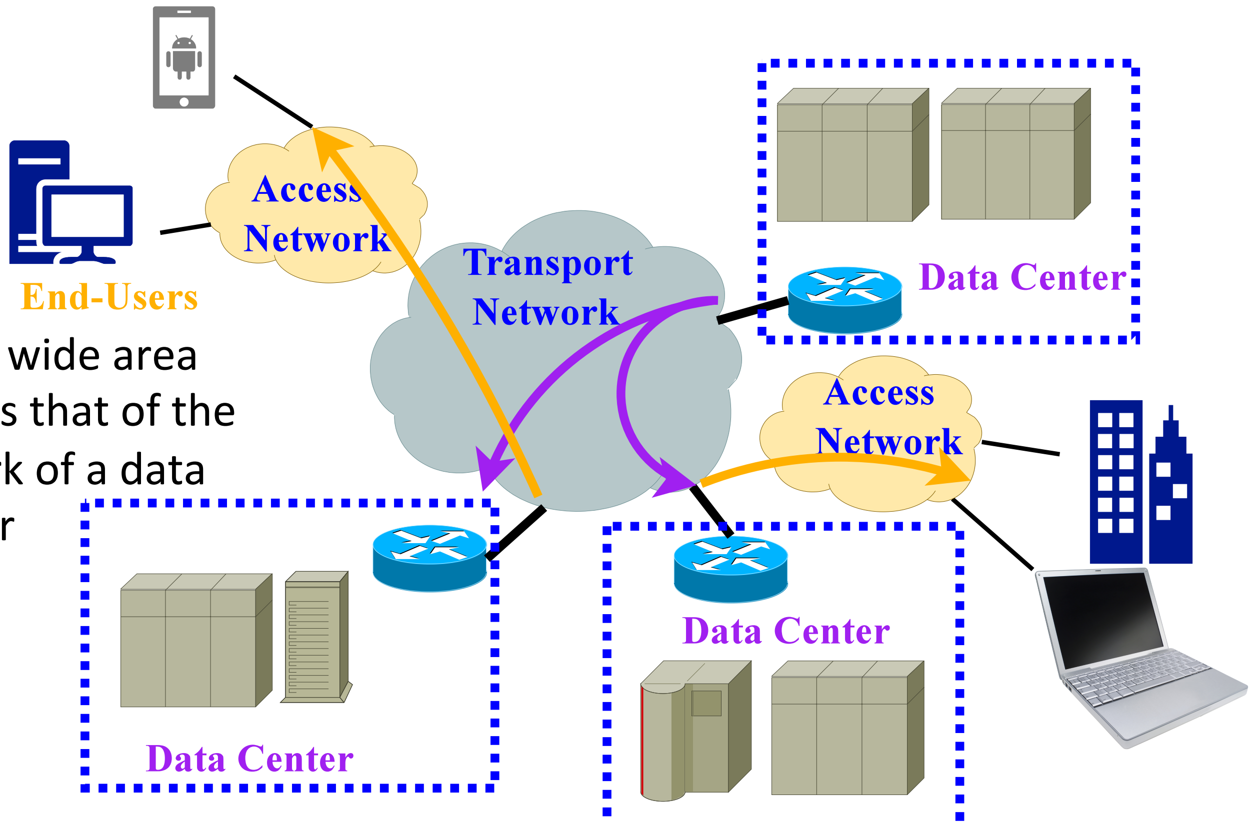
Source: "The cost of a cloud: Research problems in data center networks", *SIGCOMM* 2009.

INFORMATION CODING
Linköping University

Cloud Services Ingredients

Networking

- Connecting different data centers (inter-datacenter network).
- Connecting consumers/enterprises to the data centers.



The cost of the wide area transport exceeds that of the internal network of a data center

Challenges for Cloud Services

- Performance guarantee tends to be weaker than a dedicated facility.
- Cloud computing is totally dependent on Internet connection. When the Internet connection or network is down, cloud services are down as well. If the connection runs slow, then your services will also run slow and the output of the user/enterprise gets affected.
- Do you compromise the safety of your data with Cloud Computing? Security in the cloud is generally good and reliable, established cloud computing vendors ensure they have the latest, most sophisticated data security systems as data security is always a big concern for businesses. But there are still serious security concerns that hackers can attack and breach the system.

Challenges for Cloud Services

Availability of service is one of the top performance concerns. For example, Amazon's cloud outage in April 2011 and in June 2012 for 3 to 4 days; taking down Netflix, Instagram, and other popular services.

Other causes for unavailability are network misconfigurations, firmware bugs, and faulty components.

How to enhance the availability?

Spread the service traffic demand across multiple cloud providers or multiple regions of the same cloud provider.

Availability = (Agreed Service Time − Downtime)*100/Agreed Service Time

Why Cloud Services?

For Cloud Users

- No upfront investment: with cloud services, there is no fixed cost (hardware cost, power management cost, cooling cost)
- CapEx to OpEx conversion: instead of spending money in capital expenditure to build out dedicated facilities, a user pays rent as part of its operational expenditure to share facilities.

  Capital Expenditure - CapEx: money used to acquire or upgrade physical assets such as property, equipment.

  Operational Expenditure - OpEx: ongoing cost for running an equipment/system.
- If the resource demand varies a lot or is just hard to predict; risk of miscalculating resource demand shifts to cloud providers.

INFORMATION CODING
Linköping University

Global Cloud Readiness

The Move to the Cloud is Imminent

Netflix Moves to Cloud, Closing last Data Center

"Cloud environments are ideal for horizontally scaling architectures. We don't have to guess months ahead what our hardware, storage, and networking needs are going to be."
Netflix Representative

83% of Healthcare Organizations Are Using Cloud Based Apps Today

Top three reasons for adopting cloud solutions: [1] less cost than current IT maintenance (55.7%); [2] speed of deployment (53.2%); [3] solving the problem of not having enough internal staff/expertise to support on-premise alternatives (51.6%).
2014 HIMSS Analytics Cloud Survey

NASA Launches Massive Cloud Migration

"...we were able to see how you optimize legacy applications. It can't be business as usual. There's a whole set of different ways to think about the cloud. We can really start creating a strategy so you have better access to information anytime and anywhere."
Roopangi Kadakia, NASA's Web services executive

Apple Pay (via iCloud) in the U.S., Coming to the U.K. and Abroad

It's more secure than contactless card payments because it requires the use of the iPhone's fingerprint sensor during payment. It also creates fewer queues in stores as the faster and more efficient payment method reduces the time needed to input Pins or get cards out of wallets.
Apple CEO, Tim Cook

INFORMATION CODING
Linköping University

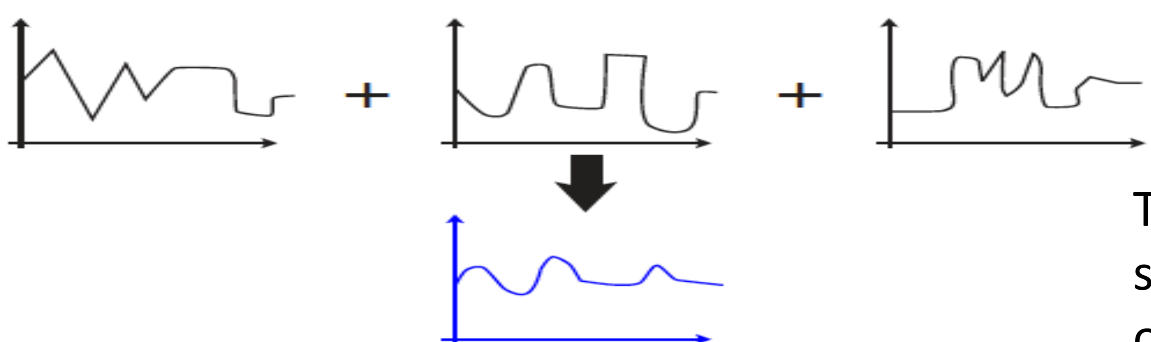
Why Cloud Services?

For Cloud Providers

Economy-of-scale on the supply side: a cloud provider can procure the servers, switches, labor, land, and electricity at significantly discounted price because of its large scale.

Economy-of-scale on the demand side: scale helps through statistical multiplexing; fluctuations of demand for each user is absorbed into a large pool of users.

Cloud is all about scale



Statistical multiplexing

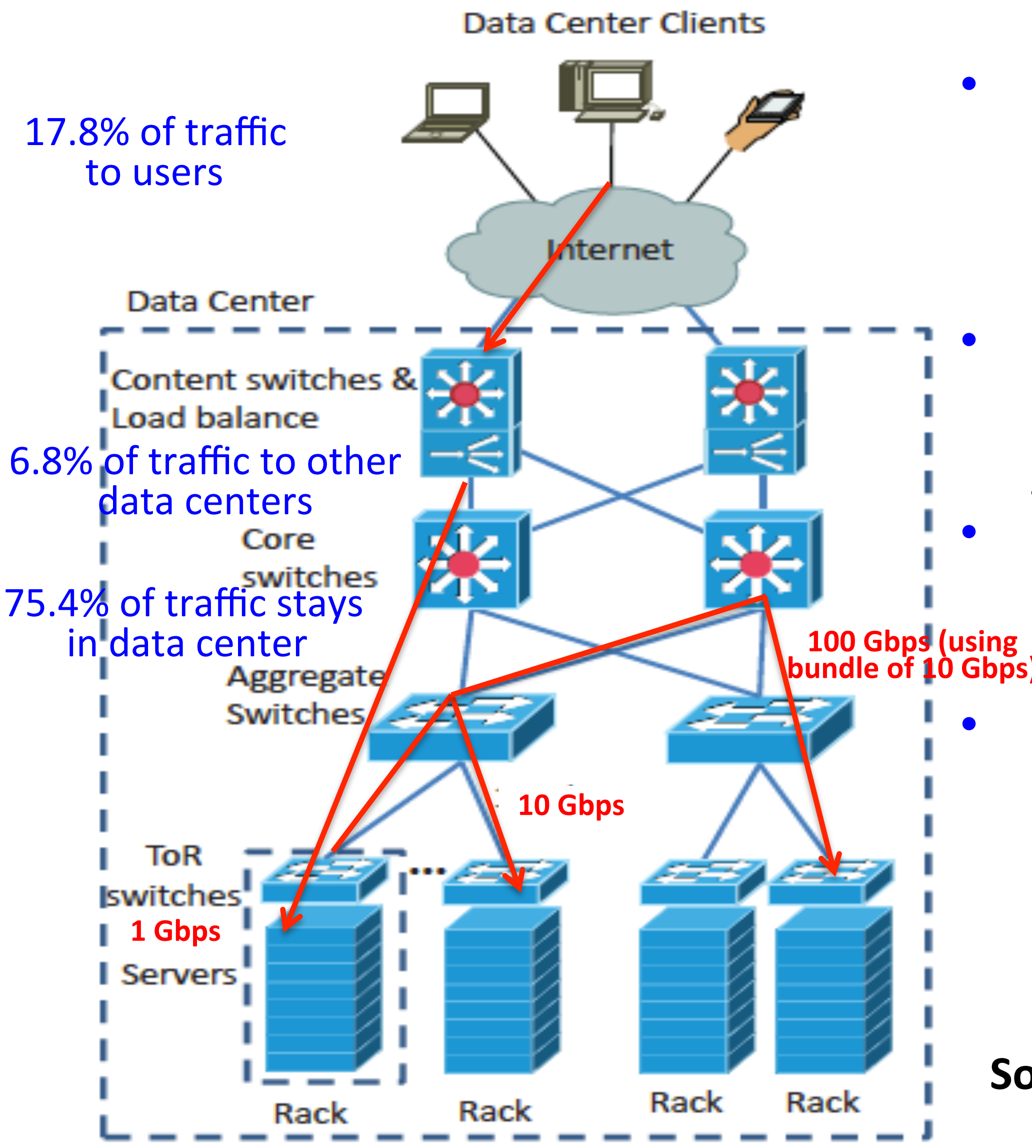
The average utilization of servers in a data center is often below 20% today

Is Scaling up Feasible?

- Many applications hosted in a data center require the transfer of data (and control packets) across the servers at different locations in a big building.
- **500,000 servers need to be connected**.
- High throughput connectivity per node does not scale up beyond a certain point.
- Cannot make a very fast and a very large switch inside the data center.

How to achieve the advantages of scale for a network without suffering the limitation of scale per node?

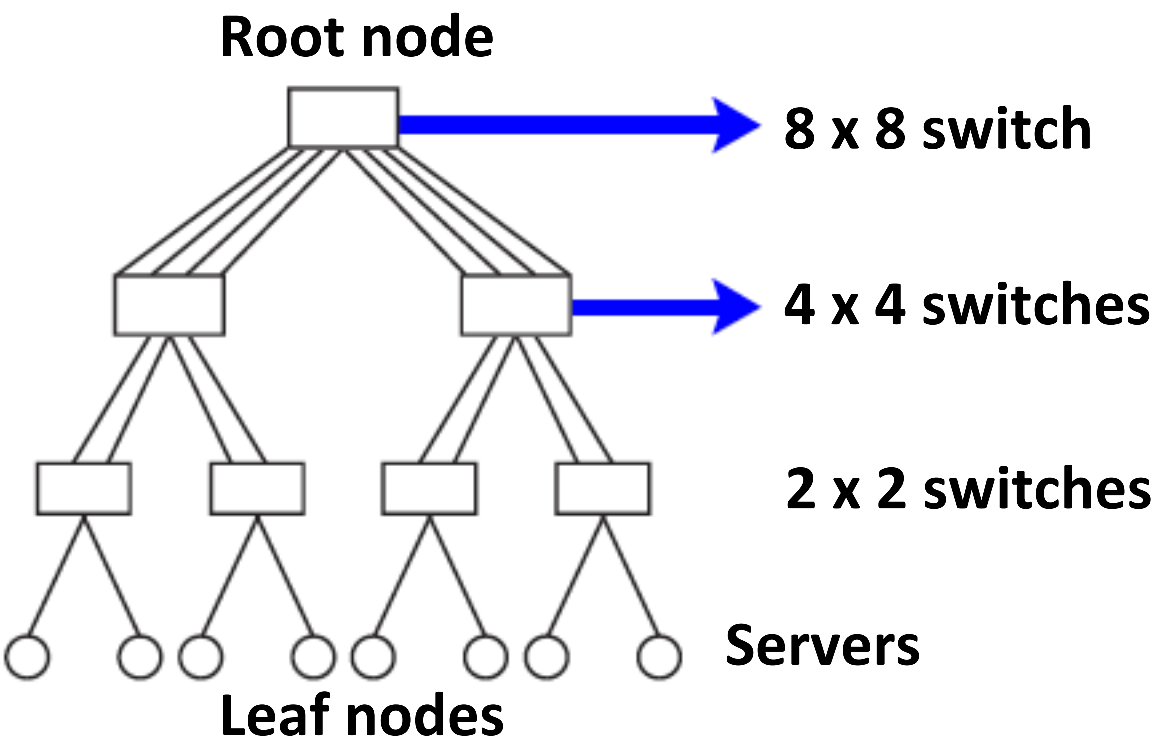
Architecture of Current Data Center Network



- When a request is issued by a client, then a <u>packet is forwarded</u> through the Internet to the front end of the data center.
- The content switches and the load balance devices are used to <u>route the request to the appropriated server</u>.
- A request may require the communication of this server with many other servers.
- For example, a simple web search request may require the communication and synchronization between the web, the application, and the database server.

Source: Cisco global cloud index, 2014-2019.

Tree Topology for Data Center Network

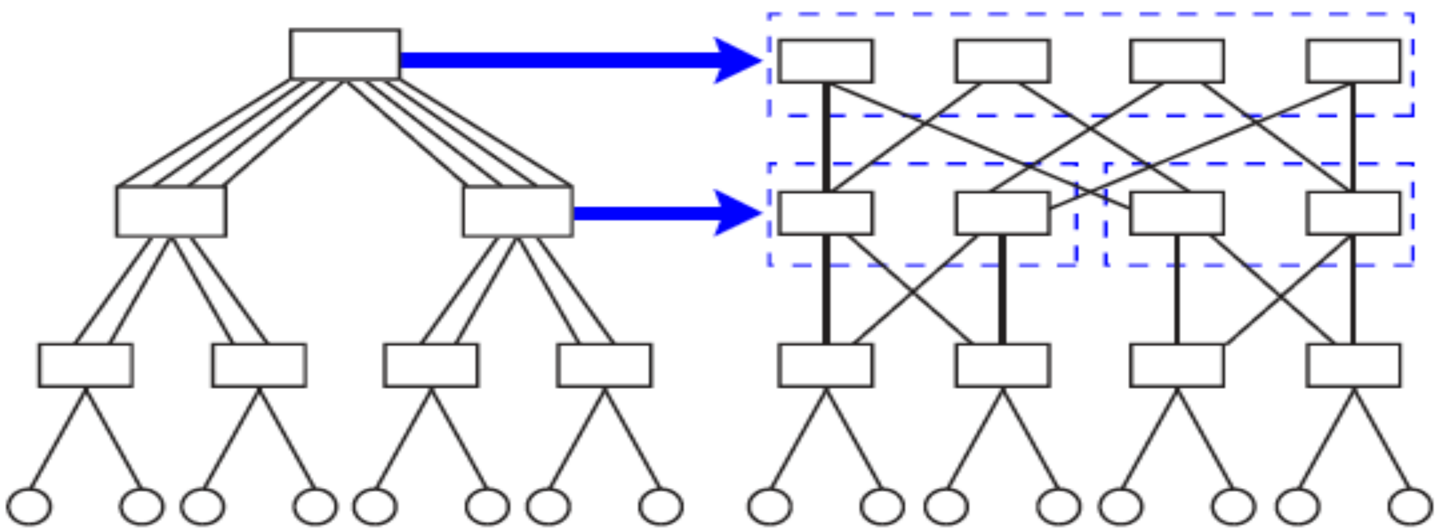


- Imagine 500,000 instead of 8 servers.
- Need a <u>huge root switch</u> with each port being 10 Gbps or more.
- A high end switch today can only support 1280 servers.
- Big switches are quite expensive.

Building a Large Network From Small Switches

Multi-stage switched network

- Connect many small switches in a smart way to form a network of connectivity (interconnection network).
- Rather than scaling up the number of ports per switch scale out the topology.



Applications of Interconnection Networks

- Circuit switched networks: for PSTN networks, it became difficult to build a large enough single switch for handling all phone calls back in 60's.
- Parallel computation: many mainframe computers connected together into a supercomputer.
- Multi-core processor: multiple processing cores connected together on a single chip.

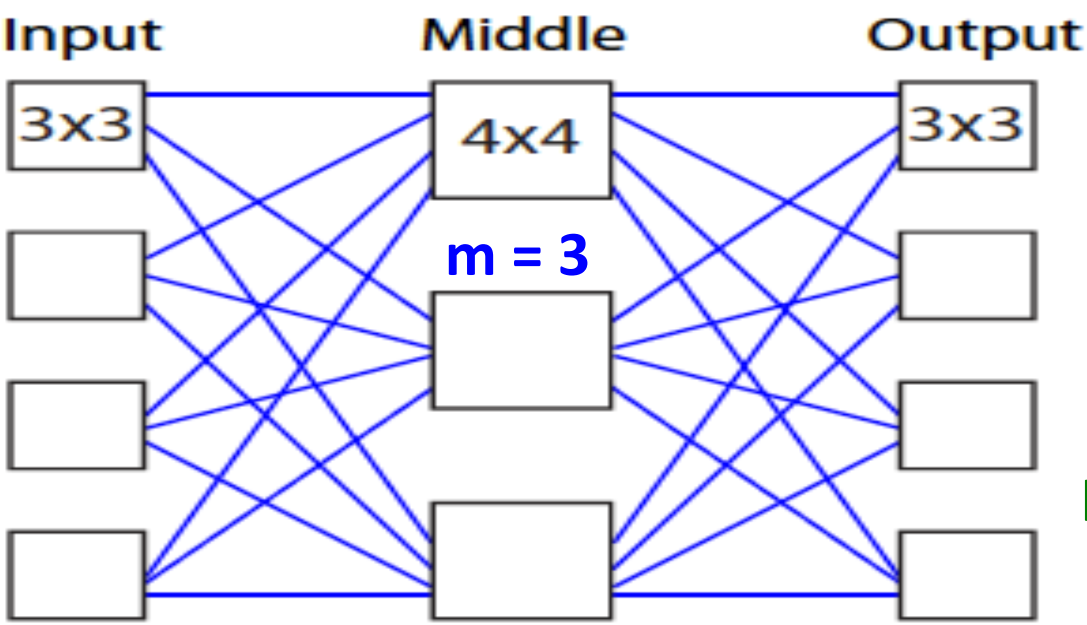
- Data center networking

Connectivity pattern itself is a resource that need to be built carefully.

Interconnection Network Topology
Clos Network (invented by Charles Clos in 1953)

- Clos network is specified by three integers, i.e., **(n,m,r)**.
- Each <u>input switch</u> is **n x m**, and there are r input switches.
- Each <u>output switch</u> (by symmetry) is **m x n**, and there are r output switches.
- Each <u>middle switch</u> is **r x r**, and there are m middle switches.

r = 4



(3,3,4) Clos network

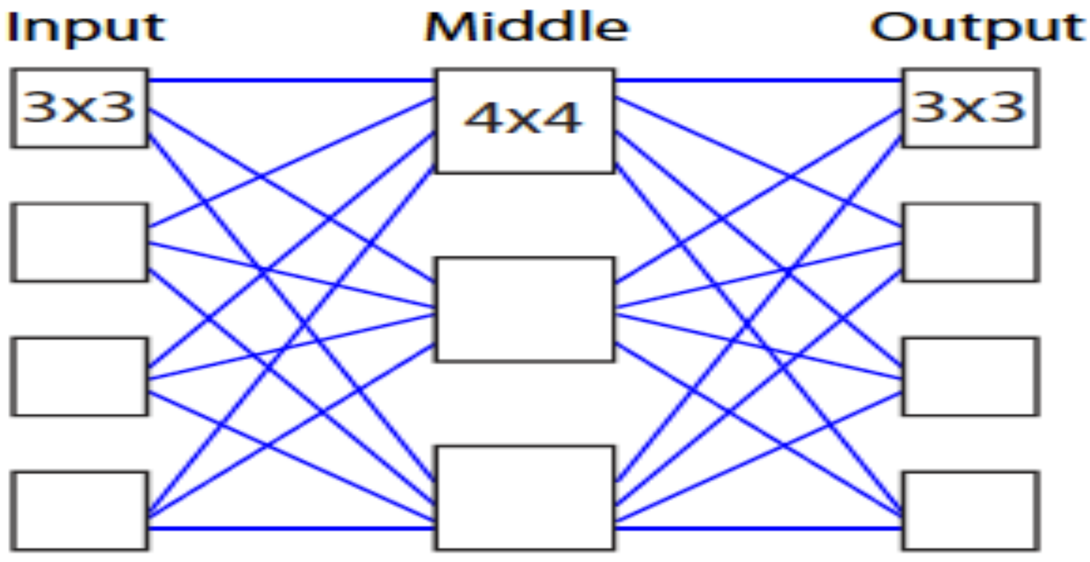
r = 4

12 (r x **n)** input-output pair of ports built from small switches

3 stage Clos network

Interconnection Network

- <u>A typical setting</u>: The servers connected to the left need to get help with storage/computations from the servers connected to the right.
- Within the switches, any connection is possible.
- The main difficulty: after setting up some connections between servers to the left and servers to the right, it may be difficult to connect additional pairs.



Metric for Comparing different kinds of Interconnection Networks

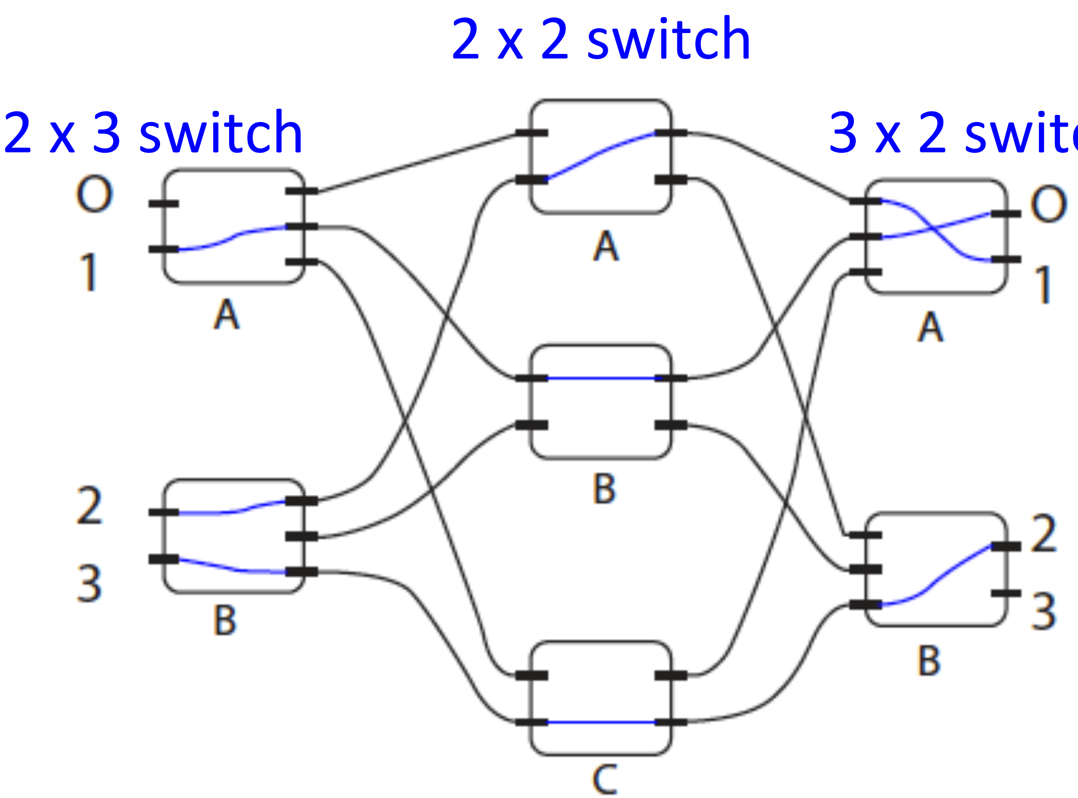
Strict-sense non-blocking: Any unused input and any unused output can always be connected; the traffic will always find a way to pass through the multi-stage switched network.

Re-arrangeably non-blocking: an unused input can always be connected to an unused output, but for this to take place, existing connections may have to be rearranged by assigning them to different middle-stage switches in the network.

Non-blocking Condition for Clos Network

- The middle switches (m) must be greater than or equal to 2n-1;

$$\mathbf{m \geq 2n-1 = 2(n-1) + 1}$$



Example

The network is already configured (denoted by blue lines).

If there is a new incoming traffic, let say at input 0 which wants to go to output 3.

Is it possible to connect them?

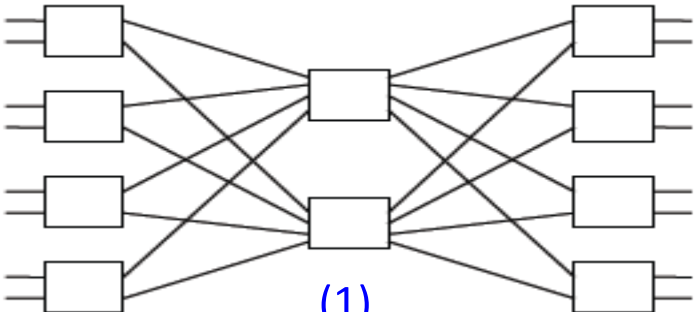
Non-blocking Condition for Clos Network

- The condition for non-blocking: **$m \geq 2n-1$.**
- Where is **r**? For Clos network, there are **n** times **r** input ports and **n** times **r** output ports.
- What if a large Clos network is build by making **r** large?
- The problem with a large **r** is that larger middle switches (**r** x **r**) will be required.
- Recursively apply the Clos network concept and replace the large middle switch with small multi-stage switches.

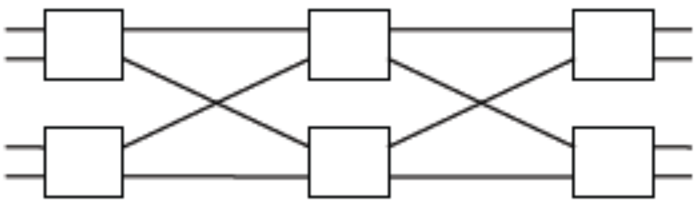
- Condition for **re-arrangeably** non-blocking: **$m \geq n$**

Recursive Expanding – For 8 x 8 Network

(2,2,4) Clos network

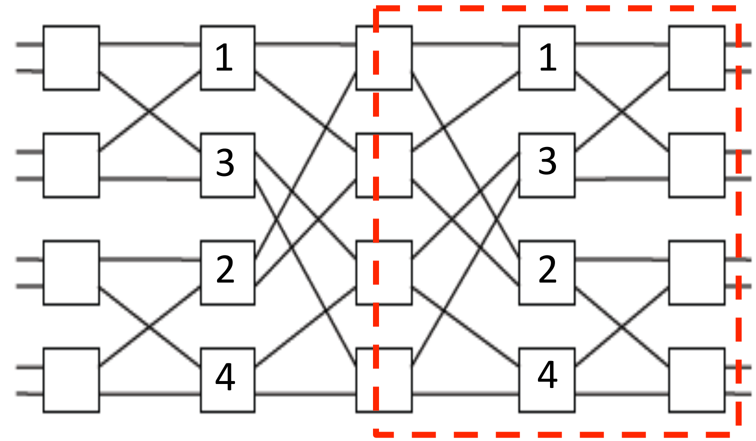


(1)



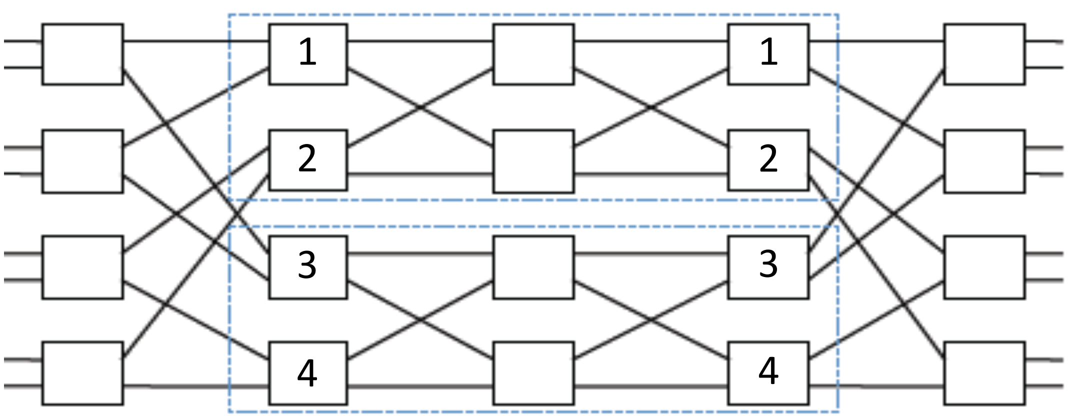
(2)

Rearrange stage 2 and 5



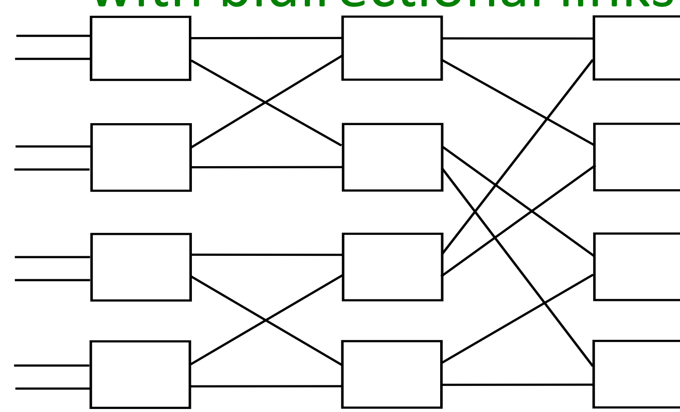
(4) Folding

Composition of two Clos networks



(3)

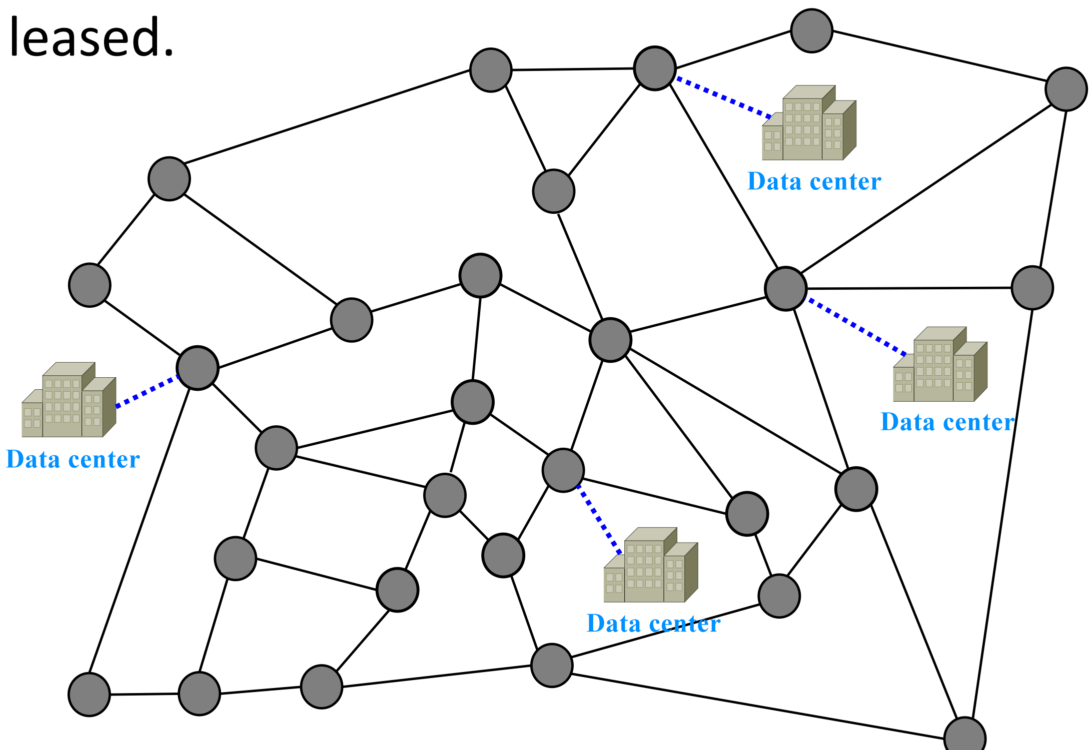
Fat tree (folded Clos network) with bidirectional links



(5)

Inter-Datacenter Network

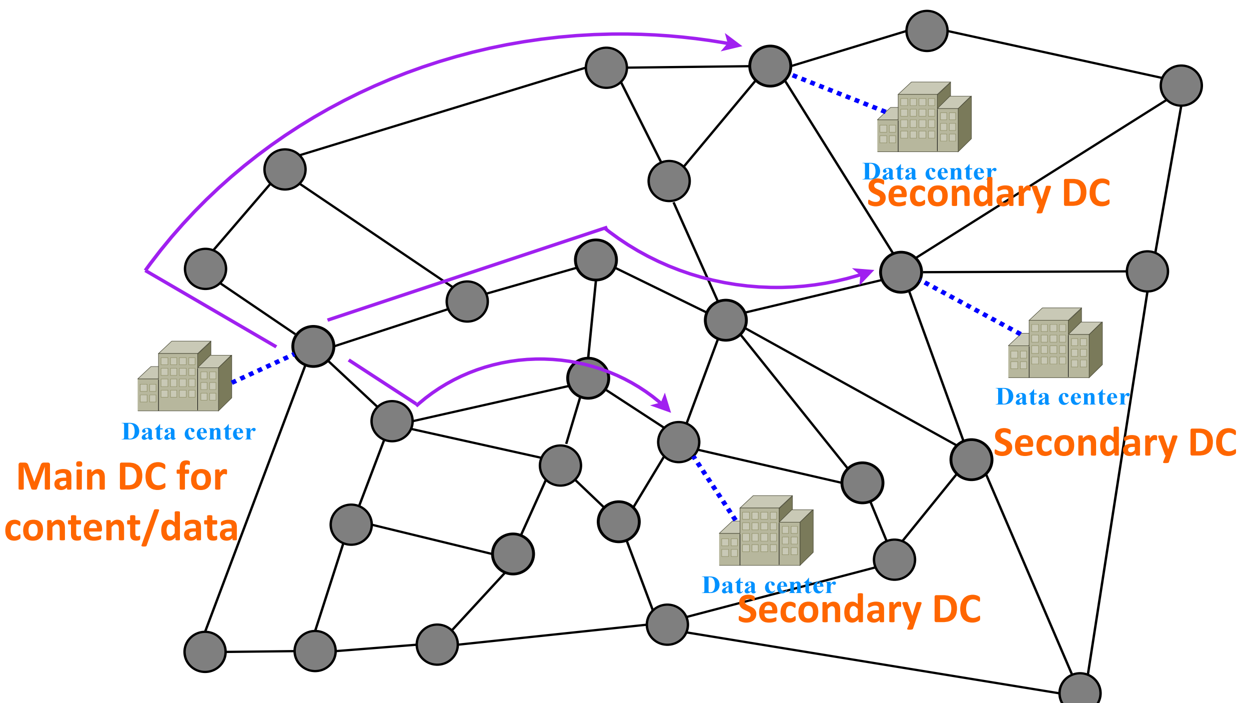
- To meet the local demand and reliability requirement, cloud service providers (CSPs) such as Amazon, Google, Facebook, etc, replicate their contents among geographically distributed datacenters using wavelength division multiplexing (WDM) based inter-datacenter networks.
- The transport resources (network resources) are either own by CSPs or leased.



Inter-Datacenter Network

Lightpaths Based on Unicast Routing

- Currently, lightpaths are established to connect data centers for end-to-end connections, where a lightpath is an all-optical channel from a source to a destination without any O-E-O conversion at intermediate nodes.
- Unicast communication: communication between a single sender and a single receiver over a network, i.e., one-to-one.

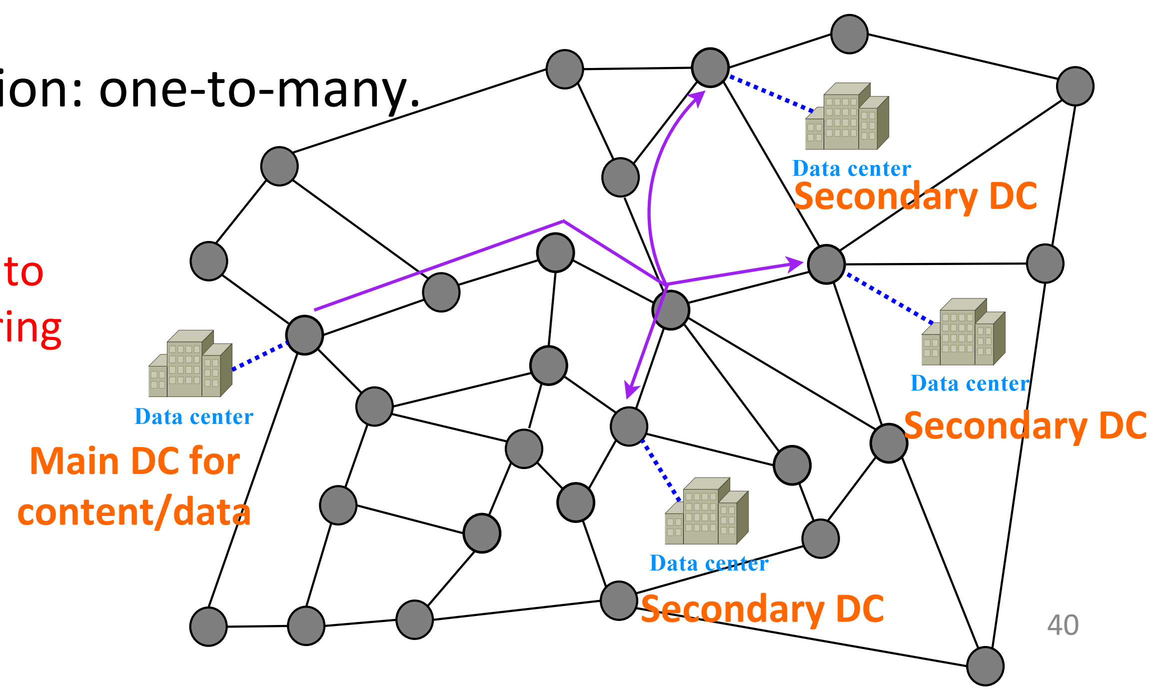


Inter-Datacenter Network

Light-Tree Based on Multicast Routing

- A light-tree is a generalization of a lightpath to be a tree topology, where traffic is sent from a root (main DC) to all its associated leaves (secondary DCs) in the optical domain.
- The optical splitter is needed for implementation of a light-tree.
- An optical splitter can split an optical signal into multiple copies in the optical domain.
- Multicast communication: one-to-many.

Power amplifiers are also needed to compensate for the power lost during the splitting.



Anycast Routing Paradigm for End-user Traffic

- An end-user request for content/service can be served by any data center that supports the specified (by end-user) content and/or service.
- **Anycast routing**: sending a request for data processing, storage, or service delivery without explicitly specifying an intended destination.
- Anycast routing can be used to find an appropriate data center for serving the end-user demand.