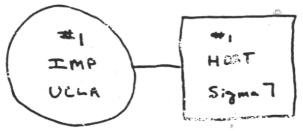
Concept of Computer networks

History of Internet
Concept of computer networks
Network architecture
Packet switching vs. circuit switching

Reading: Chapter 1, Computer Networks, Tanenbaum

History of the Internet



- Originated from an experiemental project of ARPA
- Intially having only two nodes (IMP atUCLA and IMP at SRI).

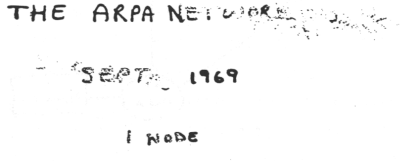
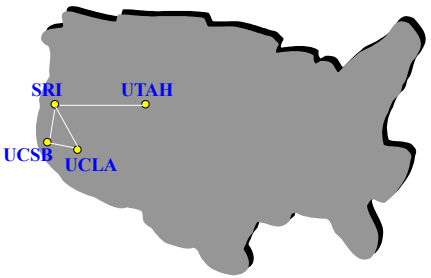


FIGURE 6.1 Drawing of September 1969 (Courtesy of Alex McKenzie)

ARPA: Advanced Research Project Agency
UCLA: University California Los Angeles
SRI: Stanford Research Institute
IMP: Interface Message Processor

Source: http://www.cybergeography.org/atlas/historical.html

In 12/1969, after 3 months



A network with 4 nodes, 56kbps

UCSB:University of California, Santa Barbara
UTAH:University of Utah

source: http://www.cybergeography.org/atlas/historical.html

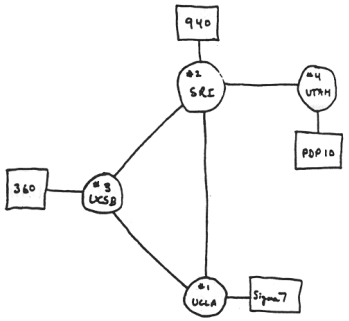
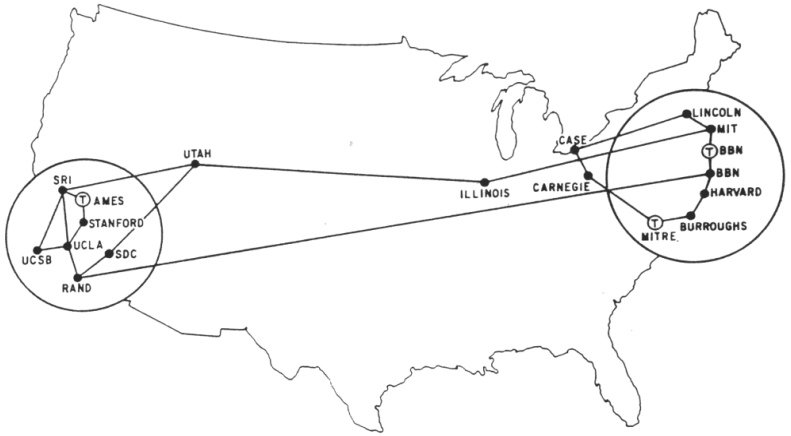


FIGURE 6.2 Drawing of 4 Node Network (Courtesy of Alex McKenzie)

ARPANET, 1971



Source: MAP 4 September 1971 http://www.cybergeography.org/atlas/historical.html

One node was added each month

Years 70s: Interconnection, new network architecture and private architectures

Expansion of ARPANET, 1974

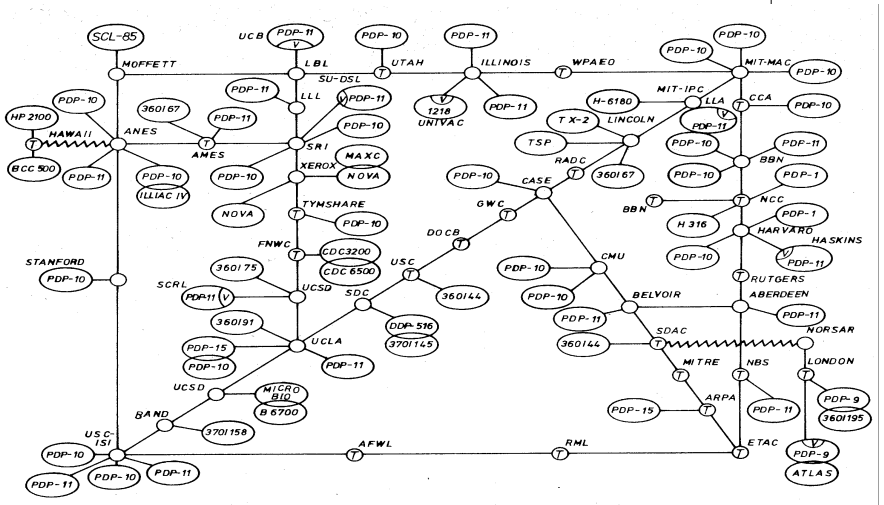


Abb. 4 ARPA NETwork, topologische Karte. Stand Juni 1974.

source:
http://www.cybergeography.org/atlas/historical.html

Traffic each day not more than 3.000.000 package

Years 70s

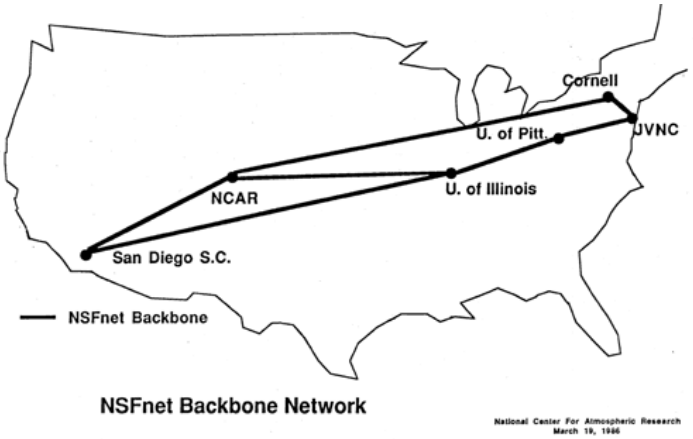
- Since 1970, new networks private architectures appear:
  - ALOHAnet in Hawaii
  - DECnet, IBM SNA, XNA
- 1974: Cerf & Kahn – principles of interconnection of open systems (Turing Awards)
- 1976: Ethernet, Xerox PARC
- End of 1970s: ATM

Years 80s: New protocols, more expansion

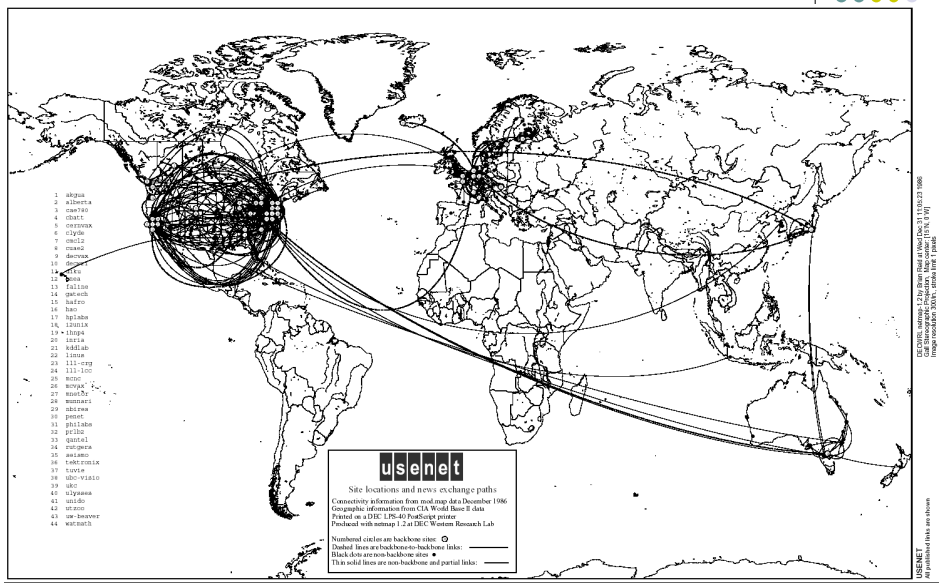
1981: Beginning of NSFNET

NSF: National Science Foundation

NSF network is separated from ARPANET for academic research uniquely



1986: Connect USENET and NSFNET



Source: http://www.cybergeography.org/atlas/historical.html

More network to join and more protocol

- More networks join in: MFENET, HEPNET (Dept. Energy), SPAN (NASA), BITnet, CSnet, NSFnet, Minitel …
- TCP/IP is standardized and becomes popular in 1980
- Berkeley integrate TCP/IP in BSD Unix
- Services: FTP, Mail, DNS …

Years 90s: Web and E-commerce over Internet

Years 90s

- Begining of 90s: Begining of Web
  - HTML, HTTP: Berners-Lee
  - 1994: Mosaic, Netscape
- End of 90s: Commercialized the Internet

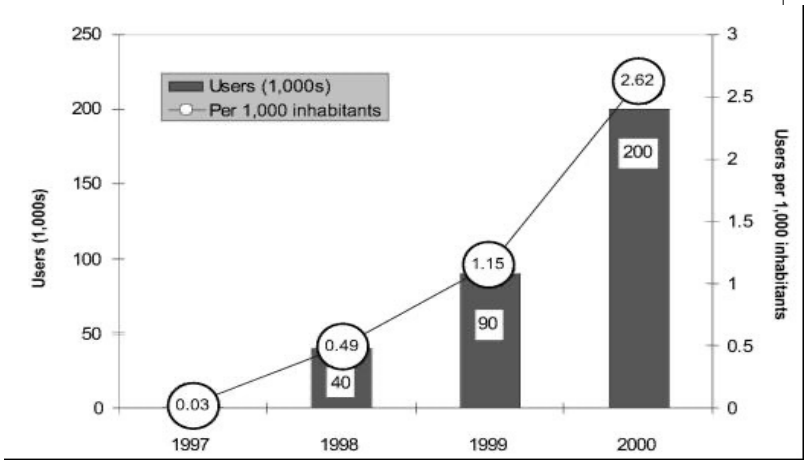
End of 1990's – 2000's:

- Many new Internet applications was introduced:
  - Chat, file sharing P2P…
  - E-commerce, Yahoo, Ebay, Amazon, Google…
- > 50 millions hosts, > 100 millions users.

Internet in Việt Nam

- 1996: Preparation for the Internet infrastructure
  - ISP: VNPT
  - 64kbps, 01 connection to the world, few end users.
- 1997: Việt Nam connects to the Internet officially
  - 1 IXP (Internet Exchange Point): VNPT
  - 4 ISP (Internet Service Provider) : VNPT, Netnam (IOT), FPT, SPT
- 2007: After 10 years
  - 20 ISPs, 4 IXPs: VNPT, FPT, Viettel, EVN Telecom
  - 19 mil. users, 22.04% population

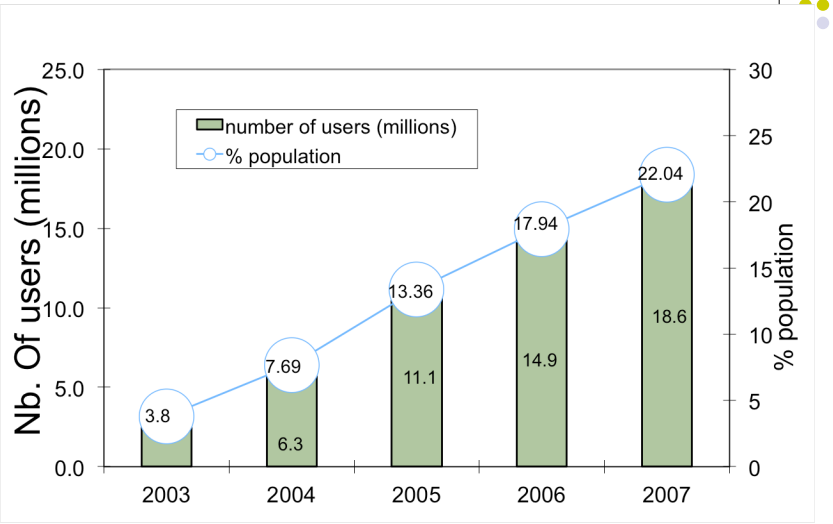
Development of the Internet in Vietnam



The numbers of users are estimated by 2 times the number of subscribers

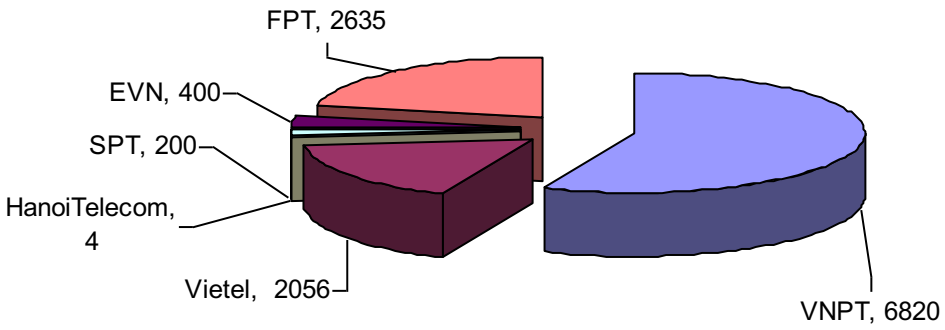
Source: Vietnam Internet Case Study, http://www.itu.int/asean2001/reports/material/VNM%20CS.pdf

Statistics until 2007



Source: *Vnnic*, http://www.thongkeinternet.vn

Bandwidth to the world (Mbps), 3rd Quarter 2007



Total: 12115.0 Mbps

Internet subscription, 2019

5 2019 Xem

Tình hình phát triển thuê bao băng rộng cố định tháng 5/2019	
Số thuê bao truy nhập Internet qua hình thức xDSL:	182,853
Số thuê bao truy nhập Internet qua kênh thuê riêng:	22,929
Số thuê bao truy nhập Internet qua hệ thống cáp truyền hình (CATV):	868,039
Số thuê bao truy nhập Internet qua hệ thống cáp quang tới nhà thuê bao (FTTH):	12,606,506
Tổng số thuê bao băng rộng cố định:	13,680,327

Statistics are provided by Department of Telecommunication, Ministry of Information and Communication.
http://vnta.gov.vn/thongke/Trang/dulieuthongke.aspx

Some fixed internet connection technologies to ISP

- Dial-up:
  - 56kbps,
  - use public telephone lines,
  - Data are transmitted over the same frequency with voice,
  - Old technology, popular before 2000
- ADSL, xDSL:
  - few Mbps,
  - use public telephone lines,
  - Data are transmitted over the different frequency with voice,
  - popular between 2000-2010

Some fixed internet connection technologies to ISP

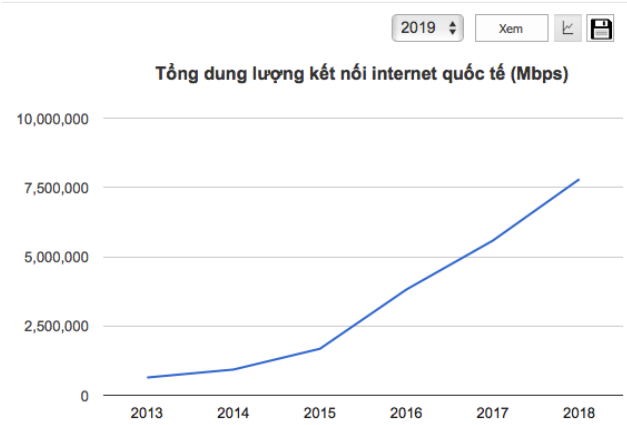
- Internet over TV cable
  - Use TV cable to carry data
- FTTH
  - several dozen Mbps,
  - Use optical fiber
  - Popular nowadays.

Internet usage on Mobile phone 2019

Tình hình phát triển thuê bao điện thoại di động tháng 5/2019	
▸ Tổng số thuê bao điện thoại di động có phát sinh lưu lượng:	133,877,535
▸ Tổng số thuê bao điện thoại di động đang hoạt động chỉ sử dụng thoại, tin nhắn:	75,216,569
▪ Thuê bao trả trước:	70,448,710
▪ Thuê bao trả sau:	4,767,859
▸ Tổng số thuê bao điện thoại di động đang hoạt động có sử dụng dữ liệu:	58,660,966
▪ Thuê bao trả trước:	54,158,129
▪ Thuê bao trả sau:	4,502,837

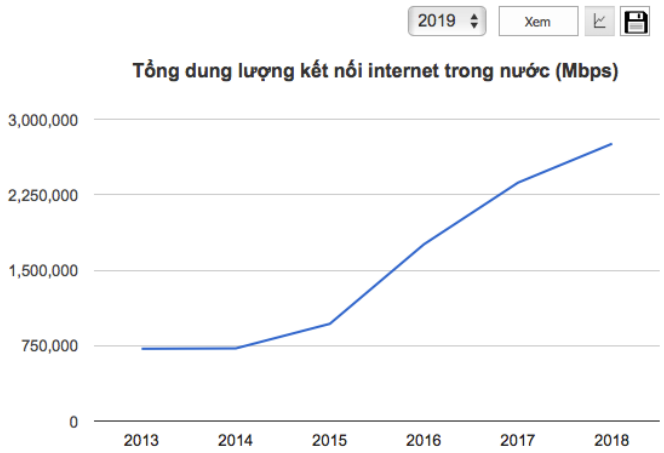
Statistics are provided by Department of Telecommunication, Ministry of Information and Communication.
http://vnta.gov.vn/thongke/Trang/dulieuthongke.aspx

International Internet data volume 2019



Statistics are provided by Department of Telecommunication, Ministry of Information and Communication.
http://vnta.gov.vn/thongke/Trang/dulieuthongke.aspx

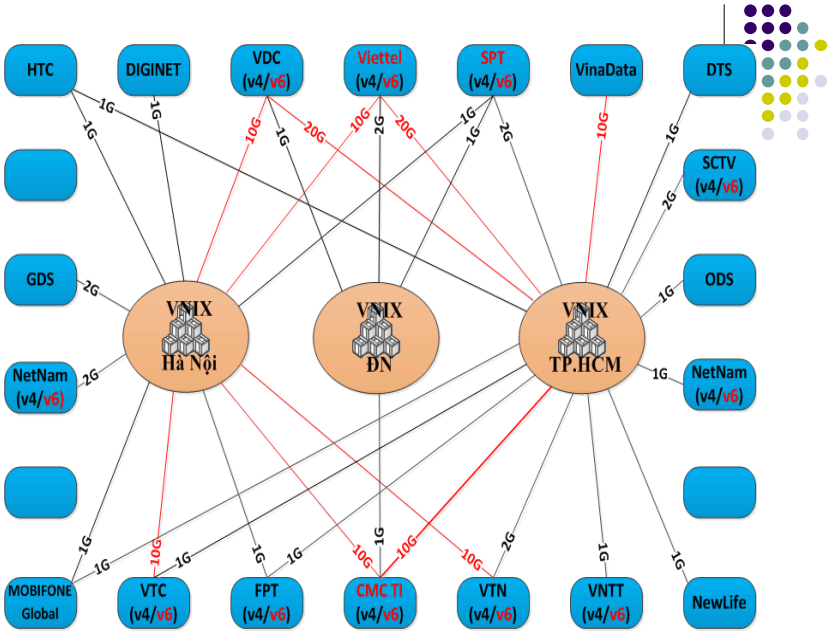
Domestic Internet data volume



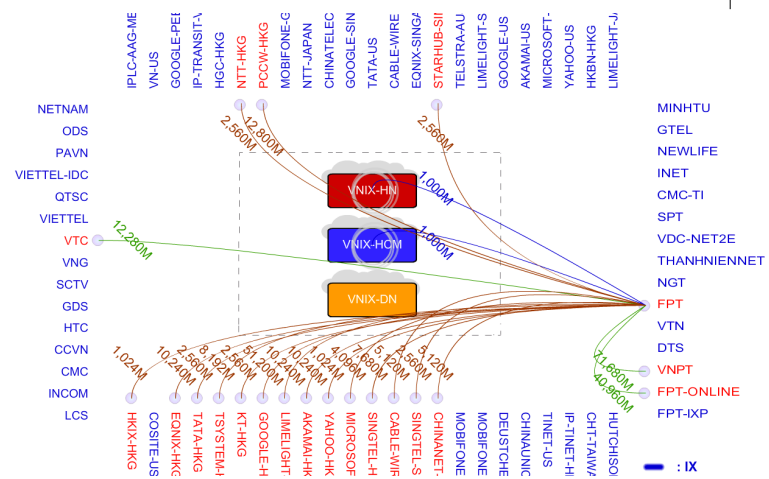
Statistics are provided by Department of Telecommunication, Ministry of Information and Communication.
http://vnta.gov.vn/thongke/Trang/dulieuthongke.aspx

Internet management in Việt Nam

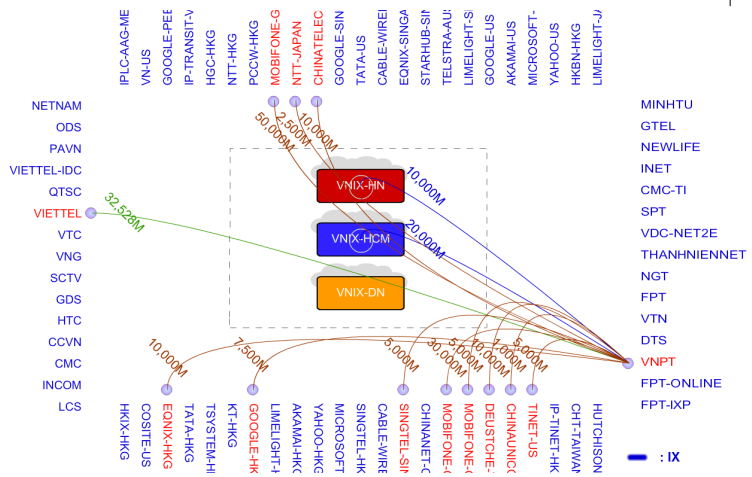
- VNNIC
  - Is responsible for managing the Internet domain name, address in Việt Nam;
  - Provides guidelines, statistics about Internet and participates in international activity about Internet.
- **VNIX: V**ietnam **N**ational **I**nternet e**X**change
  - switching system between national ISP.



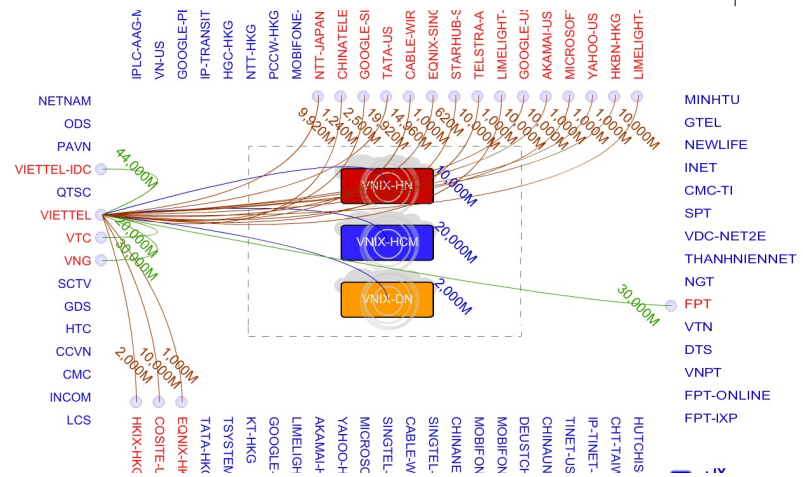
International connections



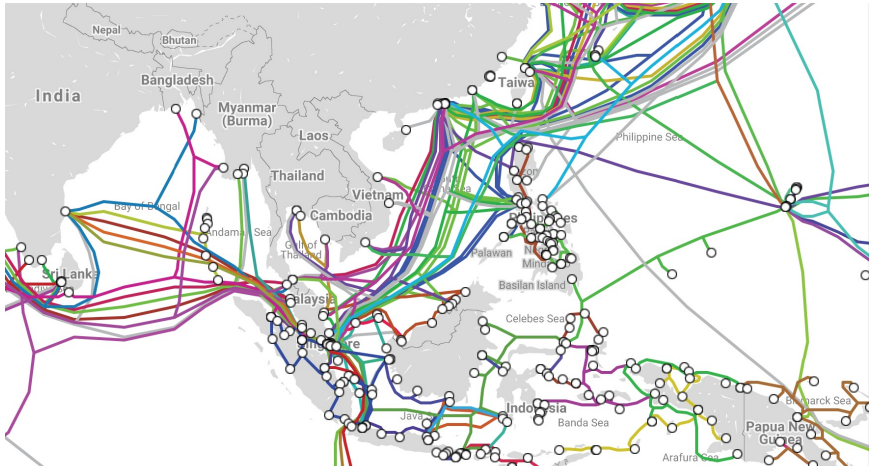
International connections



International connections

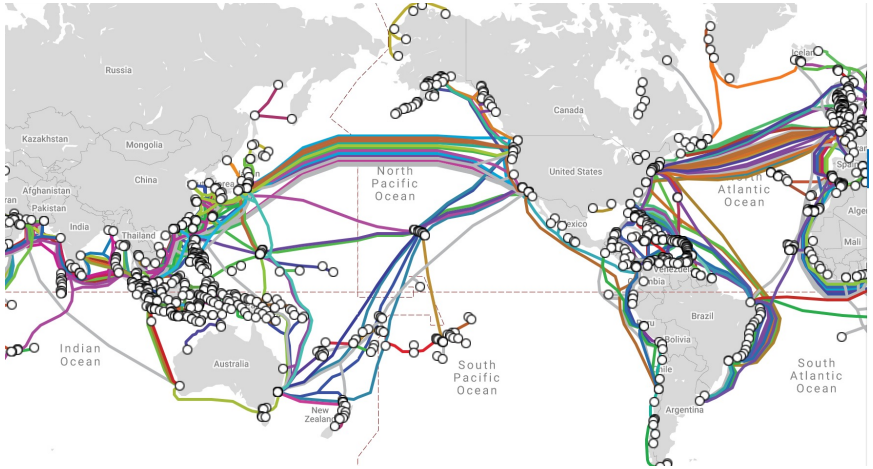


Optical fiber under the ocean



Source: https://www.submarinecablemap.com

Optical fiber under the ocean

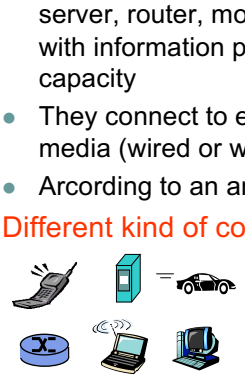


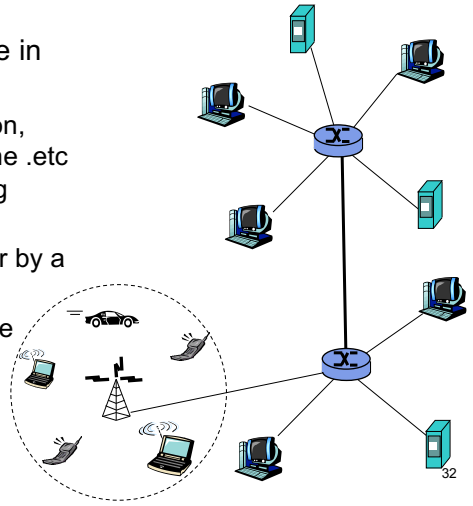
Source: https://www.submarinecablemap.com

Concepts of computer networks

Concepts

- A set of computers/nodes connecting to each other according to an architecture in order to exchange data
  - Computer/node: workstation, server, router, mobile phone .etc with information processing capacity
  - They connect to each other by a media (wired or wireless)
  - Arcording to an archirecture
- Different kind of computers



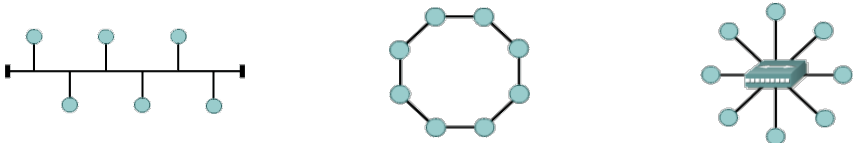


Example of computer networks

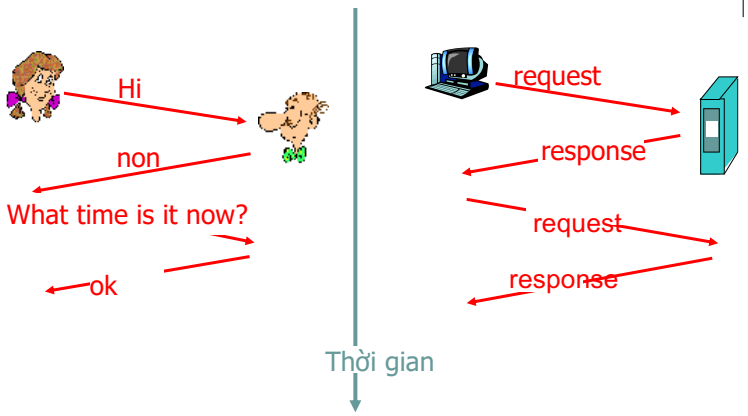
- The Internet
- A local network using Ethernet
- An wireless LAN in a cafe: using 802.11 standard
- A network connecting ATMs

Network architecture

- Network architecture contain 2 aspects:
  - topology: the form that network nodes connects to each other
  - Protocol: language and procedure of communication between nodes.
- Topology
  - Bus, Ring, Star…



What is a protocol?



Protocol between human being: vocabulary, procedure

Protocol between machines

Network protocol

- A protocol defines communication rules between nodes
- Protocol defines:
  - Format of messages/ information to be exchanged between nodes.
  - Order of messages sending between entities/nodes
  - Action should be performed when an entity receives a message.
- Example of protocols running on the Internet: TCP, UDP, IP, HTTP, Telnet, SSH, Ethernet, …

Communication medium

- Physical medium that can carry signal
- Classification:
  - Wired media: twisted pair, coaxial cable, optical fiber,…
  - Wireless media: radio wave, microwave, infrared wave,…
- Some characteristics:
  - Bandwidth (băng thông): width of the frequency band could be used for carrying signal
    - f_{min}: minimum frequency, f_{max}: maximum frequency
    - Bandwidth = f_{max} - f_{min}
  - BER – Bit Error Rate= nb of error bits/nb of transmitted bits)
  - Attenuation (suy hao): signal power decrement level

Computer network classification

- PAN – Personal Area Network
  - Scope: several metres
  - #users: few
  - To serve an individual
- LAN – Local Area Network):
  - Scope: few km
  - #users: few to hundreds of thousands
  - To serve an individual, house hold, organization

Computer network classification

- MAN – Metropolitian Area Network
  - Scope: hundreds of km
  - #users: Millions
  - To serve a metro, area
- WAN – Wide Area Network
  - Scope: thousands of km
  - #users: billions
- GAN – Global Area Network: over the world (ex: Internet)

LAN

- LAN (Local Area Networks):
  - Scope: a building, an office, an organization
  - Wireless LAN
    - VD: WIFI
  - Wired LAN
    - VD: Ethernet

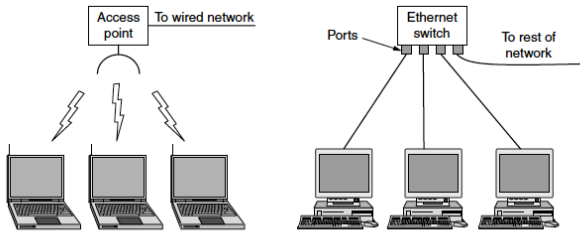


Figure 1-8. Wireless and wired LANs. (a) 802.11. (b) Switched Ethernet.

MAN

- Metropolitan Area Networks
  - Cover a city
  - Ex:
    - Television network
    - Backbone networks of ISP.

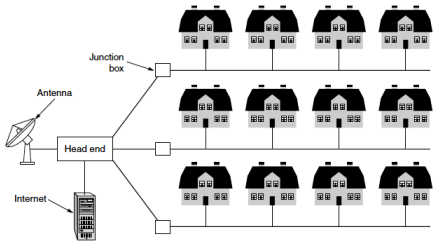


Figure 1-9. A metropolitan area network based on cable TV.

WAN

- Wide Area Networks
- Cover a large scope such as a country
  - Ex: network connecting different branchs of the same company
- Technology characteristics:
  - Using long distant lines to connect different parts of the network
    - Ex: Using PSTN network, using optical cable.

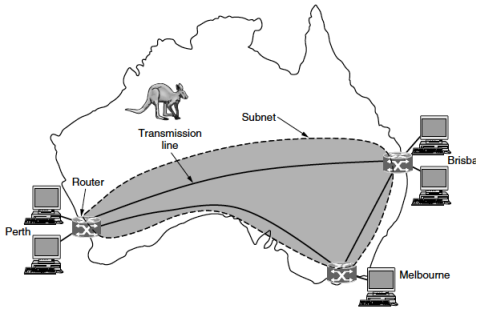


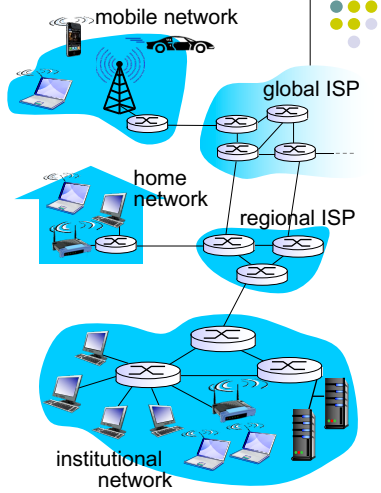
Figure 1-10. WAN that connects three branch offices in Australia.

Mạng GAN

- Global Area Networks
- Interconnect different networks
- Cover many continentals.

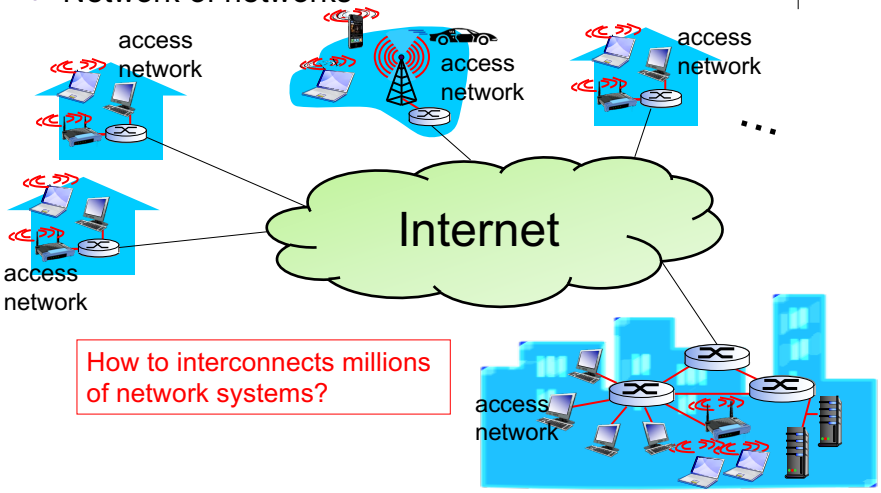
Internet

- Contain more than 5 billions devices
- 3.2 billion users (40%)
- Medium: optical fiber, twisted pair, Wimax, 3G…
- Transport ~3x109 GB data per day
- Services: Web, email, social networks, …



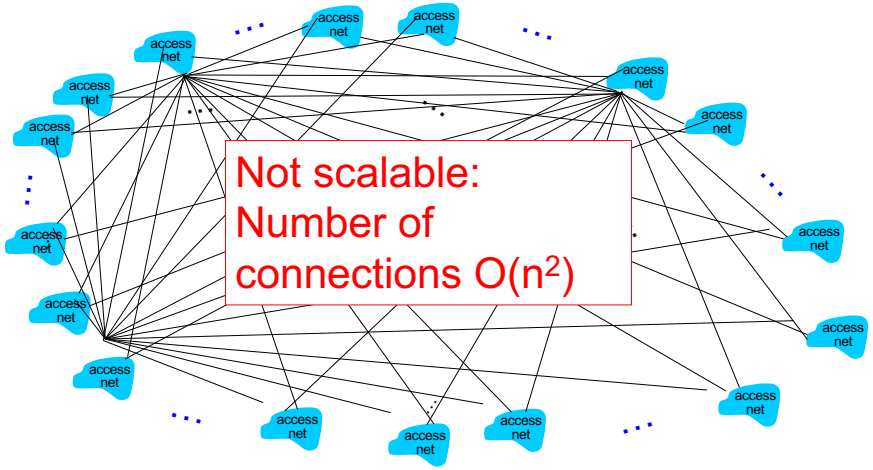
Internet

- Network of networks



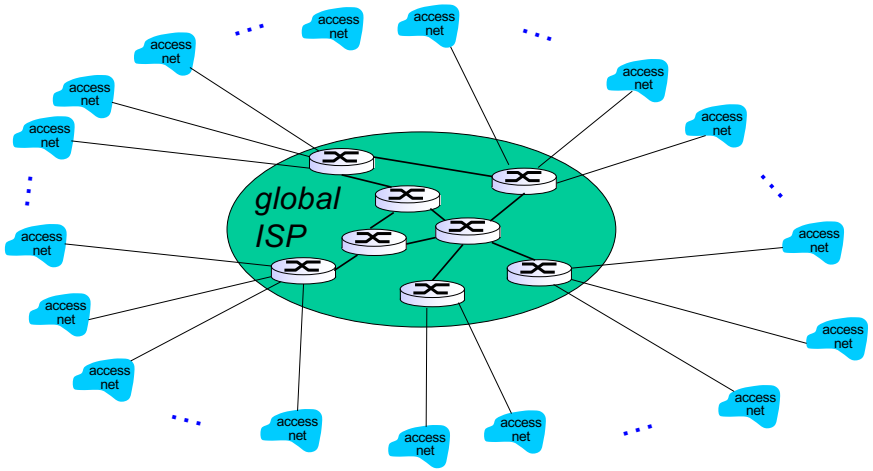
Internet: network of networks

- *Directly connect one network with all others?*



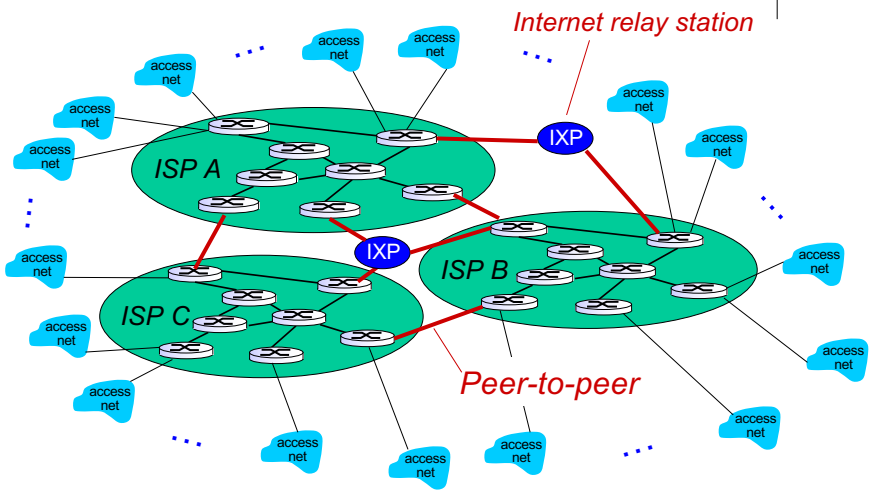
Internet: Network of networks(2)

- *Connect each access network to a relay stations of a global ISP*



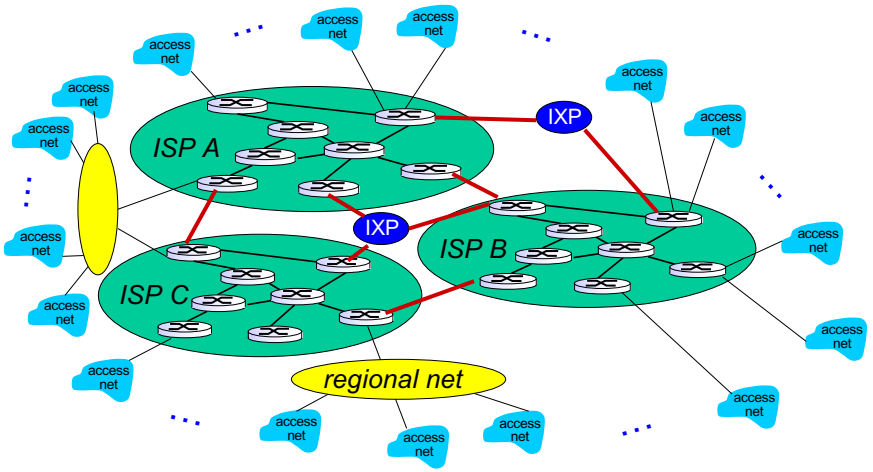
Internet: Network of networks(3)

- *Add more ISP...*



Internet: Network of networks(4)

- *Add regional networks...*



Transmission models

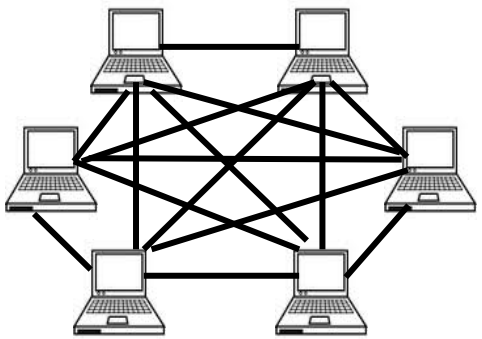
Packet switching vs. Circuit switching

Connection oriented vs. Connectionless

Connecting hosts

Direct links model

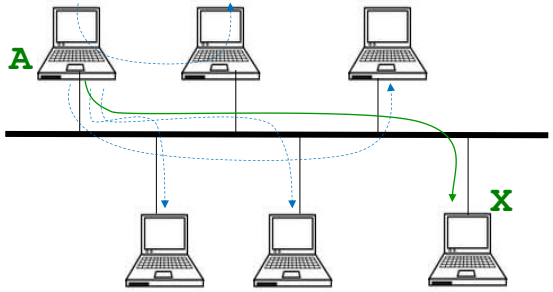
- Using direct links between all pairs of hosts
  - A link: a segment of medium without any processing unit in the middle
  - Weakness: too many links, distance limitation.



Connecting hosts

Bus model

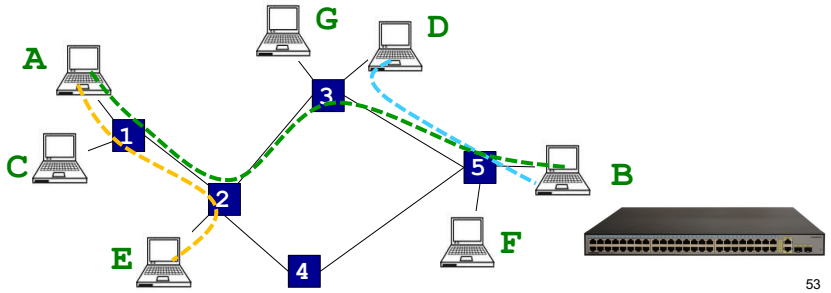
- Point-to-multipoint:
  - Single communication medium is used for all hosts → broadcast communication
  - Weakness: long physical link, few hosts can communicate simultaneously



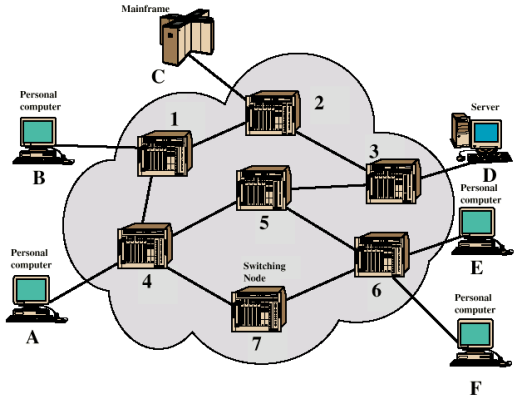
Connecting hosts

Switching model

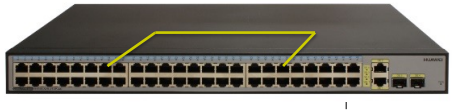
- Solution: using switch
  - Switch: device with multiple ports
  - A host links to a switch
  - Switches link together point-to-point
  - Switch forwards data/signal between ports toward destination.



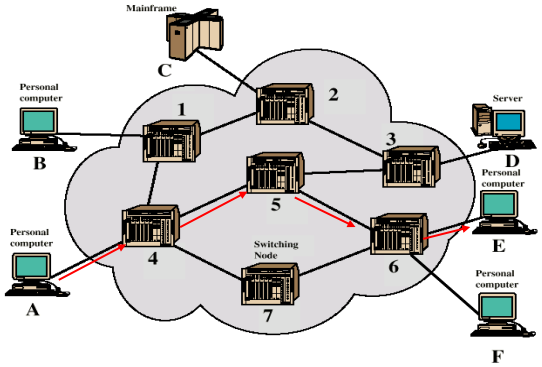
Data switching network



Circuit switching



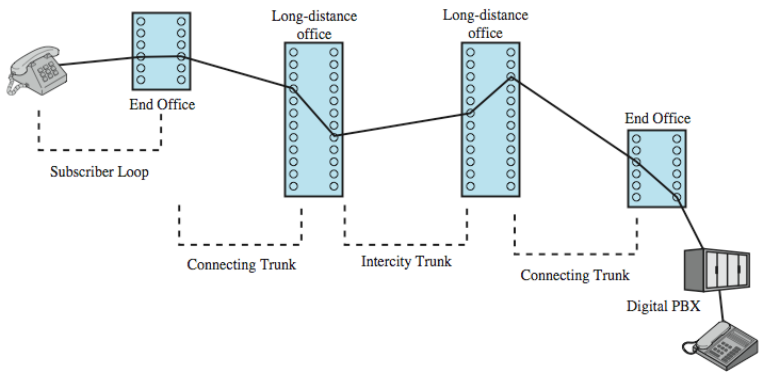
- A switch closes two ports together, making data from in-port to flow to out-port.
- Circuit is a path/channel, going through several switches, over which data flows



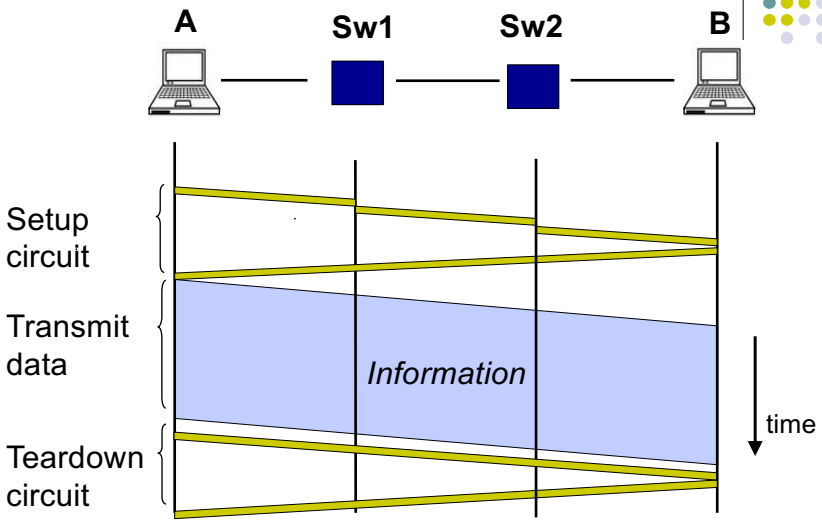
Circuit switching

- Resources (ex: bandwidth over a link) is dedicatedly assigned to each circuit. When the circuit is unused (no data is transmitted), no other circuit can use the resources.
- 3 phases of data transmission
  - Setup circuit
  - Transmit data
  - Teardown the circuit
- Circuit switching guaranties that the circuits uses the whole available the bandwidth over each link for data transmission (good for audio/video transmission)
- Waste of bandwidth if the data transmission process does not consume the whole capacity of each link of the circuit.

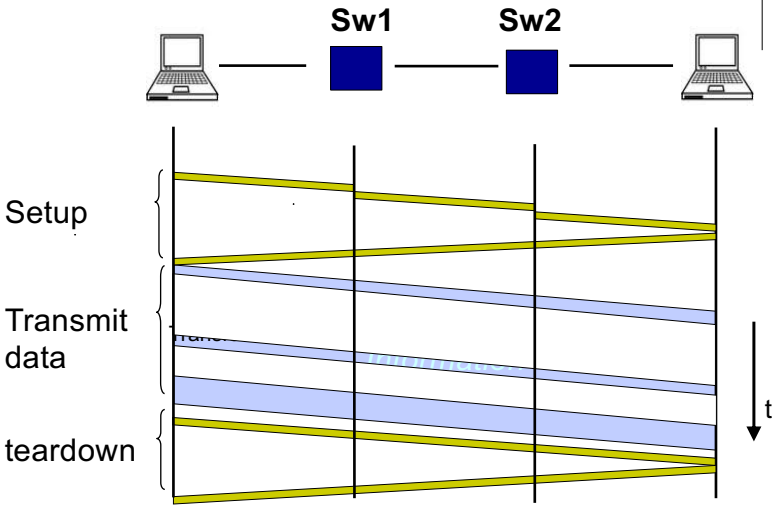
Example of circuit switching : Public Switched Telephone Network PSTN



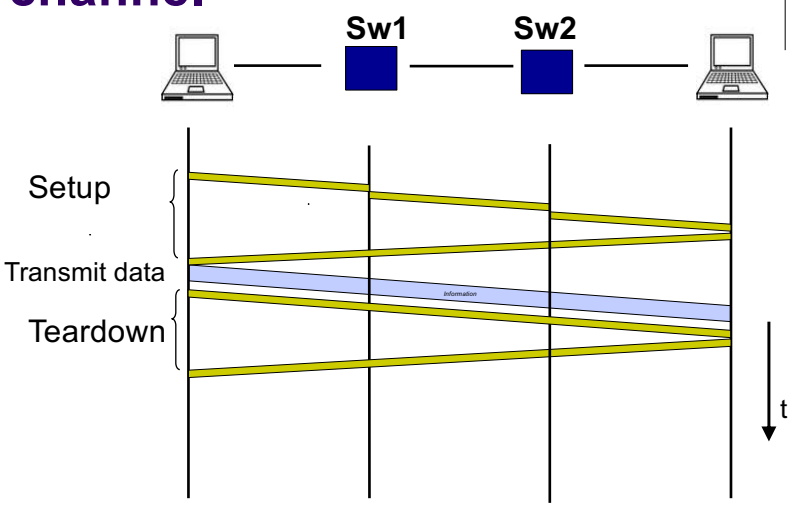
End-to-end data transmission time in circuit switching



Weakness: case of idle chanel



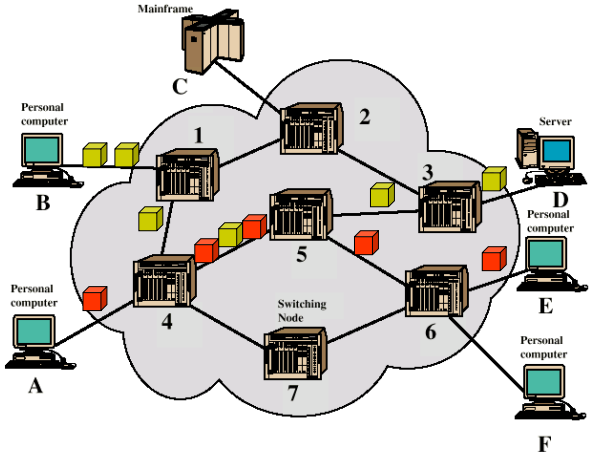
Weakness: case of small channel



Packet switching

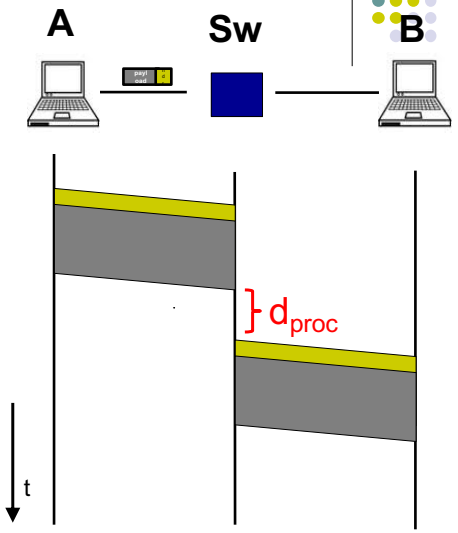
- Data is divided into small packets and transferred through the network
- Switch does not close one port to another but just copies a packet between ports.
- Multiple connections can share a single channel
  - Increase bandwidth utilization efficiency
- Each packet is forwarded individually

Example of packet switching



Transmission time in packet switching

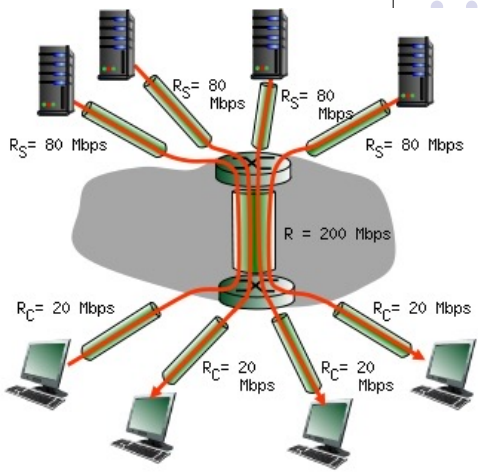
- Switch forwards a packet only after receiving all the packet (store and forward)
- Switch need time to process a packet (d_{proc}):
  - Check error
  - Decide which ports to forward packet out
  - d_{proc} is usually smaller than transmission delay



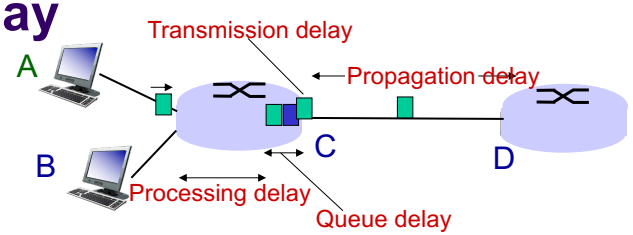
Bandwidth vs throughput

- Bandwidth - R
  - In telecommunication: bandwidth= f_{max} - f_{min}
  - In computer networks: Maximum amount of data can be transmitted in a unit of time over a link (bps – bit per second).
  - Ex: optical fiber has bandwidth of 1000Mbps.
- *Throughput:* actual data transmission speed (bits/sec)

Bottle neck



Delay



$d_{nodal} = d_{proc} + d_{queue} + d_{trans} + d_{prop}$

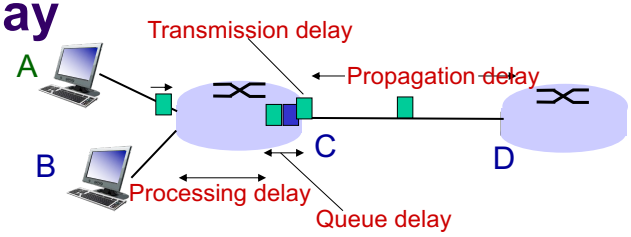
d_{trans}: transmission delay:
Time to send data out of a node

- L: data size(bits)
- R: bandwidth(*bps*)
- $d_{trans} = L/R$

d_{prop}: propagation delay
Time to propagate data from one end of link to the other

- d: length of link
- s: signal propagation speed in medium (ex: ~$2x10^8$ m/sec)
- $d_{prop} = d/s$

Delay



$d_{nodal} = d_{proc} + d_{queue} + d_{trans} + d_{prop}$

d_{proc}: processing delay

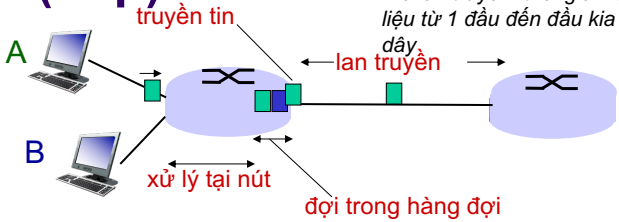
- Error check
- Identify out port
- Usually < µsec

d_{queue}: queue delay

- Time data stay in queue waiting for processing
- Depending on the amount of data in the queue.

Độ trễ (tiếp)

Trễ trên thiết bị đầu cuối
Trễ trên thiết bị trung gian
Trễ truyền tin: thời gian cần để phát dữ liệu
Trễ lan truyền: thời gian lan truyền dữ liệu từ 1 đầu đến đầu kia của đường dây



$d_{nodal} = d_{proc} + d_{queue} + d_{trans} + d_{prop}$

d_{proc}: trễ xử lý

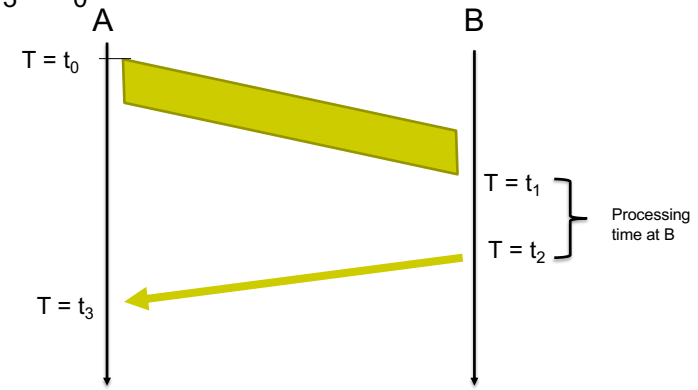
- Kiểm tra lỗi bit
- Xác định liên kết ra
- Thường < µsec

d_{queue}: trễ hàng đợi

- Thời gian dữ liệu nằm trong hàng đợi chờ xử lý
- Phụ thuộc vào lượng dữ liệu trong hàng đợi

Round Trip Time (RTT)

- RTT: = $t_3 - t_0$



- One way delay: $t_1 - t_0$

Connection oriented transmission vs. connectionless

- Connection oriented transmission:
  - Data are transmitted over a connection already established
  - 3 working phases: Establishing a connection, data transmission, teardown the connection.
  - Reliable
- Connectionless transmision
  - No connection establishing phase
  - Only data transmission phase
  - Not reliable - “Best effort”

Summary

- Introduction to the course
- History of the Internet
- Concept of Computer Networks
- Architecture
  - Topology
  - Protocol
- Circuit switching vs. packet switching
  - Pros & cons

Next week…

- Layering architecture
- OSI reference model
- IP, MAC address, port number
- DNS service.

Lecture 2:
Basic concepts of computer networks

Last lecture

- Introduction of the course
- History of the Internet
- Concept of Computer Networks
- Some fundamental concepts: switching, connection oriented, connection less.

Content of this lecture

- Layer architecture
- OSI & TCP/IP reference model
- Addressing
- Domain name and conversion/resolution of domain name

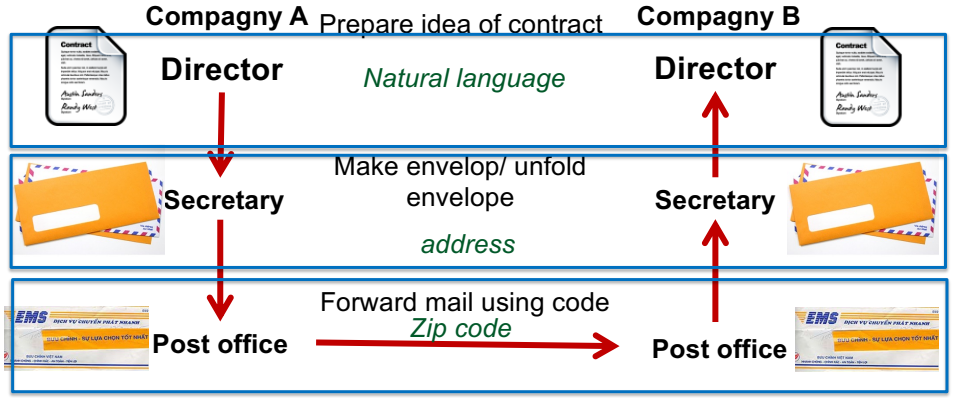
Layer architecture

Devide and conquer principle

- Big work is divided into small tasks
- Assign some tasks to individuals
- Ex: Compagny A & B needs to discuss about a contract
  - Director of A,B: Identify the main points of the contracts & ask secretary to write down the contract.
  - Secretary:
    - Format the contract, put contract to envelope, write down the address of company B
    - Ask post office (VNPT) to send to company B
  - Post office:
    - Forward the envelop through several hub of post then to B

Example

- Parties at the same level performs similar tasks and use the same information communication methods.



Advantage of layering systems

- For the complex system: principle of *" devide and conquer"*
- Allow to determine the responsibility of each layer and the relationship amongst them
- Allow to maintain and upgrade easily the system
  - Changes in some parts do not influence the other parts.
  - Ex: upgrade a media lecture from CD lecture to DVD lecture without the need to change speakers.

Example of layers

Architecture with layers

Sound system

Player
Speaker
Amplifier

Architecture without layers

Cassette

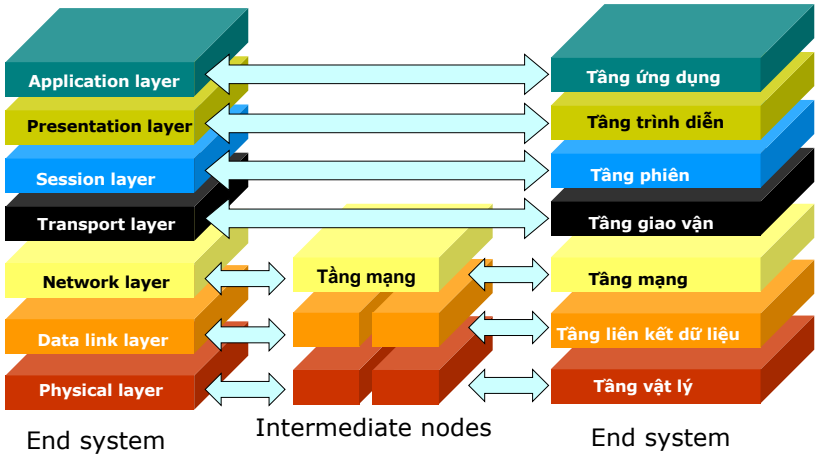
All functionalities are put on the same box
When we want to upgrade: Upgrade the whole box

Reference models

OSI

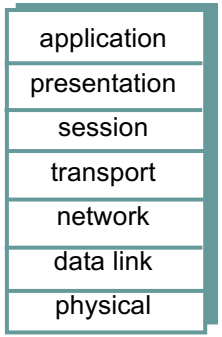
TCP/IP

OSI - Open System Interconnection: 7 layers



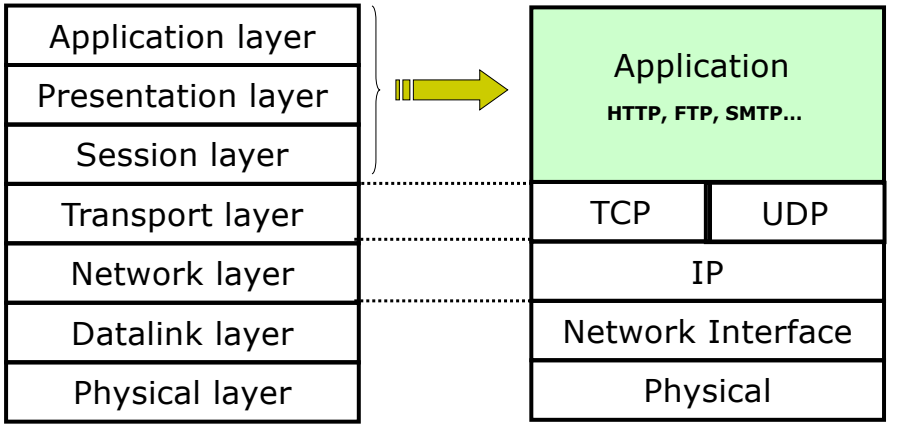
The main functionality of each layers

- Physical layer: Transfering bits "over medium"
- Datalink layer: Transfering data between direct connected elements in the networks.
- Network layer: Routing, forwarding data from the source to the distant destination
- Transport: Transmitting data between applications
- Session : synchronization, check-point, recovery of transmission process
- Presentation: data encoding, compression, data conversion…
- Application: Supporting communications between distant parts of an application.



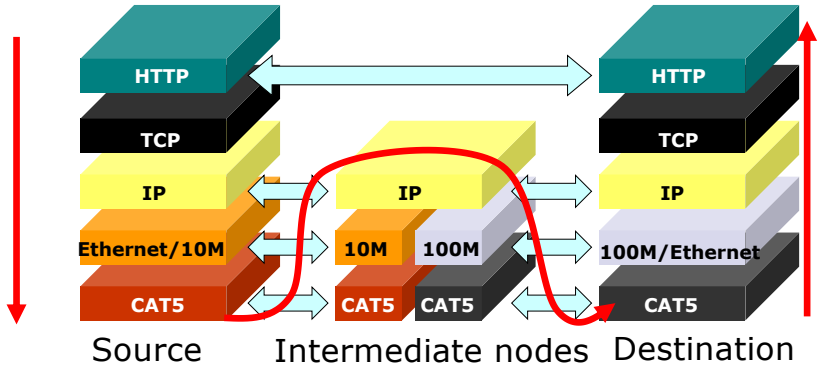
Models OSI and TCP/IP

In the TCP/IP model of the Internet, the functionalities of 3 first layers are combined in a single layer.



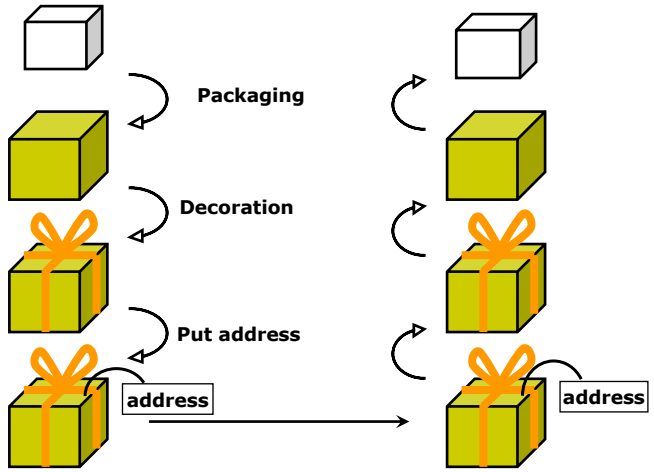
Layering model of the Internet

Example of data transmission from a source to a destination through intermediate nodes (router)



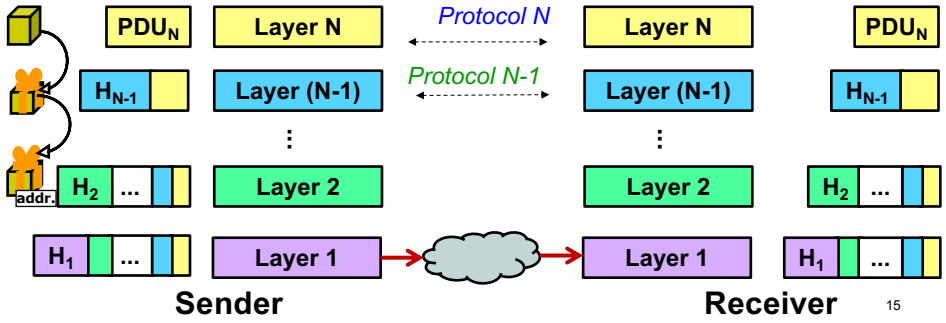
Data Encapsulation

Data encapsulation is similar to a packaging process for a gift.

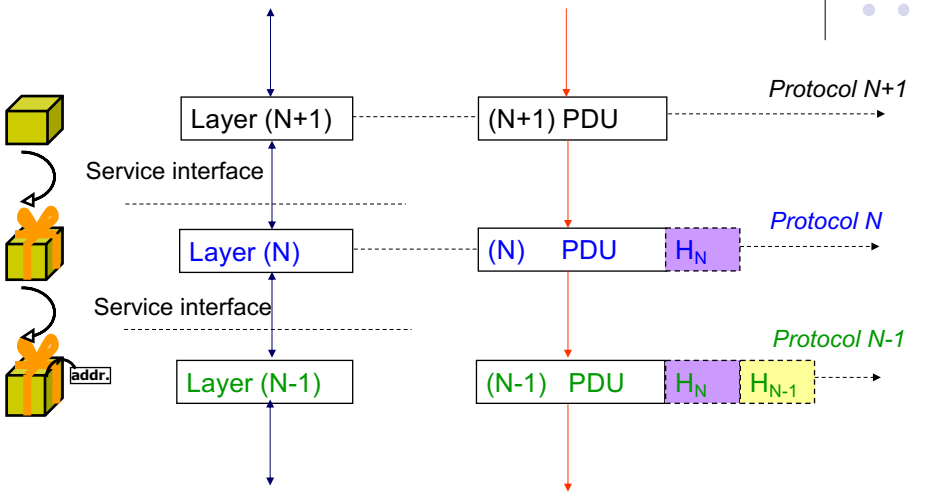


Data Encapsulation

- Sender side: Add header containing the information necessary for package processing at that layer, then send packet to the lower layer.
- Receiver side: Process data in the package according to information in the header, remove the header and send data to the upper layer.

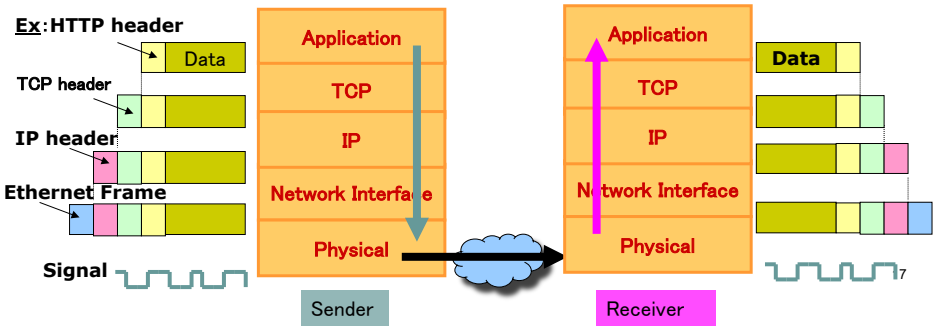


PDU: Protocol Data Unit

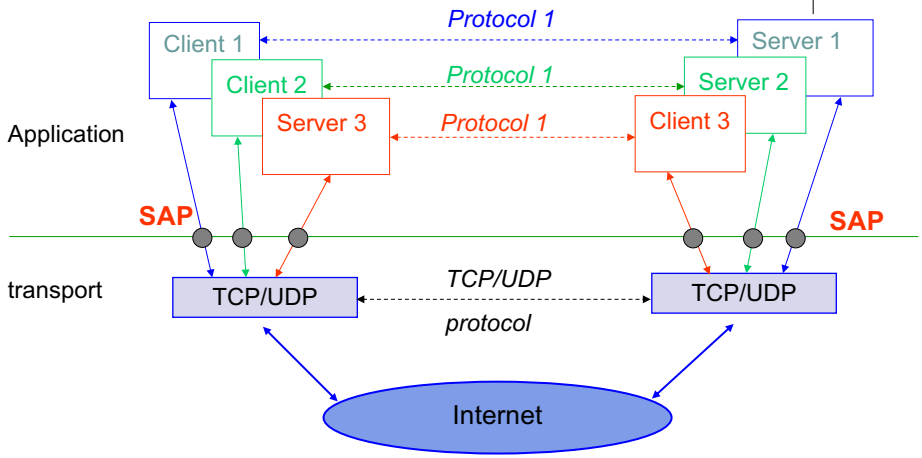


The protocols TCP/IP and encapsulation process

- At sender
  - Each layer add control information to the header of packet and transfer to the lower layer.
- At receiver
  - Each layer process packet according to the information of the header, then remove the corresponding header and deliver the remaining data to the upper layer.

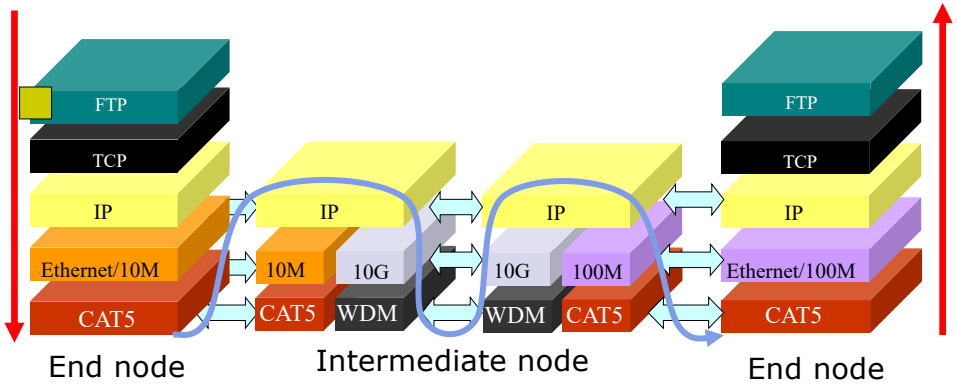


SAP: Service Access Point



18

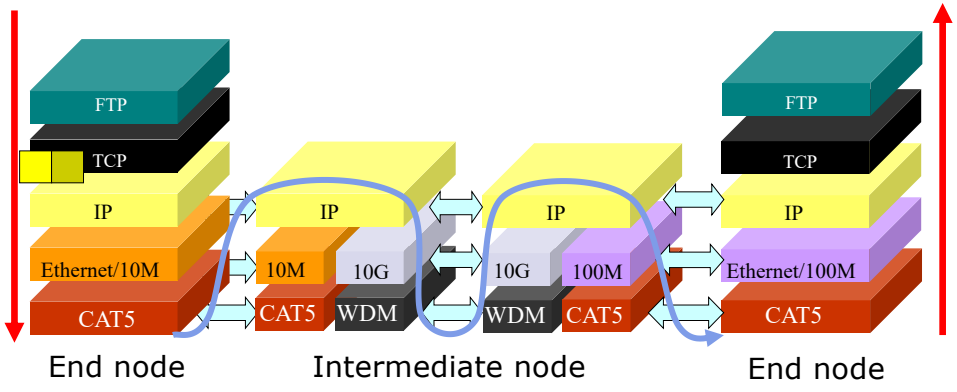
Protocol stack and encapsulation



Dữ liệu - payload

19

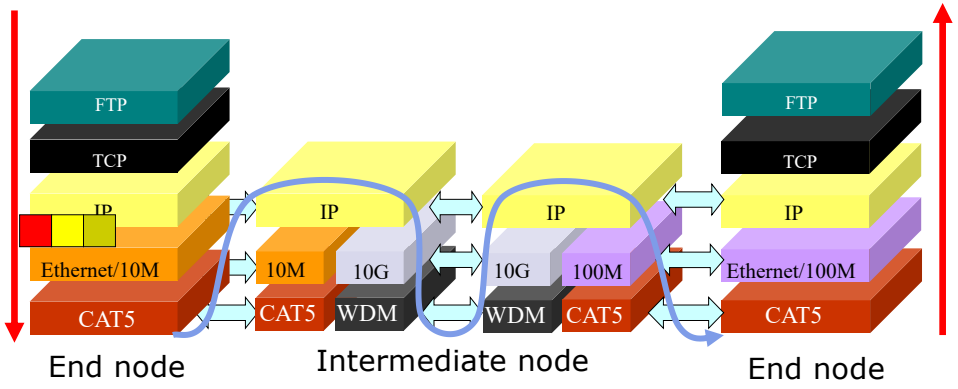
Protocol stack and encapsulation



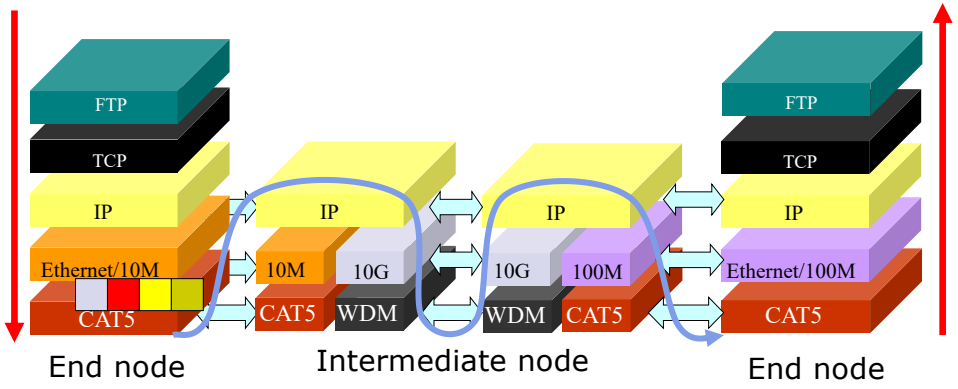
TCP header | Dữ liệu - payload

20

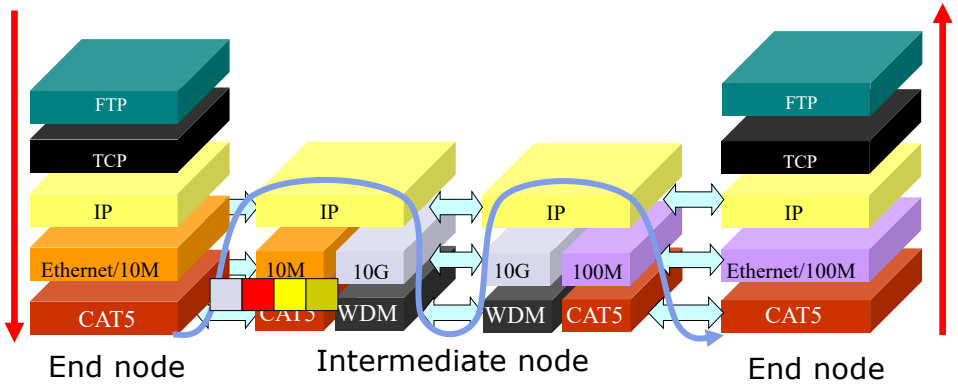
Protocol stack and encapsulation



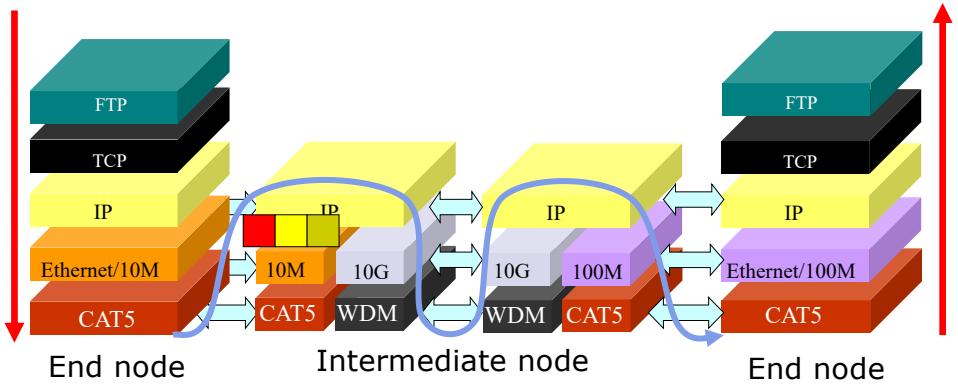
Protocol stack and encapsulation



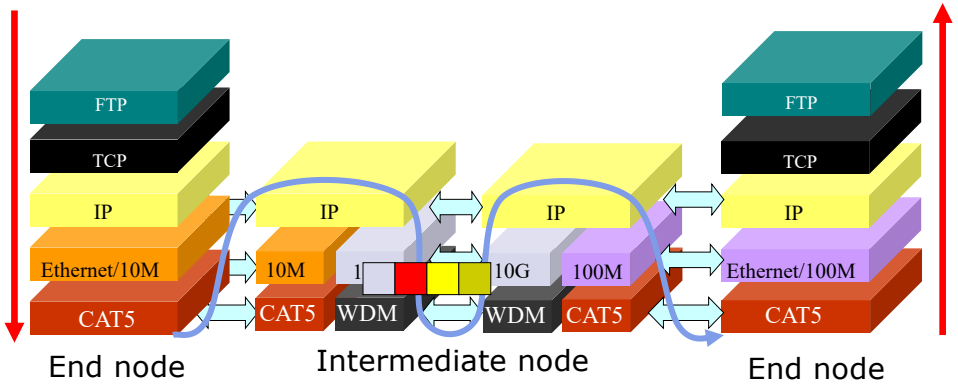
Protocol stack and encapsulation

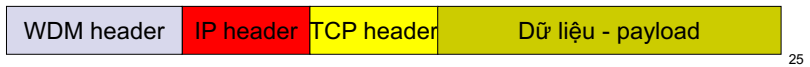


Protocol stack and encapsulation

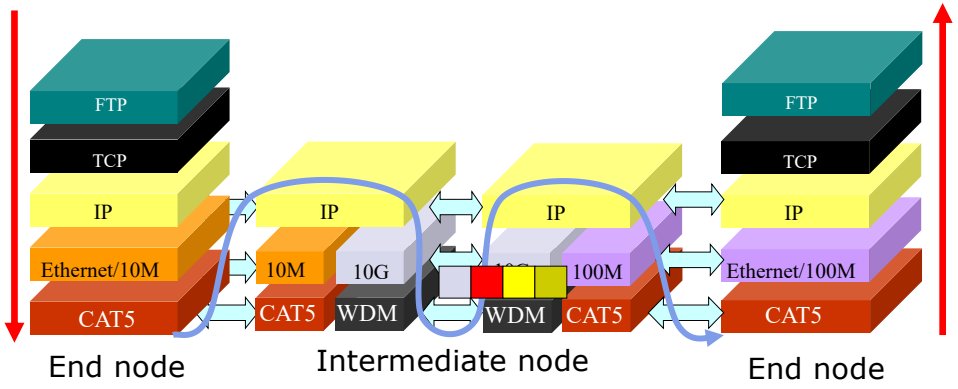


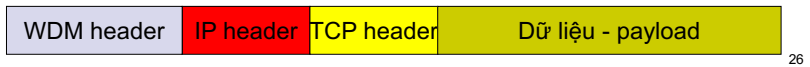
Protocol stack and encapsulation



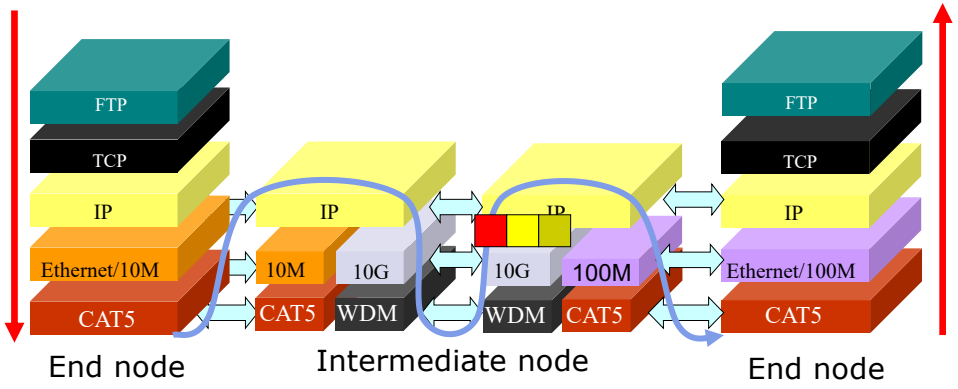


Protocol stack and encapsulation

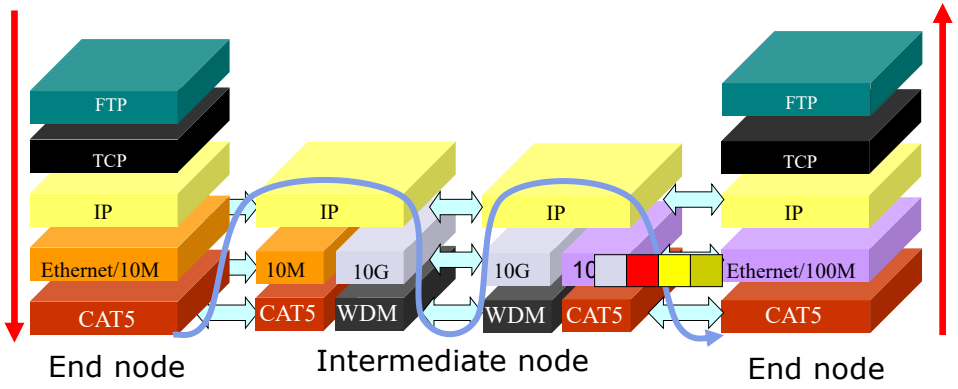


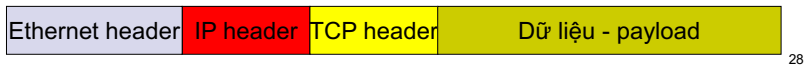


Protocol stack and encapsulation

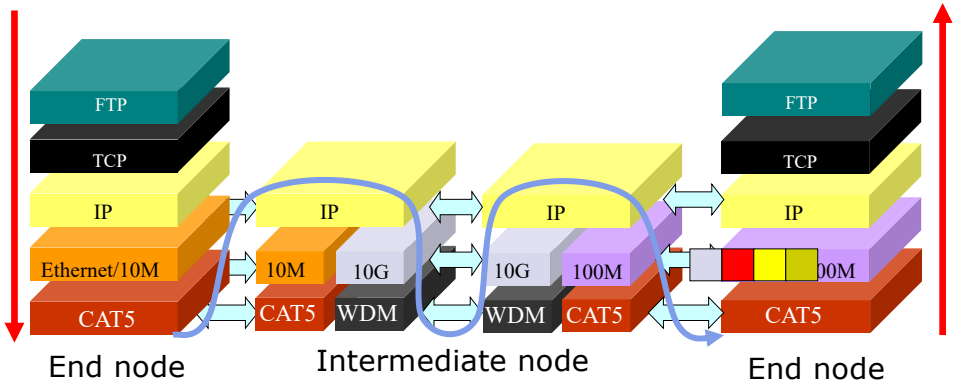


Protocol stack and encapsulation

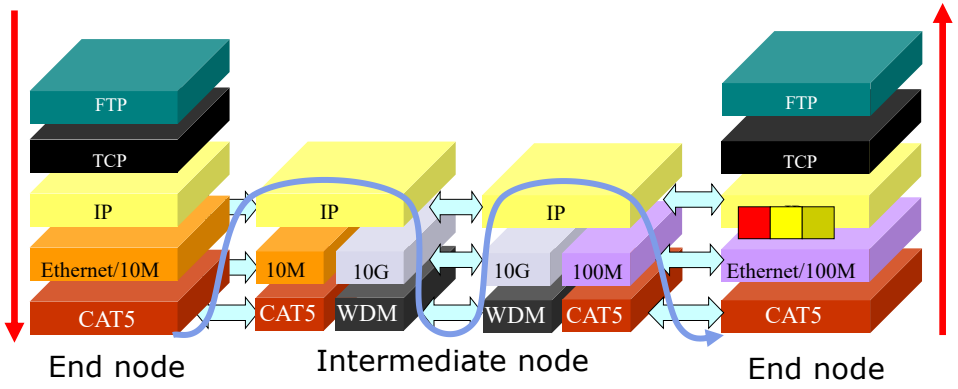




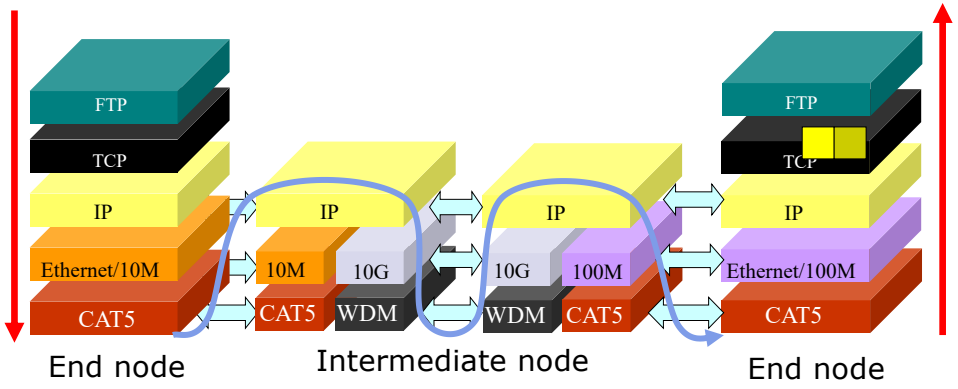
Protocol stack and encapsulation



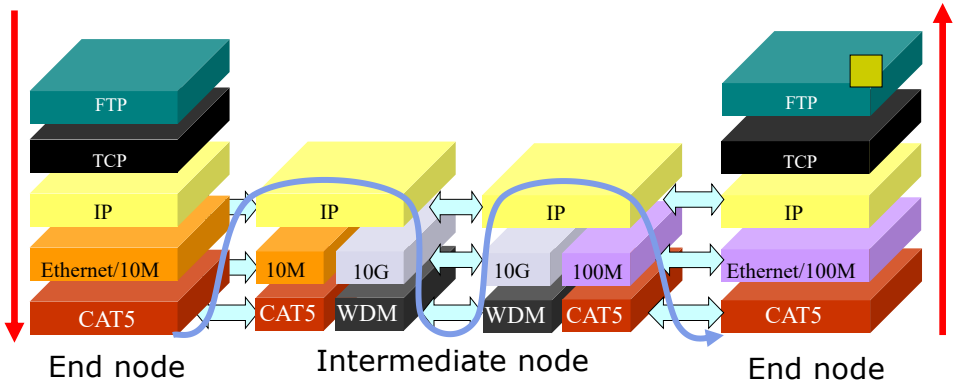
Protocol stack and encapsulation

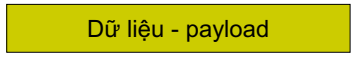


Protocol stack and encapsulation



Protocol stack and encapsulation





Summary: Advantage of layering architecture

- Layering architecture allows to divide the functionalities of networks into small components
- Layers are independent:
  - An upper layer makes use of the functionality of its right bellow layer but does not care about further layer.
- Extensibility/Scalability
- Flexible
  - It is possble to upgrade the communication system by upgrading the technology of some layers: Ex:
    - ADSL→FTTH
    - IPv4→IPv6
- Without layering:
  - Any change in the system requires changing the whole systems.

Unicast, Multicast, Broadcast protocols

- Unicast protocol: control data to send to one destination node
- Multicast protocol: control data to send to multiple destination nodes
- Broadcast protocol: control data to send to all nodes

Identification in the Internet

MAC Address

IP Address

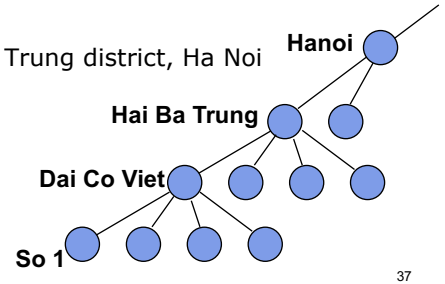
Port number

Identification

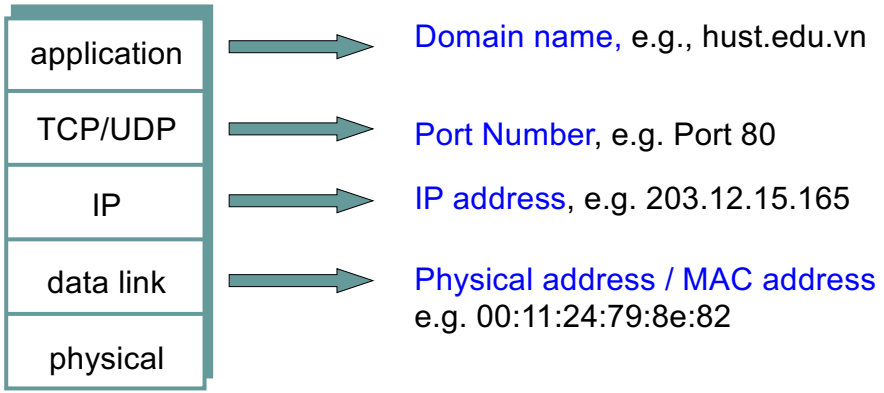
- Identification allows identify a person or an object
  - Name
    - Nguyen Thuc Hai
  - Address
    - 1 Dai Co Viet, Hai Ba Trung, Ha Noi
  - Telephone number
    - 8680896
  - Email
    - hai--xxx@it.hut.edu.vn

Identification

- Identification usually has hierarchical structure
  - Allow to manage efficiently a large addressing space
  - Scalability
- Example of hierarchy
  - Address
    - 1 Dai Co Viet street, Hai Ba Trung district, Ha Noi
  - Telephone number
    - +84-(4) 868-08-96



Identification in the Internet and the relationship between layers



Addressing in the Datalink layer

- Physical address/ MAC address
  - Using in Datalink layer
  - Fixed on NIC (Network Interface Card)
  - Used for identifying machine in broadcast network segment.

HEX 00:11:24:79:8e:82

BIN 00000000 00010001 00100100 01111001 10001110 10000010

OUI — Assigned by manufacture

OUI (Organizationally Unique Identifier):
Each Manufacture have an some OUI unique

Addressing in the Internet

- IP address
  - Used in IP-Internet Protocol (network layer)
  - Value depends on the networks. Each network interface card should be assigned an IP address.
  - Used for identifying a machine in an IP network, example:
    - 133.113.215.10 (ipv4)
    - 2001:200:0:8803::53 (ipv6)

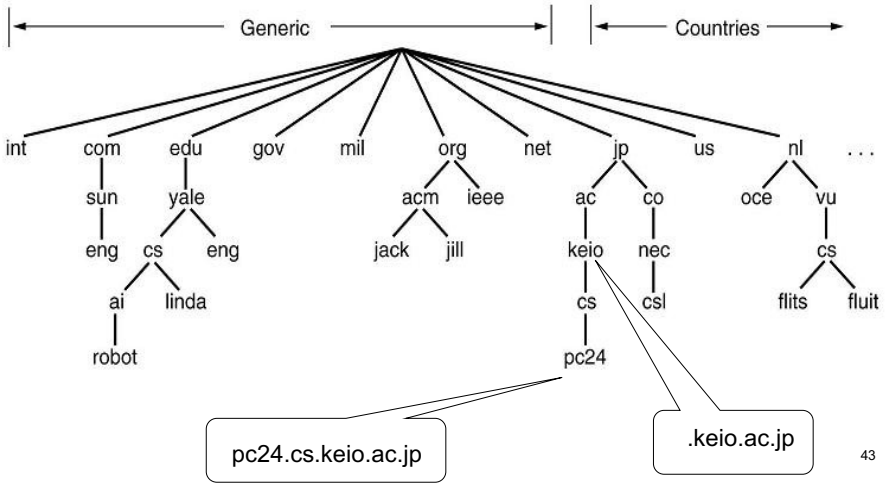
Addressing in transport layer

- Port number
  - On each machine, there may be several applications running.
  - Applications of the same machine are distinguished by port number.
  - An application instance in the internet is identified by the IP address of the host and port number on which it runs
  - Similar to the address of a room in a building
    - Buiding address: B1 Building, 1 Dai Co Viet, Ha Noi => similar to IP address
    - Room number 325 => Similar to port number
- E.g. HTTP runs on port 80, FTP runs on ports 20, 21 …
- http://bidv.vn:81

Addressing in Application layer

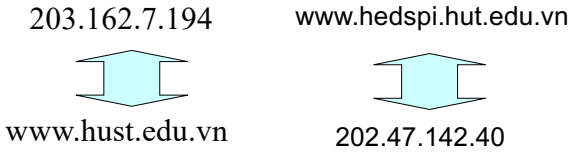
- Domain Name
  (FQDN: Fully Qualified Domain Name)
  - Domain name is the name given to a computer or a network using alphabet and numbers
    - www.keio.ac.jp
    - soict.hut.edu.vn

Domain name space



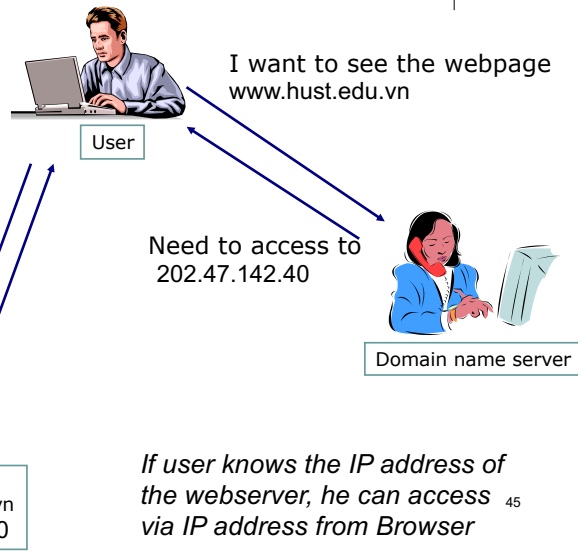
Domain name and IP address

- For sending data to a host/machine, the host must be identified
  - By an IP address
  - By a domain name (easy to be memorized by human)
- name
  - Variable length
  - easy to be memorized by human
  - Nothing to do with the location of the host
- IP address
  - Fixed length (32 bits or 128 bits)
  - Computer process address more easily
  - Used for routing purpose



Conversion/resolution of address

- Computer prefers numbers
- Human prefers names

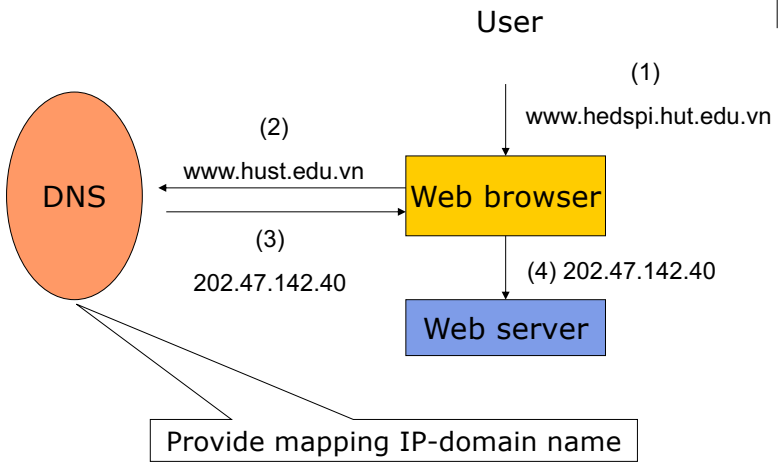


If user knows the IP address of the webserver, he can access via IP address from Browser

Address resolution/conversion

- Concept
  - Mechanism finding address IP from a domain name and vice versa.
  - There is no mathematical formula for this conversion.
- Domain name server (DNS)
  - Store the mapping of IP address and Domain name of the same host in a database
  - Answer requests to resolve IP addresses or domain names from users.
  - Widely used in the Internet

Example



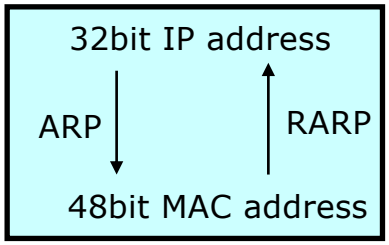
Nslookup tool on Windows, Linux

- nslookup www.soict.hust.edu.vn
- Conversion "name⇔ IP address"

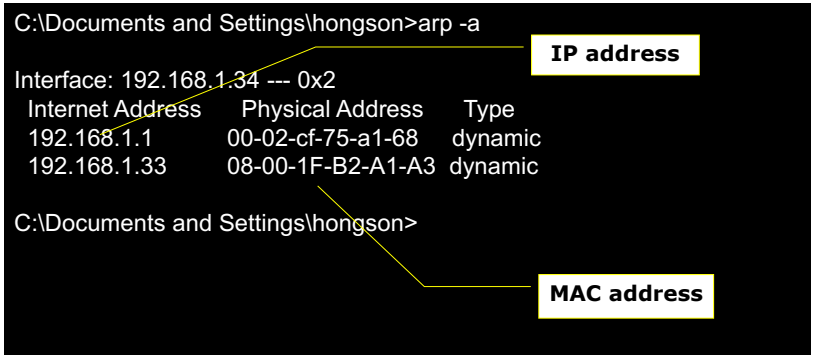
```
C:\>nslookup www.hedspi.hut.edu.vn
Server:
Address:  192.168.1.1

Non-authoritative answer:
Name:    www.hedspi.hut.edu.vn
Address:  202.191.56.68

C:\>
```

ARP Conversion of Mac address and IP address

- Address Resolution Protocol
- MAC and IP are both used for identifying a NIC.
- ARP allows to find MAC address from IP address



Example: ARP table (on Windows)



Summary

- Layer architecture
  - Why layering
  - Model TCP/IP vs. Model OSI
  - Encapsulation, PDU. SAP
- Addressing on Internet
  - Adress IP, MAC, domain name, port
  - Address conversion

Quizz

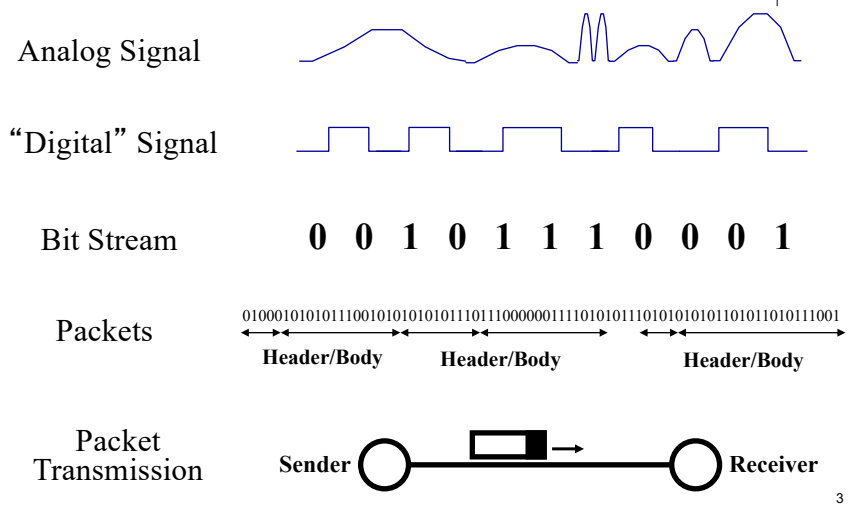
- What do the following objects identify
  - IP
  - Transport port
  - Mac address
  - Domain name
- What identifies uniquely an application.
  - IP of the host running the application?
  - Transport port of the application?

Physical layer

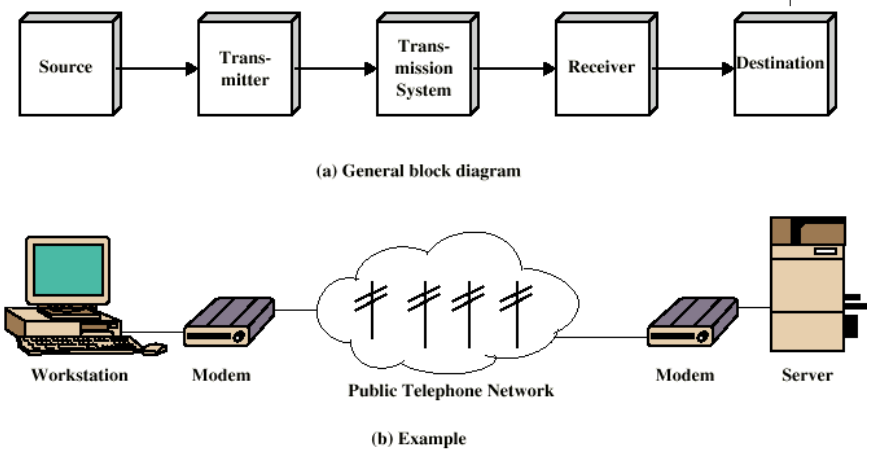
Overview

- Physical layer is responsible for transmission of a stream of bits
  - Put bits from a machine to a medium
  - Pick bits from the medium give to receiver
- Some issues
  - Medium
  - Line Encoding: representing the digital logic levels using the physical attributes associated with the media.
  - Multiplexing

From signal to packet



Model of data transmission system



Data Communication networks

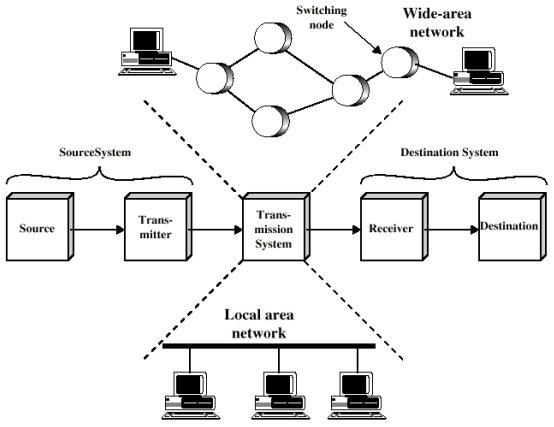
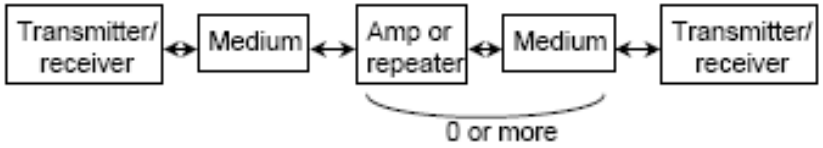


Figure 1.3 Simplified Network Models

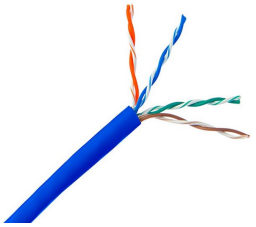
Direct Data transmission system



Media

- Wired
  - Twisted Pair
  - Coaxial Cable
  - Fiber Optics
- Wireless
  - Radio
  - Infra red
  - Light
  - …

Twisted pair



(a)

(b)

(a) Category 3 UTP.
(b) Category 5 UTP.

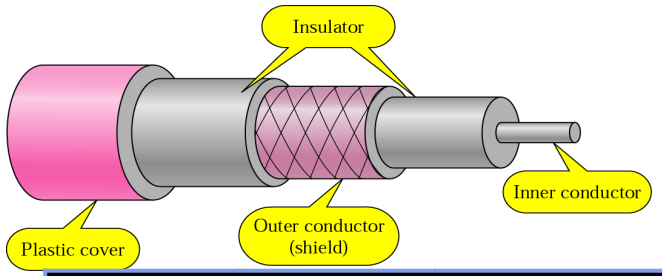
Twisted pair

- Contains several pairs of copper, cable in the one pair is twisted together.
- Two kinds of twisted pair:
  - STP-Shielded Twisted Pair:
    - There is a metal coat, not popular
  - UTP-Unshielded Twisted Pair:
    - No metal coat, popular

Evaluation

- Cheap, simple
- Widely used
- Weak resistance to noice
- Short Transmission distance
- Need amplification after each 5km in analog transmission
- In digital transmission
  - Need repeater after each 2 km
- In Ethernet LAN deployment < 100m
- Limited speed (100Mbps)

II. Coaxial



Category	Impedance	Use
RG-59	75 Ω	Cable TV
RG-58	50 Ω	Thin Ethernet
RG-11	50 Ω	Thick Ethernet

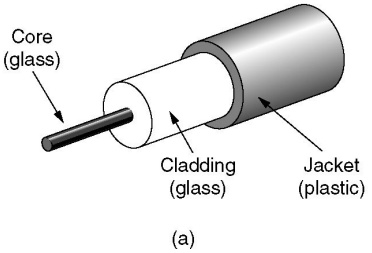
Coaxial

- Structure:
  - Inner conduct is coated by an insulator enviroment
  - Shielded by a metal grill
  - A plastic cover for protection.

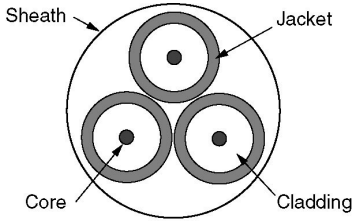
Application

- Using in TV transmission
- For transmission of telephone signal
  - 10,000 calls in the same time
  - Is being replaced by fiber optics
- Linking the computers of the short distance
- LAN 10BaseT, 100BaseT, …
- For digital transmission
  - Repeater should be used after each 1km
  - More repeater is needed for high speed transmission

Optical fiber



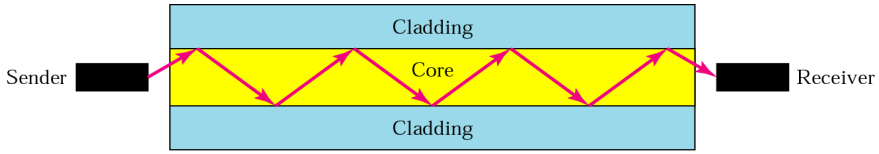
(a)

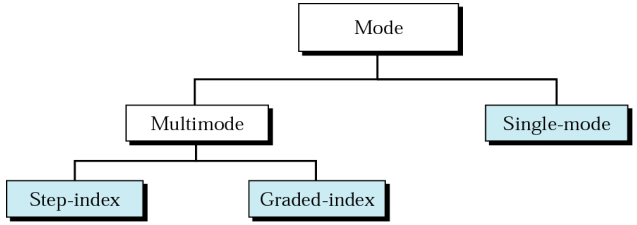


(b)

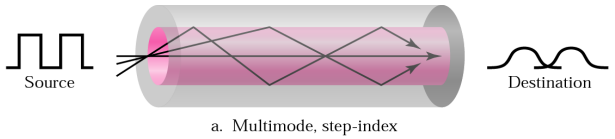
(a) Single core
(b) Cable with 3 cores

Optical fiber transmission mode

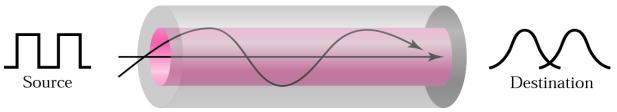




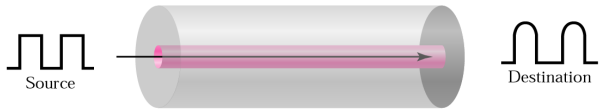
Optical fiber



a. Multimode, step-index



b. Multimode, graded-index



c. Single-mode

Optical fiber

- Multimode stepped index :
  - Several beam travel in slightly different direction
  - Beams arrive in different delay
  - Pulse can easily distort
- Multimode graded index:
  - Index reduce gradually from the center to cladding.
  - Beams closed to center travel slower than cladding. Beams travel in curve form.
  - Reduce pulse distortion.

Optical fiber

- Single mode:
  - Index change less from center to cladding in comparison with multimode.
  - Beams travel along the center axe.
  - Pulses experience less distortion.

Application of optical fiber

- Used for long distance transmission
- Used for communication in metropolitan networks
- Used for connecting routers of ISP
- Used in backbone part of a LAN
- Advantage in comparison with other cables
  - Large data rate
  - Small and light cable
  - Low attenuation
  - Better isolation from electromagnetic environment
  - Large distance between repeaters
    - Multimode →10km
    - Singlemod →40 km

Wireless media

- Terrestrial microwave
  - Used for metropolitan connection, for cellular network
- Microwave satellite
  - Used in TV, Long distance telephone communication
- Radio broadcast
- Infrared
  - Small scope, low data rate, unable to travel through the wall

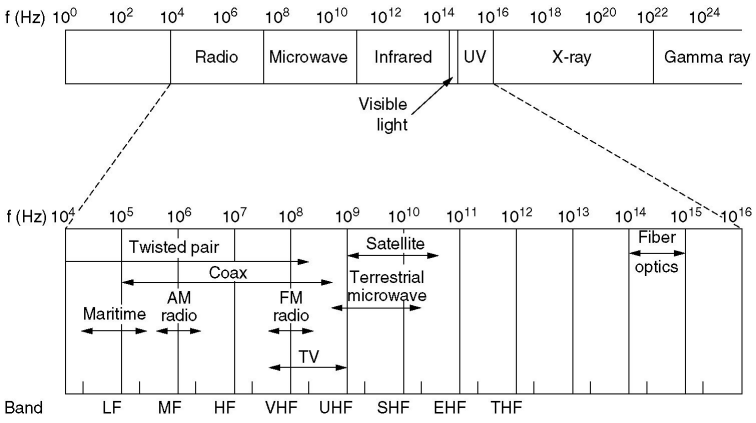
Wireless media

- Radio wave:
  - Wavelengths: 1mm – 100.000km
  - Frequencies: 3 Hz – 300 GHz
  - Ex: Bluetooth, WIFI
- Microwave:
  - Wavelengths: 1mm-1 m
  - Frequencies: 300 MHz-300 GHz
  - Terrestrial microwave : metro connection, cellular communication
  - Satelite microwave: TV, long distance telephone

Wireless media

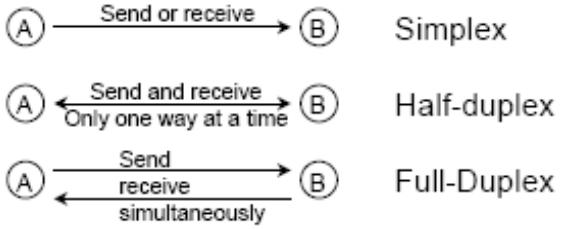
- Infra red:
  - Wavelengths: 700 nm- 1 mm
  - frequency: 300 GHz-430 THz
  - Small scope, no wall penetration
  - Ex: use in remote controls
- Free Space Optics
  - Wavelengths: 850nm, 1300nm, 1550 nm

Frequency range of transmission chanels



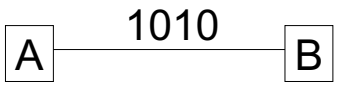
Transmission methods

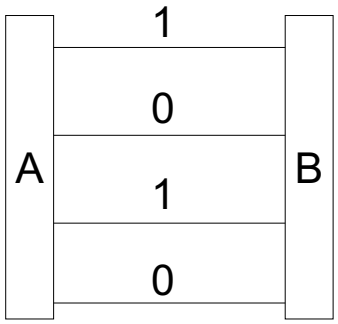
- Simplex: Data is trasmitted in one direction
- Full Duplex: Data can be transmitted in both directions in the same time
- Half duplex: Data can be transmitted in both directions but one direction at a time.



Transmission format

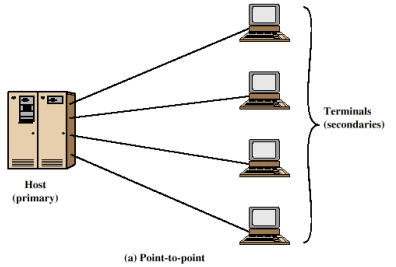
- Sequent transmission: Transmit 1 bit at a moment (over a signal line)
- Parallel transmission: Trasmit multiple bits in the same time (over multiple signal lines)

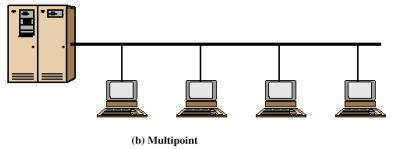




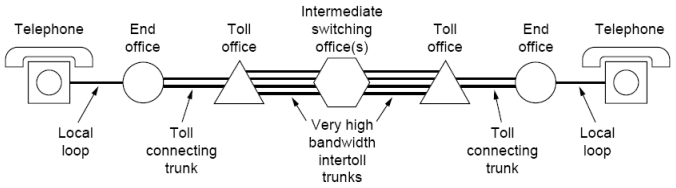
Topology

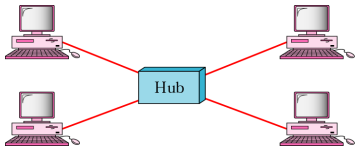
- Point-to-point
  - Star
  - Ring
  - Mesh
- Point-to-multipoint
  - Bus
  - Ring
  - Star

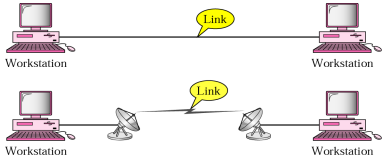


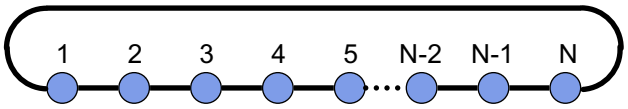


Point -to-Point





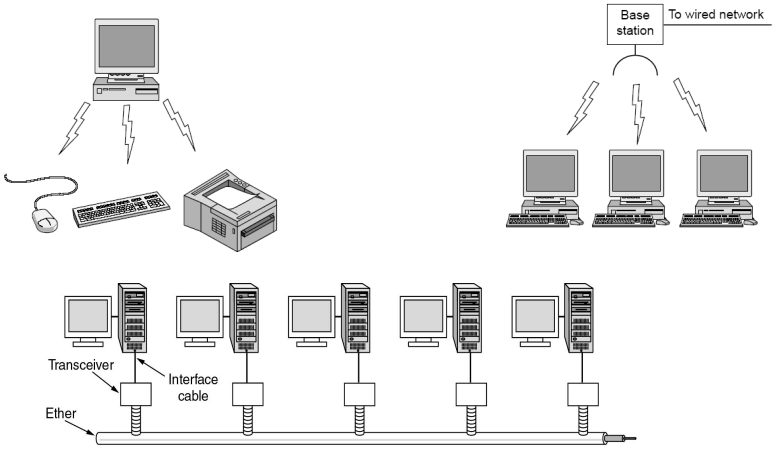




Point-to-point

- A transmission line connects two devices
- Link between two devices:
  - 1 line (half duplex) or
  - 2 lines (duplex)
- *In case of half duplex transmission, there may be collision if two devices on the same link send data in the same time*

Point-to-multipoint



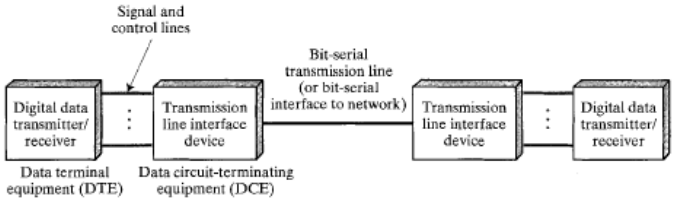
Point-to-multipoint

- Common character of point-to-multipoint topo is to use an unique medium to connects multiple nodes.
- Data is broadcasting over the medium
- Collision when two nodes transmit signal in the same time
- Need a control mechanism to allow a single node to transmit →*multiple access method* → see in Datalink layer.

Medium interface

- Data terminal equipment (DTE)
  - Have data to transmit but has no feature for transmission
  - Need an additional device for accessing the media
- Data circuit terminating equipment (DCE)
  - Transmit bits on the media
  - Transmit data and control information with DCE through connection the media
- Need a clear interface standard between DTE, DCE

DTE-DCE



(a) Generic interface to transmission medium

(b) Typical configuration

Media interface

- Mechanism
  - Define the form of the interface, number of pins for assuring the interfaces match together
- Electrics
  - Define the level of voltage to be used
  - Define the length of pulse (frequency)
  - Define enconding method
- Functionalities
  - Functionality of each pins
  - There are 4 groups of pins: data, control, synchronization, ground
- Procedure
  - Lists of events to perform for transmitting data

Example: EIA-232-E/RS-232

- Define for serial communication
- Mechanism: ISO 2110
- Electrics: V. 28
- Functionality: V. 24
- Procedure: V. 24

Example: V.24 /EIA-232-E

- Mechanic:
  - 25 or 15 pins
  - Transmission distance 15m
- Electrics
  - Digital data
  - 1=-3v, 0=+3v (NRZ-L)
  - Data rate 20kbps
  - Transmission distance< 15m

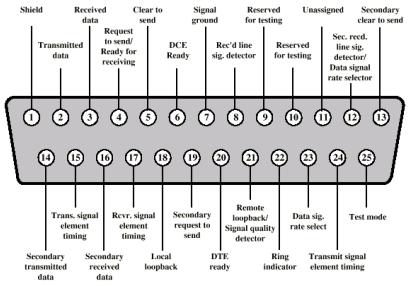


Figure 6.5 Pin Assignments for V.24/EIA-232 (DTE Connector Face)

Data encoding

- Use different discrete signal, different voltage level for representing bit 0 and 1.
- Data transmission should be synchronized between sender and receiver: clock synchronization
- Encoding could be performed by bit or by a group of bit e.g., 4 or 8 bits.
- There are many way to represent 0 and 1 → See data transmission technique.

Data Encoding

Introduction
Encoding digital data to digital signal
Encoding digital data to analogical signal
Encoding analogical data to digital signal
Encoding analogical data to digital signal

Digital data- Digital signal

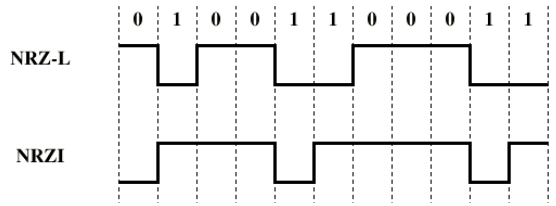
- Data unit: 1 bit
- Digital data is a digital signal
  - Each pulse is considered as a signal unit.
- Data encoding: mapping data to signal
- Set of mapping is called Encoding scheme
- Ex: Mark:1, Space:0

Line encoding method

- NRZ
  - NRZ-L,NRZI
- Bipolar
  - Bipolar alternate mark inversion
  - Pseudoternary
- Phase encoding
  - Manchester
  - Manchester vi sai

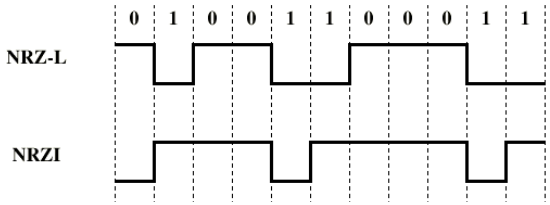
NRZ-L Non Return to Zero Level

- During bit time, signal does not go back to 0 level
- Signal level is not changed during bit time.
- NRZ-L Non return to zero level
  - Bit 1 signal is in low/high level
  - Bit 0 signal is in high/low level



NRZ-I Non return to zero invert

- Bit 0: signal level is not changed at the begining of bit time
- Bit 1: signal level is changed at the begining of bit time
- A differential encoding method :
  - 0 and 1 represent by the signal level change, not by the level itself.
  - Reliable/ simple.



Line encoding consideration

- Two aspects should be considered in any encoding method:
  - **Clock recovery on receiver side**: If the clock recovery is not ideal, then the signal to be decoded will not be sampled at the optimal times. This will increase the probability of error in the received data.
  - **DC-component**: Directed Current vontage component.
    - DC-component makes recepter mistakenly detect level of signal
    - Encoding should avoid DC-component by having signal mean altitude to be around 0.

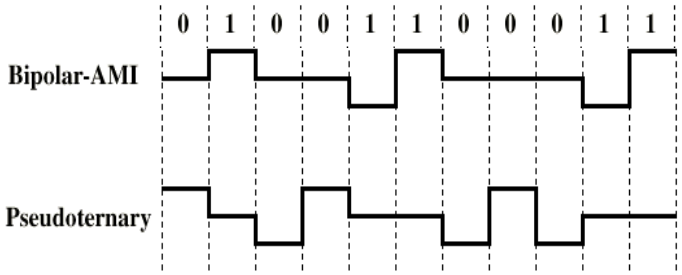
NRZ

- NRZ Advantage
  - Simple, utilise the maximum capacity of the line
- NRZ Weakness
  - NRZ does not contain element supporting clock synchorization
    - Example: when sending a suit of 1 or 0
  - Contain DC-component when sending a suit of 1.
- Application
  - Encoding data on magnetic storage
  - Not popular in data transmission

Bipolar AMI

- Use more than 2 signal level for 1 bit
- Bipolar alternate mark inversion
  - 0 : No signal
  - 1: Presence of signal. Two consequent 1 have two different signal levels
- pseudoternary
  - 1 : No signal
  - 0 : Presence of signal. Two consequent 0 have two different signal levels

Bipolar-AMI



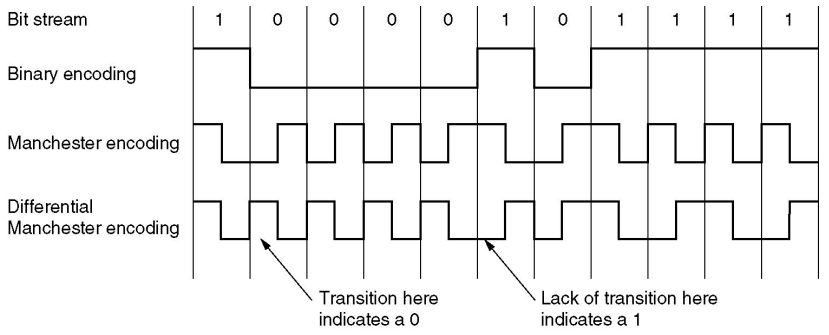
Bipolar AMI

- DC component =0
- Good synchronization when there are many bit 1(0), lost of synchronization when there are many bit 0(1)
- 3 possible signal levels for 1 bit:
  - Not optimal in using transmission line.
  - Receipter needs to distingush 3 levels of signal

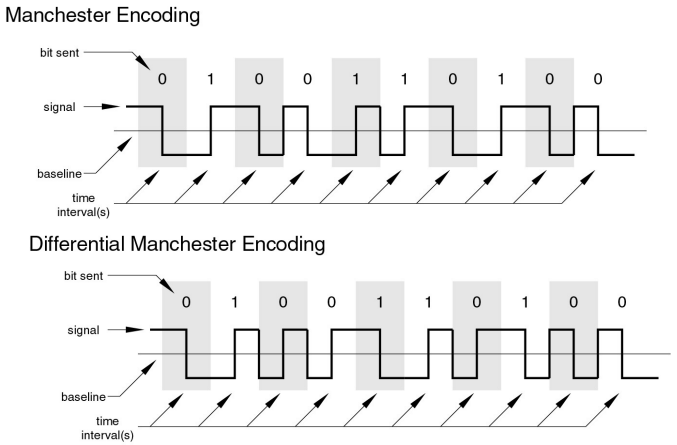
Biphase: Manchester

- Manchester: Always change signal level in the middle of bit time
  - Bit 1: Signal change from low level to high level
  - Bit 0: Signal change from high level to low level
  - Level change provide synchonisation mechanism.
- Differential Manchester:
  - 0: signal level change at the begining of bit
  - 1: **no** signal level change at the begining of bit
  - Always change signal level in the middle of bit time for synchronization purpose

Manchester encoding



Manchester encoding



Units in transmitting digital data in digital transmission

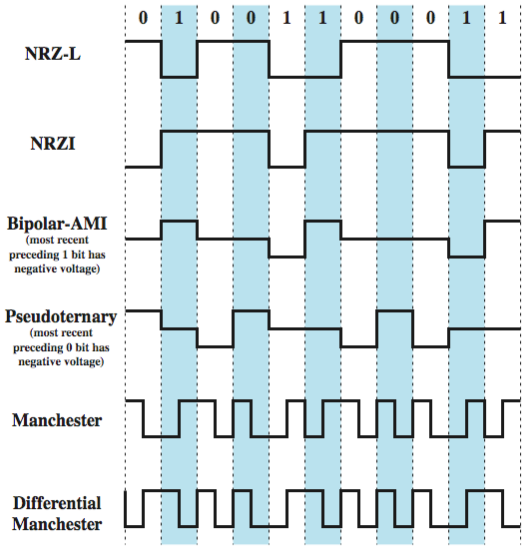
Term	*Units*	*Definition*
Data unit	bit	A single bit, Value 0 or 1
Data rate	bit/s	Rate transmitting bit
Signal unit	Pulse a sinus	Part of the signal correspond to the smallest duration of a symbol
Symbol rate/ Rate of modulation	Number of symbol/s (baud)	Number of symbols generated in a unit of time

Encoding rate: Baud rate

- Number of symbol changes, waveform changes, or signaling events across the transmission medium per unit of time
- Unit: Baud/s = symbol/s

	Minimum	101010…	Maximum
NRZ-L	0 (all 0's or 1's)	1.0	1.0
NRZI	0 (all 0's)	0.5	1.0 (all 1's)
Binary-AMI	0 (all 0's)	1.0	1.0
Pseudoternary	0 (all 1's)	1.0	1.0
Manchester	1.0 (1010…)	1.0	2.0 (all 0's or 1's)
Diff Manchester	1.0 (all 1's)	1.5	2.0 (all 0's)

Line encoding

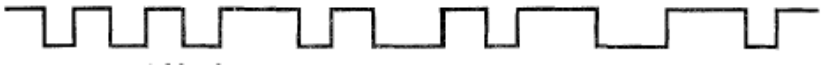


Bài tập-01

- Biểu diễn các tín hiệu mã hóa chuỗi dữ liệu sau đây bằng các phương pháp mã hóa đã học
  - 11000000 00000010 11001101 01010101

Bài tập-02

- Dữ liệu mã hóa bằng mã manchester (không vi sai) cho tín hiệu
  - Xác định thời gian của từng bít
  - Xác định dữ liệu ban đầu



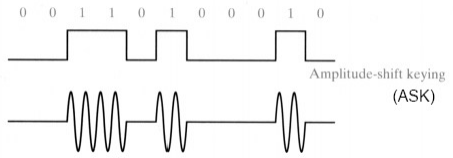
2. Điều chế số-liên tục

- Ví dụ: truyền số liệu thông qua hệ thống điện thoại
  - Hệ thống điện thoại truyền, chuyển tiếp tín hiệu điện có tần số 300Hz đến 3400Hz
  - Tại nguồn và đích, dữ liệu số cần được điều chế thành tín hiệu liên tục để truyền trên đường điện thoại
- Căn cứ vào tính chất của tín hiệu, chúng ta có 3 kỹ thuật điều chế
  - Điều chế khóa dịch biên độ
  - Điều chế khóa dịch pha
  - Điều chế khóa dịch tần số

Điều chế khóa dịch biên độ (ASK)

- 0 và 1 tương ứng với hai biên độ tín hiệu, thông thường một trong hai biên độ=0
- Dễ bị ảnh hưởng bởi nhiễu (1200bps cho đường thoại)
- Khó đồng bộ
- Thường được dùng trong cáp quang (LED hoặc laser

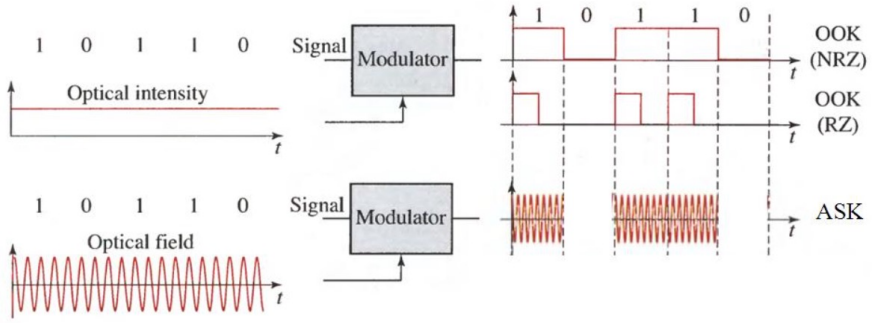
$$s(t) = \begin{cases} A\cos(2\pi ft) \ cho\ 1 \\ 0 \ cho\ 0 \end{cases}$$



Mã On-Off Keying (OOK)

- Dùng trong cáp quang
- Là một loại điều chế dịch biên độ.
  - 1: có xung ánh sáng trong thời gian bit (bật nguồn sáng).
  - 0: không có xung ánh sáng trong thời gian bit (tắt nguồn sáng).
- OOK có thể dùng nhiều định dạng tín hiệu khác nhau:
  - NRZ: xung ánh sáng chiếm toàn bộ độ dài bit 1.
  - RZ (return-to-zero): chỉ phát xung ánh sáng trong một phần thời gian của bit 1.

Mã On-Off Keying (OOK)



On off key nhìn từ phương diện cường độ sáng (hình trên) và tín hiệu quang học (hình dưới)

Điều chế khóa dịch tần số (FSK)

$$s(t) = \begin{cases} A\cos(2\pi f_1 t) \\ A\cos(2\pi f_2 t) \end{cases}$$

- Hai giá trị nhị phân được biểu diễn bởi hai tín hiệu tần số khác nhau
- Ví dụ về điều tần song công
- Tỷ suất lỗi thấp hơn
- Dùng trong truyền số liệu qua đường điện thoại (tần số thấp), hoặc trong mạng không dây (tần số cao)

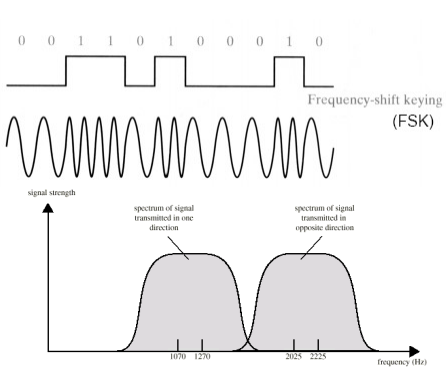


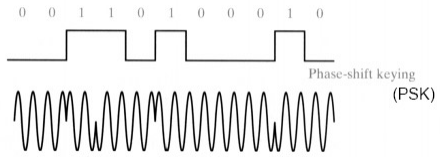
Figure 5.8 Full-Duplex FSK Transmission on a Voice-Grade Line

Điều chế khóa dịch pha (PSK)

$$s(t) = \begin{cases} A\cos(2\pi f_c t + \pi) & \text{binary 1} \\ A\cos(2\pi f_c t) & \text{binary 0} \end{cases}$$

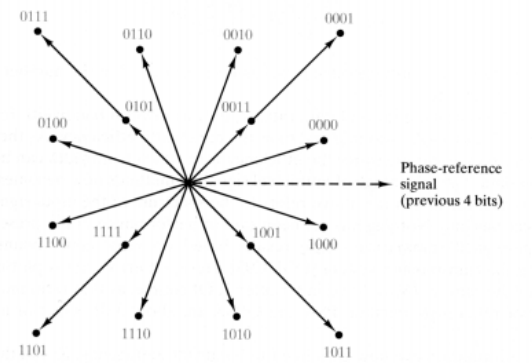
$$s(t) = \begin{cases} A\cos(2\pi f_c t + 45°) & 11 \\ A\cos(2\pi f_c t + 135°) & 10 \\ A\cos(2\pi f_c t + 225°) & 00 \\ A\cos(2\pi f_c t + 315°) & 01 \end{cases}$$

- 0,1 tương ứng với hai độ lệch pha khác nhau
- 0,1 tương ứng với chuyển pha (vi sai)
- Có thể sử dụng giải thông một cách hiệu quả hơn khi mã hóa cùng lúc nhiều bít
- Có thể kết hợp với điều biên
- Nếu tốc độ dữ liệu là 9600 bps, tốc độ điều chế là ?

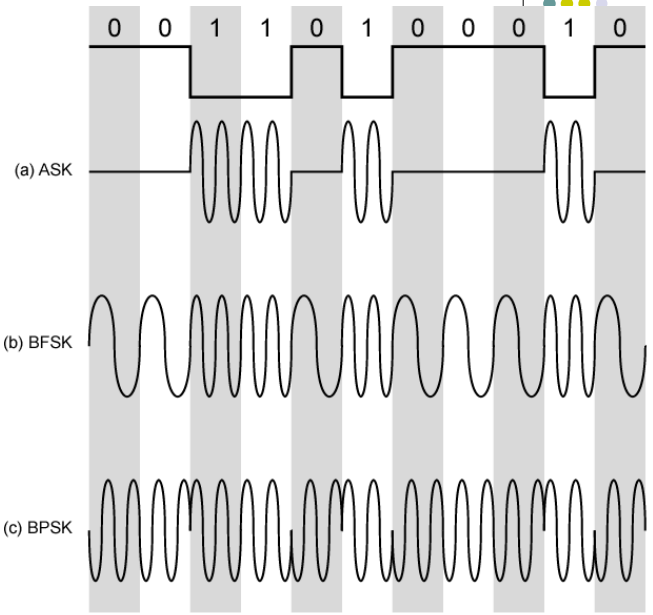


Kết hợp với điều biên

– 9,600 bps modem (2,400 baud x 4)



Điều chế số/liên tục

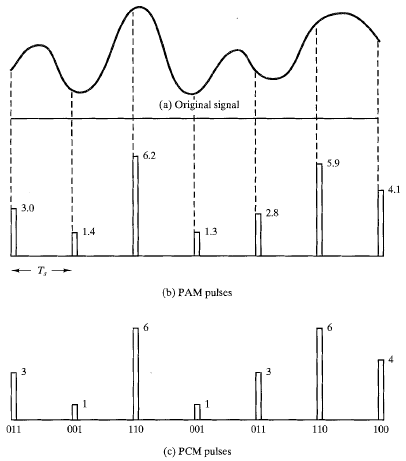


3. Điều chế dữ liệu liên tục- số

- Điều chế dữ liệu liên tục thành dữ liệu số, sau đó
  - Điều chế thành tín hiệu số
    - Mã hóa trực tiếp bằng NRZ-L
    - Sử dụng phương pháp mã hóa tín hiệu số khác
  - Điều chế thành tín hiệu liên tục
    - Sử dụng các biện pháp điều chế số-liên tục đã học
- Có hai phương pháp chính điều chế dữ liệu liên tục thành dữ liệu số
  - Điều chế mã xung
  - Điều chế Delta

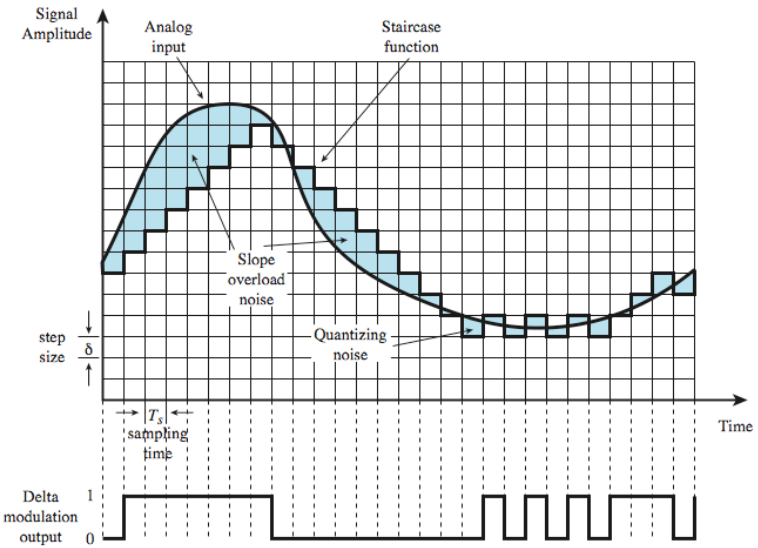
Điều chế mã xung (PCM)

- Pulse Code Modulation
- Lấy mẫu tín hiệu dựa trên định luật lấy mẫu của Shannon
  - Nếu tần số lấy mẫu >= 2 lần tần số (có ý nghĩa) cao nhất của tín hiệu, phép lấy mẫu bảo toàn thông tin của tín hiệu
  - Vd: Tiếng nói tần số tối đa 4300Hz, cần lấy mẫu với tần số min 8600Hz
- Tiến hành theo hai bước
  - Lấy mẫu (PAM)
  - Lượng tử hóa



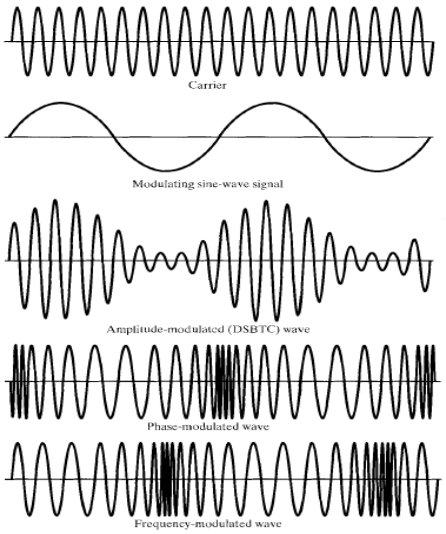
Điều chế delta (Delta Modulation)

- Sử dụng hàm bậc thang
  - Khi hàm số tăng, xung=1
  - Khi hàm số giảm, xung=0
- Tổng quát
  - Biểu diễn giá trị của đạo hàm theo bít
- Tham số
  - Bậc thang
  - Tốc độ lấy mẫu
- Sai số
  - Khi tín hiệu thay đổi chậm: nhiễu lượng tử
  - Khi tín hiệu thay đổi nhanh: nhiễu tràn



Dữ liệu liên tục tín hiệu liên tục

- Kết hợp tín hiệu m(t) và sóng mang có tần số Fc thành một tín hiệu tập trung xung quanh Fc
- Cho phép chuyển tín hiệu trên một tần số khác phù hợp với kênh truyền
- Cho phép dồn kênh bằng các tần số sóng mang khác nhau
- 3 phương pháp chính dựa vào đặc điểm của tín hiệu
  - Điều biên
  - Điều tần
  - Điều chế góc pha

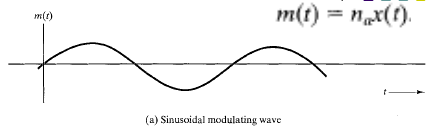


Điều biên

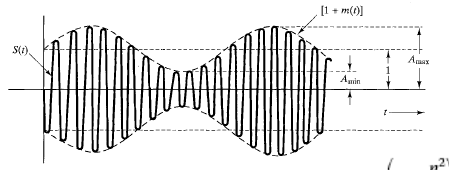
- Biến đổi biên độ sóng mang theo đầu vào
- Nếu đầu vào cũng là hình sin
  - Tín hiệu đầu ra sẽ có hai thành phần lệch với tần số sóng mang một khoảng bằng tần số đầu vào
  - Na<1 điều biên hợp lệ
  - Na>1 mất thông tin
- Giải thông=2 lần giải thông đầu vào
- Điều biên một chiều: 1 lần giải thông

$$s(t) = [1 + m(t)]\cos 2\pi f_c t$$

$$m(t) = n_a x(t)$$

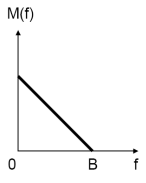


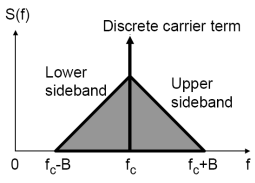
(a) Sinusoidal modulating wave



(b) Resulting AM signal

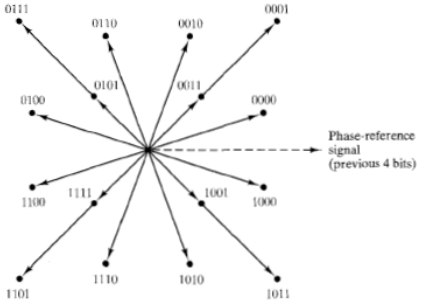
$$P_t = P_c\left(1 + \frac{n_a^2}{2}\right)$$





Bài tập-04

- Biểu diễn phương pháp điều chế pha-biên độ sau bằng công thức
- Tốc độ ký hiệu là 2400 baud. Tốc độ dữ liệu là bao nhiêu?



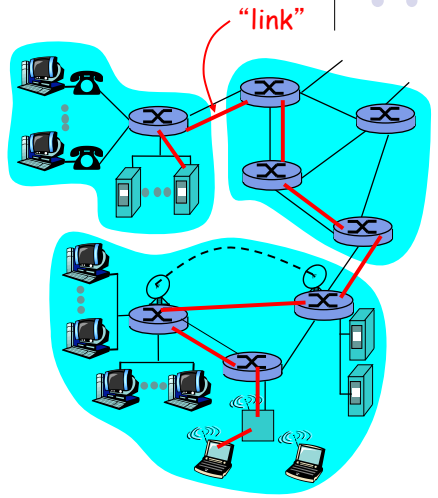
Lecture 4: Datalink layer

- Functionalities:
  - Encapsulation, addressing
  - Error detection and correction
  - Flow control
  - Media access control

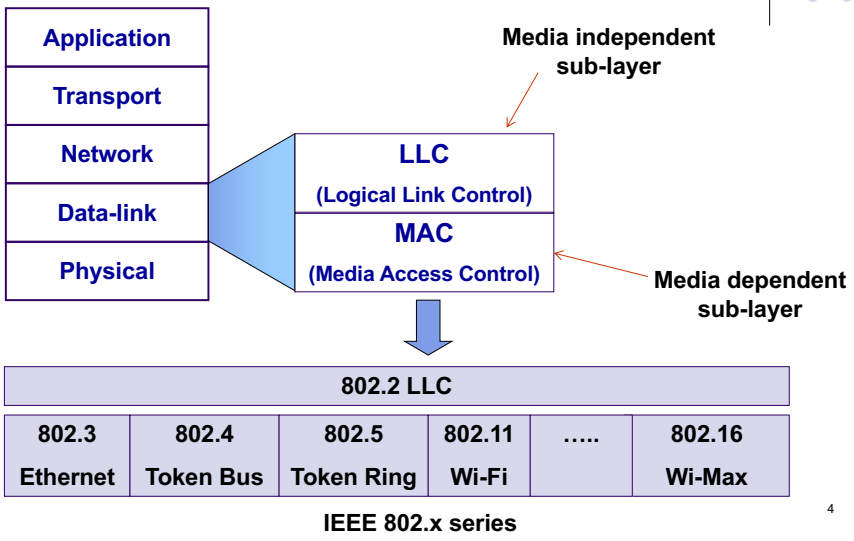
Overview of Data link layer

Network nodes and links

- Network nodes:
  - PCs, Laptop, Routers, Server…
- Links:
  - Communication chanel between **adjacent nodes**
  - Wired link: Ethernet LAN, ADSL, fiber optic…
  - Wireless link: Wi-fi, FSO, Satellite,…
- Datalink layer responsibility:
  - Transmit data between adjacent elements.

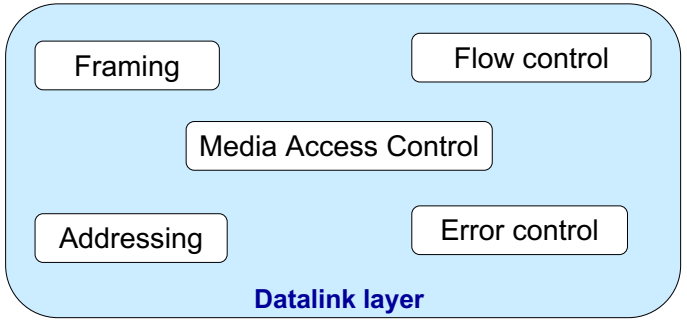


Datalink layer in Layer architecture



IEEE 802.x series

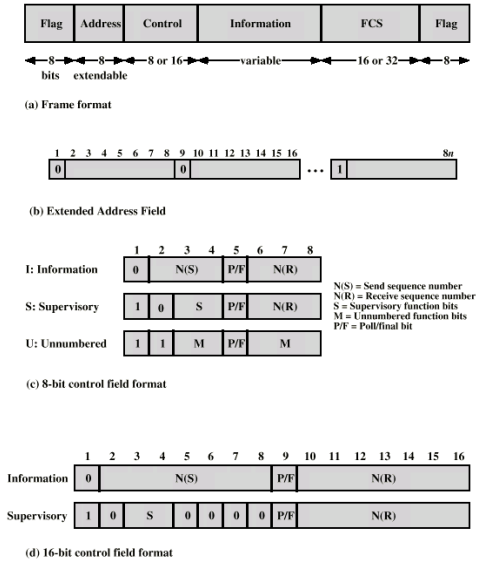
Functionalities



Functionalities

- Framing:
  - Sender: place the network layer packet into the frame, add header, tail
  - Receiver: Remove the header, tail for extracting the network packet.
- Addressing:
  - Physical address in the header of the frame for identifying the source and the destination.

Framing-Example of HDLC frame



Functionalities (2)

- Media access control:
  - If the nodes in the network share common media, a Media access control protocol is required.
- Flow control:
  - Control the transmission speed of the sender so that the receiver does not overloaded.
- Error control:
  - Detect and correct errors
  - e.g. parity check, checksum, CRC check

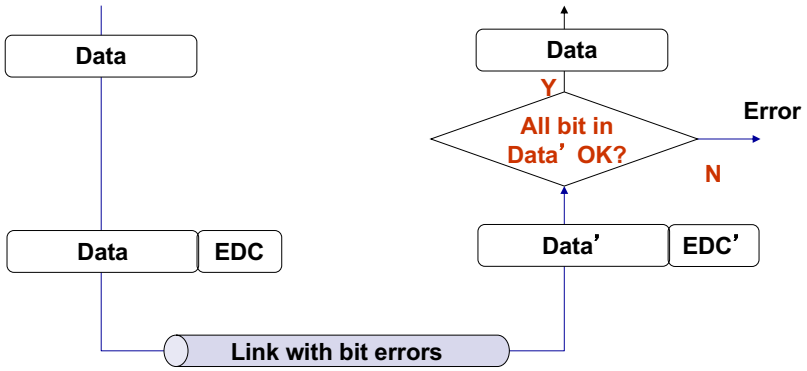
Error control

Error detection

Error correction

Principle of error detection

EDC= Error Detection Code (redundancy)
EDC is added to data before sending to the destination.



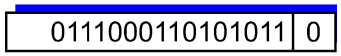
Principal of error detection

- Condition for all error detection
  - $f_{edc}(D_{send}) != f_{edc}(D_{receive})$ $\forall D_{send} != D_{receive}$
- Space of codes (number of possible codes) must at least equals to space of data (number of possible data)
  - ➔ code length (bits) >= data length (bits)
  - Transmission efficiency <= %50

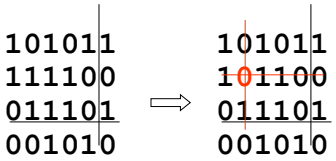
Parity code

A check bit is added to the original data to ensure that the total number of bit 1 is even (even parity code) or odd (odd parity code)

- Single code
  - Able to detect single bit error



- Two-dimension code
  - Detect and correct single bit error



- Application: mainly on hardware, ex: while sending data on PCI and SCSI bus

Parity code

- Sent data with Odd code:
  - 01010101 →Code: 1
- Case 1: Received data 01110101 Received code: 1
  - → Total number of 1 : 6→ even number→ Code does not match with data
  - → Error
- Case 2: Received data 01110100 Received code: 1
  - Total number of bit 1 → 5→ code matches with data
  - → No error
- Data of m bit long→ space of data is 2^m → expected to have different code for different data → codes must be >=m bit long.

Checksum code

- Sender:
  - Divides data into small parts of n bit
  - Calculates binary sum of all parts. If there are some overflow bit, add the overflow bit to the result.
  - Alters all bits (two's complement) to get the checksum
  - Sends the checksum with data
- Receiver:
  - Extracts data and checksum
  - Divides data into block of n bits
  - Calculates the sum as in the sender side including with the checksum received
  - If result contains at least one bit 0→ error.

Checksum: Example

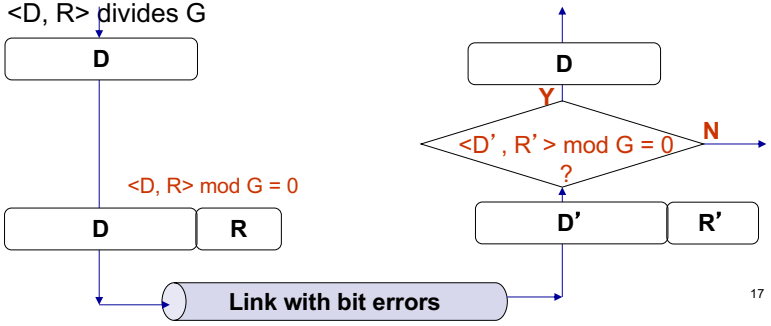
```
Data: 0011 0110 1000
Calculate checksum 4 bit:
    0011
  + 0110
    1000
   10001
Overflow
bit      1
    0010
Alter bit -> checksum code: 1101
Bits to send: 0011 0110 1000 1101
```

Checksum: Processing on receiver

```
Bits received: 0011 0110 1000 1101
Verification:
     0011
     0110
  +  1000
     1101
    11110
Overflow bit  1
     1111 → no bit error
```


CRC: Cyclic Redundancy Check

- Data is considered as a binary string: D
- We wants to generate a error code with length r
- Choose another binary string of (r+1) bit, G (Generator)
- Find a string R with length r bits such that the concatenation of D and R is a binary number that divides G (modulo 2)
  - <D, R> divides G



CRC: How to find R

- <D, R> = $D.2^r$ XOR R
- Since <D, R> divides G then
  - $D.2^r$ XOR R = n.G
  - → $D.2^r$ = n.G XOR R (associativity)
- This means, R is the remainder of the division $D.2^r$ by G (division modulo 2)

R= $D.2^r$ mod G

R=110, the string to send is 10101001110

D R

- Ex: D= 10101001
- r= 3 bits
- G=1001

```
10101001000   1001
1001          1011110
  1110
  1001
   1110
   1001
    1111
    1001
     1100
     1001
      1010
      1001
       110
```

G

D

R

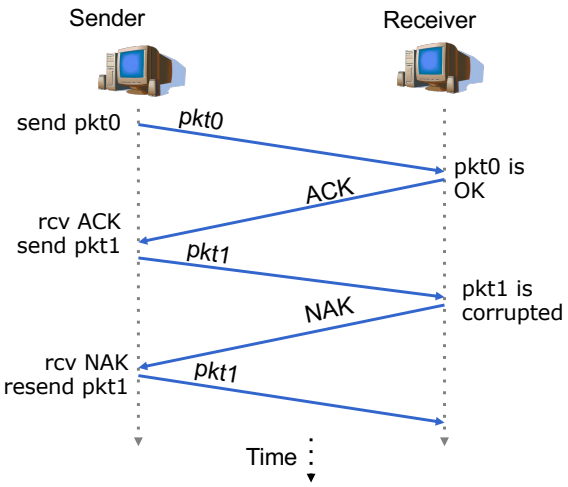
CRC under polynomial form

- 1011 <-> $x^3 + x + 1$
- Example of some CRC generators using in the pratice:
  - CRC-8 = $x^8 + x^2 + x + 1$
  - CRC-12 = $x^{12}+x^{11}+x^3+x^2+x$
  - CRC-16-CCITT = $x^{16} + x^{12} + x^5 + 1$
  - CRC-32 = $x^{32} + x^{26} + x^{23} + x^{22} + x^{16} + x^{12} + x^{11} + x^{10} + x^8 + x^7 + x^5 + x^4 + x^2 + x + 1$
- The longer G is, the more possible that CRC detects errors.
- CRC is widely used in the practice
  - Wi-fi, ATM, Ethernet…
  - Operation XOR is implemented in hardware
  - Capable to detect less than r+1 bits errors

Reaction when errors detected

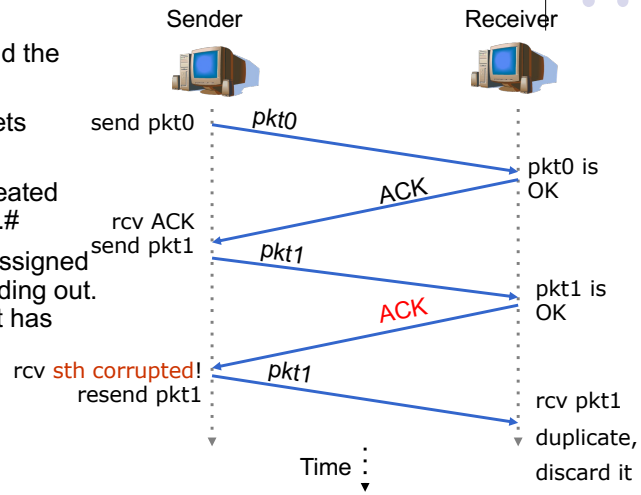
- Objective: assure that data are received correctly even though the chanel is not realiable.
- Constraint
  - Data frame must be correctly received
  - Negligible transmission delay.
- Possible errors
  - Whole frame loss
  - Error frame
  - Loss of error warning message
- Popular techniques:
  - Error detection (as we seen)
  - Acknowledgement/confirmation
  - Retransmis after a clear confirmation that frame is not arrived
  - Retransmis after timeout
- ARQ technique: automatic repeat request). There are 3 versions:
  - Stop and Wait ARQ
  - Go Back N ARQ
  - Selective Reject ARQ
- Similar to techniques used in flow control.

Stop-and-wait ARQ *Normal case*

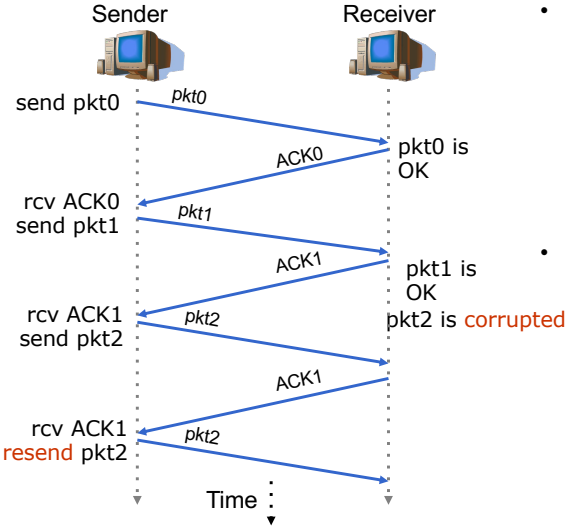


Stop-and-wait ARQ *Error ACK/NAK*

- ACK error, resend the previous packet
- Duplicated packets problem.
- To eliminate repeated packet: Use Seq.#
- All packets are assigned Seq# before sending out. Repeated packet has identical Seq#



Stop-and-wait ARQ *not using NAK*

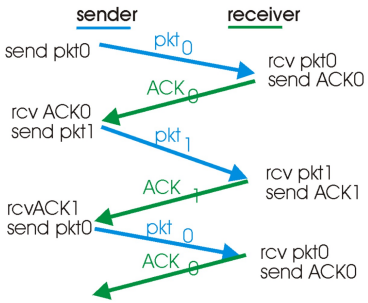


- ACK packet carries #Seq of the packet to be acknowledged. This number is called acknowledgment number
- An ACK with acknowledgment number n implicitly confirms that all packet with #seq number < =n have been well received

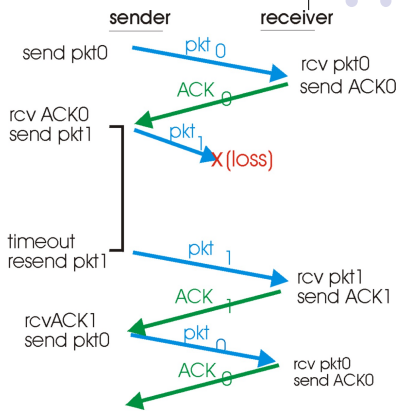
Stop-and-wait ARQ: When ACK is lost

- Data packet and ACK packet may be lost
  - No ACK is received at sender side
  - How a sender decides to resends data or not?
- Solution:
  - After sending out a packet, sender starts a timer specifying maximum waiting time (timeout) for an ACK of the packet.
  - When timeout expired sender re-sends the packet
- How long a Timeout should be?
  - At least 1 RTT (Round Trip Time)
- If a packet arrives at the destination but its ACK is lost, the packet is still resent because associated timeout expired.
  - The duplicated packets are eliminated at the receiver side according to repeated #seq.

ARQ with timeout

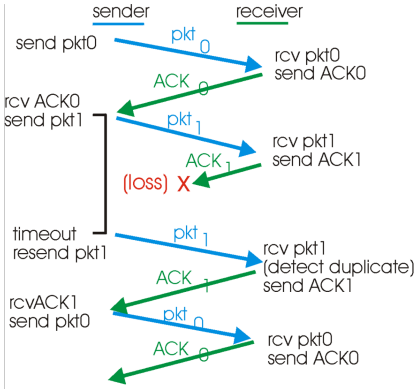


(a) operation with no loss

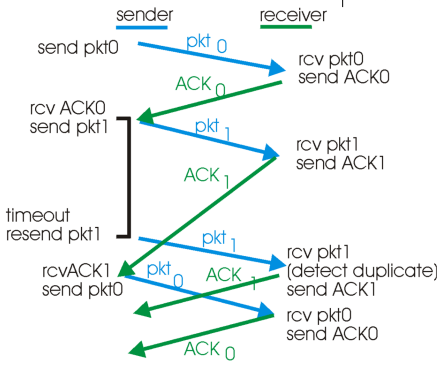


(b) lost packet

ARQ with timeout



(c) lost ACK



(d) premature timeout

Media access control

Connection types

- Point-to-point
  - ADSL
  - Telephone modem
  - Leased Line….
- Broadcast
  - LAN using bus topology
  - Wireless LAN
  - HFC:
  - …
- Broadcast networks need media access control protocol in order to avoid collision when nodes try to send data.

Classification of MAC protocol

- Chanel division:
  - Resources of the media is divided into small parts (time - TDMA, frequency- FDMA, Code- CDMA)
  - Distribute a part to each nodes
- Random access:
  - Chanel is not divided, all nodes are allowed to access simultaneously with collision possibility
  - Need a mechanism to avoid collision
  - e.g. Pure Aloha, Slotted Aloha, CSMA/CD, CSMA/CA…
- Sequent access:
  - Nodes can send data one after the other.
  - Token Ring, Token Bus….

Channel division

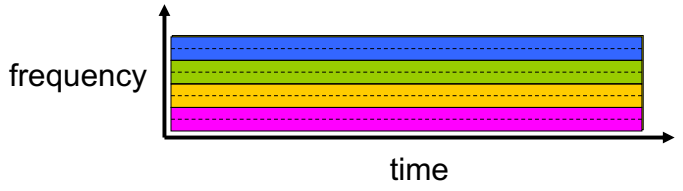
- FDMA: frequency division multiple access
- TDMA: time division multiple access
- CDMA: code division multiple access

TDMA và FDMA

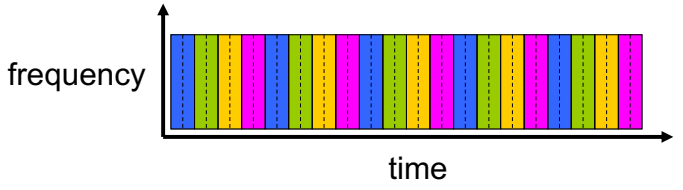
ex

4 stations

FDMA



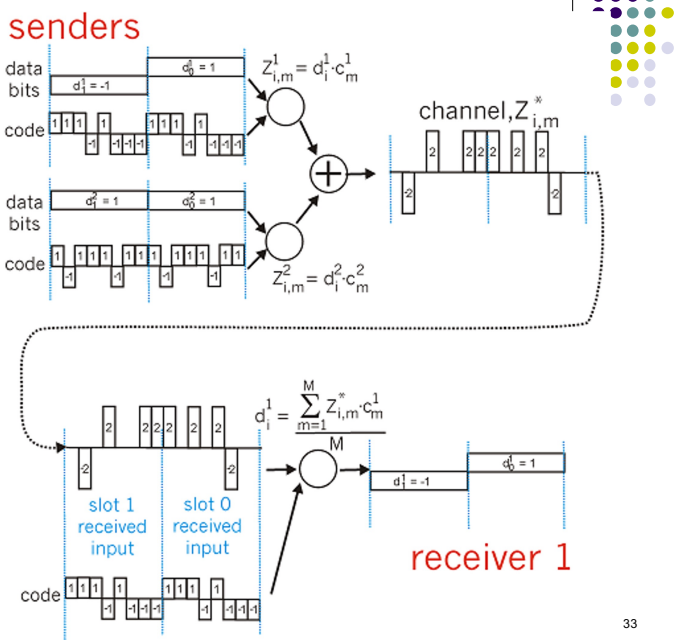
TDMA:



CDMA

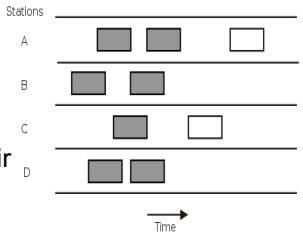
- Several senders can share the same frequency on a single physical channel.
- Signals come from different senders are encoded (multiplied) with different random code. Those code must be orthogonal.
- Encoded signals are mixed and then transmit on a common frequency.
- The signals are recovered at the receiver by using finding the correlation with the same codes as at sender side.
- **CDMA** shows a lot of advantages that other technology cannot achieve. For example, the same frequency can be used in adjacent mobile cell without interference as if TDMA or FDMA are used

CDMA



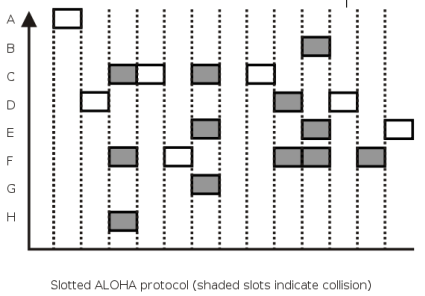
Random access: Pure Aloha

- Aloha is used in mobile network of 1G, 2.5G, 3G using GSM technology .
- Pure Aloha:
  - When one sender has data to send, just sends it
  - If while sending, the senders receive data from other stations → there is collision. All stations need to resend their data.
    - There are possibility to have collision when retransmit.
  - Problem: Sender does not check to see if the chanel is free before sending data
  - Grey package are having overlap in time→ causing collision



Random access: Slotted Aloha

- Times axe is divided into equal slots.
- Each station sends data only at the beginning of a time slot.
- ➔ Collision possibility is reduced
- Still have collision in grey package

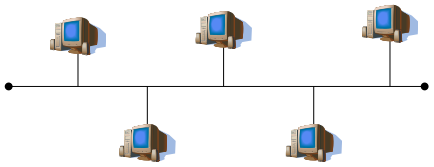


Slotted ALOHA protocol (shaded slots indicate collision)

Random access: CSMA

- CSMA: Carrier Sense Multiple Access
- CSMA idea is similar to what happens in a meeting.
- CSMA:
  - The sender "Listen before talk"
  - If the channel is busy, wait
  - If the chanel is free, transmit

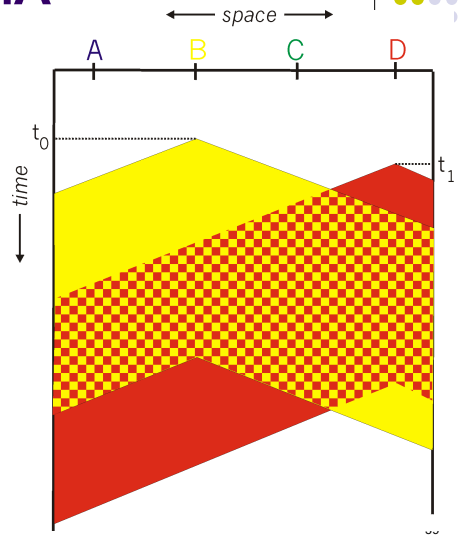
CSMA



- **CSMA**: Sender listens before transmission:
  - If the channel is free, send all the data
  - If the channel is busy, wait.
- Why there are still collision?
  - Due to propagation delay

Collision in CSMA

- Assume that there are 4 nodes in the channel
- The propagation of the signal from one node to the other requires a certain delay.
- Ex:
  - Transmissions from B and D cause collision



CSMA/CA (Collision Avoidance)

- CSMA/CA is used WIFI standard IEEE 802.11
- If two stations discover that the channel is busy, and both wait then it is possible that they will try to resend data in the same time.
  - → collision
- Solution CSMA/CA.
  - Each station wait for a random period → reduce the collision possibility

CSMA/CD

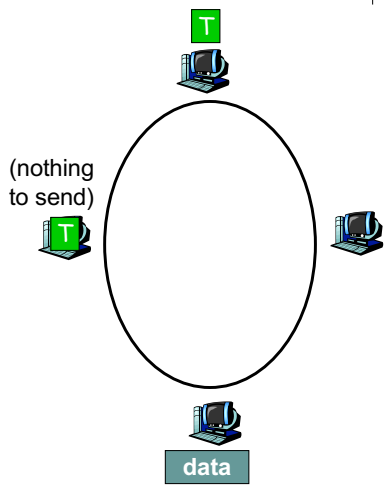
- Used in Ethernet
- CSMA with Collision Detection:
  - “Listen while talk”.
- A sender listen to the channel,
  - If the channel is free then transmit data
    - While a station transmit data, it listens to the channel. If it detects a collision then transmits a short signal warning the collision then stop
    - Do not continue the transmission even in collision as CSMA
  - If the channel is busy, wait then transmit with probability p
- Retransmit after a random waiting time.

Comparison between channel division and random access

- Channel division
  - Efficient, treat stations equally.
  - Waste of resources if one station has much smaller data to send than the others
- Random access
  - When total load is small: Efficient since each station can use the whole chanel
  - When total load is large: Collision possibility increases.
- Token control: compromise between the two above methods.

Token Ring

- A “token” is passed from one node to the other in a ring topo
- Only the token holder can transmit data
- After finishing sending data, the token need to be passed to next nodes.
- Some problem
  - Time consuming in passing token
  - Loss of token due to some reasons



Summary on Media access control mechanisms

- Channel division
- Random access
- Token
- What do you thinks about their advantages and weaknesses

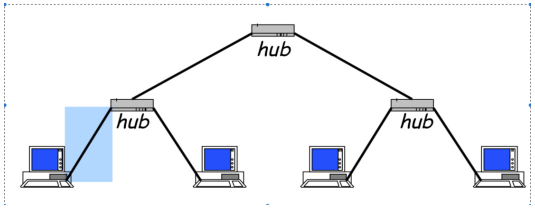
Point-to-Point forwarding mechanism

Hub, Switch, Bridge

Devices of LAN

- Repeater, Hub, bridge and switch
  - All are LAN devices with many ports
- Repeater:
  - Repeats the bits received in one port to the other port
  - One network with repeaters = one collision domain
  - Repeater is a physical layer system.
- Hub:
  - Receive the signal from one port (amplify) and forward to the remaining ports
  - Do not offer services of datalink layer
  - Layer 1 intermediate system

Hub

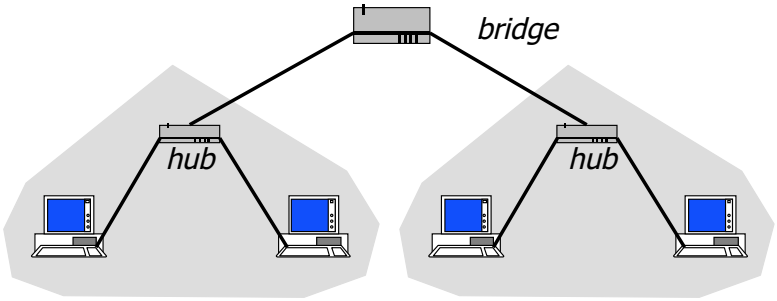


Hub=Multiple port repeater
Single collision domain

Devices of LAN (cont.)

- Bridge
  - More intelligent than hub
  - Can store and forward data (Ethernet frame) according to MAC address.
  - Bridge breaks the network into two collision domains.
  - Layer 2 intermediate system
- Switch
  - More ports than bridge
  - Can store and forward data according to MAC address
    - Receive full frame, check error, forward

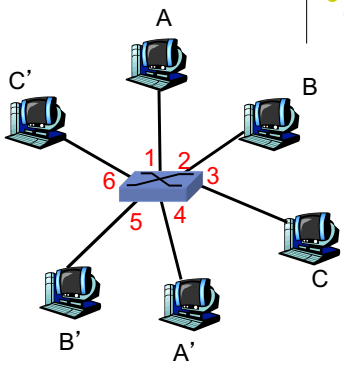
Bridge



Two ports systems
- Forward frames from one port to the other based on MAC address
- Create two collision domains

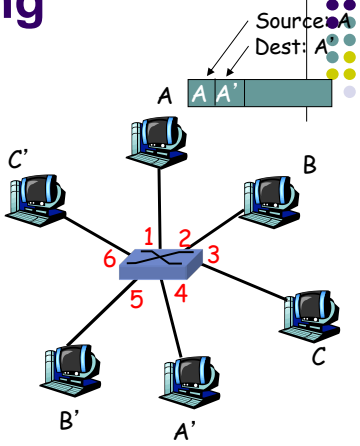
Switch

- Allows multiple node pairs sending data in the same time
  - E.g. A-to-A' and B-to-B' without collision
  - Each link is an independant collision domain
- Switch has a table of MAC addresses showing which node connects to which port
  - (MAC address of host, port index, TTL)



Switch: Self learning mechanism

- Switch learns the MAC addresses of all hosts connected to the switch
- Each time switch receive a frame, it may update the source MAC address of the frame and corresponding connected port
  - if the address is not in MAC table
  - If the address is in the MAC table, but the corresponding pỏt is different to that in the table
- Forwarding table



MAC addr	interface	TTL
A	1	60

Switch: forwarding mechanism

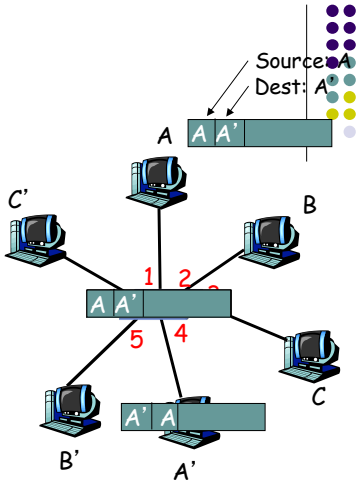
When receiving a frame

1. The incoming port and MAC associated is learnt
2. Looking for outgoing port based on destination MAC and forwarding table
3. if outgoing port is found
   then {
   if incomming port== outgoing port
   then destroy the frame
   else forward the frame to outgoing port
   }
   else broacast the frame

Ex:

- Outgoing port unknown: *Broadcast*
- Know A:

Direct transfering

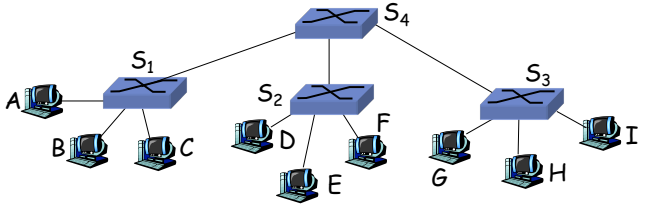


MAC addr	interface	TTL
A	*1*	*60*
A'	*4*	*60*

Forwarding table (empty initially)

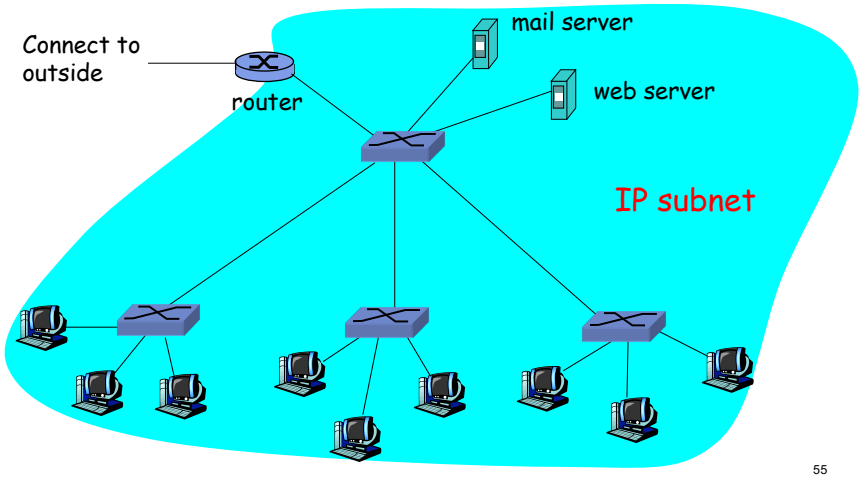
Connecting switch in cascade

- Switches could be connected to eachother



- Switches in cascade uses also self learning mechanism
  - A sends a frame to I, S1 learns the address of A and connected port, broadcasts the frame to: B, C, S4
  - S4: learns A from S1, broadcasts the frame to: S2, S3
  - S3: learns A from S4, broadcasts the frame to: G, H, I

A typical LAN



Flow control

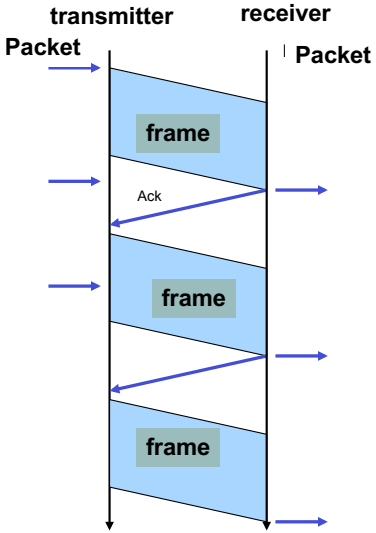
What is flow control

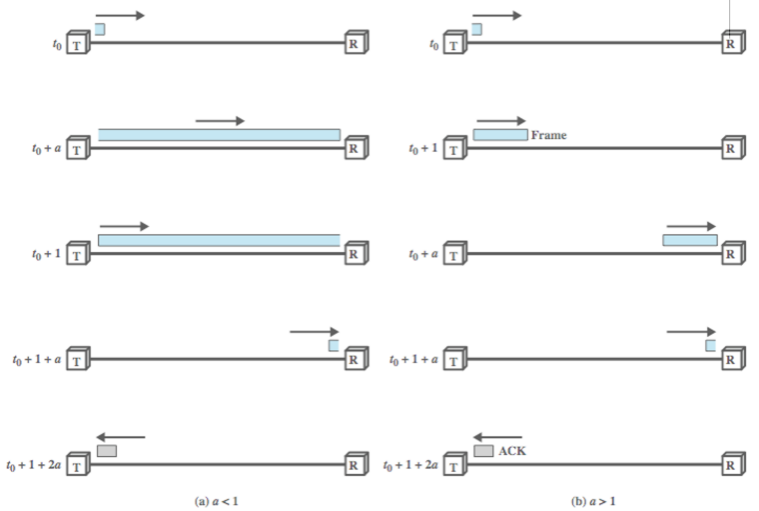
- Goal: Make sure that the sender does not overload the receiver
- Why overloading?
  - The receiver stores data frame in buffer.
  - Receiver performs some processing before deliver data to the upper level.
  - Buffer could be full, leaving no space for receiving more frame → some data fram must be dropped.
- Problem of errors in transmission is excluded
  - All frames are transmitted to correct receiver without error
  - Propagation time is small and could be ignored
- Solution
  - Stop-and-wait mechanism
  - Sliding window mechanism

Stop-and-wait

- Principles
  - Transmitter sends a single frame
  - Receiver receives the fram, process and then informs the transmitter that it is ready to receives next frames by a clear acknowledgement (ACK).
  - Transmitter waits until reception of the ACK before sending next frames.

Stop-and-wait





Stop-and-wait

- Advantage
  - Simple, suitable for transmission of big size frames
- Weakness
  - When frames are small, the transmission chanel are not used efficiently.
  - Cannot use often for big size frame due to
    - Limitation in buffer size
    - Big size frame prones to bigger error probability
    - In shared medium, it is not convenient to leave one station using medium for long time

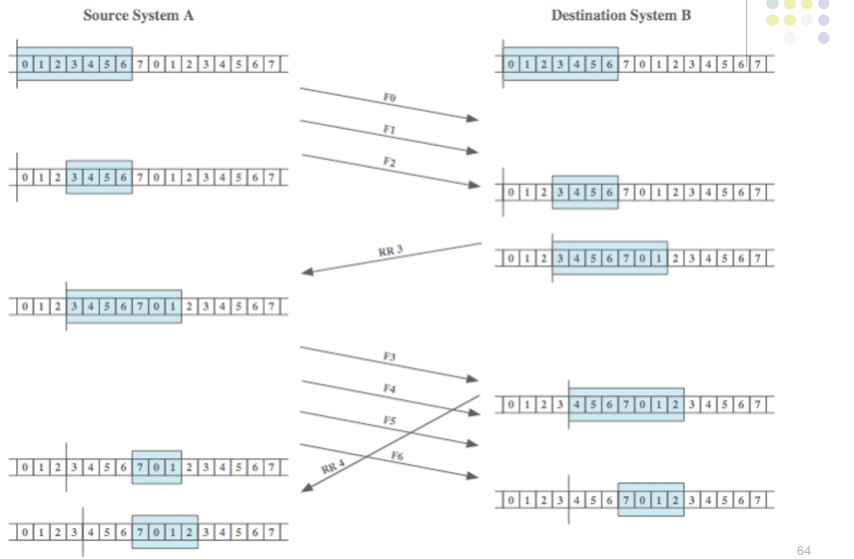
Sliding window: principle

- Transmitter sends more than one frame without waiting in order to reduce waiting time
- Transmitted frame without ACK ➔stored in a buffer of the transmitter.
- Frame arrives to receiver ➔ put in buffered and get out one by one for processing and send back ACK/NAK
- Buffer of transmitter = buffer of receiver
- Number of frames to be transmitted without ACK depends on the size of free buffer
- When transmitter receives ACK, it realises the succesfully transmitted frame from buffers
- Transmitter continues sending a number of frame equivalent to the number of succesfully trasmitted frames.

Sliding window: principle

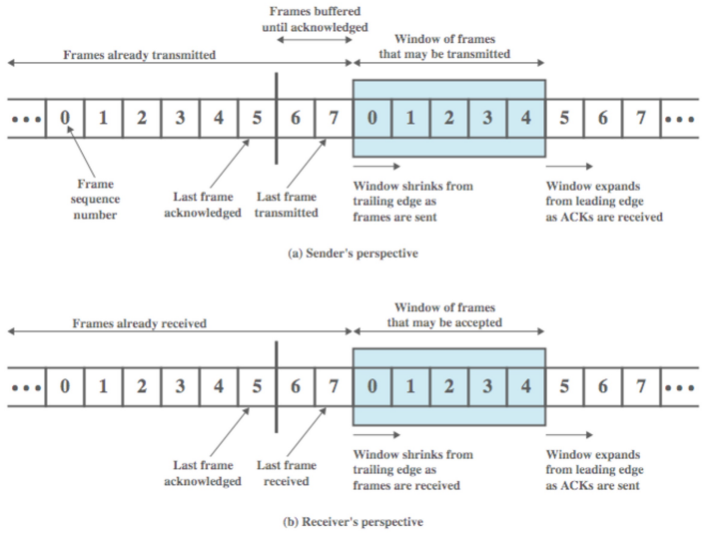
- Assume that A and B are two stations connected by a full duplex media
  - B has a buffer size of n frame.
  - B can receives n frame without sending ACK
- Acknowledgement
  - In order to keep track of ACKed frames. It is neccessary to number frames.
  - B acknowledge a frame by telling A which fram B is waiting for (by number of frame), implicitely saying that B receives well all other frame before that.
  - One ACK frame serves for acknowledes several frames.

Sliding windows: principle



Window list the frames to transmit **Window list the frames in waiting to receive**

Sliding windows



Sliding windows

- Frame are numbered. The maximum number must not be smaller than the size of the window.
- Frame are ACKed by another message with number
- Accumulated ACK: If frame 1,2,3,4 are well receive, just send ACK 4
- ACK with number k means all frame k-1, k-2 …already well received.

Sliding windows

- Transmitter needs to manage some information:
  - List of frames transmitted sucessfully
  - List of frames transmitted without ACK
  - List of frames to be sent immediatly
  - List of frames NOT to be sent immediately
- Receiver keep tracks of
  - List of frames well received
  - List of frames expected to receive

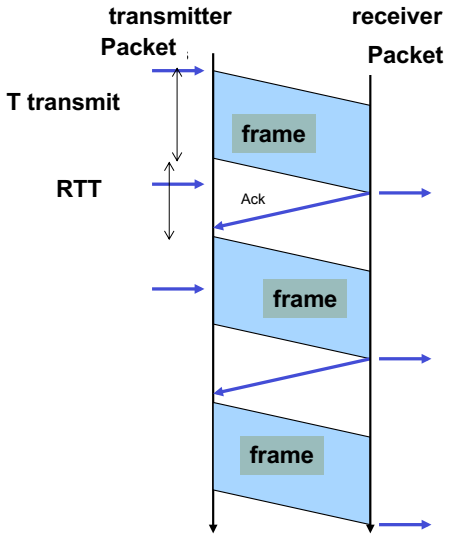
Piggy backing

- A and B transmitte data in both sides
  - When B needs to send an ACK while still needs to send data, B attaches the ACK in the Data frame: Piggybacking
  - Otherwise, B can send an ACK frame separatly
  - After ACK, if B sends some other data, it still put the ACK information in data frame.
- Sliding window is much more efficient than Stop-and-Wait
- More complicated in management.

Exercices

- Given a link with rate R=100Mbps
- We need to send a file over data link layer with file size L=100KB
- Assume that the size of a frame is: 1KB, header size is ignored
- Round trip time (RTT) between 2 ends of the link is 3ms
- An ACK message is sent back from receiver whenever a frame is arrived. Size of ACK message is negligible
- What is the transmission time required if using Stop-and-wait mechanism?
- Transmission time with sliding window if the window size is =7?
- Which size of window allow to obtain the fastest transmission?

Transmission time with Stop-and-wait

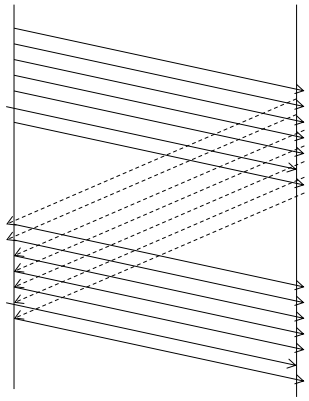


Transmission time with Stop-and-wait

- T total= Nb.frame * (T_transmit + RTT)
- T_transmit (F) = L(Frame)/ R
- Nb. frame = L/L(frame)

- With the given parametters
- Nb. frame =100 KB/1KB =100
- T_transmit (F) = 1KB/100 Mbps =10^3*8/10^8 =8. 10^-5 (s)=0.08 (ms)

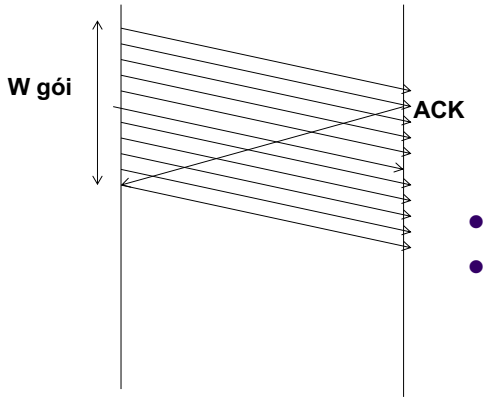
Sliding windows



Trasmission time with window size 7

- T fastest= (T transmit 7 frames+ wait) * Nb. Waiting time.
- 1 waiting= (T transmit 1 frame+ RTT) – T transmit 7 frames
- Nb. Waiting time= Nb frame /7

Fastest transmission time with sliding window



- Fastest transmission time obtained if transmitter receives ACK of the first frame when it finishes transmitting the last frame of the sliding window.
- Window size:W
- T transmit(W fram) >= T transmit first frame + RTT

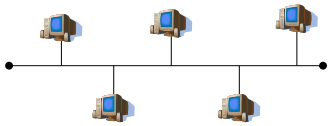
Fastest transmission time with sliding window

- T transmit (W frame) = W * 1KB/R
- => (W-1)*1KB/R >= RTT
- => W >= RTT*R/1KB +1
- W>= 3ms * 100 Mbps/ 1KB + 1
- W>=38.5
- Smallest value of W = 39
- Time to transmit all data L = L/R + RTT =8 ms +3ms =11 ms

LAN: Local Area Network

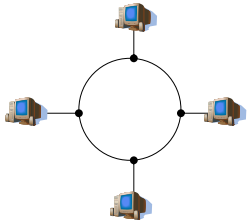
Reading: 4.3 Computer Networks, Tanenbaum

LAN topology

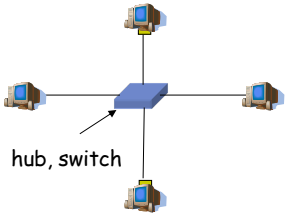


Traditional bus topo

WLAN



Ring



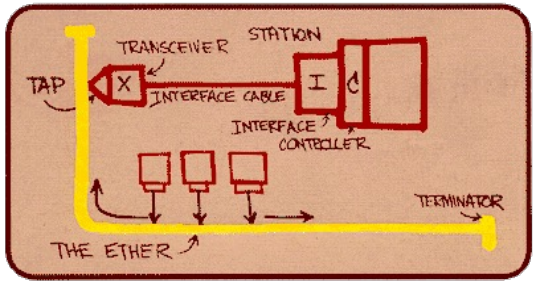
Star

LAN Standards

- IEEE 802 contains many standards for LAN technology.
  - 802.3: Ethernet
  - 802.4: Token bus
  - 802.5: Token ring
  - 802.11 a/b/g/n: Wireless LAN (Wifi)
  - 802.16: WiMax.

Ethernet LAN

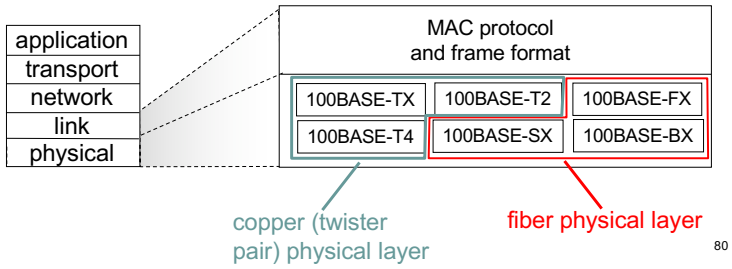
- Layer 2 technology for communication in LAN, invented in 1976
- Standardized in IEEE 802.3
- Ethernet LAN could have different speeds: 3 Mbps – 10 Gbps
  - Ethernet: 10BaseT, 10Base2…



Metcalfe's Ethernet sketch

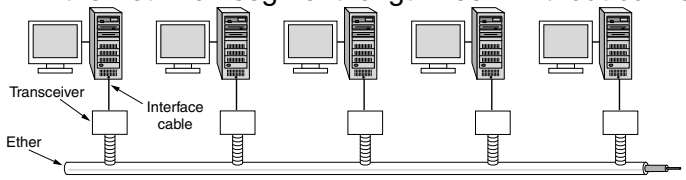
IEEE 802.3 and Ethernet Standards

- Datalink & Physical Layers
- Datalink= LLC + MAC
- MAC: CSMA/CD in classical Ethernet
- Several type of Ethernet
  - Same MAC and frame structure
  - Different rate: 2 Mbps, 10 Mbps, 100 Mbps, 1Gbps, 10G bps
  - Different cable: Optical fiber, coaxial, twisted pair



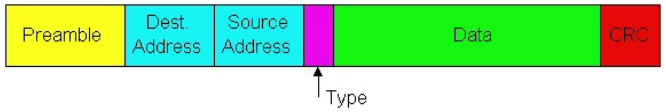
Classical Ethernet

- Bus topology was popular in the past
- All nodes share the same communication medium. Could used a central hub for connecting nodes.
- Use CSMA/CD for media access control.
- Use Manchester encoding at Physical layer
- Use coaxial cable
- Thick Ethernet: Max segment length 500m without converter
- Thin Ethernet: Max segment length 185m without converter



Ref: Computer Network, Tanenbaum

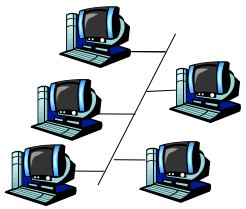
Structure of Ethernet frame



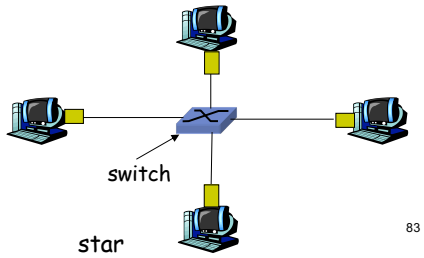
- Preamble: Marking the starting of a frame
- Address: Physical addresses of source and destination
  - 6 bytes
- Type: Uppper layer protocol (IP, Novell IPX, AppleTalk, …)
- Checksum: Error detection code. CRC

Switched Ethernet

- Switched Ethernet (nowdays):
  - Star topology,
  - Use a central switch Ethernet
  - The switch outputs a frame only to the port linking to the destination ➔ independent connection for each pair of two nodes
  - No collision
  - No media access control is needed.



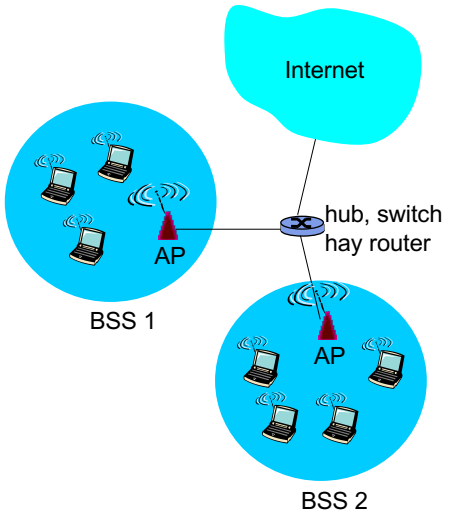
bus: coaxial cable



star

Wireless LAN

Overview of 802.11 LAN



- Include base station = access point) and stations with wireless network interfaces
- Base station mode
  - Basic Service Set (BSS)
    - wireless hosts
    - access point (AP): base station
- Ad hoc mode:
  - Stations pay also the role of AP

Standards

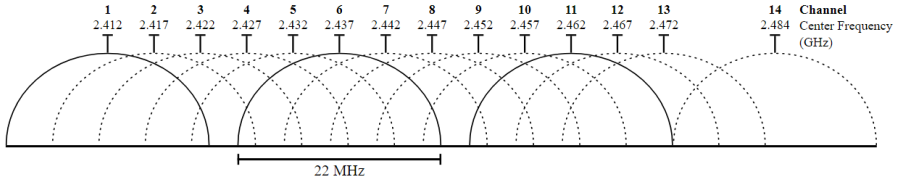
- 802.11b
  - Band 2.4-5 GHz (unlicensed spectrum)
  - Maximum speed 11 Mbps
- 802.11a
  - Band 5-6 GHz
  - Maximum speed 54 Mbps
- 802.11g
  - Band 2.4-5 GHz
  - Maximum speed 54 Mbps
- 802.11n: use multiple antennas (MIMO)
  - Band 2.4-5 GHz
  - Maximum speed 200 Mbps

- Employ CSMA/CA for multiple access control
- Working in 2 modes : base-station and ad hoc

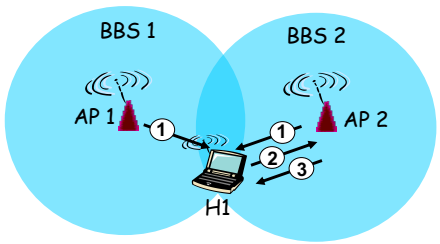
802.11: Chanel and connection

- Band is divided into 14 chanels spaced 5MHz apart. Europe uses 13 channels, America uses 11 channels, Japan uses 14 channels.
  - Admin chooses a working frequency for AP (may leave AP to choose automatically)
- Station: need to connect to an AP
  - Scan channels, listen to initial frames (*beacon frames)* containing the ID (SSID) and MAC address of the AP
  - Choose one AP.

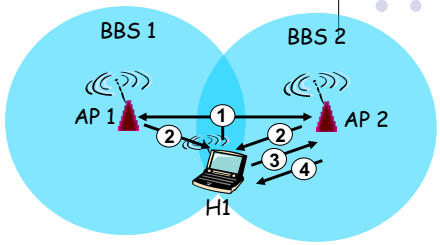
802.11: Kênh, liên kết



Scanning mechanism: active/passive



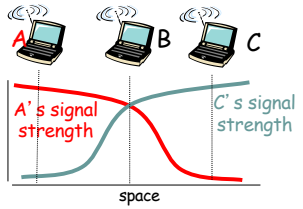
Passive Scanning:
(1) Beacon frames are sent from APs
(2) H1 send a connection request to AP2
(3) AP2 accepts the request



Active Scanning:
(1) H1 broacast the request to find an AP
(2) APs reply with their information
(3) H1 send a connection request to AP2
(4) AP2 accepts the requests

IEEE 802.11: Multiple access control

- 802.11: CSMA
- 802.11: CA – Collision Avoidance
  - It is difficult to implement Collision detection (CD) in wireless environment.
  - In some cases, it is even impossible to detect the collision : hidden terminal, fading



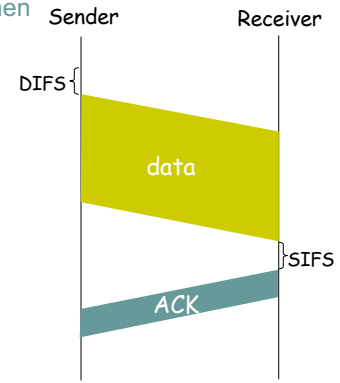
IEEE 802.11 MAC Protocol: CSMA/CA

Sender

1 If the channel is available during **DIFS** time then
  Send the entire frame (no CD)
2 if channel is busy then
  Starting random back-off (waiting)
  At the end of back-off time, send data
  If no ACK is received, double the back-off time and try again.

Receiver

- If receive well a frame then
  reply by an ACK after **SIFS**



DIFS: Distributed Inter Frame Space

SIFS: Short Inter Frame Space

Why need ACK?

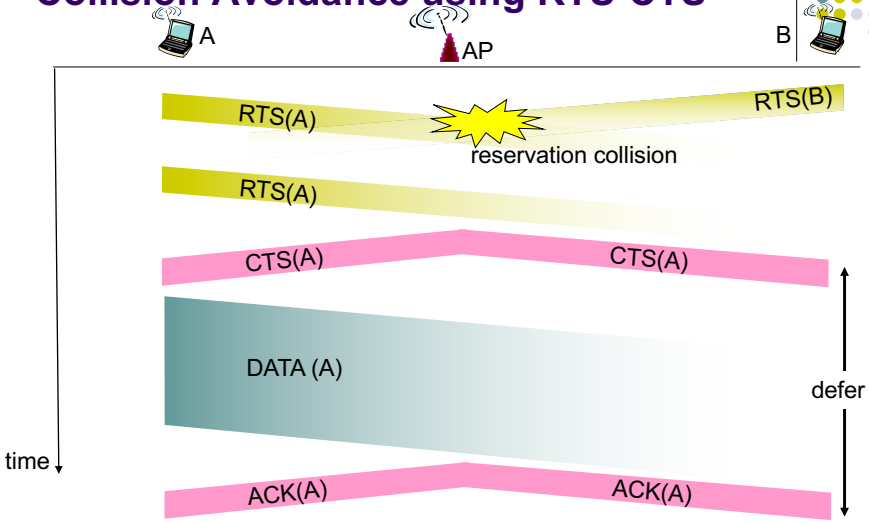
Avoid Collision mechanism

Idea: Sender can reserve channel without random access→ avoid collision for long frame

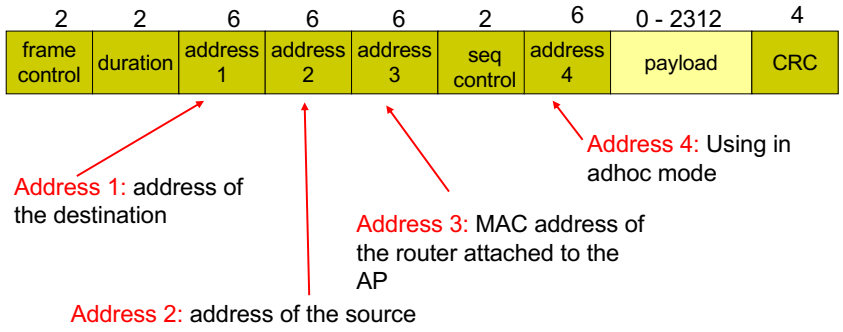
- Sender send frame RTS (request-to-send) to BS using CSMA
  - RTS may meet a collision (with low probability because the frame is short)
- BS broadcast the frame CTS (clear-to-send CTS) to answer
- All stations receive CTS
  - Sender send data frame
  - All other stations has to cancel the intention to send frames.

Avoid collision thanks to the reservation made by small size control frames

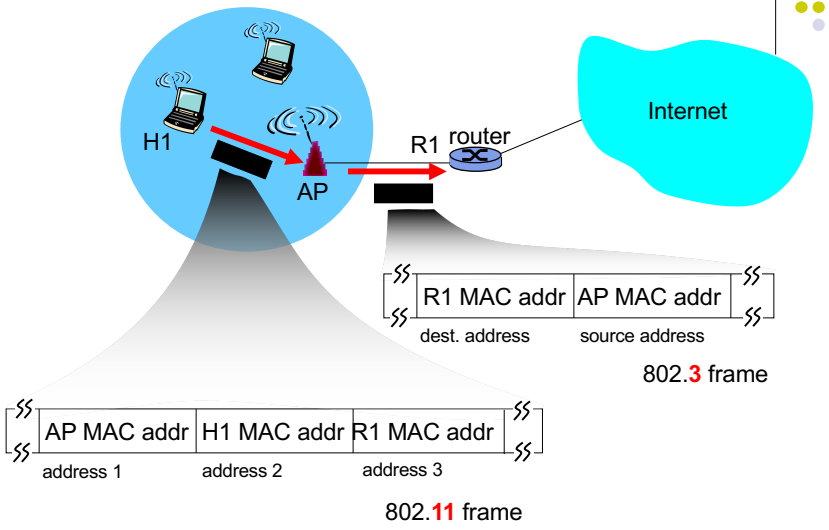
Collision Avoidance using RTS-CTS



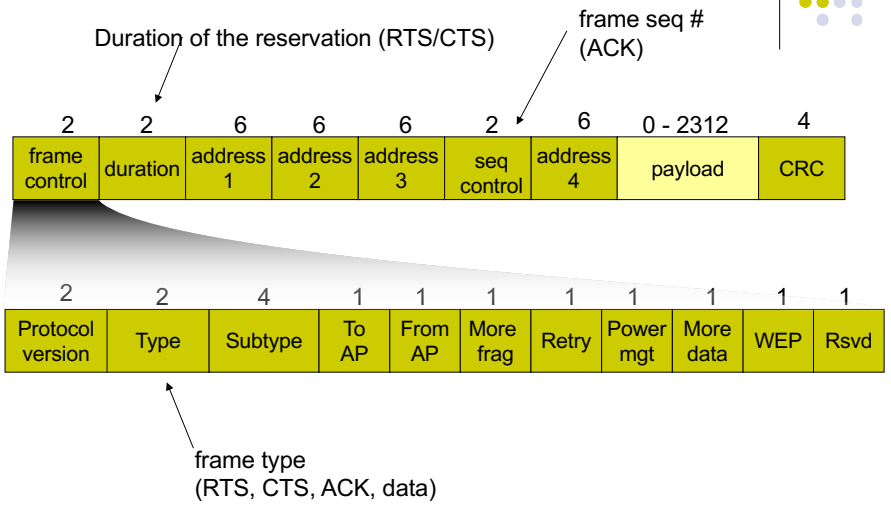
802.11 frame: Addressing



802.11 frame: Addressing



802.11 frame



Access networks using optical fiber

Access networks

- Access networks gather data from users to feed to core network
- Popular access networks for providing services to users
  - Public telephone network
  - TV Cable network
  - Internet to home network.

Architecture of access network

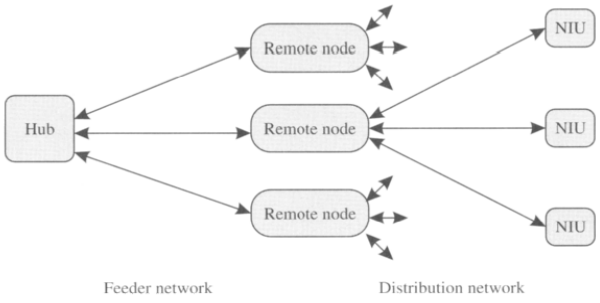


Figure 11.1 Architecture of an access network. It consists of a hub, which is a telephone company central office or cable company head end, remote nodes deployed in the field, and network interface units that serve one or more individual subscribers.

Architecture of access network

- Hub
  - Device on the service provider side receiving data
- Network Interface Unit (NIU)
  - Device on the user side connecting an user or an organization
- Remote Node (RN)
  - In broadcasting networks, RN distribute data from Hub to NIUs
  - In switched networks, RN receive data from Hub and distribute different flows to NIUs

Development of technologies for connecting to ISP using cable

- Dial-up:
  - speed 56kbps,
  - Using telephone line
  - Data is transmitted using the same frequency with voice → either data or voice communication available
  - Obsolete technology, used before 2000
- ADSL technology:
  - Speed few Mbps,
  - Using telephone line
  - Data is transmitted in different frequency than voice, technology used in 2000-2010
- Technology using TV cable
- FTTH technology:
  - Speed dozens Mbps,
  - Using optical fiber
  - Popular technology nowadays

Optical access network: FTTx

- Data is distributed on the fiber cable in the distribution network until ONU (Optical Network Unit)
  - Expectation: fiber approaches the customers
- **FTTCab** (*Fiber To The Cabinet*): Optical fiber ends at a cabinet in less than 1 km distance to the subscriber using copper cable.
- **FTTC** (*Fiber To The Curb*) / **FTTB**(*Fiber To The Building*); ONU serves some subscribers (8 to 64); from ONU to NIU using copper cable (< 100m)
- **FTTH** (*Fiber To The Home*); ONUs performs the functionality of NIUs;

Optical access network: FTTx

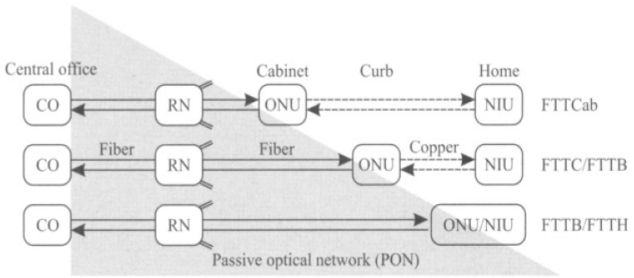
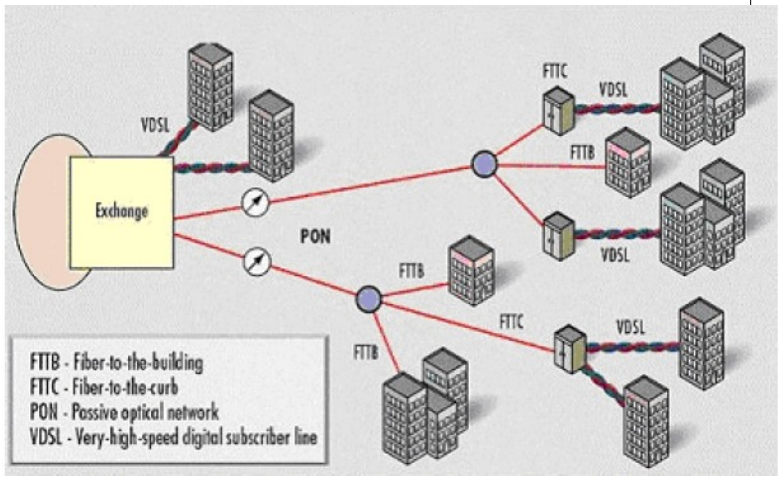


Figure 11.5 Different types of fiber access networks, based on how close the fiber gets to the end user. In many cases, the remote node may be located at the central office itself. The ONUs terminate the fiber signal, and the links between the ONUs and the NIUs are copper based.

- ONU: ex: optical modem.

Optical access networks: FTTx

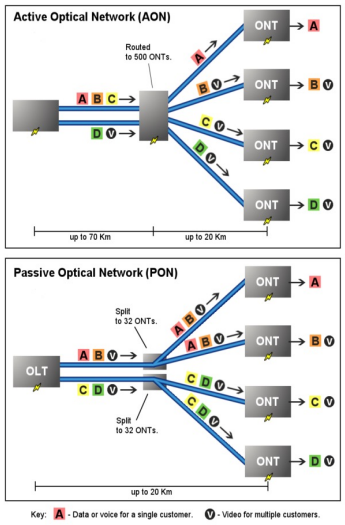


AON vs. PON

Remote Note (Distribution nodes) chia dữ liệu về các đích.

AON: Active Optical Network
- Network using active technology
(Remote Node consume electricity)
-Remote node analyse and forward packets to destination according to addresses
-Cable distance can go up to 100 km.

PON: PassIve Optical Network
- network using passive technology
(Remote Node does not consume electricity)
-Remote node (Splitter) does not analyze but repeat signal to all out ports
-Upstream: MUX from different sources using TDM (TDM PON) or WDM (WDM PON)
-Cable distance is limited within 20km



EPON: Ethernet PON

- EPON: PON transport Ethernet frames
- Down stream
  - Broadcast common data

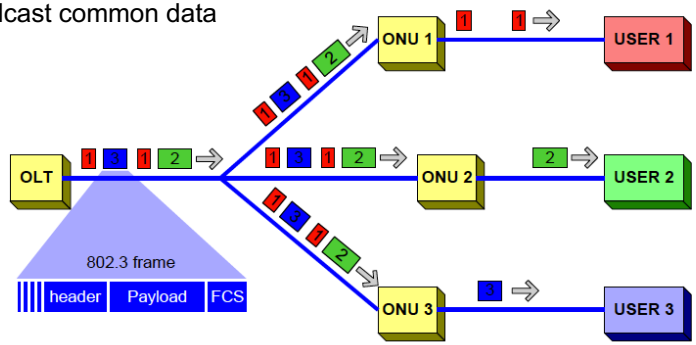


Figure 8-6. Downstream traffic in EPON.

EPON

- Upstream: Mux Ethernet frames from users to the common link OLT-RN using TDM

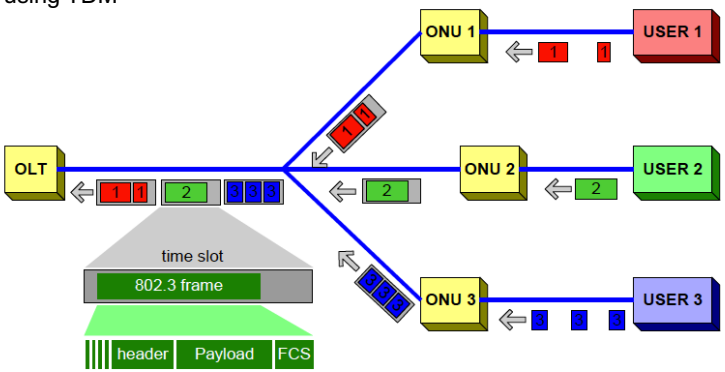


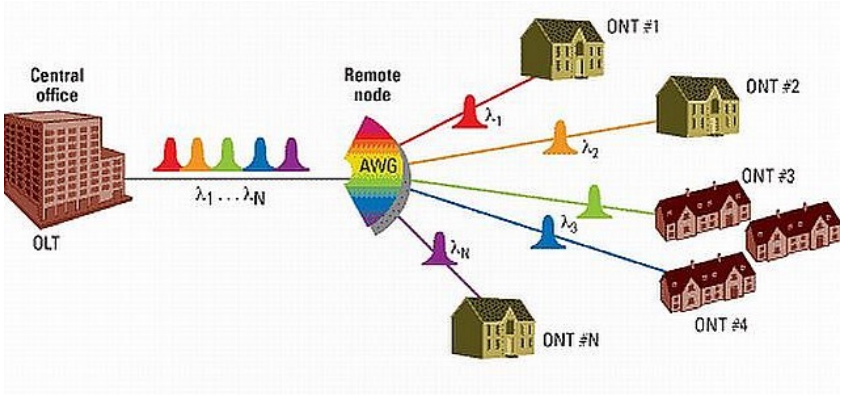
Figure 8-7. Upstream traffic in EPON.

GPON: Gigabit Capable PON

- GPON may be used to carry different data: Ethernet, ATM, voice …
- Data from OLT to users share common chanel between OLT or RN
  - Downstream broadcast
  - Upstream TDM
  - Data are encapsulated in GPON frames with ID of the receiver (downstream direction), sender (upstream direction)

WPON (WDM PON)

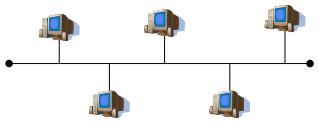
- Developed by compagnies and has not been standardized
- Each ONT use a wavelength to transmit data
- Remote note is AWG (arrayed waveguide grating). The AWG is capable to MUX/DEMUX wavelengths from up and down streams.



Lecture 5
LAN: Local Area Network

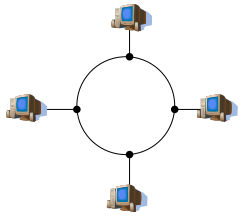
Reading: 4.3 Computer Networks, Tanenbaum

LAN topology

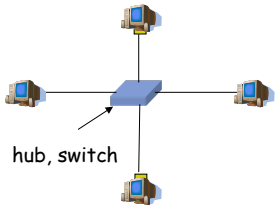


Traditional bus topo

WLAN



Ring



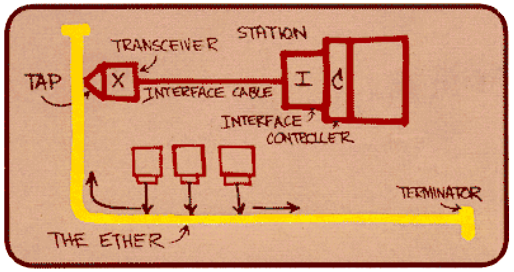
Star

LAN Standards

- IEEE 802 contains many standards for LAN technology.
  - 802.3: Ethernet
  - 802.4: Token bus
  - 802.5: Token ring
  - 802.11 a/b/g/n: Wireless LAN (Wifi)
  - 802.16: WiMax.

Ethernet LAN

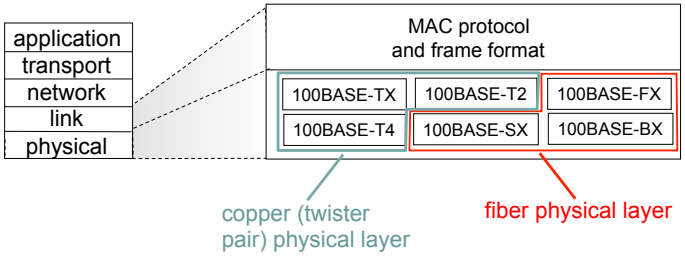
- Layer 2 technology for communication in LAN, invented in 1976
- Standardized in IEEE 802.3
- Ethernet LAN could have different speeds: 3 Mbps – 10 Gbps
  - Ethernet: 10BaseT, 10Base2…



Metcalfe's Ethernet sketch

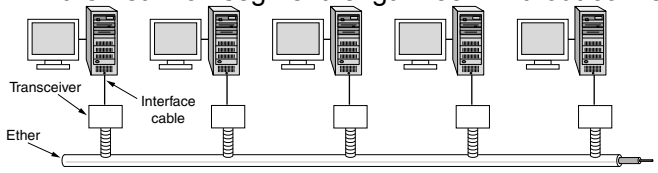
IEEE 802.3 and Ethernet Standards

- Datalink & Physical Layers
- Datalink= LLC + MAC
- MAC: CSMA/CD in classical Ethernet
- Several type of Ethernet
  - Same MAC and frame structure
  - Different rate: 2 Mbps, 10 Mbps, 100 Mbps, 1Gbps, 10G bps
  - Different cable: Optical fiber, coaxial, twisted pair



Classical Ethernet

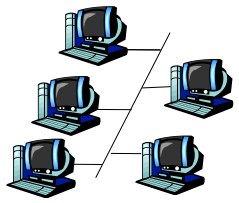
- Bus topology was popular in the past
- All nodes share the same communication medium. Could used a central hub for connecting nodes.
- Use CSMA/CD for media access control.
- Use Manchester encoding at Physical layer
- Use coaxial cable
- Thick Ethernet: Max segment length 500m without converter
- Thin Ethernet: Max segment length 185m without converter



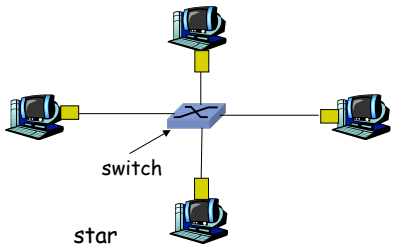
Ref: Computer Network, Tanenbaum

Switched Ethernet

- Switched Ethernet (nowdays):
  - Star topology,
  - Use a central switch Ethernet
  - The switch outputs a frame only to the port linking to the destination ➔ independent connection for each pair of two nodes
  - No collision
  - No media access control is needed.



bus: coaxial cable



star

Classical Ethernet

- Ethernet frame

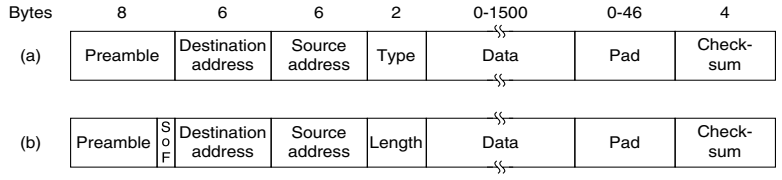
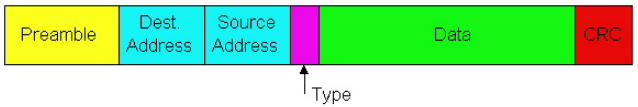


Figure 4-14. Frame formats. (a) Ethernet (DIX). (b) IEEE 802.3.

Structure of Ethernet frame



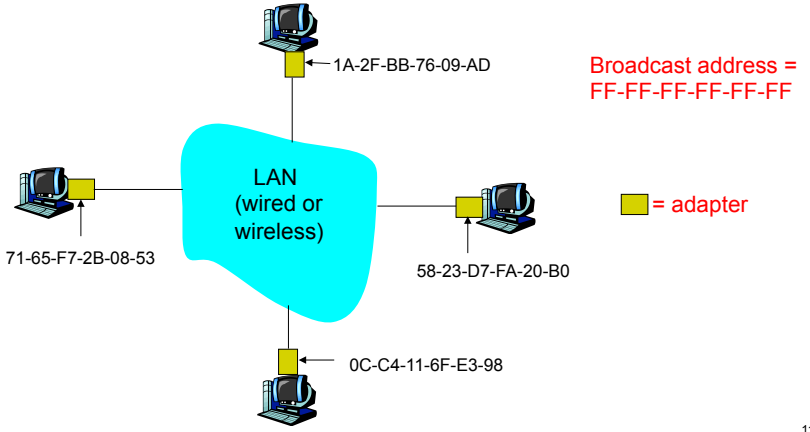
- Preamble: Marking the starting of a frame
- Address: Physical addresses of source and destination
  - 6 bytes
- Type: Uppper layer protocol (IP, Novell IPX, AppleTalk, …)
- Checksum: Error detection code. CRC??

MAC address and ARP

- IP Address :
  - 32-bit
  - Used in Network layer
- MAC address:
  - Used in Data link layer
  - 48 bit

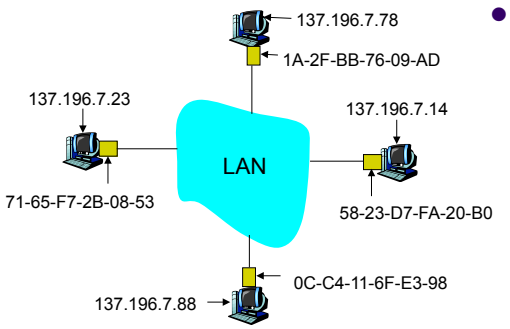
ARP and MAC address

Each network adapter has a MAC address



ARP: Address Resolution Protocol

Question: Identify MAC address from an IP address



- Each network node (host, router) has an ARP table
- ARP table: contain mapping IP/MAC of some nodes

< IP address; MAC address; TTL>

- TTL (Time To Live): ~20 min.

ARP : Work on a network segment

- A wants to send data to B on datalink layer but do not know MAC of B
- A broadcast an ARP package stating the IP address of B
- B receives the package with is address and reply to A with MAC of B
- A saves the MAC address of B

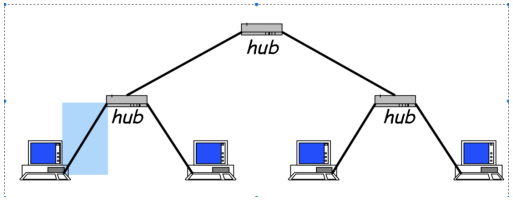
LAN (cont.)

Hub, Switch, Bridge

Devices of LAN

- Repeater, Hub, bridge and switch
  - All are LAN devices with many ports
- Repeater:
  - Repeats the bits received in one port to the other port
  - One network with repeaters = one collision domain
  - Repeater is a physical layer system.
- Hub:
  - Receive the signal from one port (amplify) and forward to the remaining ports
  - Do not offer services of datalink layer
  - Layer 1 intermediate system

Hub

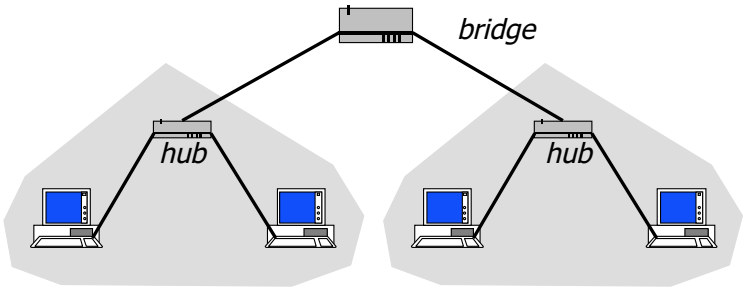


Hub=Multiple port repeater
Single collision domain

Devices of LAN (cont.)

- Bridge
  - More intelligent than hub
  - Can store and forward data (Ethernet frame) according to MAC address.
  - Bridge breaks the network into two collision domains.
  - Layer 2 intermediate system
- Switch
  - More ports than bridge
  - Can store and forward data according to MAC address
    - Receive full frame, check error, forward

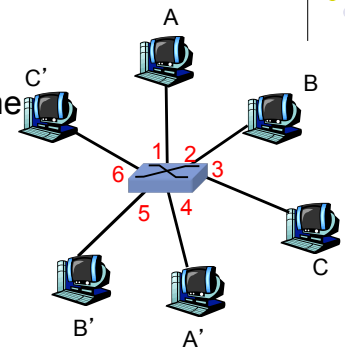
Bridge



Two ports systems
- Forward MAC frame from one port to the other based on MAC address
- Create two collision domains

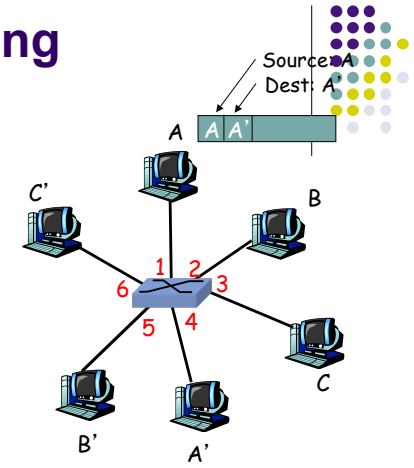
Switch

- Allows multiple node pairs sending data in the same time
  - E.g. A-to-A' and B-to-B ' without collision
  - Each link is an independant collision domain
- Switch has a table of MAC addresses showing which node connects to which port
  - (MAC address of host, port index, TTL)



Switch: Self learning mechanism

- Switch learns the MAC address of all hosts connected to the switch
- Forwarding table



MAC addr	interface	TTL
A	*1*	*60*

Switch: forwarding mechanism

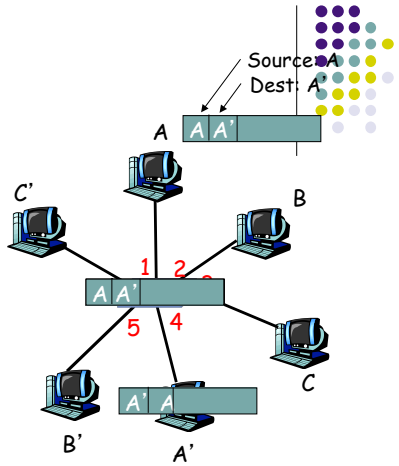
When receiving a fram

1. The incomming port and MAC associated is learnt
2. Looking for outgoing port based on destination MAC and forwarding table
3. **if** outgoing port is found
   then {
   if incomming port== outgoing port
   then destroy the frame
   else forward the frame to outgoing port
   }
   else broacast the frame

Ex:

- Outgoing port unknown: *Broadcast*
- Know A:

Direct transfering

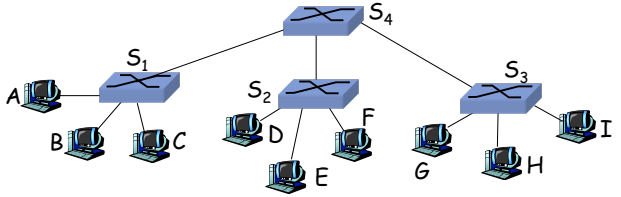


MAC addr	interface	TTL
A	*1*	*60*
A'	*4*	*60*

Forwarding table (empty initially)

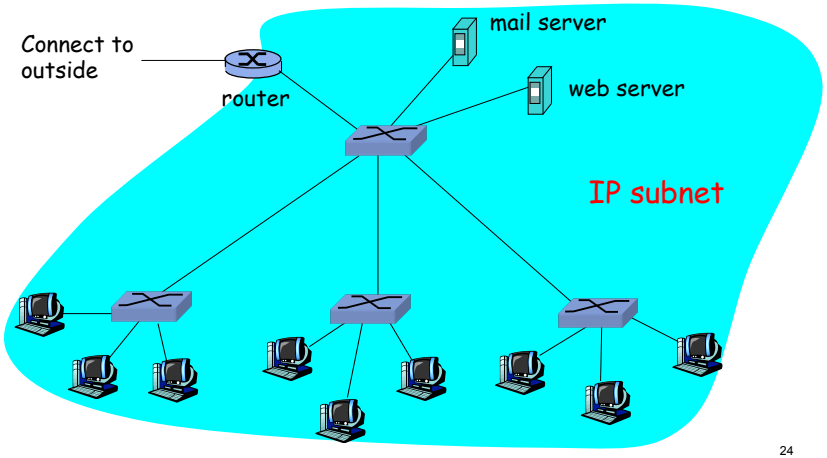
Connecting switch in cascade

- Switches could be connected to eachother



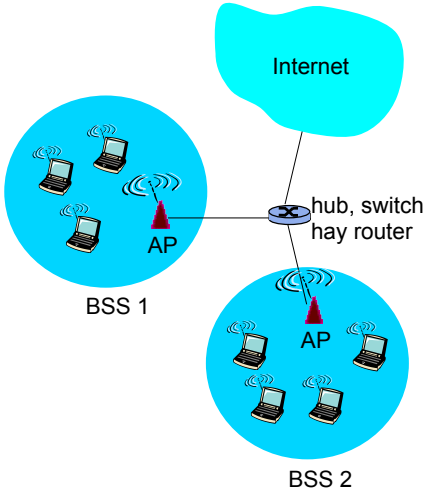
- Switches in cascade uses also self learning mechanism

A typical LAN



Wireless LAN

Overview of 802.11 LAN



- Include base station = access point) and stations with wireless network interfaces
- Base station mode
  - Basic Service Set (BSS)
    - wireless hosts
    - access point (AP): base station
- Ad hoc mode:
  - Stations pay also the role of AP

Standards

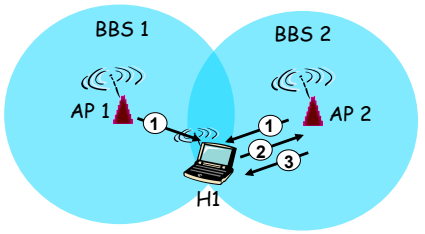
- 802.11b
  - Band 2.4-5 GHz (unlicensed spectrum)
  - Maximum speed 11 Mbps
- 802.11a
  - Band 5-6 GHz
  - Maximum speed 54 Mbps
- 802.11g
  - Band 2.4-5 GHz
  - Maximum speed 54 Mbps
- 802.11n: use multiple antennas (MIMO)
  - Band 2.4-5 GHz
  - Maximum speed 200 Mbps

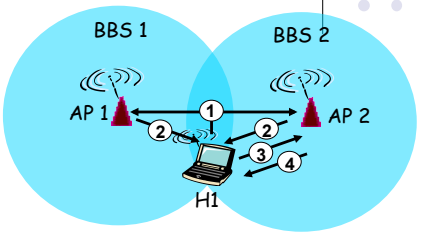
- Employ CSMA/CA for multiple access control
- Working in 2 modes : base-station and ad hoc

802.11: Chanel and connection

- Band is divided into 14 chanels spaced 5MHz apart. Europe uses 13 channels, America uses 11 channels, Japan uses 14 channels.
  - Admin chooses a working frequency for AP (may leave AP to choose automatically)
- Station: need to connect to an AP
  - Scan channels, listen to initial frames (*beacon frames)* containing the ID (SSID) and MAC address of the AP
  - Choose one AP.

Scanning mechanism: active/passive



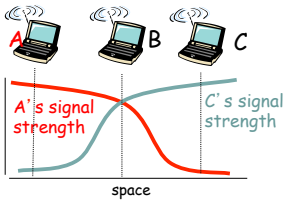


Passive Scanning:
(1) Beacon frames are sent from APs
(2) H1 send a connection request to AP2
(3) AP2 accepts the request

Active Scanning:
(1) H1 broacast the request to find an AP
(2) APs reply with their information
(3) H1 send a connection request to AP2
(4) AP2 accepts the requests

IEEE 802.11: Multiple access control

- 802.11: CSMA
- 802.11: CA – Collision Avoidance
  - It is difficult to implement Collision detection (CD) in wireless environment.
  - In some cases, it is even impossible to detect the collision : hidden terminal, fading



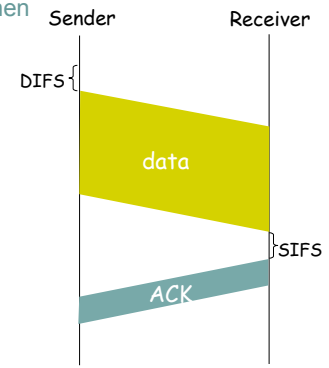
IEEE 802.11 MAC Protocol: CSMA/CA

Sender

1 If the channel is available during **DIFS** time then
  Send the entire frame (no CD)
2 if channel is busy then
  Starting random back-off (waiting)
  At the end of back-off time, send data
  If no ACK is received, double the back-off time and try again.

Receiver

- If receive well a frame then
  reply by an ACK after **SIFS**



DIFS: Distributed Inter Frame Space

SIFS: Short Inter Frame Space

Why need ACK?

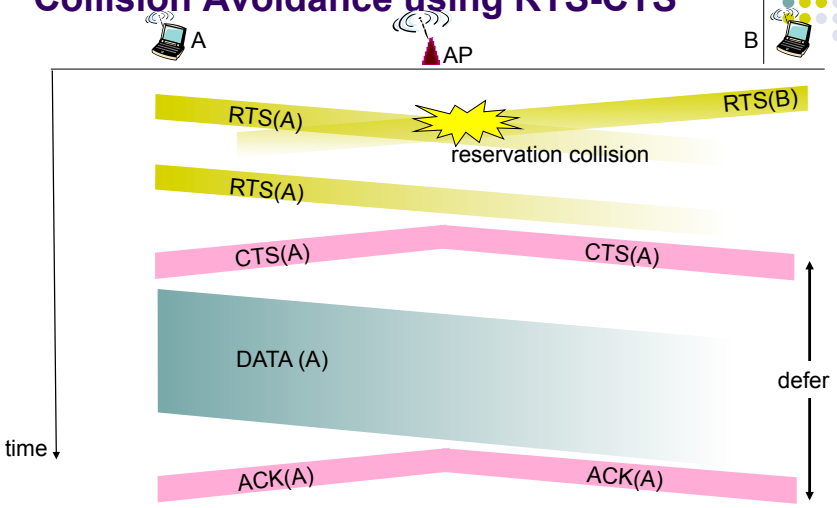
Avoid Collision mechanism

Idea: Sender can reserve channel without random access→ avoid collision for long frame

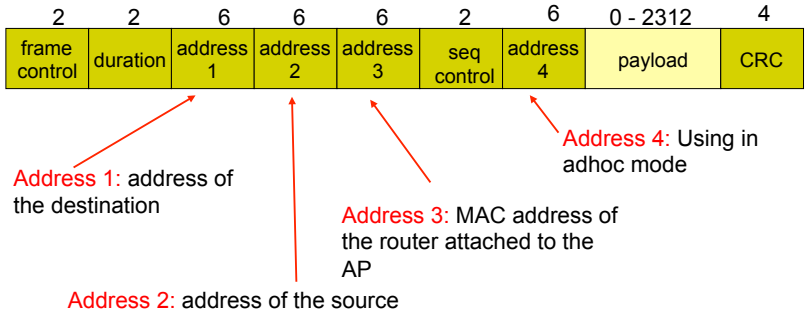
- Sender send frame RTS (request-to-send) to BS using CSMA
  - RTS may meet a collision (with low probability because the frame is short)
- BS broadcast the frame CTS (clear-to-send CTS) to answer
- All stations receive CTS
  - Sender send data frame
  - All other stations has to cancel the intention to send frames.

Avoid collision thanks to the reservation made by small size control frames

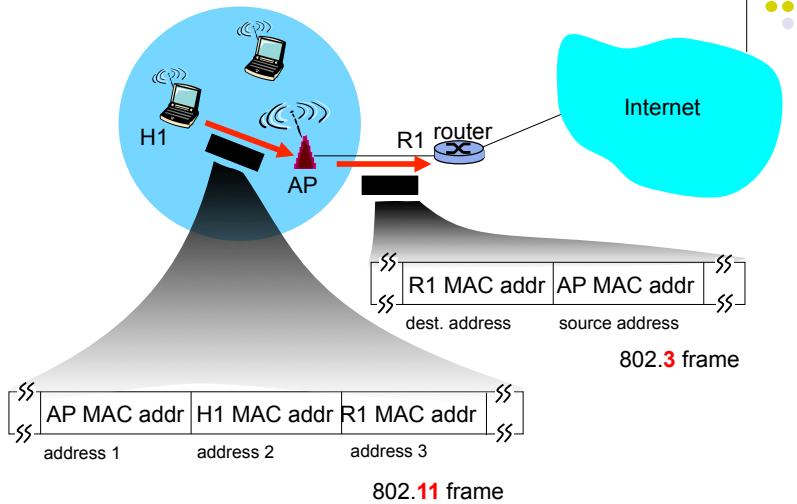
Collision Avoidance using RTS-CTS



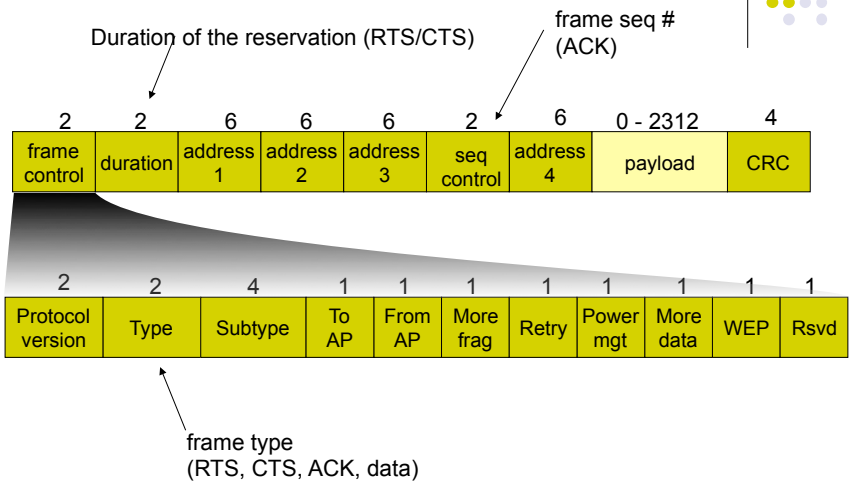
802.11 frame: Addressing



802.11 frame: Addressing



802.11 frame



Mạng truy nhập sử dụng cáp quang

Mạng truy nhập

- Mạng truy nhập thu thập dữ liệu từ phía người dùng và cung cấp cho mạng lõi
- Các dịch vụ phổ biến từ phía người dùng
  - Điện thoại
  - Mạng tryền hình cáp
  - Truyền dữ liệu. Ví dụ trên nền đường truyền điện thoại (xDSL) hoặc cáp quang.

Kiến trúc của mạng truy nhập

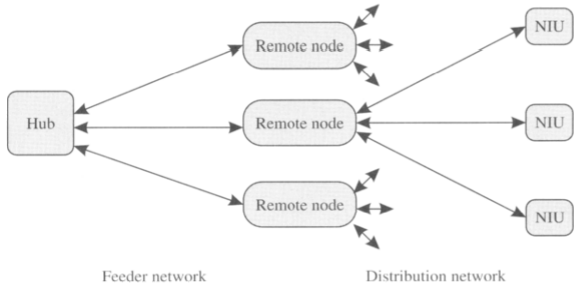


Figure 11.1 Architecture of an access network. It consists of a hub, which is a telephone company central office or cable company head end, remote nodes deployed in the field, and network interface units that serve one or more individual subscribers.

Kiến trúc mạng truy nhập

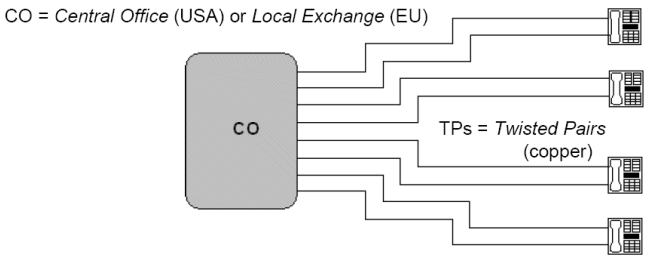
- Hub
  - Nằm phía nhà cung cấp
- NIU: Network Interface Unit
  - Nằm phía người sử dụng
  - Nối với 1 người dùng hoặc 1 doanh nghiệp
- Remote Node
  - Trong mạng broadcast, RN phân phối dữ liệu từ Hub đến mọi NIU
  - Trong mạng switched, RN nhận dữ liệu từ Hub và phân phối các luồng khác nhau đến các NIU

Phân loại mạng truy nhập

- Các loại mạng truy cập phổ biến:
  - Mạng điện thoại
  - Mạng truyền hình cáp
  - Mạng dữ liệu sử dụng cáp quang

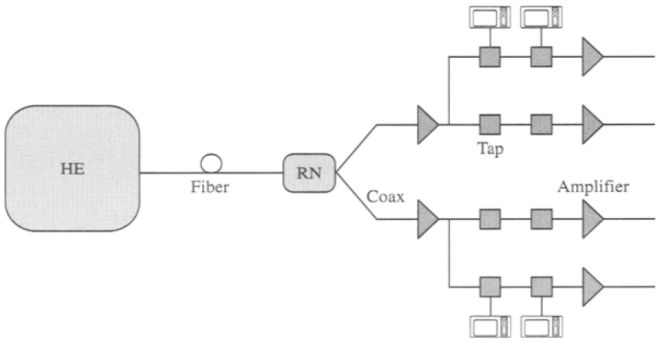
Mạng điện thoại nội bộ

- Sử dụng cáp xoắn



Mạng truyền hình cáp

- Dùng cả cáp đồng trục và cáp quang
  - Hybrid Fiber Coaxial cable: HFC
- HE: headend



Mạng truy nhập quang: FTTx

- Dữ liệu được truyền trên cáp quang trong mạng phân phối (distribution network) cho đến ONU (Optical Network Unit)
  - Mong muốn: Cáp quang đến gần thuê bao nhất
- **FTTCab** (*Fiber To The Cabinet*): Cáp quang kết thúc ở một cabinet, dưới 1km cuối đến thuê bao dùng mạng phân phối cáp đồng.
- **FTTC** (*Fiber To The Curb*) / **FTTB**(*Fiber To The Building*); ONU phục vụ một số thuê bao (8 to 64); từ ONU đến NIU dùng cáp đồng (dưới 100m)
- **FTTH** (*Fiber To The Home*); ONUs thực hiện chức năng của NIUs;

Mạng truy nhập quang: FTTx

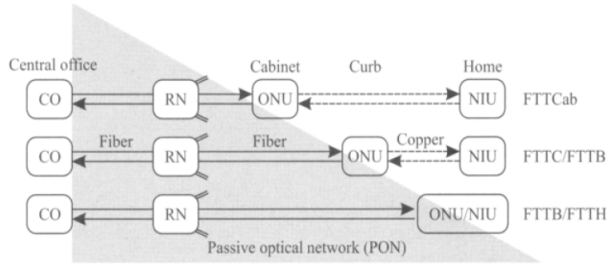
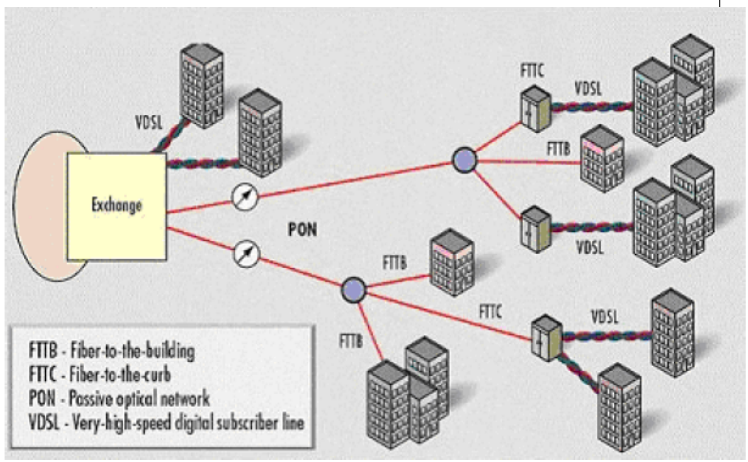


Figure 11.5 Different types of fiber access networks, based on how close the fiber gets to the end user. In many cases, the remote node may be located at the central office itself. The ONUs terminate the fiber signal, and the links between the ONUs and the NIUs are copper based.

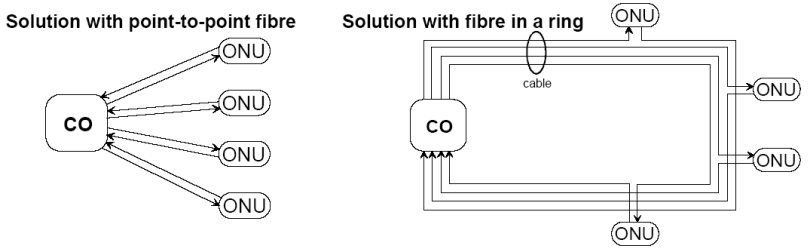
- PON: Passive Optical Network: giữa CO và ONU
- ONU: có thể là modem quang.

Mạng truy cập FTTx



Kiến trúc AF (all fiber)

- Một cặp cáp dành riêng nối CO trực tiếp với mỗi ONU
- Giá thành tỉ lệ với số ONU và chi phí bảo trì cáp
- Sử dụng trong phạm vi nhỏ như doanh nghiệp

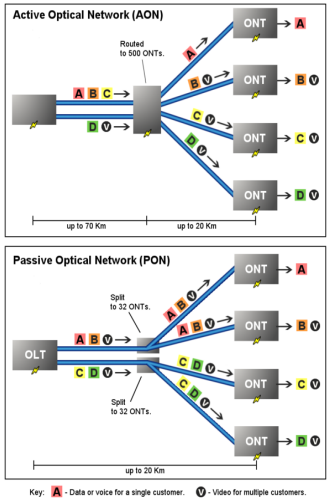


AON vs. PON

Remote Note (Distribution nodes) chia dữ liệu về các đích.

AON: Active Optical Network
- là mạng sử dụng công nghệ chủ động (Remote Node tiêu thụ điện)
- Remote node phân tích và định tuyến riêng các gói tin theo địa chỉ đích
- Khoảng chạy cáp có thể dài đến 100km

PON: PassIve Optical Network
- Là mạng sử dụng công nghệ thụ động, (Remote Node không tiêu thụ điện)
- Remote node (Splitter) không phân tích mà chỉ lặp tín hiệu trên tất cả các cổng ra
- Upstream: MUX từ các nguồn khác nhau bằng TDM (TDM PON) hoặc WDM (WDM PON)
- Khoảng chạy cáp giới hạn 20km



EPON: Ethernet PON

- EPON: PON vận chuyển dữ liệu là các frame Ethernet
  - Chiều xuống (down stream)
    - Quảng bá dữ liệu chung

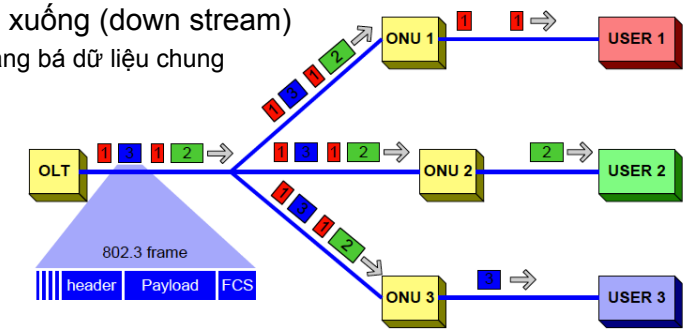


Figure 8-6. Downstream traffic in EPON.

EPON

- Chiều lên (Upstream): dồn kênh theo thời gian (TDM) trực tiếp các gói Ethernet của người dùng từ các nguồn khác nhau vào kết nối chung OLT-RN
- EPON thuộc loại TDM PON

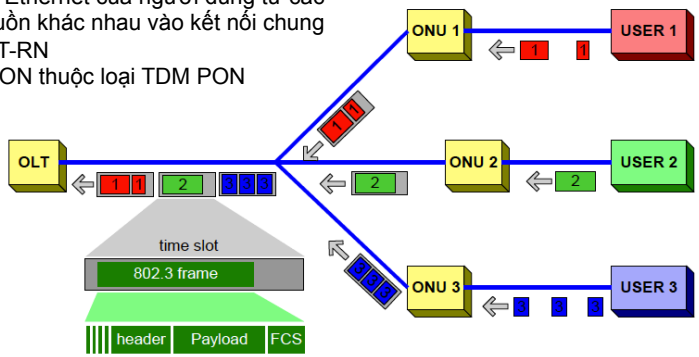


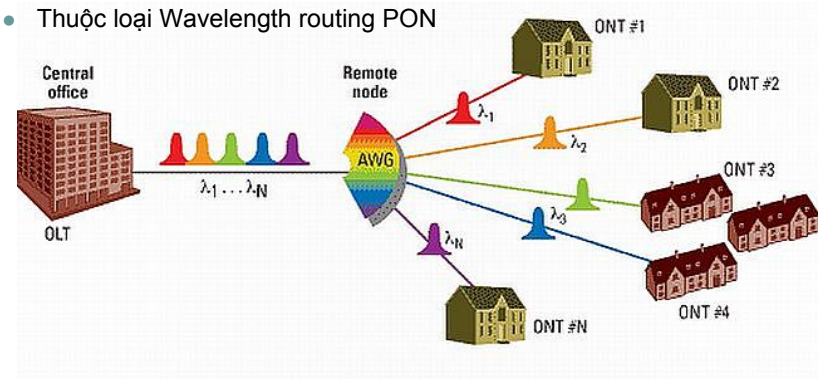
Figure 8-7. Upstream traffic in EPON.

GPON: Gigabit Capable PON

- GPON có thể dùng để tải nhiều dữ liệu khác nhau: Ethernet, ATM, voice …
- Dữ liệu từ OLT đến người dùng chia sẻ kênh chung giữa OLT và RN
  - Downstream broadcast
  - Upstream TDM
  - Các gói được đóng trong khung dữ liệu GPON có trường định danh người nhận (chiều downstream), người gửi (chiều upstream)

WPON (WDM PON)

- Được phát triển bởi các công ty, chưa chuẩn hóa
- Mỗi ONT sử dụng một bước sóng để truyền dữ liệu
- Remote note là AWG thiết bị có khả năng tách ghép các bước sóng, thực hiện MUX/DEMUX theo bước sóng chiều xuống và lên.
- Thuộc loại Wavelength routing PON



Lecture 6: Internet Layer

Reading 5.1. and 5.6 in Computer Networks, Tanenbaum

The lecture uses materials provided by Keio University, Japan

ONE LOVE. ONE FUTURE.

Contents

- Internet Protocol
- IP address and IP packet format
- ICMP- Protocol for control message

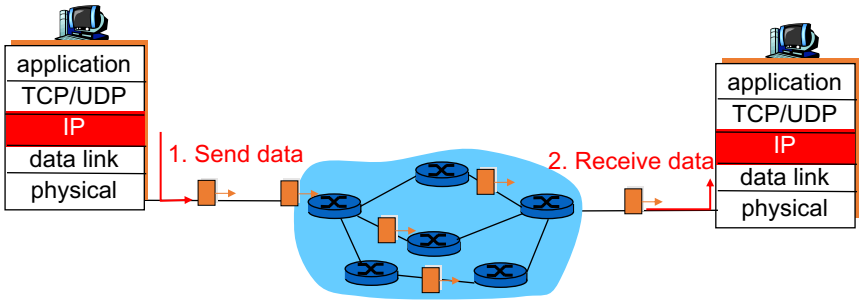
Introduction about IP

Concepts
Store and forward principles
Characteristic of IP

ONE LOVE. ONE FUTURE.

Network layer and Internet protocol

- Role of network layer: Transferring data between distant nodes
- Two main functionalities of Network layer
  - *Routing*: Determine the path for transferring data from the source to the destination nodes → Role of routing protocol.
  - *Forwarding*: Transferring data from an incoming port to an outgoing port of a node (router) according to the path defined above ➔ Role of routed protocol: Internet Protocol (IP)

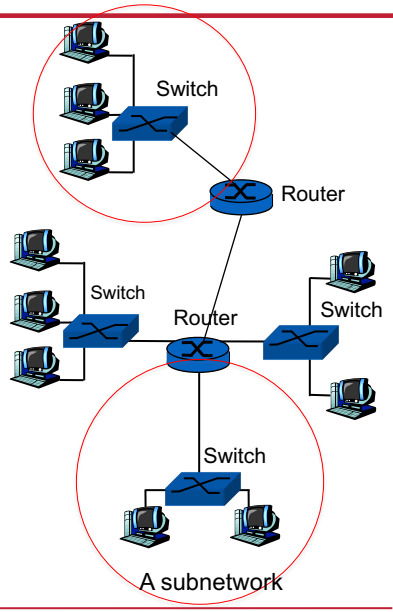


Network layer and Internet Protocol (IP)

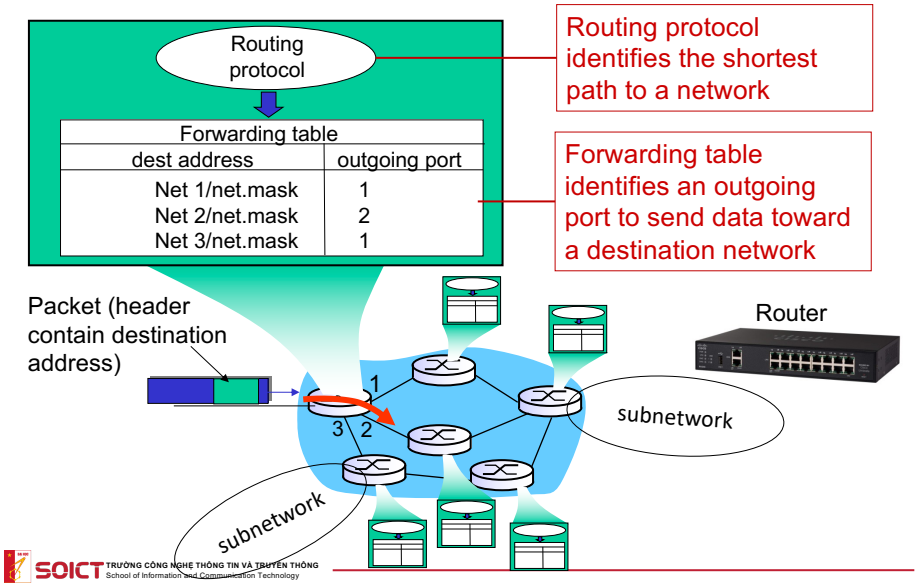
- Layer 2 devices (switches) allow to connect limited number of close hosts
- When hosts are far from each other or there are too many hosts, using switch is inefficient.
  - Forwarding table of switch becomes too big
- ➔ Need intermediates nodes with forwarding and better path finding functions → Routers
  - Finding routes according to destination Network layer address
  - Forwarding data

IP principles

- Network elements
  - **host** = end system;
  - **subnetwork** = a collection of hosts that are connected by layer-2 devices
    - Hosts of a subnetwork have similar addresses: a common prefix
  - **Routers**: intermediate nodes interconnect subnetworks
- Packet forwarding
  - **Within a subnetwork**: hosts communicate directly through layer-2 device (switch)
  - **Between subnetworks**: one or several routers forward packets based on the destination network address.



Routing and forwarding



IP address

IP address classes

CIDR – Classless Inter-Domain routing

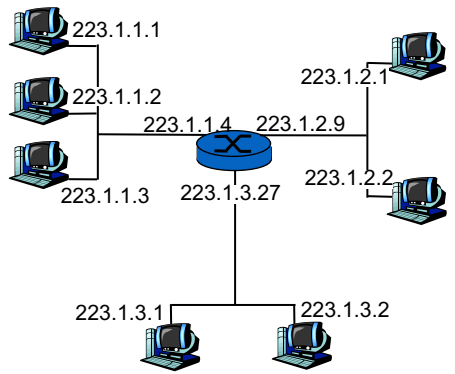
Subnet and netmask

Special IP addresses

ONE LOVE. ONE FUTURE.

IP address (IPv4)

- IP address: A 32-bit number identifying uniquely a network interface
- Interface:
  - router's typically have multiple interfaces
  - host may have multiple interfaces
  - IP addresses associated with interface, not host, router



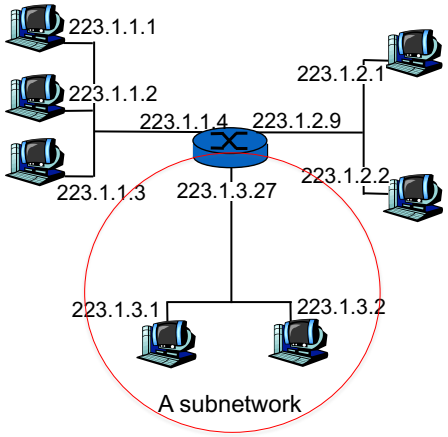
223.1.1.1 = 11011111 00000001 00000001 00000001

223 1 1 1

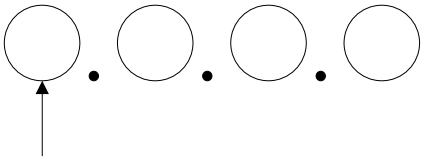
IP address (IPv4)

- For routing purpose, IP address of interfaces in a subnetwork have the same prefix.
- A subnetwork from IP address perspective is a part of network where:
  - Devices can physically reach each other without intervening router (using layer-2 only technology)

●a



Dot notation



8 bits
0 – 255 integer

Example:
203.178.136.63 o
259.12.49.192 x
133.27.4.27 o

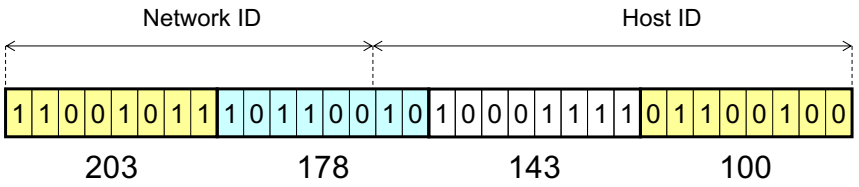
Use 4 x 8 bits describing a 32 bits address

3417476964

1	1	0	0	1	0	1	1	1	0	1	1	0	0	1	0	1	0	0	0	1	1	1	1	0	1	1	0	0	1	0	0
203								178								143								100							

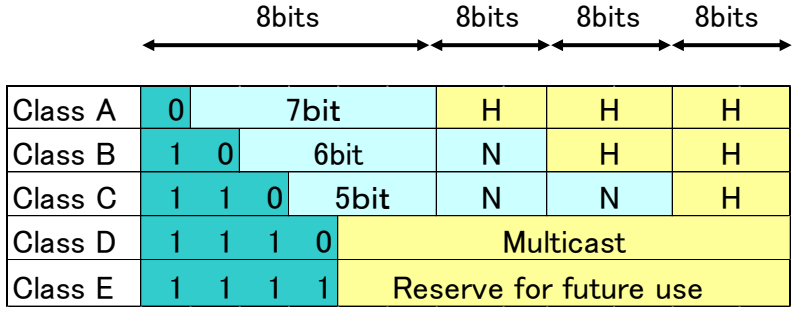
Host address, network address

- IP address contains two parts
  - Host ID – identify a host in a network
  - Network ID – identify a network



- How to know which bits belong to network ID or host ID parts?
  - Use classful IP address
  - Use classless IP address– CIDR

Classify IP addresses



	# of network	# of hosts
Class A	128	2^24
Class B	16384	65536
Class C	2^21	256

Exercise

- Determine which classes these IP addresses belong to:
  - 10.10.10.9
  - 192.168.70.5
  - 129.60.4.7
- What are network ID part of each address

Limitation of classful IP address

- Inefficient use of addressing space
  - Hard classification of addressing space into classes (A, B, C, D, E) makes it is difficult to use all the address space

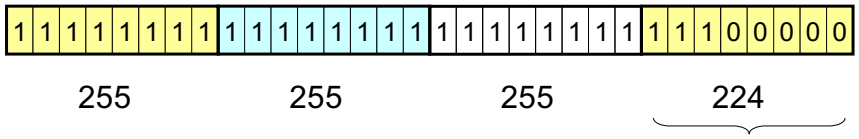
Solution…

- CIDR: Classless Inter Domain Routing
  - Network ID part will have variable length.
  - Length of Network ID part is specified in Network mask
  - Address notation: a.b.c.d/x, where x (mask) the number of bits in Network ID part.

Network mask

- Network mask = number of bits in Network ID part
- IP addresses are assigned to hosts in the same network have the same Network ID part.
- Based on a network mask, it is possible to
  - Identify the network where an IP address belongs to
  - Calculate how many IP addresses available in the network associated with the mask.

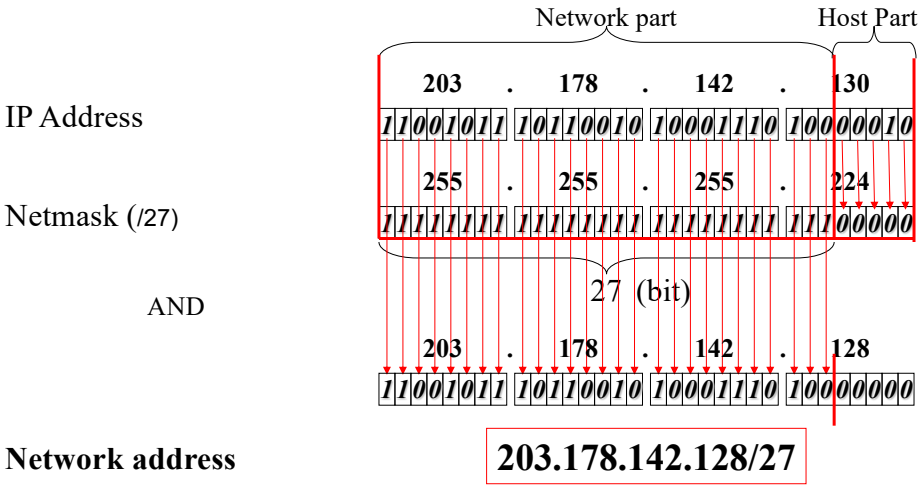
Presentation of network mask



- 255.255.255.224
- /27
- 0xFFFFFFe0

- Last byte may be:
  - 0 248
  - 128 252
  - 192 254
  - 224 255
  - 240

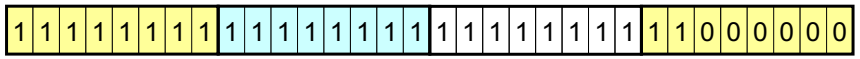
Calculation of network address



Different significations of IP address

- Network address
  - IP address assigned to a network
  - hostID contains all 0
- Host address
  - IP address assigned to a network card
- Broadcast address
  - Address used for sending data to all hosts in a network
  - All bit 1 in HostID part.

Calculation of network size



- Network size
  - Power of 2
- RFC1878

- In case of mask /26
  - Bits for Host ID = 6 bits
  - 2^6=64 possible address:
    - 0 - 63
    - 64 - 127
    - 128 - 191
    - 192 – 255
  - Including network address and broadcast address

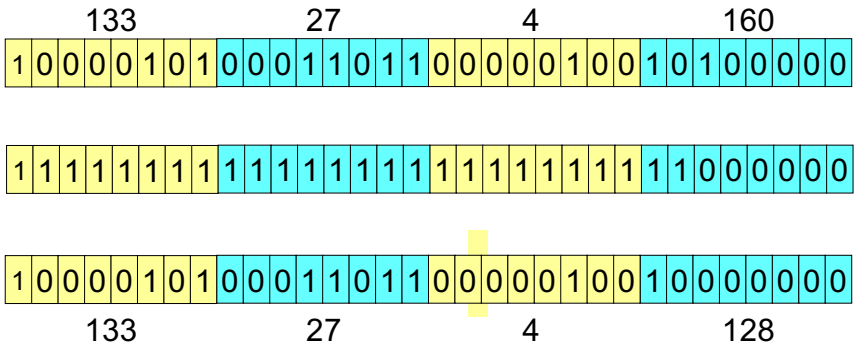
Calculation of network size

- Network mask: /n
- Network size: The maximum acceptable number of hosts in that network
- IP address :

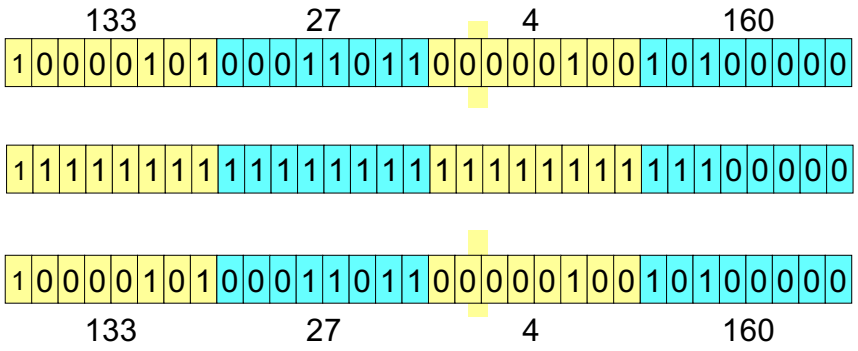
- Number of hosts (Unicast Address):

$$2^{32-n} - 2$$

Network address or host address (1)



Network address or host address (2)



Exercice: IP address and network mask

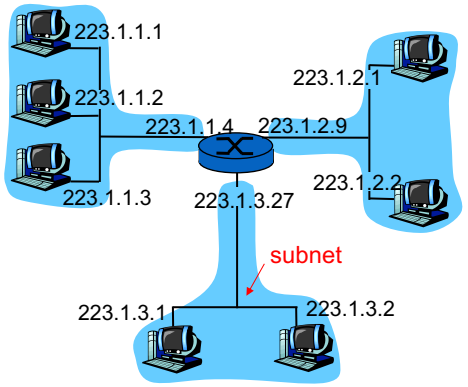
- Which of the following IP addresses are host address, network address, broadcast address? Which network they belong to?

(1) 203.178.142.128 /25
(2) 203.178.142.128 /24
(3) 203.178.142.127 /25
(4) 203.178.142.127 /24

- Attn: With CIDR addressing, IP address should always coming with a network mask

Subnet

- Subnet is a part of a network
  - Hosts of a subnet communicate directly without reaching to layer 3.
  - Usually is one department of an organization
- Design question: How to assign addresses of a network to subnets
  - Use a longer netmask



A network with 3 subnets.

Example: Divide into 2 subnets

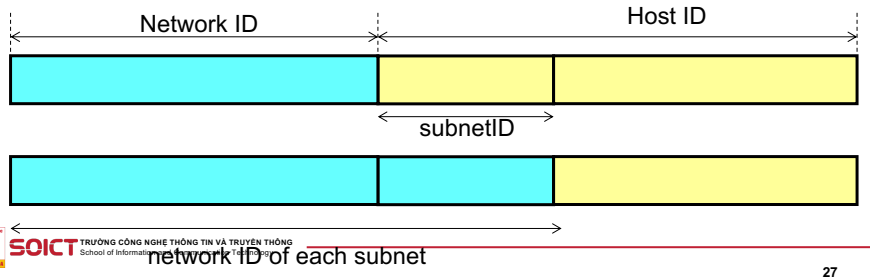
11001000 00010111 00010000 00000000
200. 23. 16. 0 /24

SubnetID

11001000 00010111 00010000 00000000
200. 23. 16. 0 /25
11001000 00010111 00010000 10000000
200. 23. 16. 128 /25

Principle

- Divide a IP range into sub-ranges of equal size
- Take some bits from HostID part to distinguish subnets
  - each subnet contains IP addresses with a fixed values of subnet ID.



Exercise: Dividing into subnets

- Given IP addresses in the range 200.23.16.0/24

1) Need to organize into 3 subnets
   - Address of each subnetwork? Mask? Number of hosts/subnetwork

2) General question: Need to create as many subnets as possible so that
   - Each network can contain 14 hosts
   - Each network can contain 30 hosts
   - Each network can contain 31 hosts
   - Each network can contain 70 hosts

   Network address? Mask?

/28
/27
/26
/25

Addressing space of IPv4

- In theory
  - All between 0.0.0.0 ~ 255.255.255.255
  - Some special IP address (RFC1918)

Private address	10.0.0.0/8 172.16.0.0/12 192.168.0.0/16
Loopback address	127.0.0.0
Multicast address	224.0.0.0 ~239.255.255.255

- Self assigned IP address: 169.254.0.0/16

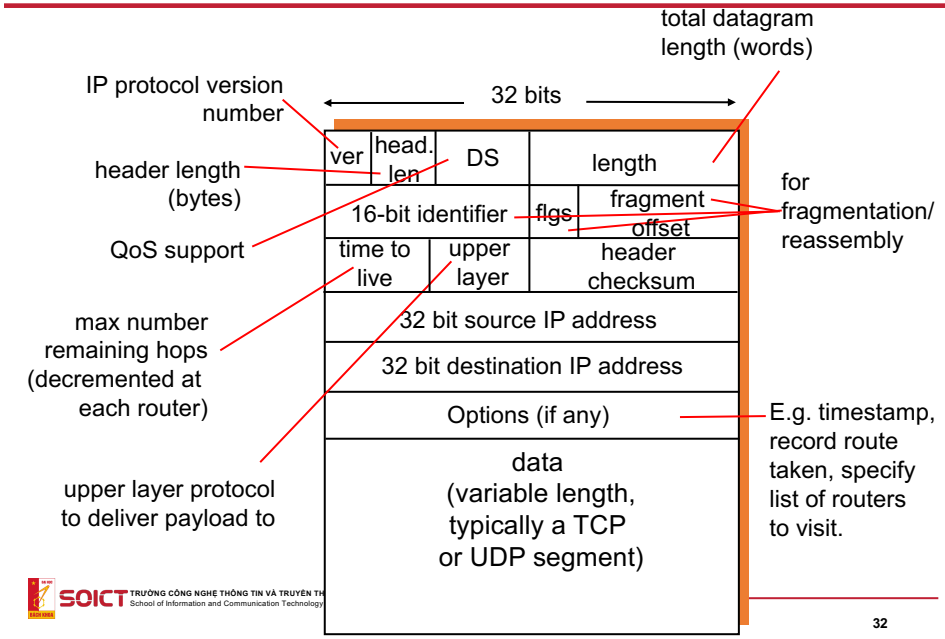
Attention about IP

- Currently IPv4: 32 bits
  - 133.113.215.10 (IPv4)
- IPv6 is also widely used: 128bits
  - 2001:200:0:8803::53 (IPv6)
  - Fix 64 first bit for subnet ID, 64 last bit belongs to interface ID.
  - Security feature is integrated

IP packet

ONE LOVE. ONE FUTURE.

Header of IP



IP header (1)

- Version (4 bits)
  - IPv4
  - IPv6
- Header length: 4bits
  - In word unit (4 bytes)
  - Min: 5
  - Max: 60

IP header (2)

- DS (Differentiated Service : 8bits)
  - Old name: Type of Service
  - Used for QoS management by some router
  - Diffserv

IP header (3)

- Length: total length including header (16 bits)
  - In bytes unit
  - Max: 65536
- 16 bits Identifier– ID of the packet
  - Used for identifying all fragments of the same packet when it is fragmented
  - Flag
  - Fragmentation offset – offset of the first byte of the fragment in its original packet

IP header (4)

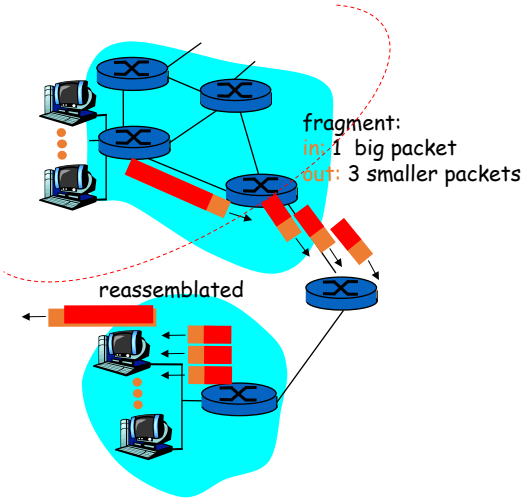
- TTL, 8 bits – Time to live
  - Maximum number of hops (router) the packet is allowed to travel
  - Max: 255
  - Router decreases TTL 1 unit when processing a packet
  - The packet will be destroyed when TTL reaches to 0
- Protocol – upper layer protocol
  - Transport protocol (TCP, UDP,...)
  - Other network layer protocols that are encapsulated in IP packet (ICMP, IGMP, OSPF)

IP header (4)

- Checksum: to detect corruption in the header of IPv4 data packets
- Source IP address
  - 32 bit, address of the sender
- Destination IP address
  - 32 bit, address of the receiver.

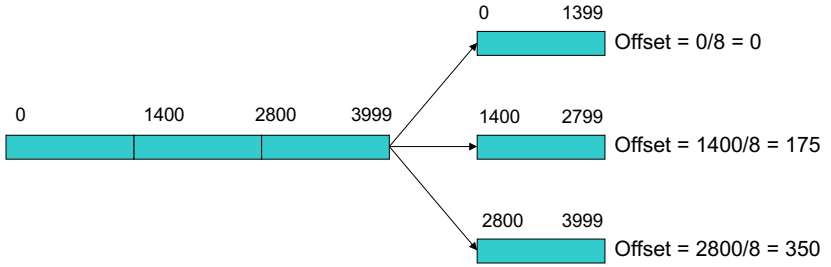
Packet fragmentation (1)

- Each link has a fixed MTU (Maximum transfering unit)
- Different media have different MTU
- If IP packet > MTU, it should be
  - Divided into small fragments
  - Gathered at the destination



Packet fragmentation (2)

- Offset
  - Position of the fragment in the original packet
  - In 8 bytes units



Network Address Translation (NAT)

ONE LOVE. ONE FUTURE.

Principal NAT

- Data communication from a LAN (using private IP address) to the Internet (using public IP address) and vice-visa

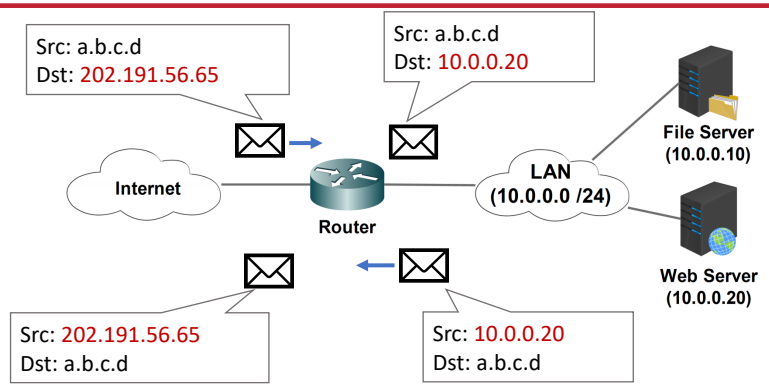
→ NAT (Network Address Translation) convert private address to public address. The task is performed on routers

- Advantage:
  - Solve the problem of limiting public IP address
  - Hide the private address inside a LAN from outside intruders
  - Avoid IP address re-assignment within a LAN when change ISP

Static NAT

- The simplest NAT
- A IP private is assigned a public IP
- NAT router store a converting table in its memory
  - Converting table map a private IP address to a public IP address for internet communication
- This mechanism is often used for servers located inside a LAN and providing public service. Ex: dk-sis, soict web server.

Static NAT - Example



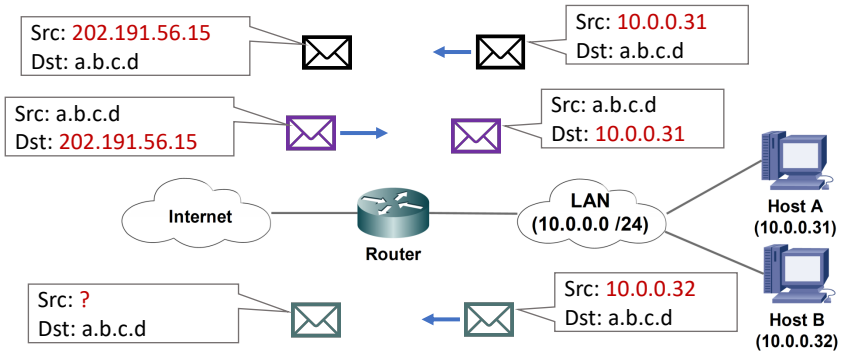
NAT Table

Inside IP	Outside IP
10.0.0.10	202.191.56.11
10.0.0.20	202.191.56.65

Dynamic NAT

- NAT router map automatically a range of private IP to a range of public IP so that computer inside a LAN can communicate to Internet when it needs
- No fix mapping
- Any private IP address will be translated automatically to one (available) public IP address from the pool of public IP addressed maaged by the NAT router

Dynamic NAT – Example



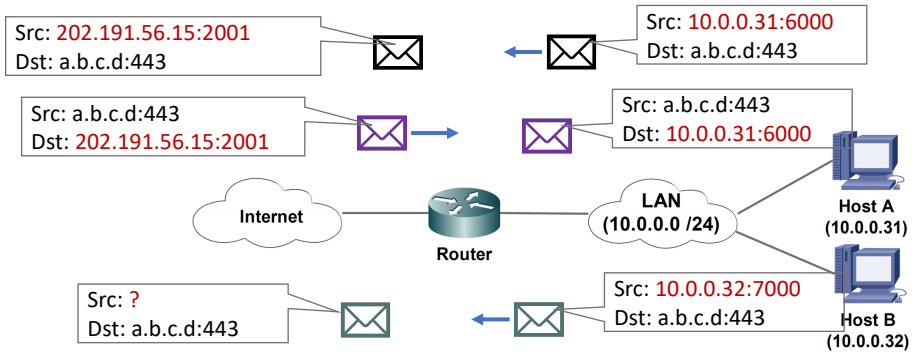
NAT Table

Inside IP	Outside IP
10.0.0.31	202.191.56.15
10.0.0.32	202.191.56.16

Port Address Translation

- Port Address Translation= PNAT=NAT **overloading**
- Special form of dynamic NAT
- Map many private IP to a single public IP public by adding a port number
  - n private IP → 1 public IP
- PAT use sockets information to map
  - (private IP : port) <-> (public IP: port)

PAT– Example



NAT Table

Inside IP	Outside IP
10.0.0.31:6000	202.191.56.15:2001
10.0.0.32:7000	202.191.56.15: 2002

Lecture 7: Routing

Reading 5.2

Computer Networks, Tanenbaum

ONE LOVE. ONE FUTURE.

What is routing?

Routing principals

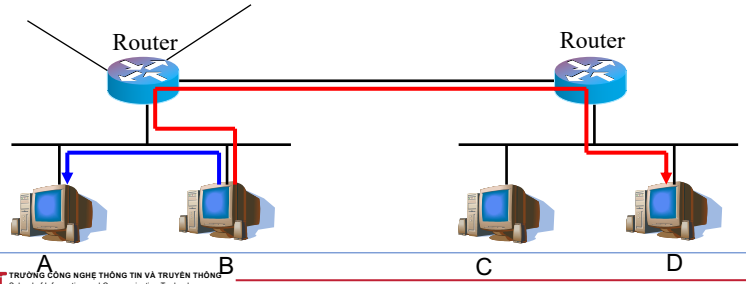
Forwarding mechanism

"Longest matching" rule

ONE LOVE. ONE FUTURE.

Routing and Forwarding principles (1)

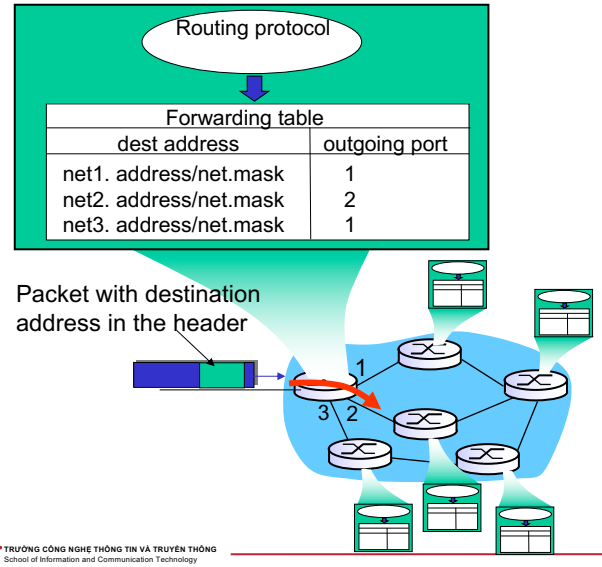
- When a host send an IP packet to another host:
  - If the destination and the source are in the same network (by IP address):
    - the packet is transferred directly to the destination by Layer 2
  - If the destination is in a different network with the source:
    - The packet is sent to a router (to choose a route)
    - The next router then forward data again until the destination



Forward IP packets

- Each router has a Forwarding Table
  - A part of Routing Table
- Forwarding table contain:
  - Destination: Network address/network mask
    - Outgoing port: label of the port on the router that connect to the next router in the part to the destination
  - Defaut network address:
    - 0.0.0.0/0 stands for any networks.

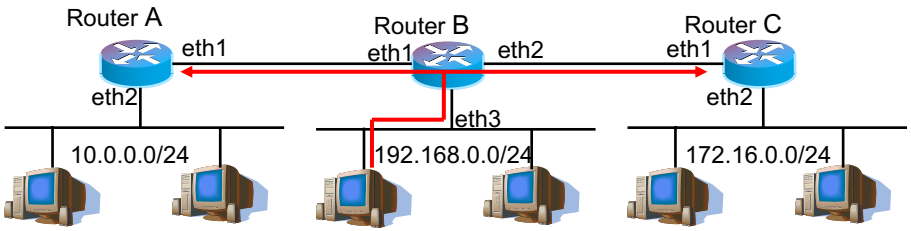
Routing table



Some examples of routers...

Routing table and forwarding mechanism (1)

Network	Next-hop	Interface
10.0.0.0/24	A	eth1
172.16.0.0/24	C	eth2
192.168.0.0/24	direct	eth3



Rule: ***No routes, no reachability*****!**

Example – Routing table on a host

```
C:\Documents and Settings\tungbt>netstat –r
Route Table
===========================================================================
Interface List
0x1 ...                    ....MS TCP Loop
0x2 ...08 00 ff b2 a1 a3 ..... Realtek RTL8139 Family PCI Fast Ethernet NIC -
===========================================================================
```

Destination

Outgoing port

Network	Netmask	Gateway	Interface	Metric
0.0.0.0	0.0.0.0	192.168.1.1	192.168.1.34	20
127.0.0.0	255.0.0.0	127.0.0.1	127.0.0.1	1
192.168.1.0	255.255.255.0	192.168.1.34	192.168.1.34	20
192.168.1.34	255.255.255.255	127.0.0.1	127.0.0.1	20
192.168.1.255	255.255.255.255	192.168.1.34	192.168.1.34	20
224.0.0.0	240.0.0.0	192.168.1.34	192.168.1.34	20
255.255.255.255	255.255.255.255	192.168.1.34	192.168.1.34	1

```
Default Gateway:       192.168.1.1
===========================================================================
```

Example- Routing table on a router

Destination

Outgoing port

Router# **show ip route**

O 203.238.37.0/24 via 203.178.136.14, FastEthernet0/1
O 203.238.37.96/27 via 203.178.136.26, Serial0/0/0
C 203.238.37.128/27 is directly connected, Serial0/0/0
O 192.68.132.0/24 via 203.178.136.14, FastEthernet0/1
C 203.254.52.0/24 is directly connected, FastEthernet0/1
C 202.171.96.0/24 is directly connected, Serial0/0/1

Longest matching rule

- When a router receive a packet: ...
- The router match n first bits of the destination address with the networks in the routing table
  - /n: Mask of the destination networks in the routing table
- If there is more than 1 matching network, apply "longest matching" rule:
  - Choose the route with the largest mask

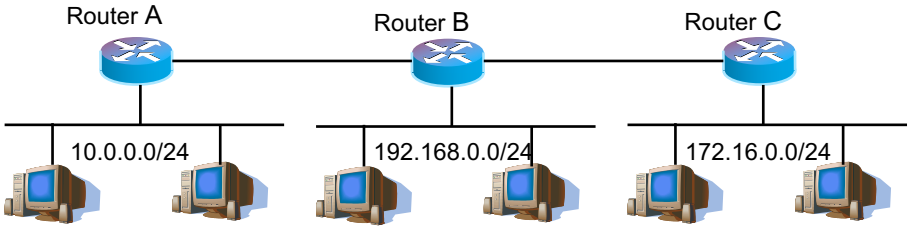
Destination address of the packet to be forwarded:
11.1.2.10

Destination	Outgoing Port
11.0.0.0 /8	Se0/1
11.1.0.0 /16	Se0/2
11.1.2.0/24	Se0/3
0.0.0.0/0	Se0/4

"Longest matching" rule (2)

Destination address:
11.1.2.5 = 00001011.00000001.00000010.00000101
Route 3:
11.1.2.0/24 = 00001011.00000001.00000010.00000000
Route 2:
11.1.0.0/16 = 00001011.00000001.00000000.00000000
Route 1:
11.0.0.0/8 = 00001011.00000000.00000000.00000000

Routing table and forwarding mechanism (2)

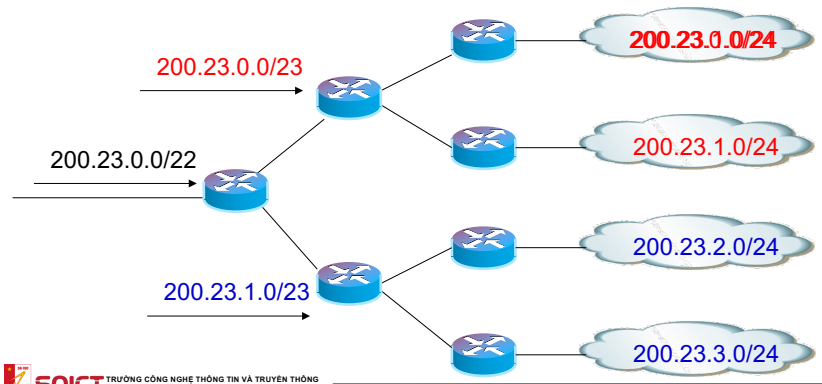


Network	Next-hop
10.0.0.0/24	A
172.16.0.0/24	C
192.168.0.0/24	Direct

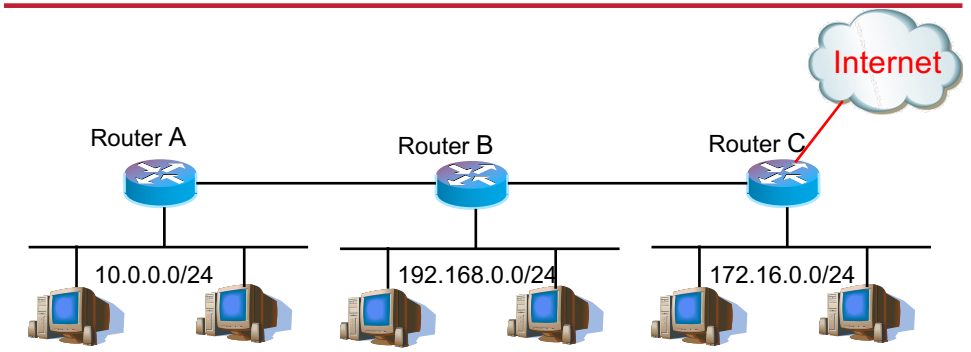
Q. What should be the routing table of router C so that all host can send data to each other?

Route aggregation

- How many networks in the Internet?
- There will be a lot of entries in the routing table?
- The entries to sub-networks of the same "big" network can be aggregated inorder to reduce the size of routing table.



Routing table and forwarding mechanism (2)



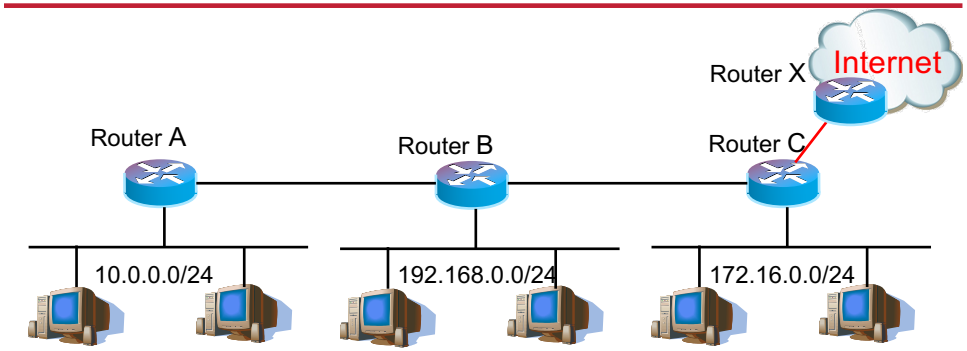
Router B

Network	Next-hop
10.0.0.0/24	A
172.16.0.0/24	C
192.168.0.0/24	Direct

Q. What is the routing table in B?

Q: What if C is connected to the Internet?

Routing table and forwarding mechanism (2)



Router B

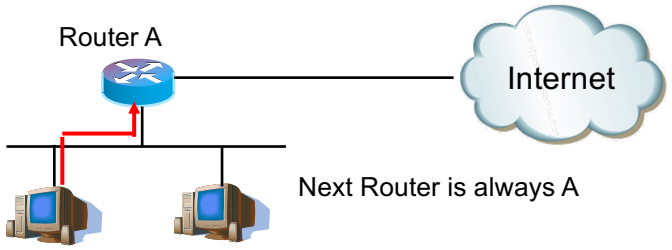
Network	Next-hop
10.0.0.0/24	A
172.16.0.0/24	C
192.168.0.0/24	Direct
0.0.0.0/0	C

Q: What if C is connected to the Internet?

What to be change in routing table of A,B, and C to communicate with the Internet

Default route

- If router does not find a route to a destination in its routing table, default route is necessary
  - Default route is defined for all destination networks that are not figured in the routing table.
- 0.0.0.0/0
  - Is a special notation for all destination networks



Next Router is always A

Packet processing on routers

- Step 1 : If TTL = 1(or TTL = 0), destroy the packet and send error message. End.
- Step 2 : If TTL >1, extract the destination IP address of the packet. Apply the mask of networks in its routing table to destination IP address to find corresponding network addresses.
- Step 3 : Compare the obtained network addresses with networks in routing table.
  - If find a matching route, forward the packet to the interface of the route, reduce TTL by 1.
  - If no route match, check if there is a default route (with network 0.0.0.0 /0).
    - If there is a default route, forward the packet to the corresponding interface and reduce TTL by 1.
    - If there is no default route: destroy the packet, send an error message back to the source.

Exercises

- A router has the following (CIDR) entries in its routing table:

Address/mask Next hop
135.46.56.0/22 Interface 0
135.46.60.0/22 Interface 1
192.53.40.0/23 Interface 2
0.0.0.0/0 Interface 3

- For each of the following IP addresses, what does the router do if a packet with that address arrives?

(a) 135.46.63.10
(b) 135.46.57.14
(c) 135.46.52.2
(d) 192.53.40.7
(e) 192.53.56.7

Solution:

Apply longest matching rule.

Solution

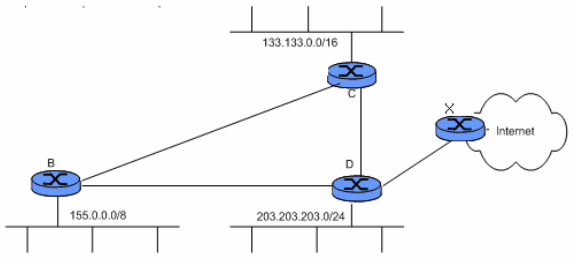
Apply longest matching rule.

(students should explain why by matching binary form of the addresses)

(a) 135.46.63.10 → Interface 1
(b) 135.46.57.14 → Interface 0
(c) 135.46.52.2 → Interface 3 (default route)
(d) 192.53.40.7 → Interface 2
(e) 192.53.56.7 → Interface 3 (default route)

Exercise

- Assume that we have a network with following topology. What should be routing table of routers B, C, D in order to assure that all hosts can send data to each other and to the Internet.



Solution

- Routing table on B

Network	Next hop
133.133.0.0/16	C
155.0.0.0/8	Direct
203.203.203.0/24	D
0.0.0.0/0	D

- Routing table on C

Network	Next hop
133.133.0.0/16	Direct
155.0.0.0/8	B
203.203.203.0/24	D
0.0.0.0/0	D

- Routing table on D

Network	Next hop
133.133.0.0/16	C
155.0.0.0/8	B
203.203.203.0/24	Direct
0.0.0.0/0	X

Internet Control Message Protocol

Packet format
Ping and Traceroute

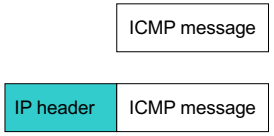
ONE LOVE. ONE FUTURE.

Idea of ICMP (1)

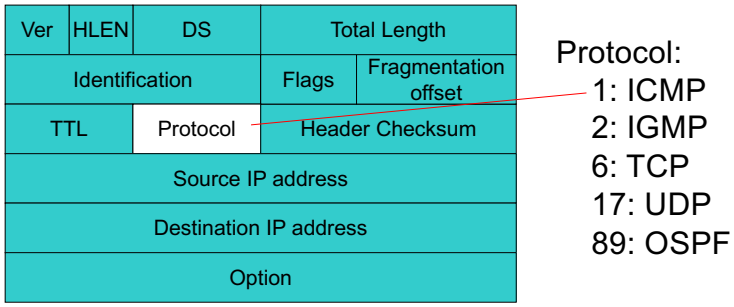
- IP is unreliable, connectionless
  - Lack of supporting and error control mechanism
- ICMP is used in network layer for providing a mechanism to exchange information between sender and receivers
  - Error information: inform that a packet cannot reach a host, a network or a port.

Idea of ICMP (2)

- Also in network layer but is “above” IP
  - ICMP message is encapsulated in IP
- ICMP message: Type, Code, with 8 first bytes of the error IP message

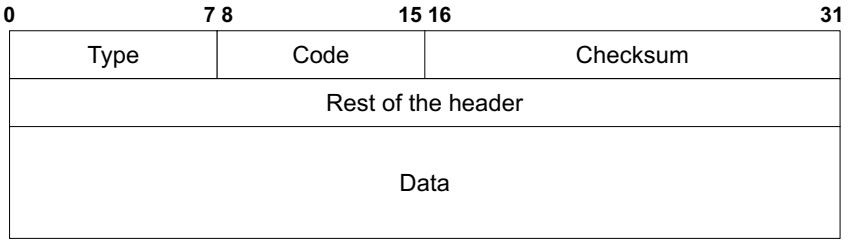


IP header and Protocol field



ICMP message format

- Type: type of ICMP message
- Code: cause of error
- Checksum
- Rest of header varies according on type



Some ICMP message types

ICMP Message Type	Error-reporting messages	3	Destination Unreachable
		4	Source quench (nguồn giảm tốc độ)
		5	Redirection
		11	Time exceeded
		12	Parameter problem
	Query messages	8 or 0	Echo reply or request
		13 or 14	Time stamp request or reply
		17 or 18	Address mask request or reply
		9 or 10	Router advertisement or solicitation

ICMP code: sub-type

- ICMP code = sub-type

3 – Destination Unreachable[6]:4	0		Destination network unreachable
	1		Destination host unreachable
	2		Destination protocol unreachable
	3		Destination port unreachable
	4		Fragmentation required, and DF flag set
	5		Source route failed
	6		Destination network unknown
	7		Destination host unknown
	8		Source host isolated
	9		Network administratively prohibited
	10		Host administratively prohibited
	11		Network unreachable for ToS
	12		Host unreachable for ToS
	13		Communication administratively prohibited
	14		Host Precedence Violation
	15		Precedence cutoff in effect

ICMP and debuging tools

- ICMP always works transparently for users
- Users can use ICMP by using some debuging tools
  - ping
  - traceroute

Ping and ICMP

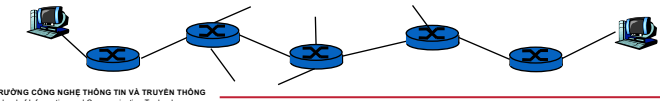
- ping
  - Test a connection
  - Sender sends packet “ICMP echo request”
  - Receiver responses with “ICMP echo reply”
- Data field contains the time stamp when the packet is sent
  - For calculating RTT (round-trip time)

Ping: Example

```
C:\Documents and Settings\hongson>ping www.yahoo.co.uk

Pinging www.euro.yahoo-eu1.akadns.net [217.12.3.11] with 32 bytes of data:

Reply from 217.12.3.11: bytes=32 time=600ms TTL=237
Reply from 217.12.3.11: bytes=32 time=564ms TTL=237
Reply from 217.12.3.11: bytes=32 time=529ms TTL=237
Reply from 217.12.3.11: bytes=32 time=534ms TTL=237

Ping statistics for 217.12.3.11:
    Packets: Sent = 4, Received = 4, Lost = 0 (0% loss),
Approximate round trip times in milli-seconds:
    Minimum = 529ms, Maximum = 600ms, Average = 556ms
```


Traceroute and ICMP

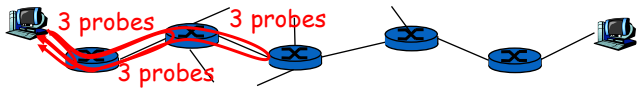
- Sender send many packets to receiver
  - First packet has TTL =1
  - Second packet has TTL=2, …
- When packet number n arrives to n^{th} router:
  - Router destroys the packer
  - Router send back an ICMP packet (type 11, code 0) containing IP address of the router
- Based on the reply message, the sender can calculate RTT



Traceroute and ICMP

Termination condition

- When ICMP echo packet arrive to the destination
- When source receives ICMP “host unreachable” (type 3, code 3)



Traceroute: Example

```
C:\Documents and Settings\hongson>tracert www.jaist.ac.jp

Tracing route to www.jaist.ac.jp [150.65.5.208]
over a maximum of 30 hops:

  1    1 ms   <1 ms   <1 ms  192.168.1.1
  2   15 ms    14 ms    13 ms  210.245.0.42
  3   13 ms    13 ms    13 ms  210.245.0.97
  4   14 ms    13 ms    14 ms  210.245.1.1
  5  207 ms   230 ms    94 ms  pos8-2.br01.hkg04.pccwbtn.net [63.218.115.45]
  6    *      403 ms   393 ms  0.so-0-1-0.XT1.SCL2.ALTER.NET [152.63.57.50]
  7  338 ms   393 ms   370 ms  0.so-7-0-0.XL1.SJC1.ALTER.NET [152.63.55.106]
  8  402 ms   404 ms   329 ms  POS1-0.XR1.SJC1.ALTER.NET [152.63.55.113]
  9  272 ms   288 ms   310 ms  193.ATM7-0.GW3.SJC1.ALTER.NET [152.63.49.29]
 10  205 ms   206 ms   204 ms  wide-mae-gw.customer.alter.net [157.130.206.42]
 11  427 ms   403 ms   370 ms  ve-13.foundry2.otemachi.wide.ad.jp [192.50.36.62]
 12  395 ms   399 ms   417 ms  ve-4.foundry3.nezu.wide.ad.jp [203.178.138.244]
 13  355 ms   356 ms   378 ms  ve-3705.cisco2.komatsu.wide.ad.jp [203.178.136.193]
 14  388 ms   398 ms   414 ms  c76.jaist.ac.jp [203.178.138.174]
 15  438 ms   377 ms   435 ms  www.jaist.ac.jp [150.65.5.208]

Trace complete.
```


Static and dynamic routing

Static routing
Dynamic routing
Advantage – Weakness

ONE LOVE. ONE FUTURE.

Problem of update routing table

- When topology change: new networks, a router is out of power
- It is necessary that routing tables are updated
  - In theory, all routers need to be updated
  - In reality, only few routers need to be updated

Network	Next-hop
192.168.0.0/24	B
172.16.0.0/24	B
10.0.0.0/24	Direct

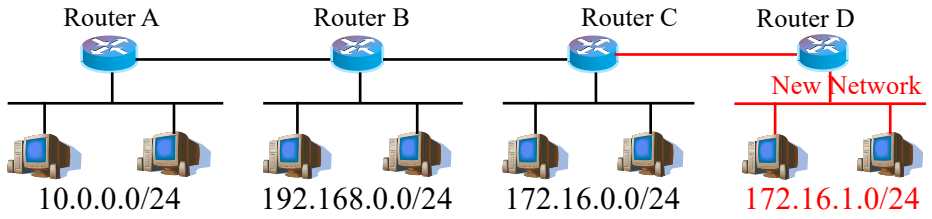
172.16.1.0/24 B

Network	Next-hop
10.0.0.0/24	A
172.16.0.0/24	C
192.168.0.0/24	Direct

172.16.1.0/24 C

Network	Next-hop
10.0.0.0/24	B
192.168.0.0/24	B
172.16.0.0/24	Direct

??

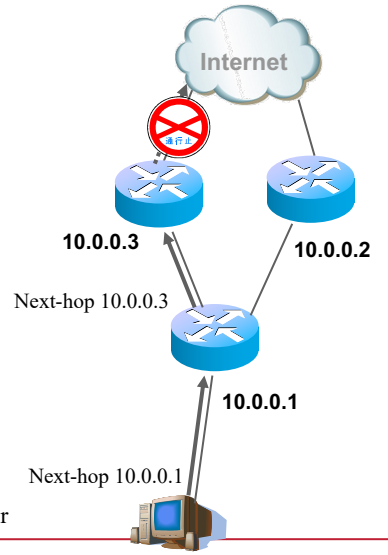


How to update routing table?

- Static routing
  - Entries in the routing tables are updated manually by network administrator.
- Dynamic routing
  - The routing table is updated automatically by some routing protocols running on routers

Static routing

- When there is some failures on a route:
  - Impossible to access to Internet even though there is an alternative route
  - Admin needs to update routing table at 10.0.0.1



Extract of routing table at 10.0.0.1

Prefix	Next-hop
0.0.0.0/0	**10.0.0.3**

Entry causing error

Dynamic routing

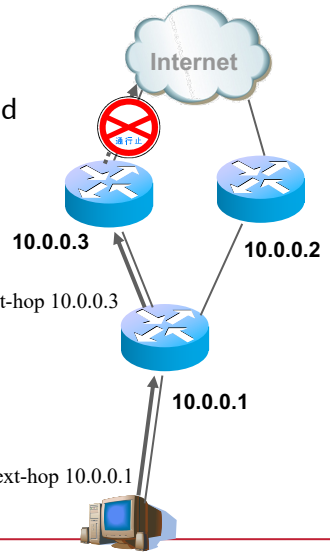
- When there is failure :
  - The entries related on the affected routes are updated automatically

Extract of routing table of 10.0.0.1

Prefix	Next-hop
0.0.0.0/0	**10.0.0.2**
~~0.0.0.0/0~~	~~10.0.0.3~~

Alternative route

Affected route



Pros/cons

- Static routing
  - Pros:
    - Stable,
    - Secure,
    - Not influence by external factor
  - Cons:
    - Not flexible,
    - It is impossible for using automatically backup routes
    - Difficult to manage
- Dynamic routing
  - Pros
    - Easy to manage
    - Backup routes are used automatically when there are failures
  - Cons
    - Not secure
    - Complicated routing protocols

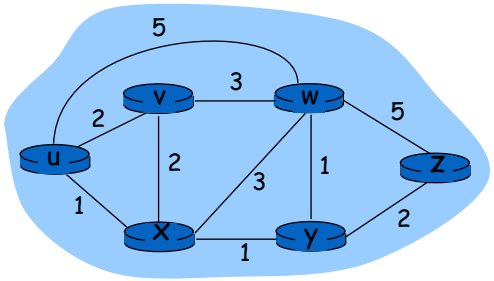
Routing algorithm and protocols

Dijkstra and Bellman-Ford Algo
link-state and distance-vector protocols

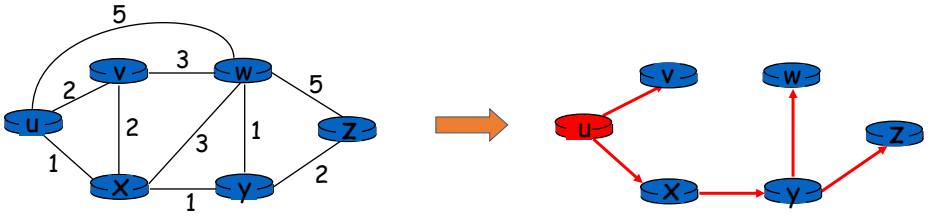
ONE LOVE. ONE FUTURE.

Graph representing the networks

- Graph with nodes (routers) and edges (links)
- Weight on each link c(x,y)
  - Weigh can be bandwidth, delay, congestion level, cost… expressing the contribution of the link in the total cost of a route
- Routing algorithm: Determine the shortest path (in term of weight) between a pair of two nodes.



Shortest path tree-SPT



- SPT – Shortest Path Tree
- Compose of shortest paths from a single source node to all other nodes.
- Each source node has it own SPT

Two classes of routing algorithms

- Link-state
  - Gathering the topology information at a node → build graph
  - Run a path calculation algorithm on the node
  - Build routing table on the node
  - OSPF routing protocol
- Distance vector
  - Each node build temporary a routing table
  - Exchange routing tables for finding better routes through the neighbors
  - RIP routing protocol

Link-state algorithms- Dijikstra

- Notations:
  - *G = (V,E)* : *Graph representing the network: V*: set of nodes, *E: set of links*
  - *c(x,y):* cost of using link *x* to *y*;
    - = ∞ f the two nodes are not linked together
  - *d(v):* current cost for going from the source node to node *v*
  - *p(v):* node right before v on the route from the source to destination
  - *T:* Set of nodes whose shortest paths have been identified.

Link-state algorithms- Dijikstra

- **Procedures:**
- Init():

  For each node v, *d[v]* = ∞, *p[v]* = NIL

  d[s] = 0
- Improve(*u,v*), where (*u,v*) is an edge of *G*

  if *d[v] > d[u] + c(u,v)* then

  d[v] = d[u] + c(u,v)

  p[v] = u

Link-state algorithms- Dijikstra

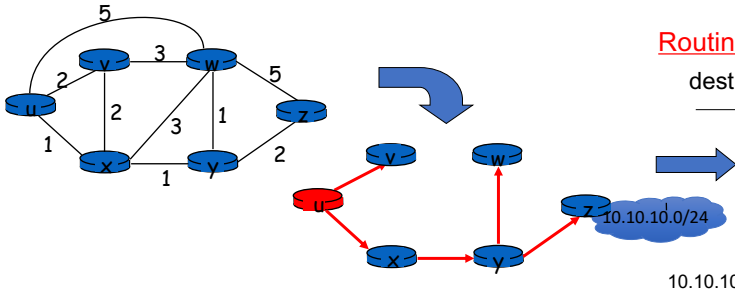
1. **Init()** ;
2. *T = Φ;*
3. **Repeat**
4. *u:* u ∉ *T* | *d(u)* is the smallest;
5. *T = T ∪ {u};*
6. **for all** *v* ∈ neighbor(u) and *v* ∉ *T*
7. *improve(u,v)* ;

8.Until *T = V*

Browse all vertexes *u* starting from those are nearest to the source, and see if it is better (shorter) to go from the source to a neighbor of *u* by going through *u*

Dijkstra's algorithm: Example

Step	T	d(v),p(v)	d(w),p(w)	d(x),p(x)	d(y),p(y)	d(z),p(z)
0	u	2,u	5,u	1,u	∞	∞
1	ux	2,u	4,x		2,x	∞
2	uxy	2,u	3,y			4,y
3	uxyv		3,y			4,y
4	uxyvw					4,y
5	uxyvwz					

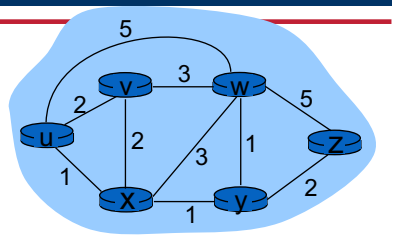


SPT of u:

Routing table of u:

destination	Next hop
v	v
x	x
y	x
w	x
10.10.10.0/24(z)	x

Distance-vector algorithm Bellman-Ford (1)



Definitions:

$d_u(z)$:= cost of the shortest path from u to z

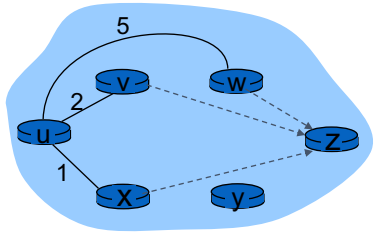
We have: Bellman-Ford equation:

$$d_u(z) = \min_a \{c(u,a) + d_a(z)\}$$

For all *a* adjacent to x

Distance-vector algorithm Bellman-Ford (2)

- Vision of u:
  - Know only neigbour and believe on the path reported by neighbors



- u choose the shortest path to a destination (ex.: z) amongst all paths via its neighbour v, x,w
  - Via x:
    - $d_u(z) = (u,x) + d_x(z)$
  - Via v:
    - $d_u(z) = (u,v) + d_v(z)$

Distance reported by x

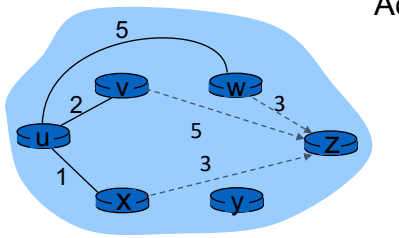
  - Via w:
    - $d_u(z) = (u,w) + d_w(z)$

Amongst all paths from u→ z, u chooses to go via the neighbor that makes the path shortest

Distance vector of u: $d_u(z)$, $d_u(y)$, $d_u(x)$, $d_u(w)$, $d_u(v)$

Distance-vector algorithm Bellman-Ford (3)

Assume that at the current step: $d_v(z) = 5$, $d_x(z) = 3$, $d_w(z) = 3$



According to B-F eq. :

$$d_u(z) = \min \{ c(u,v) + d_v(z),$$
$$c(u,x) + d_x(z),$$
$$c(u,w) + d_w(z) \}$$
$$= \min \{2 + 5,$$
$$1 + 3,$$
$$5 + 3\} = 4$$

Amongst all paths from u→ z, u chooses to go via the neighbor that makes the path shortest

Distance-vector algorithm Bellman-Ford (2)

- After choosing a path, u advertises its $d_u(z)$ to all nodes
- X, v, w... estimate their paths to all other nodes similarly ➔ shorters distances to other nodes may be found
- Advertise new distances
- Repeat best path choosing

Distance-vector algorithm (2)

Main ideas:

- Distance vector: vector of all distance from the current node to all other nodes
- Each node sends periodically its distance vector to its adjacent nodes
- When a node x receives a distance vector, it updates its distance vector by using equation Bellman-ford
- With some condition, the distance $D_x(y)$ *in each vector will converge to the smallest value of* $d_x(y)$

At each node:

Wait for a DV from neighbor

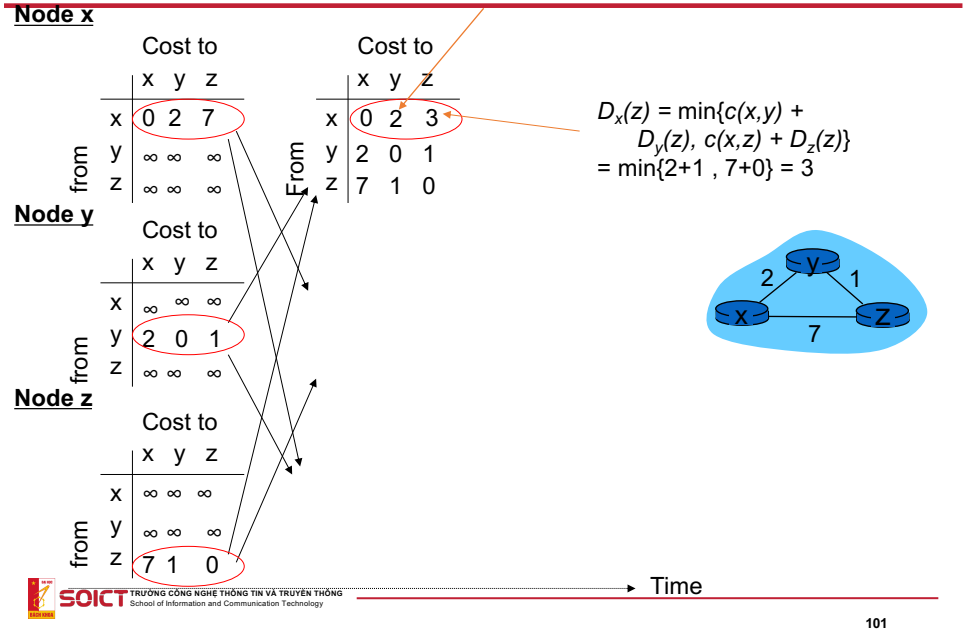
↓

Re-calculate its DV

↓

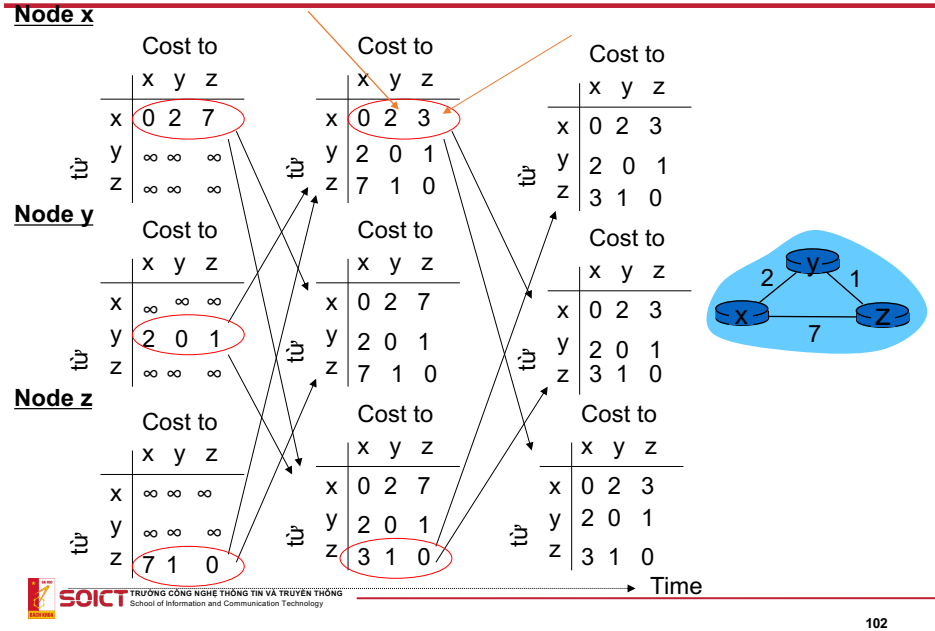
If DV changes, *Inform its neighbor*

$D_x(y) = \min\{c(x,y) + D_y(y), c(x,z) + D_z(y)\}$
$= \min\{2+0 , 7+1\} = 2$



$D_x(y) = \min\{c(x,y) + D_y(y), c(x,z) + D_z(y)\}$
$= \min\{2+0 , 7+1\} = 2$

$D_x(z) = \min\{c(x,y)+D_y(z), c(x,z)+ D_z(z)\}$
$= \min\{2+1 , 7+0\} = 3$



Comparison of Link-state and Distance vector

Number of exchange messages

- LS: n nodes, E links, O(nE) messages
- DV: Exchange only with neighbor

Convergent time

- LS: Complexity $O(n^2)$
- DV: Varies

Reliability: If one routers provide incorrect information

LS:

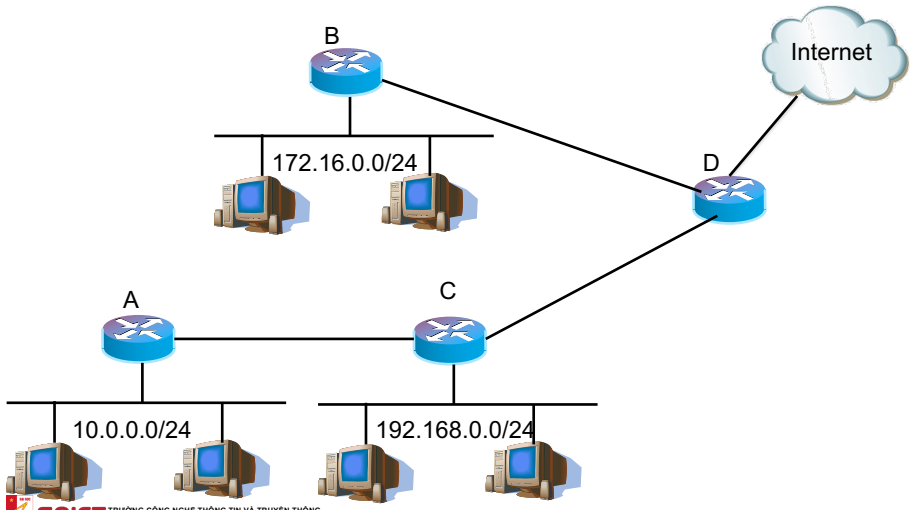
- The router may send out incorrect cost
- Each node calculate its own routing table

DV:

- Incorrect distance vector may be sent out
- Each node calculate its DV based to what receives from the neighbor
  - Error propagates in the network.

Implementation of routing protocols

- Link state protocols
  - OSPF: open shortest path first
    - Implement link information gathering phase for building topology
    - Implement Dijkstra
  - IS-IS
- Distance vector protocols
  - RIP: routing Information protocol



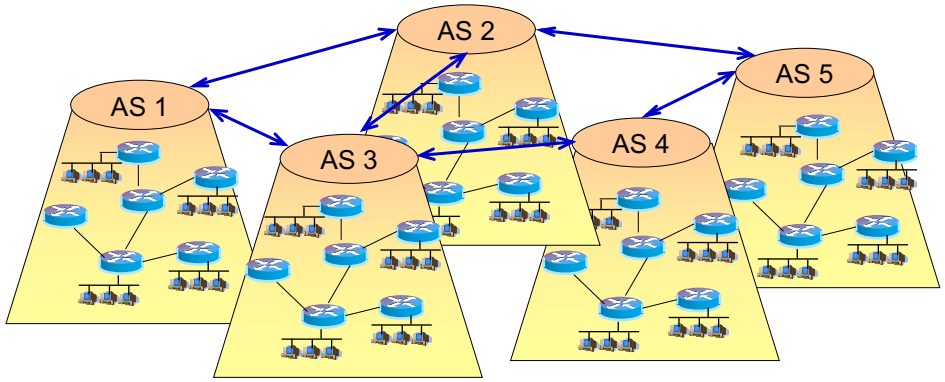
Hiarachy in routing

Autonomous system
Intra-domain routing
Inter-domain routing

ONE LOVE. ONE FUTURE.

Hierarchical architecture of the Internet

- Internet = Network of networks
- Each network may have a particular routing policy
- Such a network is an Autonomous System (AS)



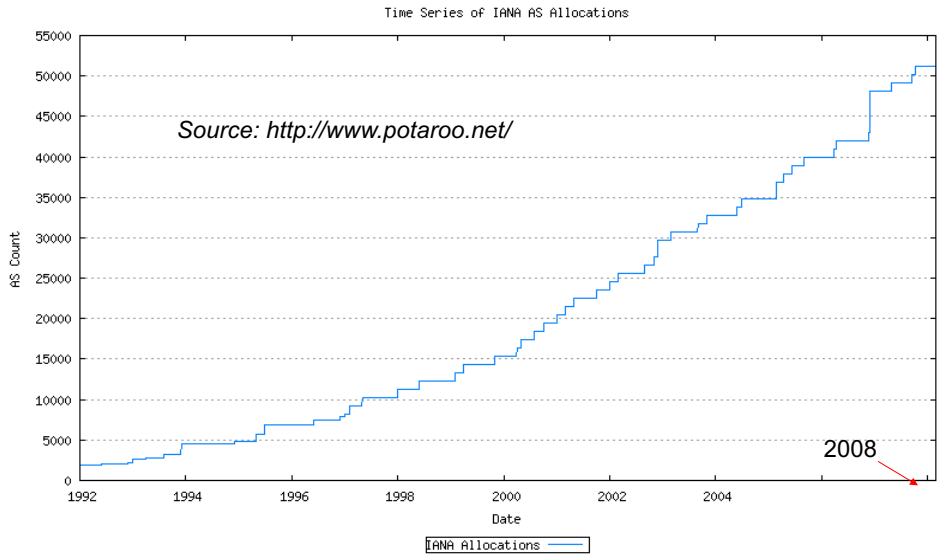
Concept of Autonomous system

- AS = Set of network nodes that follows a common routing strategy (protocol, cost convention…)
- ASes are interconnected via routers called gateways
- Each AS is assigned an AS number for identification
  - AS Number - 16 bits hay 32 bits.

2914	NTT-COMMUNICATIONS-2914 - NTT America, Inc.
3491	BTN-ASN - Beyond The Network America, Inc.
4134	CHINANET-BACKBONE No.31,Jin-rong Street
6453	GLOBEINTERNET Teleglobe America Inc.
24087	VNGT-AS-AP Vietnam New Generation Telecom
24066	VNNIC-AS-VN Vietnam Internet Network Information Center
17981	CAMBOTECH-KH-AS ISP Cambodia
.....	

Source: http://www.cidr-report.org

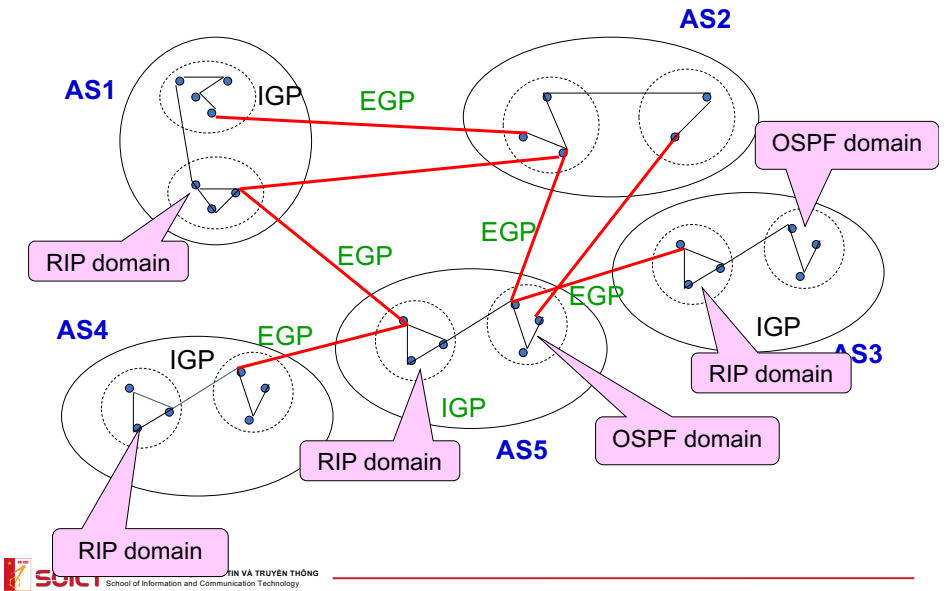
Number of ASN assigned by IANA



Hierarchical of routing protocols

- Inside an AS: Intra-domain routing protocol
  - IGP: *Interior Gateway Protocol*
    - RIP: Routing Information Protocol
    - OSPF: Open Shortest Path First
    - IS-IS, IGRP, EIGRP (Cisco)…
- Between AS: Inter-domain routing protocol
  - EGP: *Exterior Gateway Protocol*
    - BGP (v4): Border Gateway Protocol

Intra-domain and Inter-domain routing



Intradomain routing

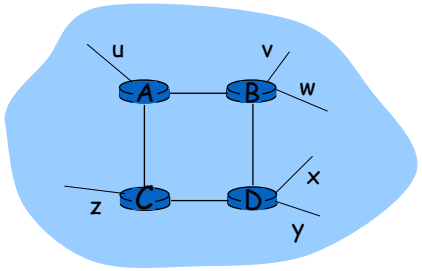
RIP

OSPF

ONE LOVE. ONE FUTURE.

RIP (Routing Information Protocol)

- RIP v.1, RIP v.2
- Distance vector protocol
- Select best routes according to the number of hop (# of hops, max = 15 hops)
- RIP v.1: RFC-1058 (www.ietf.org)

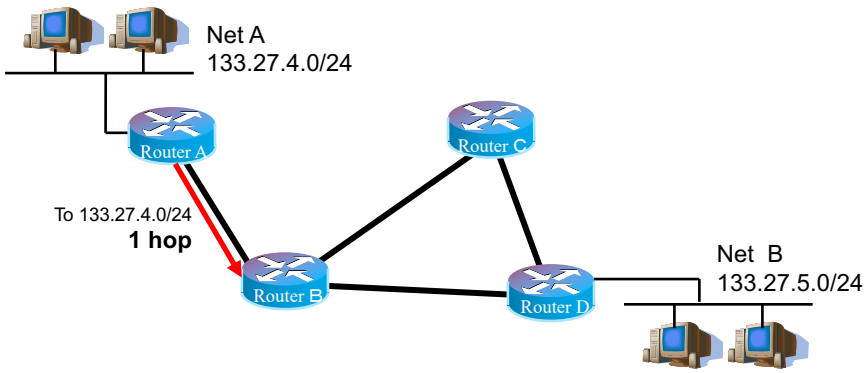


From A:

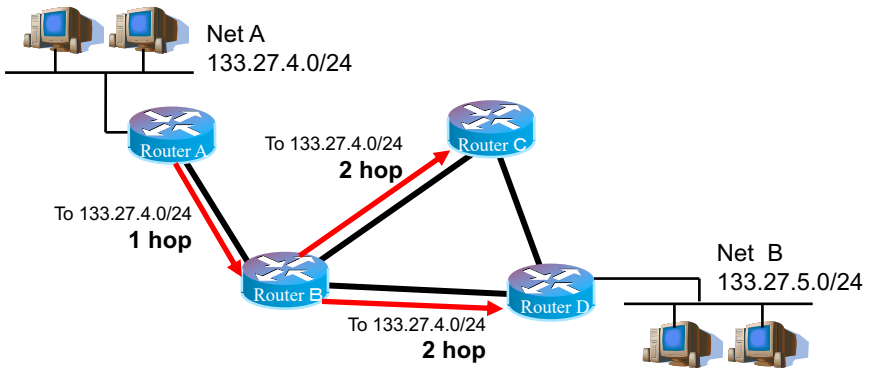
Destination	number of hops
u	1
v	2
w	2
x	3
y	3
z	2

Recall DV route calculation (1)

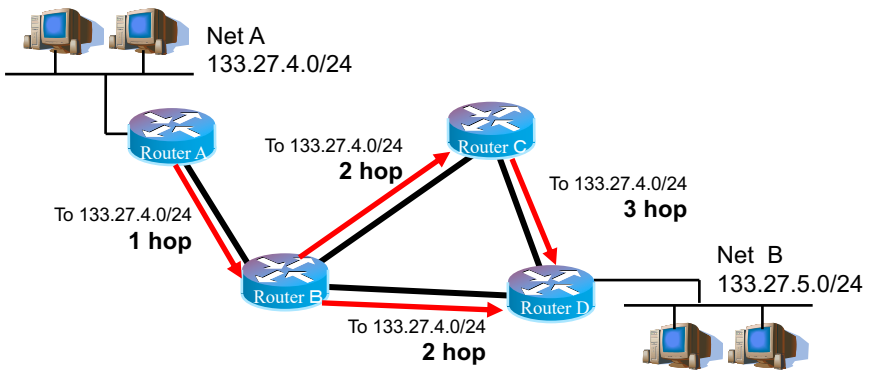
•



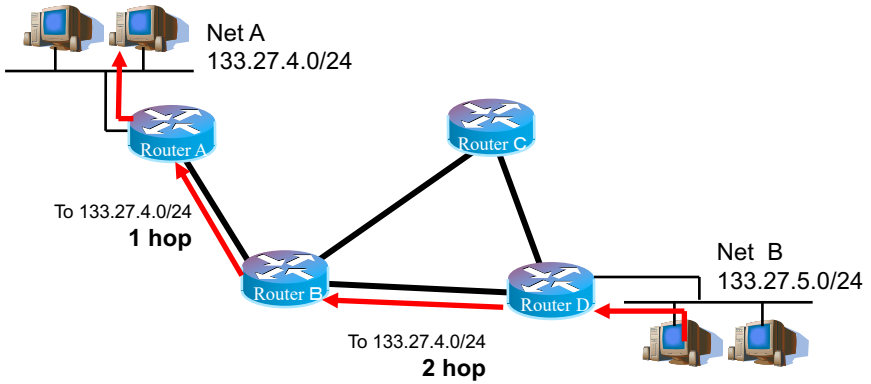
Recall DV route calculation (2)



Recall DV route calculation (3)



Recall DV route calculation (4)



RIP: Routers exchange information

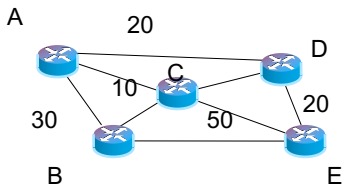
- Exchange distance vector
  - Distance vector are exchanged each 30s
  - Each message contains 25 items ➔ May have to send out more than one message if there are many elements in the vector

OSPF: Open Shortest Path First

- *Open*: Opened standard of IETF (OSPF v3 is defined in RFC 2740)
- *Shortest Path First*: Implement Dijkstra.
- Link-state protocol
- LSA (link state advertisement) is a packet describing state of some links in a network and is flooding in the network.
  - ➔ all routers in the network has the same database of link states ➔ Same topology view of the network
- Each link is assigned a weight
  - Best path is shortest weighted path.

Link state

- Link-State Advertisement (LSA): describe which nodes link each other (link) and corresponding (cost) tương ứng
- Ex: node A
  - link to B, cost 30
  - link to D, cost 20
  - link to C, cost 10
- Ex: node D
  - link to A, cost 20
  - link to E, cost 20
  - link to C, cost 50



OSPF - metric

- Weight given to a link
- By default:
  - 100Mbps / bandwidth of interface
  - Administator can assign the value

Default cost in OSPF

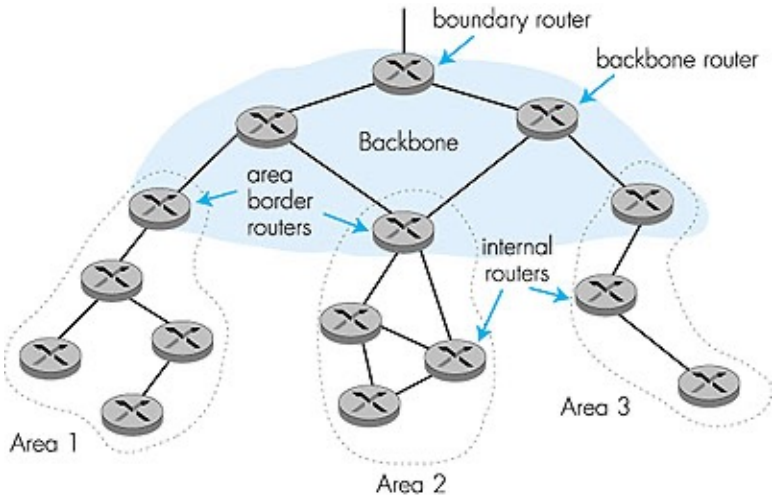
Link Bandwidth	Default OSPF cost
56Kbps serial link	1785
64Kbps serial link	1562
T1 (1.544Mbps) serial link	65
E1 (2.048Mbps) serial link	48
4Mbps Token Ring	25
Ethernet	10
16Mbps Token Ring	6
FDDI or Fast Ethernet	1
Gigabit Ethernet / 10G network	1

Hierachical OSPF

- Divide big network in smaller part to make routing more efficient?
- If there are too many routers
  - More link state messages to be circulated in the network
  - More computational effort is required
  - Bigger routing table
  - Routing table may be changed more oftenly

- Divide network into areas

Hierachical OSPF



RIP vs. OSPF comparison

	RIP	OSPF
Characteristic	• Router are equal • easy to configure • Small size network	• Hierachical • Complex to configure • Medium and big size network
Scalability	x	o
Complexity	Low	High
Convergence	Slow	Fast
Information exchange	Distant vector	linkstate
Algorithm	Distant vector	Link-state
Update neighbor	30s	10s (Hello packet)
Metrics	Hop	Bandwidth (default)

Interdomain routing protocol

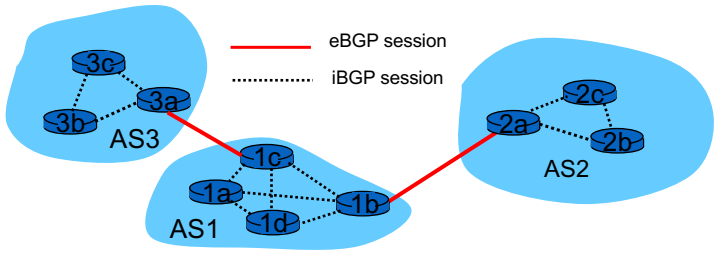
ONE LOVE. ONE FUTURE.

BGP – Border Gateway Protocol

- Exchange routing information between autonomous systems
- Information to be exchanged: NLRI (Network Layer Reachability Information)
  - Routes contains complete AS paths to other Ases
  - Routes are exchanged between border routers.
- A BGP router processes advertised routes according to policies:
  - Receives routes advertised from other AS
  - Filter the received routes according to import policies
  - Select the best routes to the same destination according to policies
  - Export the best routes to the neighbour AS according to export policies

eBGP and iBGP

- Router connects one AS to another one is called border router
- External BGP vs. Internal BGP
  - External BGP is used to exchange routing information between border router of different AS
  - Internal BGP is used to exchange routing information between border router of the same AS
- Disseminate routing informaiton
  1. 3a sends to 1c by eBGP
  2. 1c sends internal routing information to (1b, 1d, ...) within AS1 by iBGP
  3. 2a receives routing information from 1b by eBGP



Route attributes

- ORIGIN
  - Nguồn của thông tin (IGP/EGP/incomplete)
- AS_PATH
  - Liệt kê danh sách tất cả các AS cần phải đi qua để đến được một mạng đích.
- NEXT_HOP
- MED (MULTI_EXIT_DISCRIMINATOR)
- LOCAL_PREF
- ATOMIC_AGGREGATE
- AGGREGATOR
- COMMUNITY

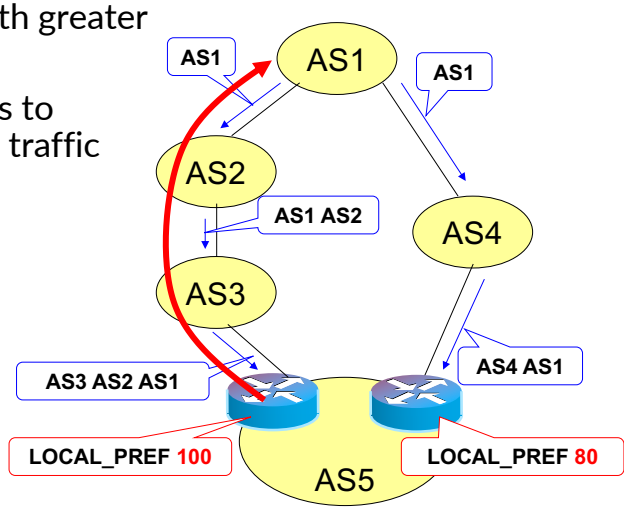
Best route selection steps

If a BGP router receives an advertised route for an known destination, it compares the new routes with the existing one on different attributes to find the best:

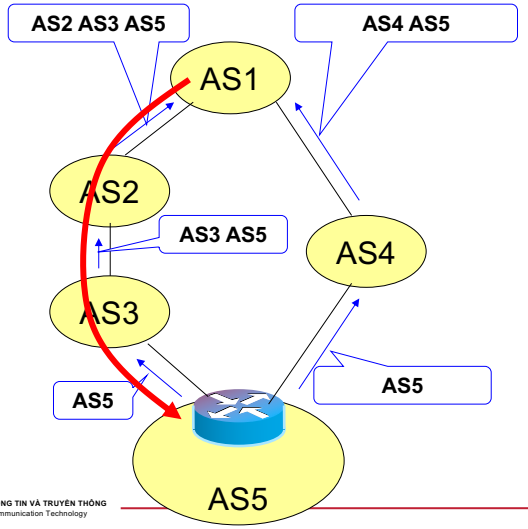
- Step 1: Compare the value of LOCAL_PREF
- Step 2: Compare the length of AS_PATH
- Step 3: Compare the value of ORIGIN
- Step 4: Compare MED
- Step 5: Compare EBGP/IBGP
- Step 6: Compare the cost to the NEXT_HOP
- Step 7: Compare Router ID

CompareLOCAL_PREF

- Choose route with greater LOCAL_PREF
- This policy allows to control upbound traffic



Compare AS_PATH Prepend

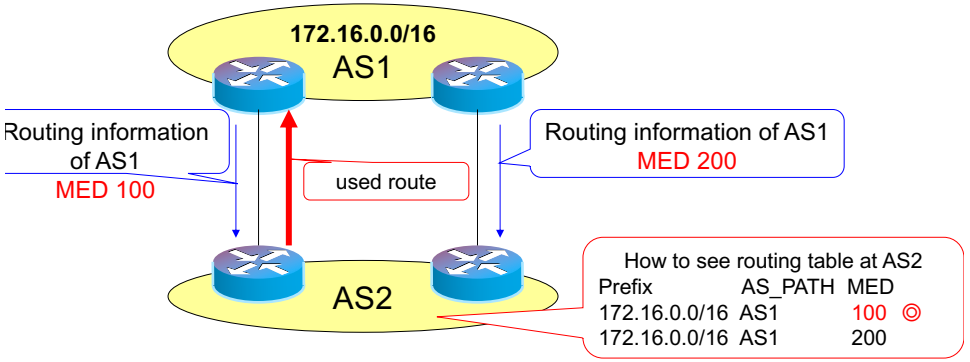


Example of AS PATH

Network	Next Hop	Metric	LocPrf	Weight	Path
4.79.201.0/26	203.178.136.29	700	500	0	7660 22388 11537 10886 40220
	203.178.136.29	700	500	0	7660 22388 11537 10886 40220
	203.178.136.29	700	500	0	7660 22388 11537 10886 40220
6.1.0.0/16	203.178.136.29	700	500	0	7660 22388 11537 668
	203.178.136.29	700	500	0	7660 22388 11537 668
	203.178.136.29	700	500	0	7660 22388 11537 668
6.2.0.0/22	203.178.136.29	700	500	0	7660 22388 11537 668

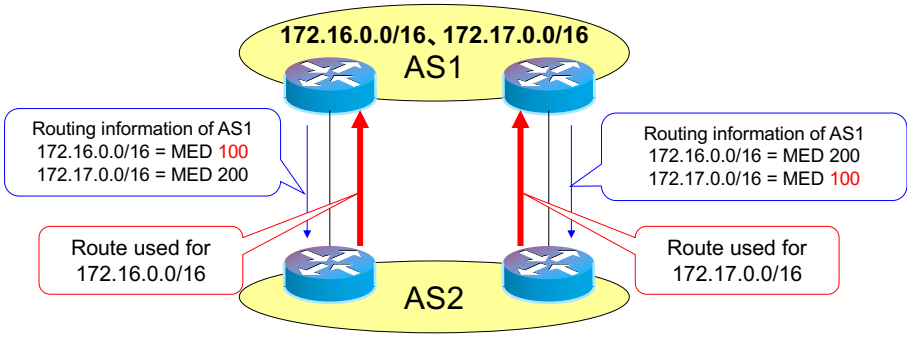
Compare MED

- In case that 2 Ases are connected to each other by multiple links
- Choose the path with smaller MED
- MED is also used to control the outbound traffic



Using MED to distribute traffic

- Set a MED value for each path



ARP protocol

ONE LOVE. ONE FUTURE.

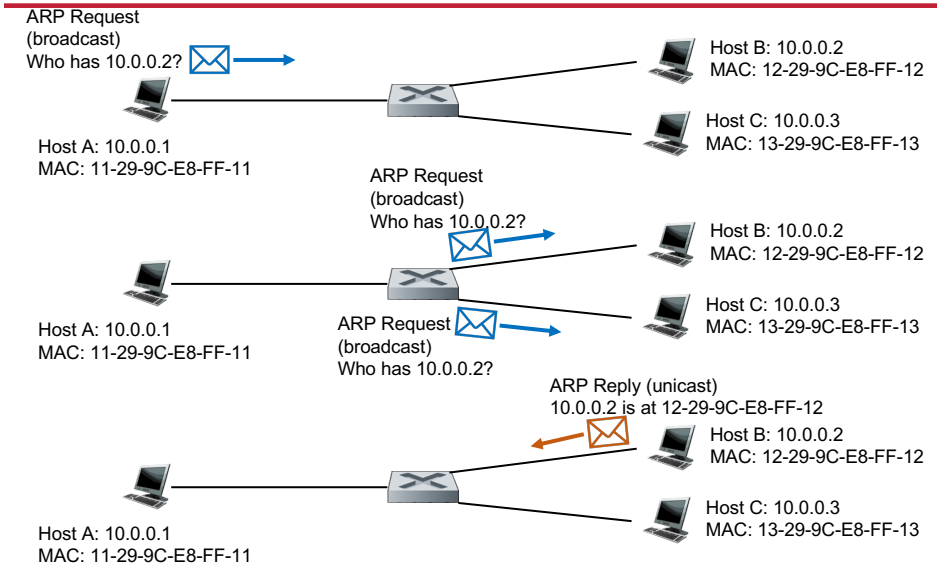
MAC address and ARP

- Address Resolution Protocol
- Identify MAC address (used by Datalink layer) of an interface given its IP address
- Why ARP is necessary?
  - Data transmission in network layer uses IP address
  - Data transmission in datalink layer uses MAC address
  - On sender side, when data is forwarded from network layer to data link layer:
    - If sending data within the same LAN : sender needs to know MAC address of destination for using layer 2 forwarding mechanism
    - If sending data outside the LAN: sender needs the MAC address of the default router.

ARP operation

- Each network node has a ARP Table:
  - Contains mapping <IP address, MAC address, TTL>
  - TTL: Time to live of the mapping information in the table (300 seconds)
- For asking MAC address of another node, the node broadcasts the ARP Request message in the network.
  - ARP Request packet contains the IP address of the node to be searched.
- The node holding the requested IP address should reply with its MAC address in an ARP Reply message

ARP operation – Example



Data transmission between LANs

Ex: A sends data to B via router R (according to routing table of A)

- A creates an IP packet with source address is A and destination address is B
- The packet is forwarded to the datalink layer of A to be packed in a layer 2 frame with MAC source address of A and MAC destination address of R

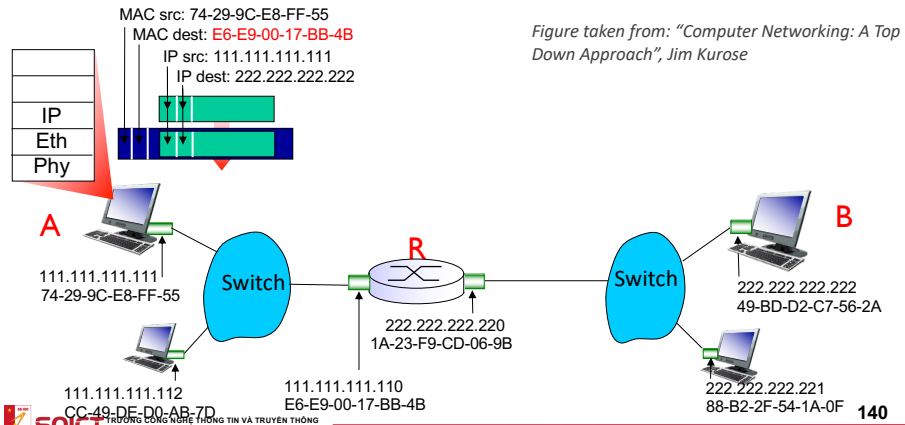


Figure taken from: "Computer Networking: A Top Down Approach", Jim Kurose

Data transmission between LANs

- Frame is forwarded from A to R by learning mechanism of switch or multiple access mechanism of layer 2
- At R: the header of the frame is removed and the content is delivered to the network layer as an IP packet

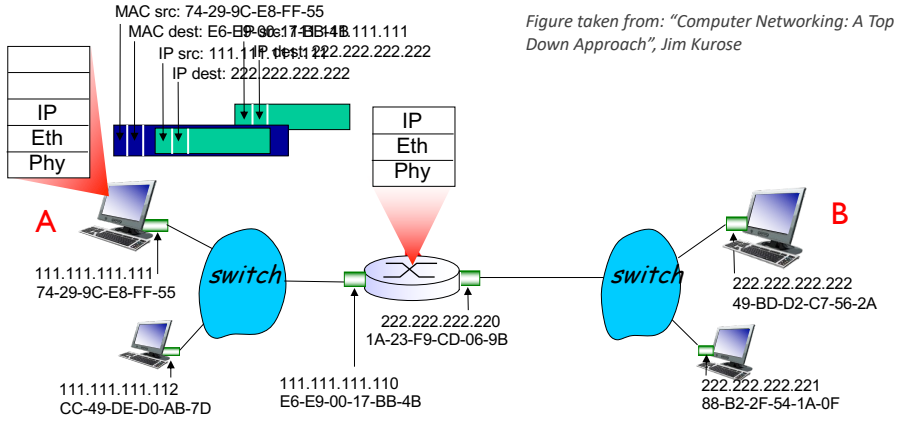
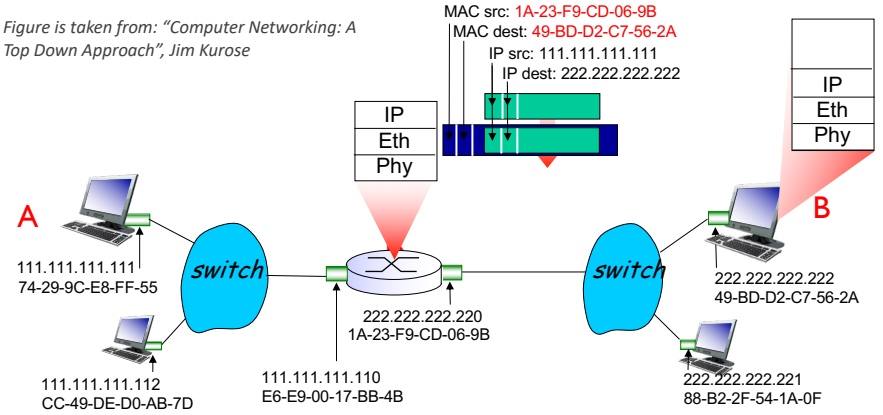


Figure taken from: "Computer Networking: A Top Down Approach", Jim Kurose

Data transmission between LANs

- R chooses the next hop to forward the IP packet with source A and destination B according to its routing table
- The IP packet is then packaged in a frame of layer 2 with MAC address of R as source and MAC address of B as destination.

Figure is taken from: "Computer Networking: A Top Down Approach", Jim Kurose



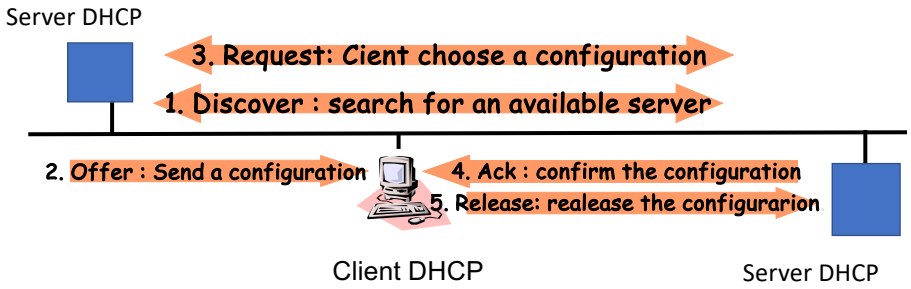
DHCP

ONE LOVE. ONE FUTURE.

Introduction

- Dynamic Host Configuration Protocol
- A service of the application layer that distributes configurations to hosts. Configuration includes
  - IP address
  - Network mask
  - default router, default gateway
  - Possibly the address of default DNS servers
- DHCP works using client/server model: DHCP client hosts use IP addresses given by DHCP servers.

DHCP : IP address distribution process



DHCP messages

Client

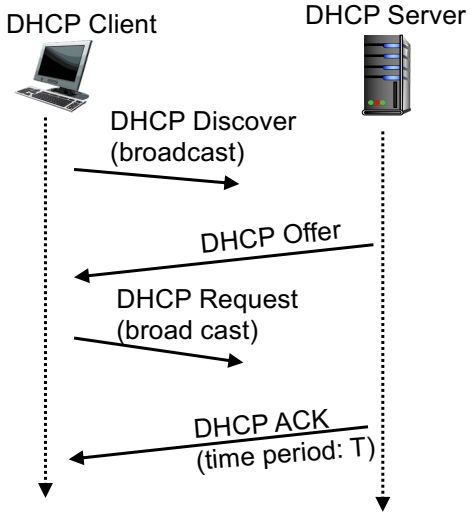
- DHCP Discover: search for DHCP Server
- DHCP Request: register an IP address
- DHCP Release: return the used IP address to the pool
- DHCP Decline: Refuse an assigned IP address

Server

- DHCP Offer: provide configuration including IP address
- DHCP ACK: Accept the registration
- DHCP NAK: Refuse the registration

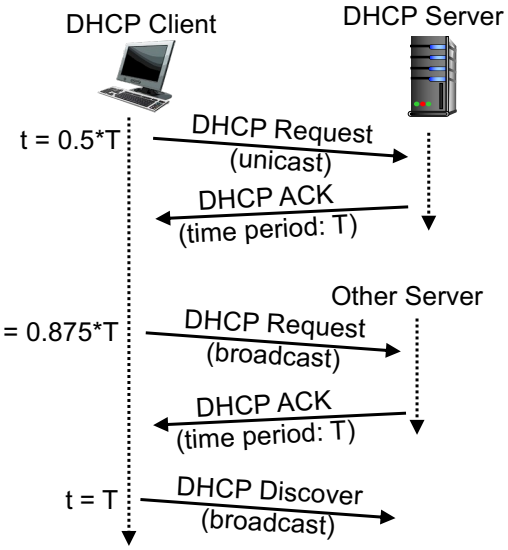
Provide new configuration

- B1: Client broadcasts DHCP Discover message to look for a Server
- B2: If there is a DHCP Server in the network,the server sends DHCP Offer with address information
- B3: Client chooses one configuration from received DHCP Offers and send DHCP Request to register the configuration
- B4: DHCP Server sends a DHCP ACK to accept.



Extend the configuration

- Each configuration is valid in a time period T → client needs to extend the validation of the configuration
- When t = 0.5*T, client sends DHCP Request to DHCP Server to extend the lease
- If there is no DHCP ACK, at t= 0.875*T, client broadcasts DHCP Request
- If there is still no DHCP ACK, at t = T, client sends DHCP Discover



DHCP Relay

- DHCP Server belong to a different network with the DHCP client → broadcasting messages are not forwarded by routers
  - → install DHCP Relay on router