



ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

Chương 4

Tầng liên mạng



ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

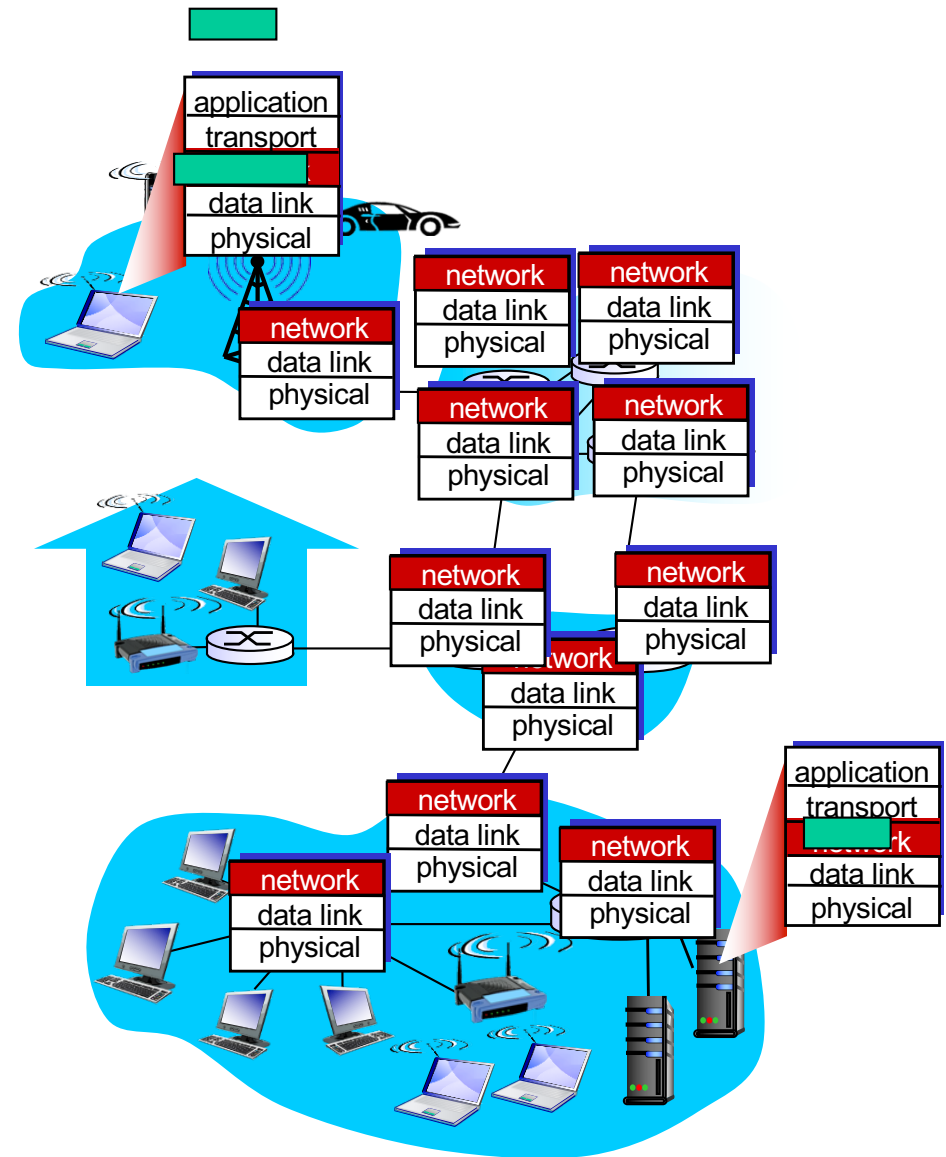
4.1. Tổng quan về tầng liên mạng

Tầng liên mạng trên kiến trúc TCP/IP

Application (HTTP, Mail, ...)	Hỗ trợ các ứng dụng trên mạng
Transport (UDP, TCP ...)	Điều khiển truyền dữ liệu giữa các tiến trình của tầng ứng dụng
Network (IP, ICMP...)	Điều khiển truyền dữ liệu sang mạng khác (truyền dữ liệu giữa các thành phần nối không trực tiếp với nhau)
Datalink (Ethernet, ADSL...)	Hỗ trợ việc truyền thông cho các thành phần kế tiếp trên cùng 1 mạng
Physical (bits...)	Truyền và nhận dòng bit trên đường truyền vật lý

Tầng mạng

- Truyền dữ liệu từ host-host
- Cài đặt trên mọi hệ thống cuối và bộ định tuyến
- Đơn vị truyền: datagram
- Bên gửi: nhận dữ liệu từ tầng giao vận, đóng gói
- Bên nhận: mở gói, chuyển phần dữ liệu trong payload cho tầng giao vận
- Bộ định tuyến(router): định tuyến và chuyển tiếp gói tin

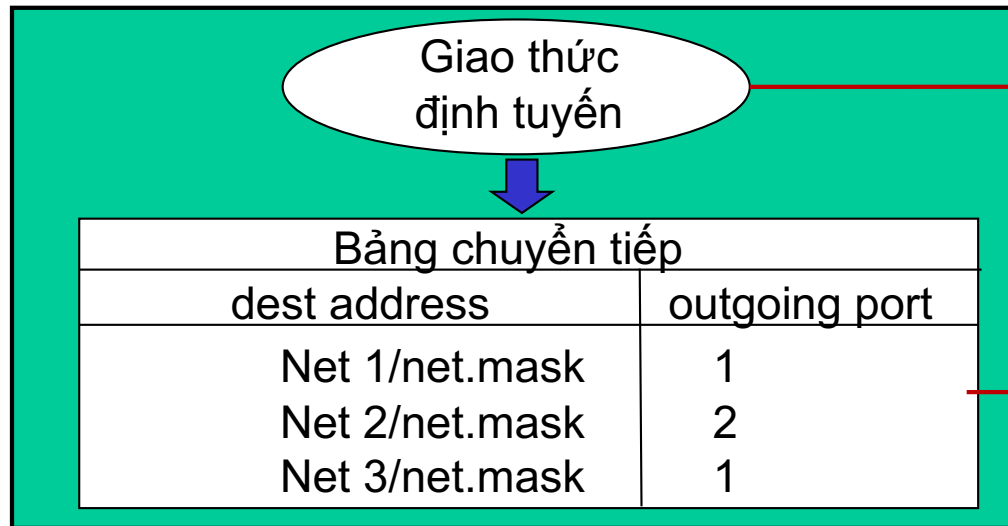


Hình ảnh từ: "Computer Networking: A Top Down Approach", Jim Kurose

Chức năng chính

- **Định tuyến (Routing)**: Tìm tuyến đường để gửi dữ liệu từ nguồn tới đích
- **Chuyển tiếp (Forwarding)**: Chuyển gói tin tới cổng ra theo tuyến đường đã có
- **Định địa chỉ (Addressing)**: Định danh cho các nút mạng
- **Đóng gói dữ liệu (Encapsulating)**
- **Đảm bảo chất lượng dịch vụ (QoS)**: đảm bảo các thông số phù hợp của đường truyền theo từng dịch vụ

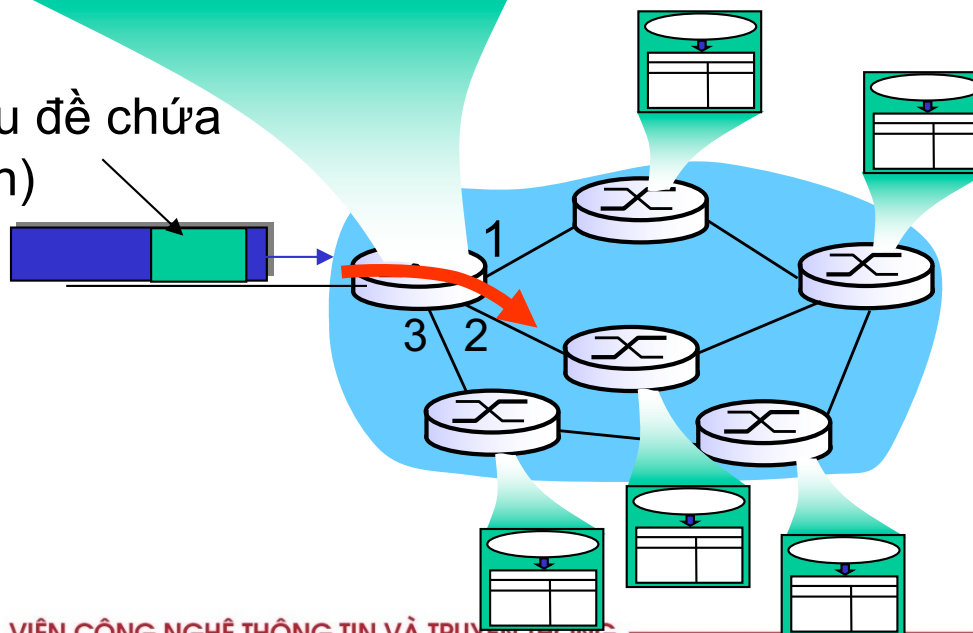
Định tuyến và chuyển tiếp



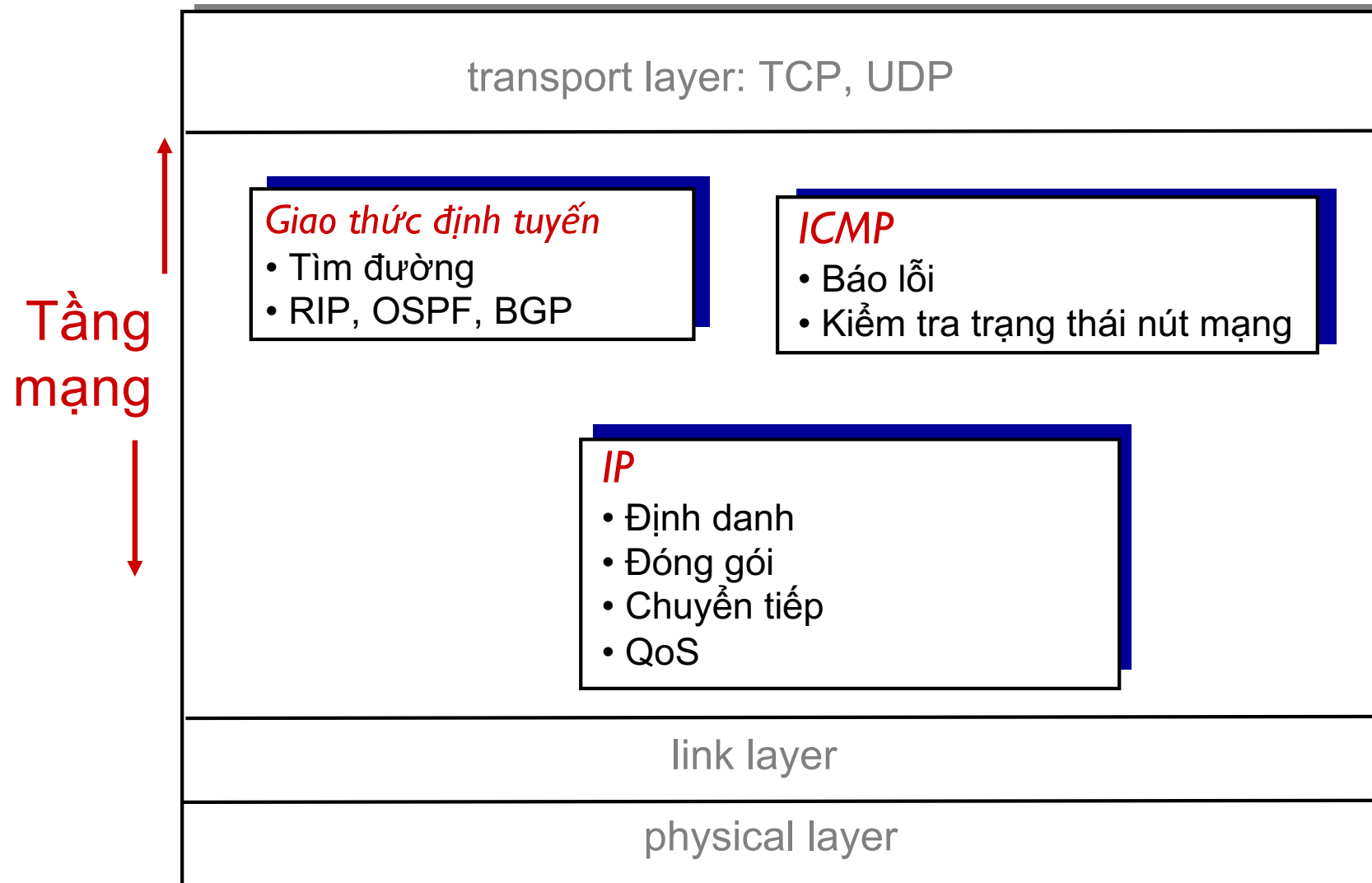
Giao thức định tuyến xác định đường đi ngắn nhất giữa 2 bên truyền tin

Bảng chuyển tiếp xác định cổng ra (outgoing port) để chuyển dữ liệu tới đích

Gói tin (tiêu đề chứa địa chỉ đích)



Các giao thức tầng mạng





ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

4.6. Giao thức IPv4

Đặc điểm cơ bản của giao thức

- Là giao thức **được** định tuyến (**routed protocol**)
 - Tập trung vào việc chuyển tiếp gói tin (forwarding)
 - Đòi hỏi phải có các giao thức định tuyến để xác định trước đường đi cho dữ liệu.
- Giao thức hướng không liên kết
- Không tin cậy / nhanh
 - Truyền dữ liệu theo phương thức “*best effort*”
 - IP không có cơ chế phục hồi nếu có lỗi
 - Khi cần, ứng dụng sẽ sử dụng dịch vụ tầng trên để đảm bảo độ tin cậy (TCP)
- Hiện nay có 2 phiên bản: IPv4 và IPv6



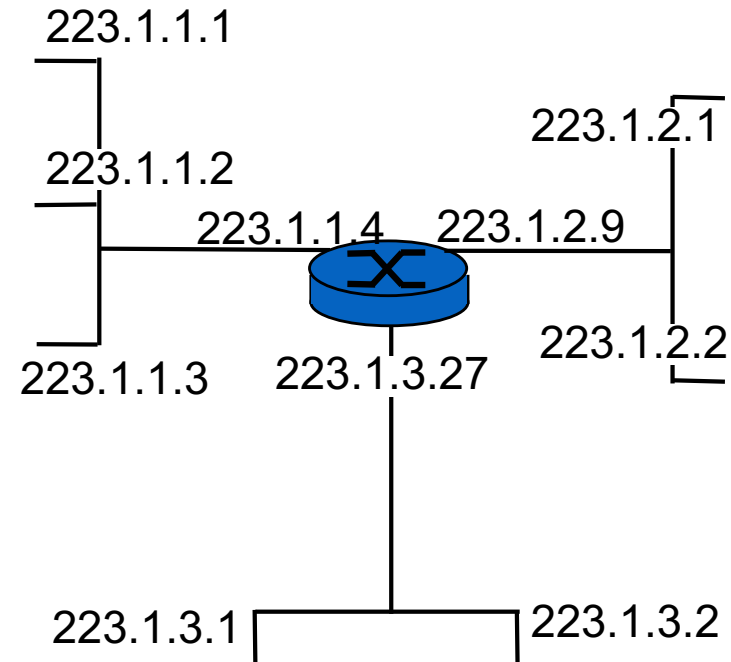
25 YEARS ANNIVERSARY
SOICT

ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

Địa chỉ IPv4

Địa chỉ IP (IPv4)

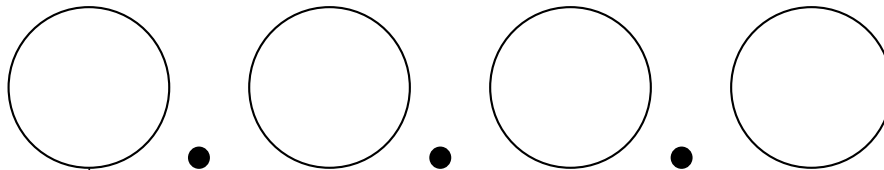
- **Địa chỉ IPv4:** Một số 32-bit để định danh một giao diện mạng (cổng kết nối mạng trên nút đầu cuối) PC, server, smart phone, bộ định tuyến
- Mỗi địa chỉ IP được gán cho một cổng duy nhất
- Địa chỉ IP có tính duy nhất trong mạng
- Các máy nằm trong cùng một đoạn mạng được gán các địa chỉ IP tương tự nhau.
 - Có các bit đầu giống nhau.



$$223.1.1.1 = \underbrace{11011111}_{223} \underbrace{00000001}_1 \underbrace{00000001}_1 \underbrace{00000001}_1$$

Biểu diễn địa chỉ IPv4

- Biểu diễn dạng thập phân có chấm



Số nguyên 8 bit

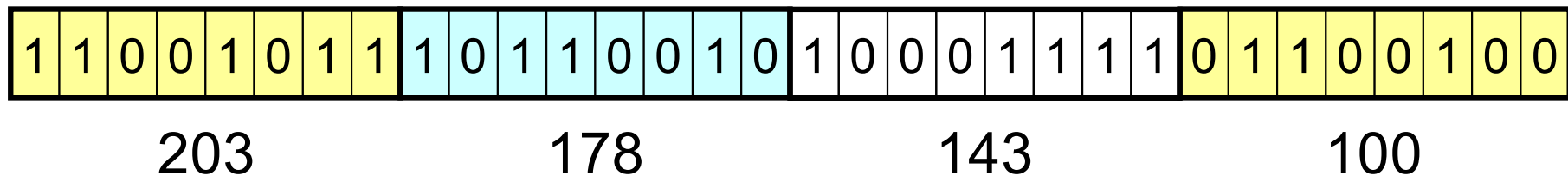
Ví dụ:

203.178.136.63 o

259.12.49.192 x

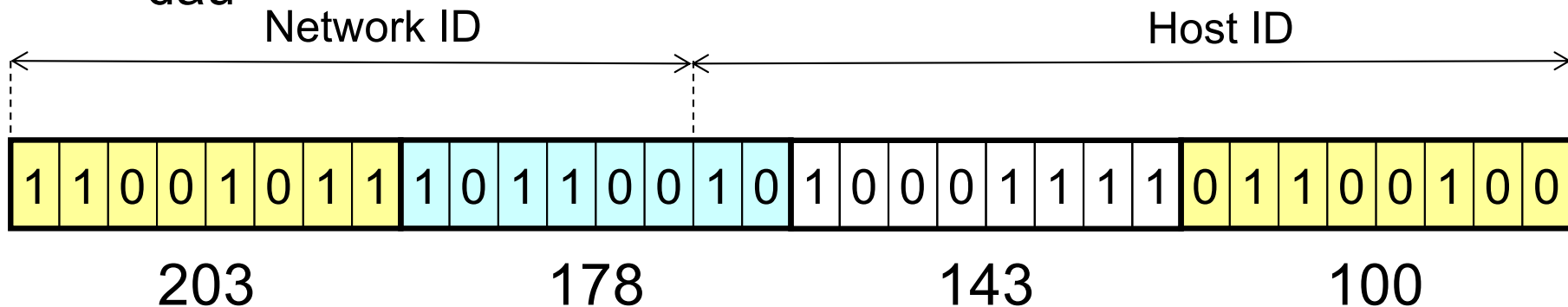
133.27.4.27 o

Sử dụng 4 phần 8 bit để miêu tả một địa chỉ 32 bits



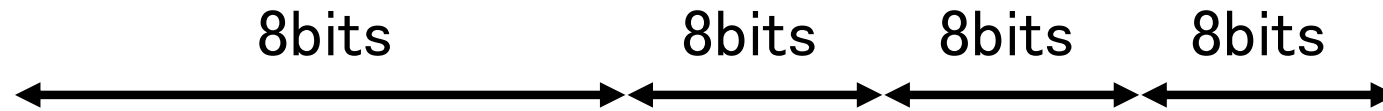
Địa chỉ IPv4

- Địa chỉ IP có hai phần
 - Host ID (định danh máy) – phần địa chỉ máy trạm: (32-n) bit cuối
 - Network ID (định danh mạng) – phần địa chỉ mạng: n bit đầu



- Làm thế nào biết được phần nào là cho máy trạm, phần nào cho mạng?
 - Phân lớp địa chỉ
 - Không phân lớp – CIDR

Phân lớp địa chỉ IP (classful)



Class A	0	7bit	H	H	H
Class B	1 0	6bit	N	H	H
Class C	1 1 0	5bit	N	N	H
Class D	1 1 1 0	Multicast			
Class E	1 1 1 1	Reserve for future use			

	# of network	# of addresses
Class A	128	2^{24}
Class B	16384	65536
Class C	2^{21}	256

Xác định lớp của địa chỉ IP

- 1) 10.1.0.1 : lớp A, byte đầu 0-127
- 2) 129.203.8.5: lớp B, byte đầu 128-191
- 3) 195.4.9.12 : lớp C, byte đầu 192-223

Bài tập:

Cho các địa chỉ sau:

154.17.4.9

193.67.2.0

Xác định địa chỉ thuộc lớp nào?

Xác định phần netID, host ID của địa chỉ.

Hạn chế của việc phân lớp địa chỉ

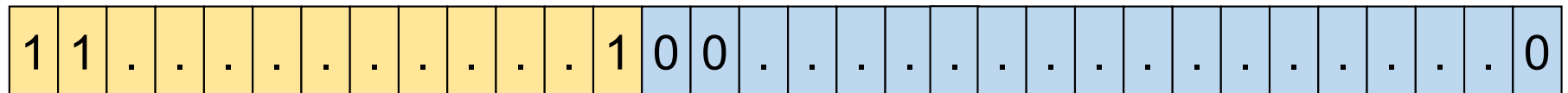
- Lãng phí không gian địa chỉ
 - Việc phân chia thành các lớp (A, B, C, D, E) làm hạn chế việc sử dụng toàn bộ không gian địa chỉ
- Phương pháp CIDR: **C**lassless **I**nter **D**omain **R**outing
 - Classless addressing
 - Phần địa chỉ mạng (NetworkID) sẽ có độ dài bất kỳ
 - Dạng địa chỉ: **m1.m2.m3.m4 /n**, trong đó n (mặt nạ mạng) là số bit trong phần ứng với địa chỉ mạng (NetworkID)

Mặt nạ mạng

- Mặt nạ mạng xác định 2 phần trong địa chỉ IPv4
 - Phần ứng với máy trạm (HostID)
 - Phần ứng với mạng (NetworkID)

- Biểu diễn:

- Mặt nạ mạng: /n. Ví dụ: /27
 - n: Số bit của NetworkID
- Nhị phân:

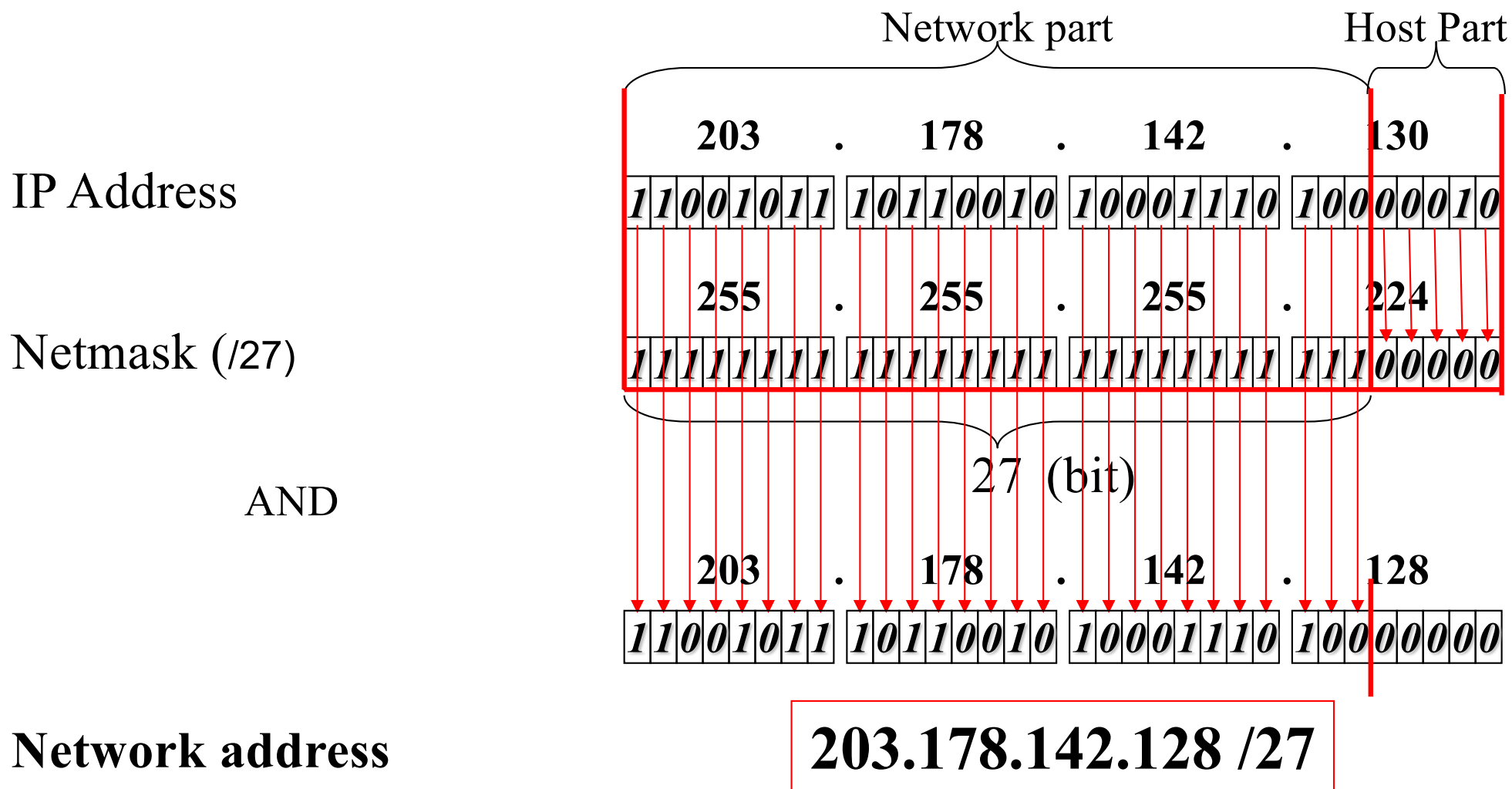


n bits

(32 - n) bits

- Ví dụ: /27 = 11111111 11111111 11111111 11100000
- Thập phân có chấm: 255.255.255.224

Cách tính địa chỉ mạng



Các dạng địa chỉ

- Địa chỉ mạng (Network Address):
 - Định danh cho một mạng
 - Tất cả các bit phần HostID là 0
 - 0.0.0.0: Địa chỉ toàn mạng, đại diện cho 1 mạng bất kỳ
- Địa chỉ quảng bá (Broadcast Address)
 - Địa chỉ dùng để gửi dữ liệu cho tất cả các máy trạm trong mạng
 - Tất cả các bit phần HostID là 1
 - 255.255.255.255: Địa chỉ quảng bá trong 1 mạng bất kỳ
- Địa chỉ máy trạm (Unicast Address)
 - Gán cho một cổng mạng
- Địa chỉ nhóm (Multicast address): định danh cho nhóm

Địa chỉ IP và mặt nạ mạng

- Tính địa chỉ mạng của các địa chỉ IP dưới đây
- Địa chỉ nào là địa chỉ máy trạm, địa chỉ mạng, địa chỉ quảng bá?

(1) 203.178.142.128 /25

(2) 203.178.142.128 /24

(3) 203.178.142.127 /25

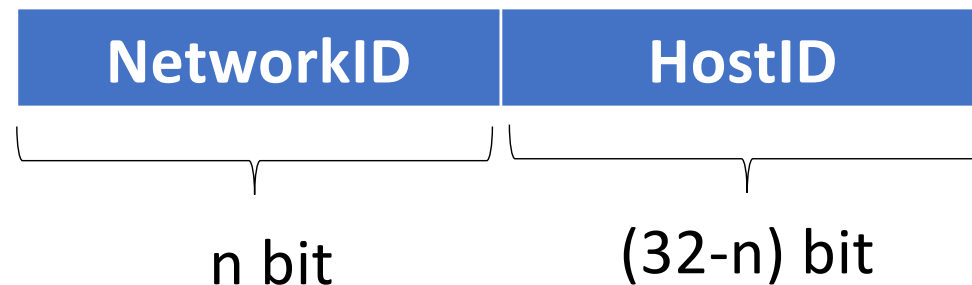
(4) 203.178.142.127 /24

- Lưu ý: Với cách địa chỉ hóa theo CIDR, địa chỉ IP và mặt nạ mạng luôn phải đi cùng nhau

Mặt nạ mạng và kích thước mạng

- Mặt nạ mạng: /n
- Kích thước mạng: Số nút mạng tối đa có trong mạng

- Địa chỉ IP:



- Số lượng địa chỉ máy trạm (Unicast Address):

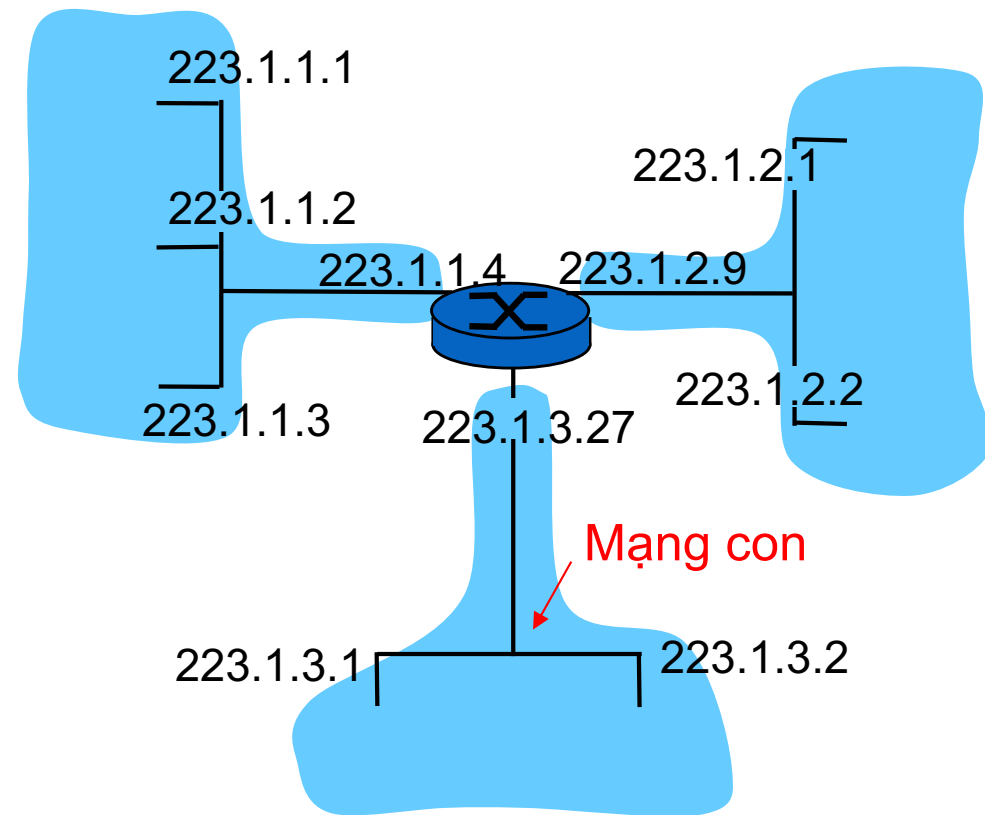
$$2^{32-n} - 2$$

Cấp phát địa chỉ IPv4

- Cấp phát cố định(Static IP):
 - Windows: Control Panel → Network → Configuration → TCP/IP → Properties
 - ipconfig
 - Linux: /etc/network/interfaces
 - ifconfig
- Cấp phát tự động: **DHCP- Dynamic Host Configuration Protocol**

Mạng con - subnet

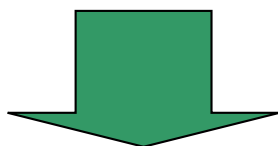
- Subnet: Là một phần của một mạng nào đó
- Vấn đề:
 - ISP thường được gán một khối địa chỉ IP
 - Một vài mạng con sẽ được tạo ra
- Điều kiện
 - Mỗi mạng con có một địa chỉ mạng riêng.
- Tạo subnet như thế nào
 - Sử dụng một mặt nạ mạng lớn hơn



Mạng với 3 mạng con
Dải địa chỉ IP gốc là 223.1.0.0/16

Ví dụ: Chia làm 2 subnets

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



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

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

SubnetID

Mặt nạ mạng con = Mặt nạ mạng cũ + Số bit SubnetID
Chia thành 4 mạng con? → Mặt nạ mạng con: /26

Ví dụ: Chia subnets

- Mạng 200.23.16.0/24

1) Cần tạo 4 mạng con

- Địa chỉ mỗi mạng con? Mặt nạ của chúng? Số lượng máy/mạng

2) Cần tạo N mạng con. Địa chỉ mạng? Mặt nạ?

- Mỗi mạng con có tối đa 14 máy tính
- Mạng với 30 máy tính
- Mạng với 31 máy tính
- Mạng với 70 máy tính



/28

/27

/26

25 /25

Không gian địa chỉ IPv4

- Theo lý thuyết
 - Có thể là 0.0.0.0 ~ 255.255.255.255
 - Một số địa chỉ đặc biệt
- Địa chỉ IP đặc biệt ([RFC1918](#))

Private address	10.0.0.0/8 172.16.0.0/16 → 172.31.0.0/16 192.168.0.0/24 → 192.168.255.0 /24
Loopback address	127.0.0.0 /8
Multicast address	224.0.0.0 ~239.255.255.255

- Địa chỉ liên kết nội bộ: 169.254.0.0/16 (tự động cấu hình)

Quản lý địa chỉ IP công cộng

- Internet Corporation for Assigned Names and Numbers (ICANN): quản lý toàn bộ tài nguyên địa chỉ IP
- Regional Internet Registries: quản lý địa chỉ IP theo vùng (châu Á-Thái Bình Dương, châu Âu và Trung Đông, châu Phi, Bắc Mỹ, Nam Mỹ)
- Cơ quan quản lý quốc gia
 - Việt Nam: VNNIC
- Nhà cung cấp dịch vụ (ISP)
- Cơ quan, tổ chức
- Ví dụ: ICANN → APNIC → VNNIC → HUST



ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

4.3. Chuyển đổi địa chỉ (NAT)

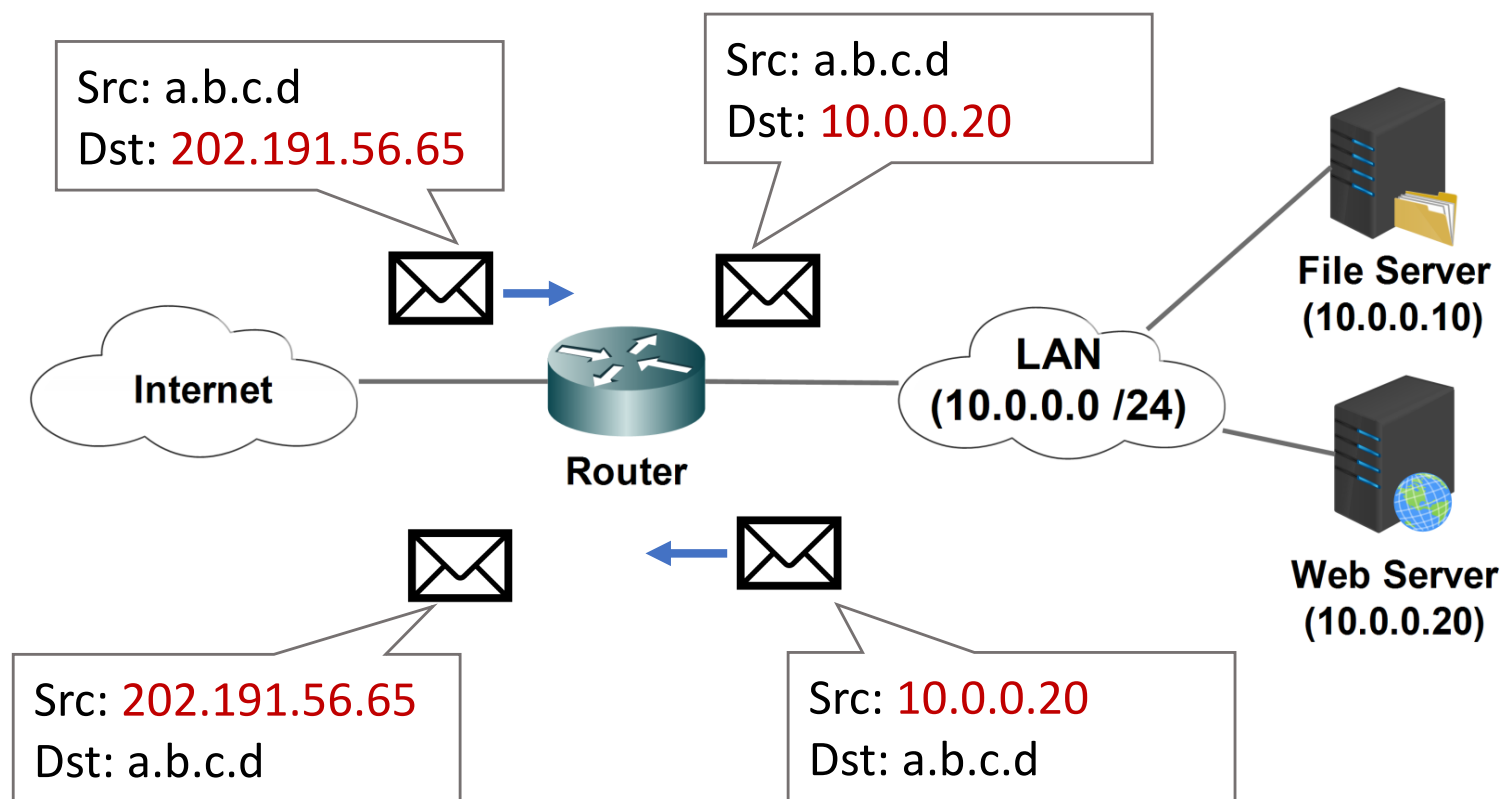
Khái niệm cơ bản NAT

- Dữ liệu chuyển tiếp từ mạng LAN (sử dụng địa chỉ cục bộ) sang mạng Internet (sử dụng địa chỉ công cộng) và ngược lại cần được chuyển đổi địa chỉ
- NAT (Network Address Translation) thực hiện trên bộ định tuyến
- Trên thực tế, có thể sử dụng NAT để chuyển đổi địa chỉ IP từ mạng LAN này sang mạng LAN khác
- Lợi ích:
 - Tiết kiệm địa chỉ IP công cộng
 - Che giấu địa chỉ riêng
 - Giảm chi phí cấu hình khi thay đổi ISP

Static NAT

- Là dạng NAT đơn giản nhất
- Một địa chỉ IP private được gán cho một địa chỉ IP public xác định
- NAT router sẽ lưu một bảng chuyển đổi trong bộ nhớ
 - Bảng này ánh xạ các địa chỉ nội bộ bên trong với địa chỉ bên ngoài Internet
- Thường sử dụng cho các server nằm trong mạng LAN để cung cấp dịch vụ cho mạng công cộng

Static NAT - Ví dụ



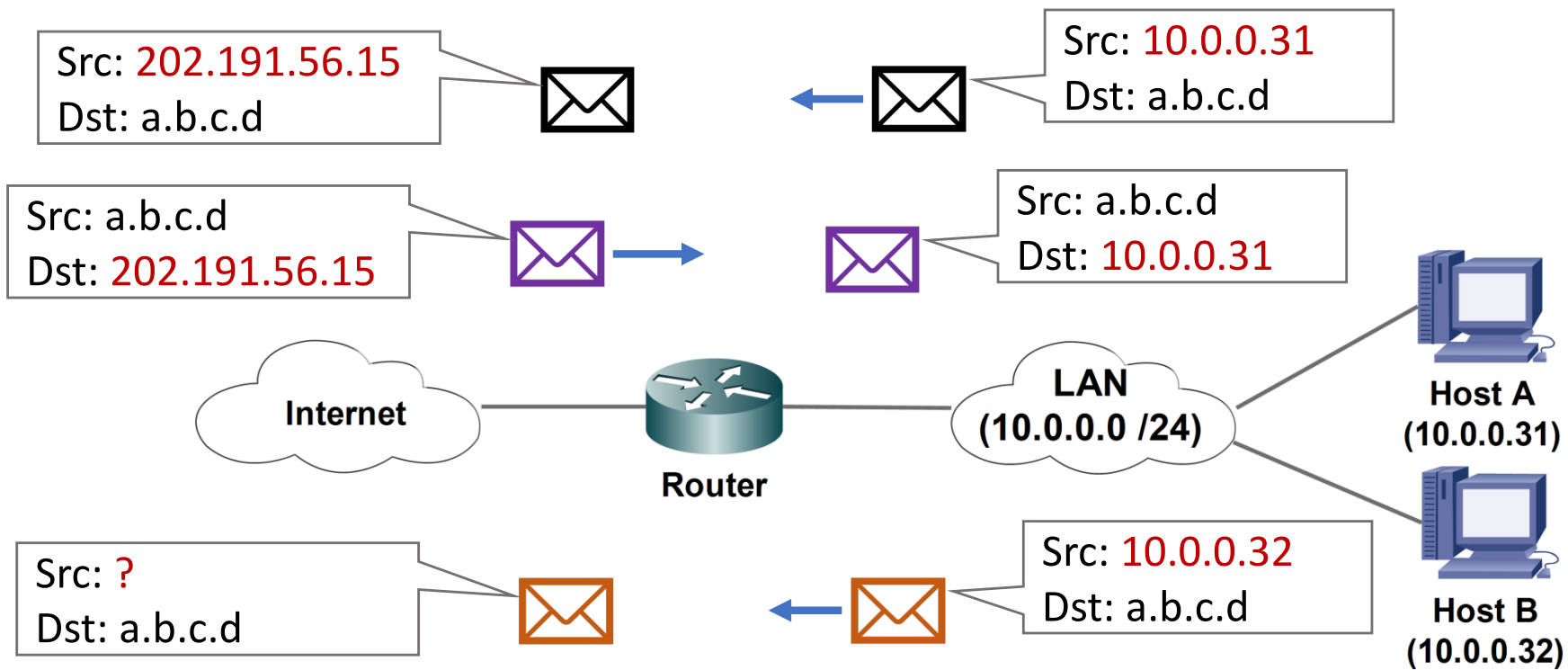
NAT Table

Inside IP	Outside IP
10.0.0.10	202.191.56.11
10.0.0.20	202.191.56.65

Dynamic NAT

- NAT router tự động ánh xạ một dải các địa chỉ IP private với một dải các địa chỉ public để các máy tính có thể đi ra ngoài Internet khi cần thiết
- Trái với static NAT, người quản trị sẽ không cần cấu hình cụ thể 1 địa chỉ IP private được gán với 1 địa chỉ IP public nào
- Bất kỳ địa chỉ IP private nào sẽ được gán và dịch tự động sang 1 trong các địa chỉ IP public đang sẵn có (đang rỗi) ở trong pool địa chỉ public bởi NAT router

Dynamic NAT – Ví dụ

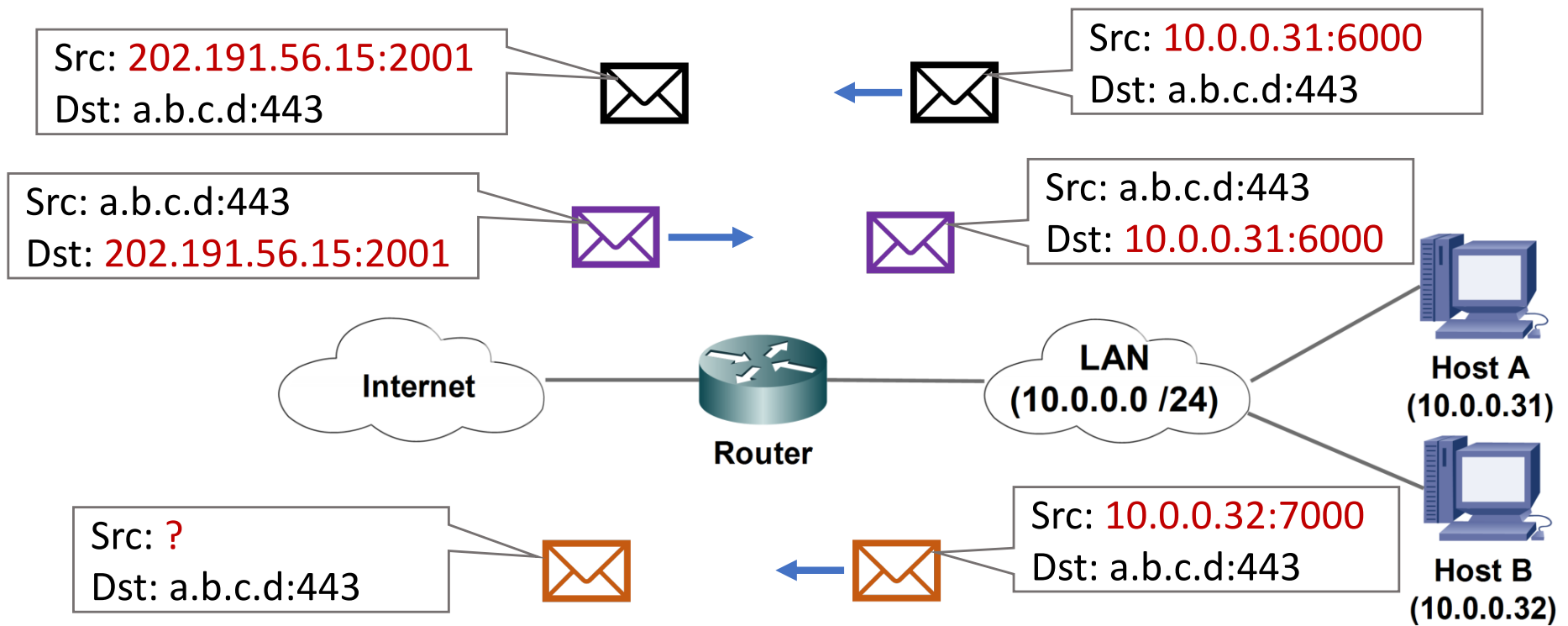


NAT Table	Inside IP	Outside IP
	10.0.0.31	202.191.56.15
	10.0.0.32	202.191.56.16

Port Address Translation

- Còn được biết là NAT **overloading**
- Là một dạng đặc biệt của dynamic NAT để tận dụng hơn nữa không gian địa chỉ IPv4
- Cho phép nhiều địa chỉ IP private ánh xạ với một địa chỉ IP public đã đăng ký kèm các số hiệu cổng khác nhau
 - n private IP \rightarrow 1 public IP
- Như vậy PAT sử dụng các sockets duy nhất để ánh xạ
 - (private IP : số hiệu cổng) \leftrightarrow (public IP: số hiệu cổng)

PAT- Ví dụ



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

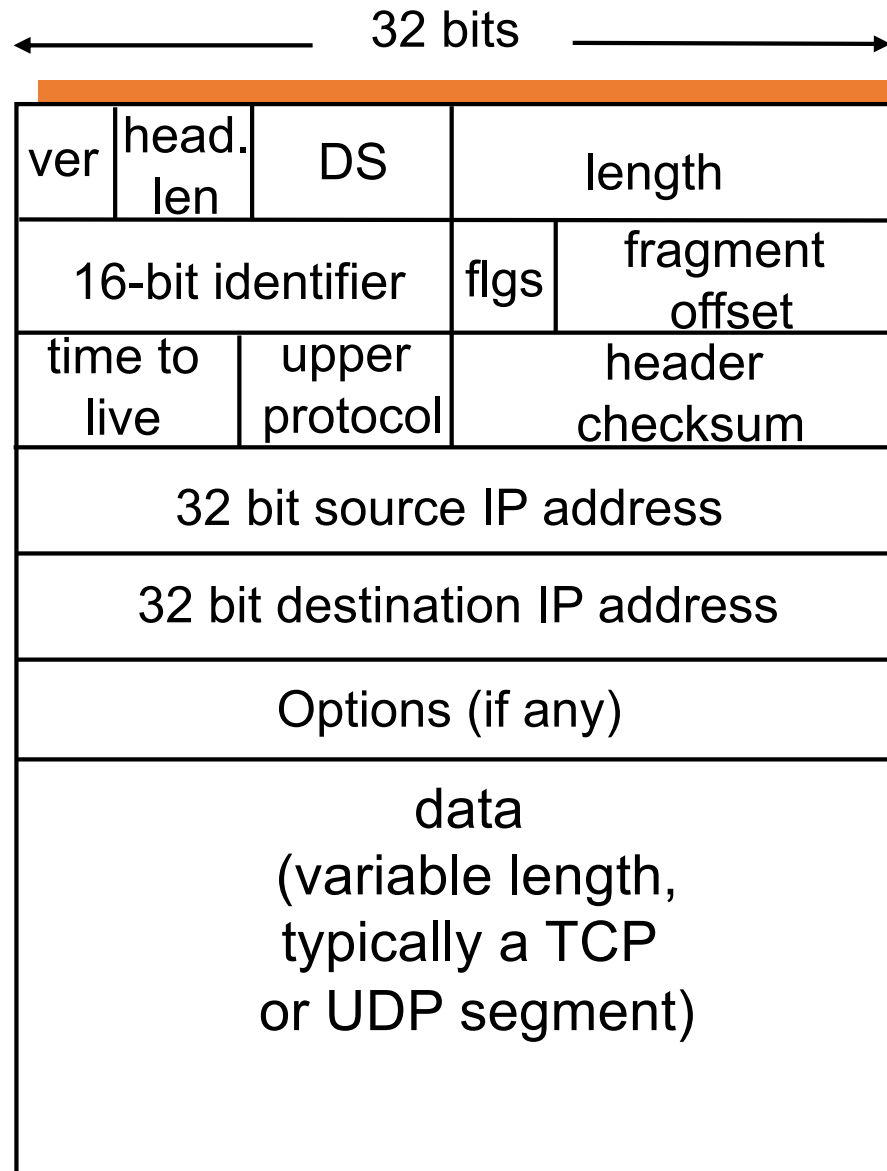


25 YEARS ANNIVERSARY
SOICT

ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

Khuôn dạng gói tin IP

Tiêu đề gói tin IP



Tiêu đề gói tin IP (1)

- Version: Phiên bản giao thức (4 bits)
 - IPv4
 - IPv6
- Header length: Độ dài phần đầu: 4bits
 - Tính theo từ (4 bytes)
 - Min: 5 x 4 (byte)
 - Max: 15 x 4 (byte)
- DS (Differentiated Service : 8bits)
 - Tên cũ: Type of Service
 - Hiện tại được sử dụng trong quản lý QoS

Tiêu đề gói tin IP (2)

- Length: Độ dài toàn bộ, tính cả phần đầu (16 bits)
 - Theo bytes
 - Max: 65536
- Identifier – Số hiệu gói tin (16 bit)
 - Dùng để xác định một chuỗi các gói tin của một gói tin bị phân mảnh
- Flag – Cờ báo phân mảnh(3 bit)
- Fragmentation offset – Vị trí gói tin phân mảnh trong gói tin ban đầu (13 bit)

Tiêu đề gói tin IP (3)

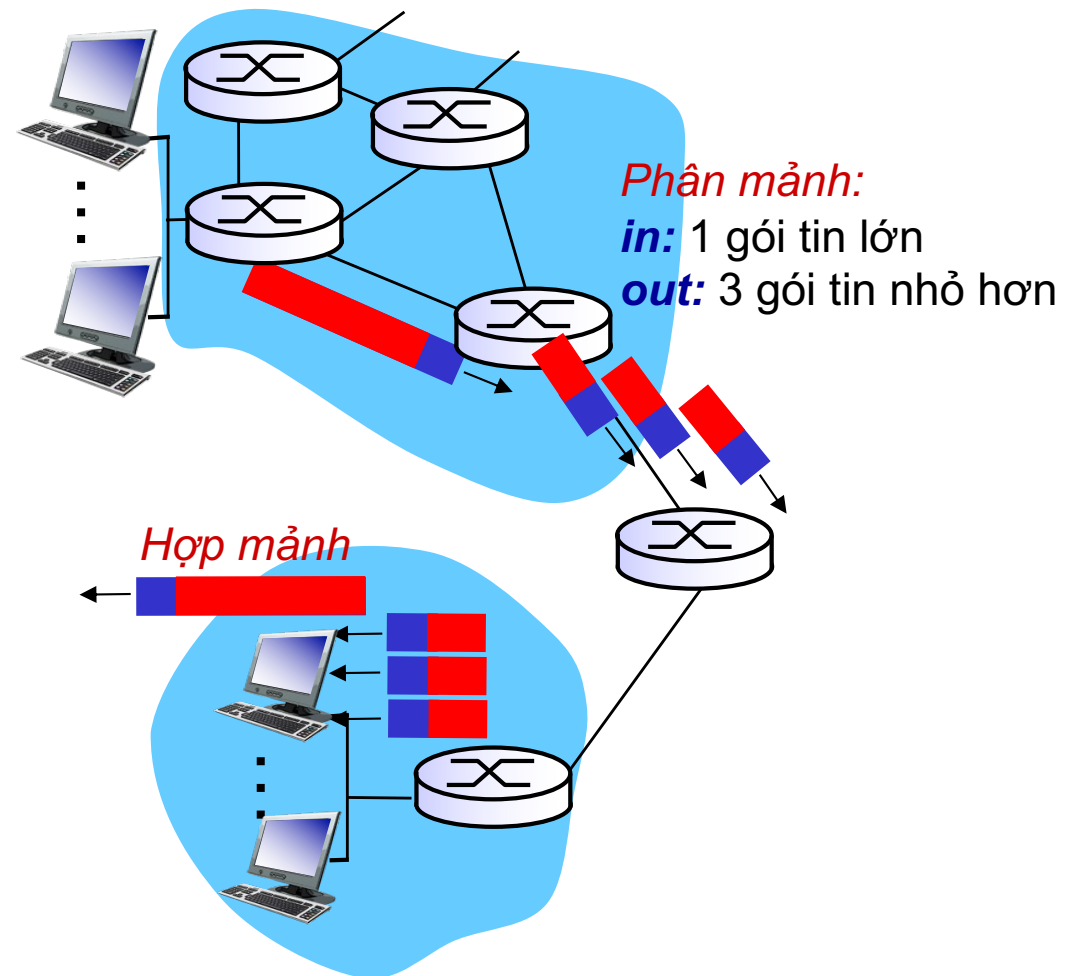
- TTL, 8 bits – Thời gian sống
 - Độ dài đường đi gói tin có thể đi qua
 - Max: 255
 - Router giảm TTL đi 1 đơn vị khi xử lý
 - Gói tin bị hủy nếu TTL bằng 0
- Upper protocol – giao thức tầng trên
 - Giao thức giao vận phía trên (TCP, UDP,...)
 - Các giao thức tầng mạng khác (ICMP, IGMP, OSPF) cũng có trường này
 - Sử dụng để dòn kênh/phân kênh

Tiêu đề gói tin IP (4)

- Checksum – Mã kiểm soát lỗi
- Địa chỉ IP nguồn
 - 32 bit, địa chỉ của trạm gửi
- Địa chỉ IP đích
 - 32 bit, địa chỉ của trạm đích

Phân mảnh gói tin (1)

- Đường truyền có một giá trị MTU (Kích thước đơn vị dữ liệu tối đa- Maximum Transmission Unit)
- Các đường truyền khác nhau có MTU khác nhau
- Một gói tin IP có kích thước lớn quá MTU sẽ bị
 - Chia làm nhiều gói tin nhỏ hơn
 - Được tập hợp lại tại trạm đích



Phân mảnh (2)

- Trường Identification
 - ID được sử dụng để tìm các phần của gói tin
- Flags – cờ (3 bits)
 - 1st bit: Dự phòng
 - 2nd bit: Không được phép phân mảnh
 - 3rd bit: Còn phân mảnh
- Độ lệch - Offset
 - Vị trí của gói tin phân mảnh trong gói tin ban đầu
 - Theo đơn vị 8 bytes

Phân mảnh (3)

Ví dụ:

- ❖ Gói tin: 4020 byte (header: 20 byte)
- ❖ MTU = 1500 bytes

	length	ID	fragflag	offset	
	=4020	=x	=0	=0	

1480 bytes in
data field

offset =
1480/8

	length	ID	fragflag	offset	
	=1500	=x	=1	=0	

	length	ID	fragflag	offset	
	=1500	=x	=1	=185	

	length	ID	fragflag	offset	
	=1060	=x	=0	=370	



ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

4.8. DHCP

Giới thiệu chung

- Dynamic Host Configuration Protocol
- Dịch vụ tầng ứng dụng cung cấp cấu hình địa chỉ IP cho các nút mạng:
 - Địa chỉ IP
 - Mặt nạ mạng
 - Địa chỉ bộ định tuyến mặc định (default router, default gateway)
 - Có thể địa chỉ máy chủ DNS phân giải tên miền
- Hoạt động theo mô hình client/server: client sử dụng các thông số địa chỉ IP do server cấp phát

DHCP : Cấp phát địa chỉ IP

Máy chủ DHCP



Các thông điệp DHCP

Client

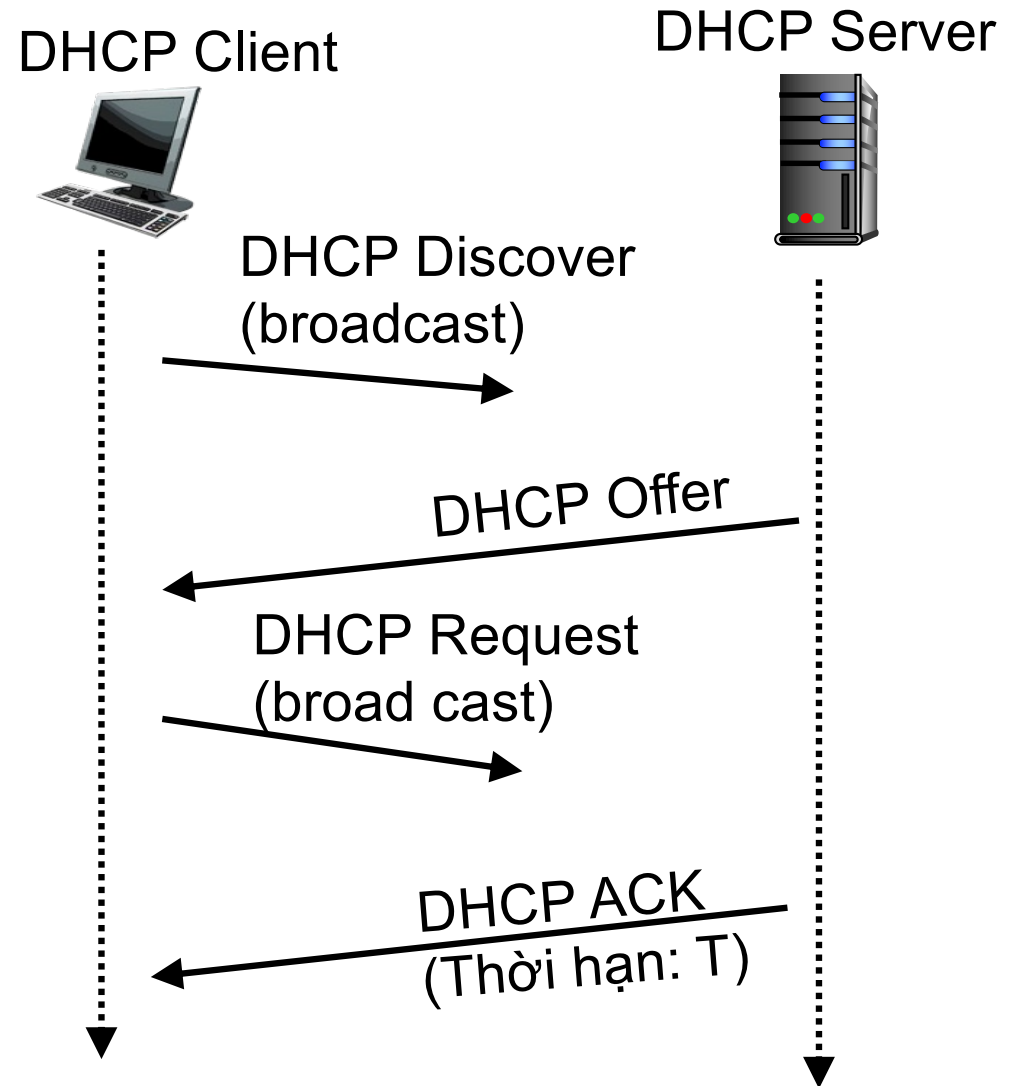
- DHCP Discover: tìm kiếm DHCP Server
- DHCP Request: đăng ký sử dụng địa chỉ IP
- DHCP Release: giải phóng địa chỉ IP đang sử dụng
- DHCP Decline: Từ chối địa chỉ IP được cấp phát

Server

- DHCP Offer: Cung cấp thông số cấu hình địa chỉ IP
- DHCP ACK: Chấp nhận đăng ký
- DHCP NAK: Từ chối đăng ký

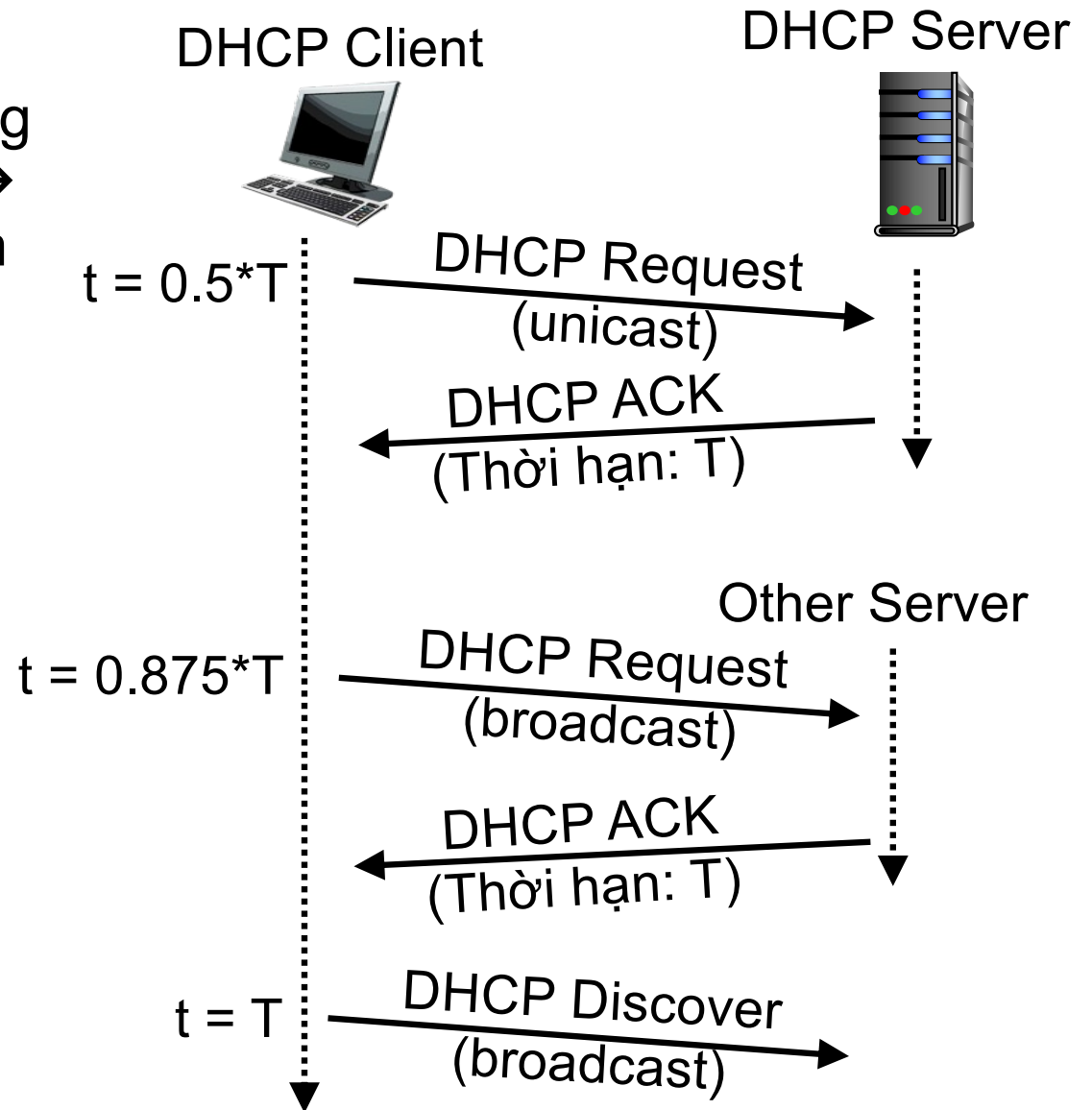
Cấp phát cấu hình mới

- B1: Client gửi thông điệp DHCP Discover quảng bá để tìm kiếm DHCP Server
- B2: Nếu có DHCP Server trong mạng, server gửi thông điệp DHCP Offer chứa các thông số địa chỉ
- B3: Client chọn một cấu hình từ các DHCP Offer nhận được và gửi thông điệp DHCP Request để đăng ký
- B4: DHCP Server gửi thông điệp DHCP ACK chấp nhận



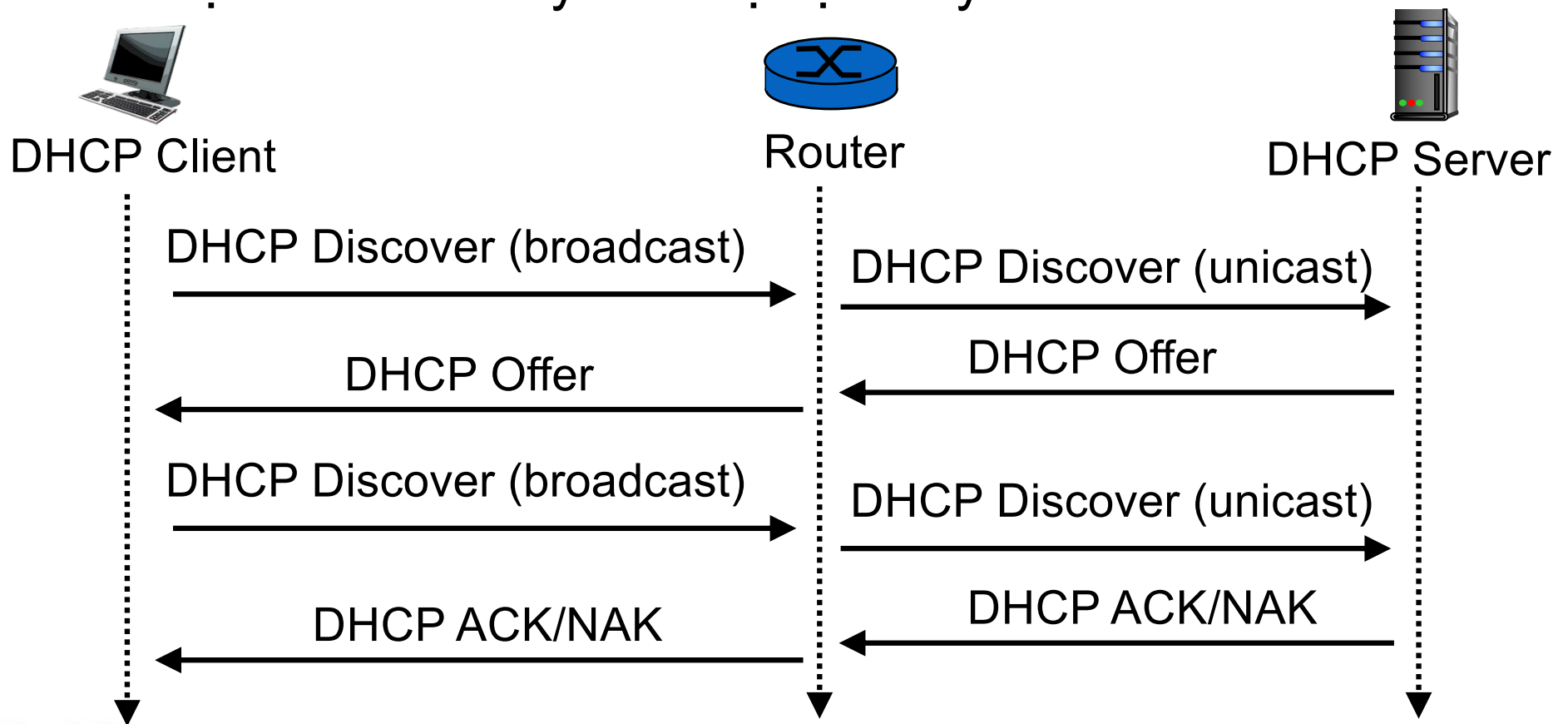
Gia hạn sử dụng

- Mỗi cấu hình được sử dụng trong khoảng thời gian T → client cần gia hạn thời gian sử dụng
- Khi $t = 0.5 * T$, client gửi DHCP Request tới DHCP Server để yêu cầu gia hạn
- Nếu không có DHCP ACK, khi $t = 0.875 * T$, client gửi DHCP Request quảng bá
- Nếu không có DHCP ACK, khi $t = T$, client gửi DHCP Discover



DHCP Relay

- DHCP Server nằm trên một mạng khác với client → các gói tin quảng bá không được bộ định tuyến chuyển tiếp
→ cài đặt DHCP Relay trên bộ định tuyến





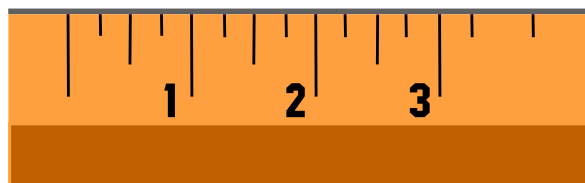
ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

4.8. Giới thiệu về IPv6

Giao thức IPv6

- Xuất phát từ nhu cầu thực tế: địa chỉ IPv4 cạn kiệt, không đủ để cấp phát
- Cải tiến trên IPv6:
 - Mở rộng không gian địa chỉ
 - Sử dụng địa chỉ có độ dài 128 bit
 - Phân vùng địa chỉ(scope)

Không gian địa chỉ IPv4



1mm

Không gian địa chỉ IPv6

84.000 lần đường kính của thiên hà

Cải tiến trên IPv6(tiếp)

- Tăng tốc độ xử lý
 - Khuôn dạng header đơn giản hơn
 - Không phân mảnh gói tin
- Hỗ trợ QoS tốt hơn
- Về an toàn an ninh
 - Sử dụng IPSec như một chuẩn
- Tự động cấu hình
 - Chuẩn hóa cơ chế tự động cấu hình



ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

Địa chỉ IPv6

Cách thức biểu diễn

- 128 bit, biểu diễn bởi số hệ 16
- Phân cách “:” giữa các nhóm gồm 4 số hexa
 - 3ffe:501:100c:e320:2e0:18ff:fe98:936d
- Bỏ qua chuỗi liên tiếp các số 0
 - 3ffe:501:100c:e320:0:0:0:1 →
3ffe:501:100c:e320::1
- Sử dụng ký hiệu mặt nạ mạng /n

Cấu trúc địa chỉ IPv6

- 2 phần: Network prefix và Interface ID
- Network prefix: 64 bit
- Host ID: 64 bit
 - Chuẩn EUI-64 (*extended unique identifier*)
 - Trong trường hợp mạng Ethernet, Host ID được xác định từ địa chỉ MAC

Phân vùng địa chỉ

- Địa chỉ toàn cục: Network prefix có
 - Global routing prefix: 48 bit với 3 bit đầu là 001
 - SubnetID: 16 bit
- Địa chỉ liên kết nội bộ: Network prefix có
 - 10 bit đầu là 1111 1110 10
 - 54 bit còn lại mang giá trị 0
- Địa chỉ cục bộ: Network prefix có
 - 10 bit đầu là 1111 1110 11
 - 38 bit kế tiếp mang giá trị 0
 - SubnetID: 16 bit

Các dạng địa chỉ IPv6

- Địa chỉ Unicast Address: gán cho một cổng giao tiếp mạng
- Địa chỉ Anycast Address: gán cho một nhóm cổng giao tiếp mạng
 - Gói tin gửi tới địa chỉ anycast addr. được chuyển tiếp cho nút gần nhất xác định bởi giao thức định tuyến
- Địa chỉ Multicast Address: gán cho một nhóm cổng giao tiếp mạng trong một scope
 - Bắt đầu bởi 1111 1111
 - Gói tin gửi tới địa chỉ multicast addr. được chuyển tới tất cả các nút trong nhóm



ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

Khuôn dạng gói tin IPv6

Khuôn dạng gói tin IPv6

Version (4 bit)	Traffic Class (8 bit)	Flow Label (20 bit)	
Payload Length (16 bit)		Next Header (8 bit)	Hop Limit (8 bit)
Source address (128bit)			
Destination address (128bit)			
Payload(gồm tiêu đề mở rộng nếu có và gói tin của giao thức tầng trên)			

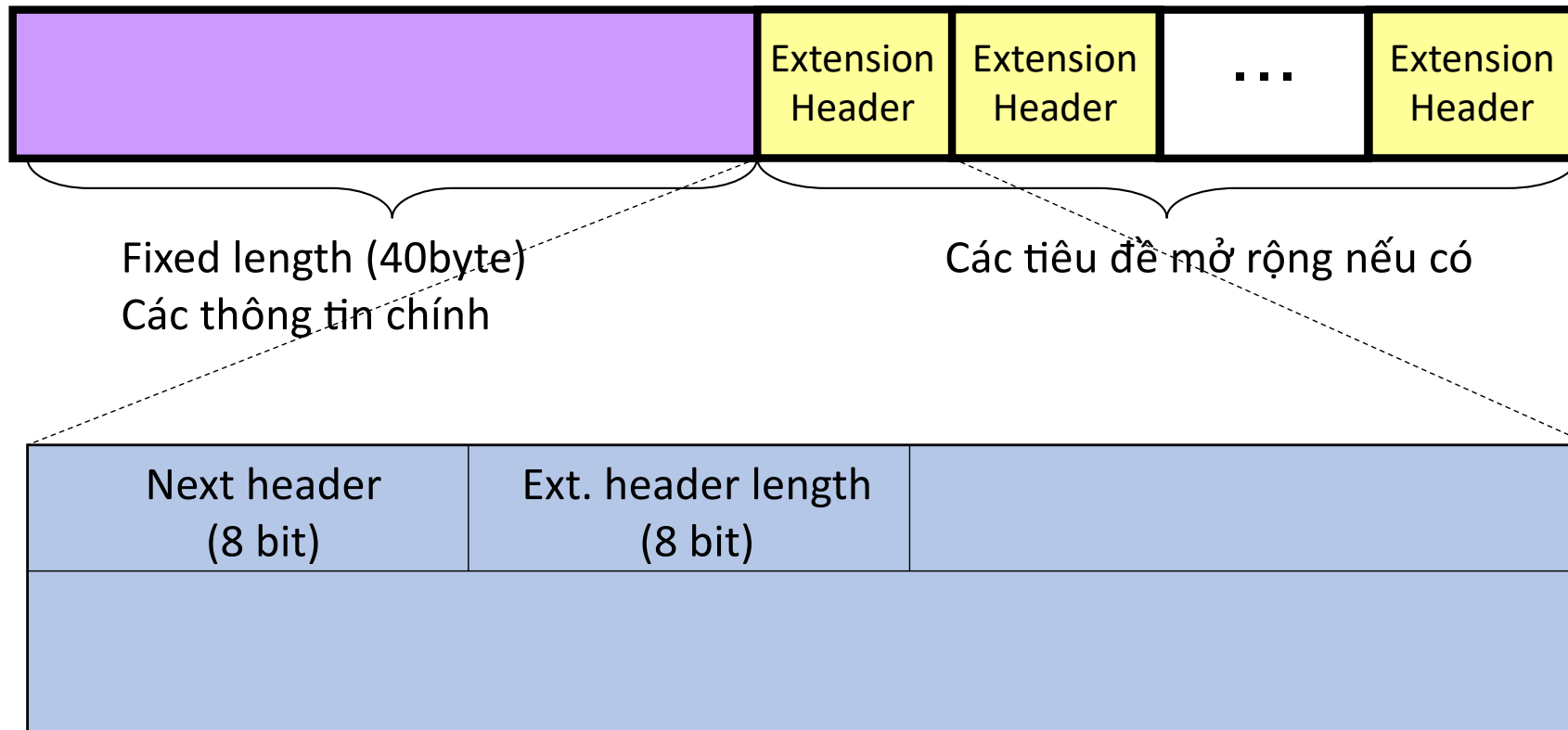
Các trường đổi tên từ IPv4

Tiêu đề gói tin IPv6 (1)

- Version: Phiên bản giao thức(=110)
- Traffic Class: số hiệu giao thức tầng trên
- Flow Label: Điều khiển QoS
- Payload Length: kích thước phần dữ liệu
- Next header: Tiêu đề tiếp theo mở rộng tiếp theo
- Hop limit: tương tự TTL

Tiêu đề gói tin IPv6 (2)

- Trong IPv6, tiêu đề mở rộng đặt trong phần payload





ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

Chọn đường là gì?

Các nguyên lý chọn đường

Cơ chế chuyển tiếp gói tin

Quy tắc “Longest matching”

Bộ định tuyến?

- Thiết bị chuyển tiếp các gói tin giữa các mạng
 - Là một máy tính, với các phần cứng chuyên dụng
 - Kết nối nhiều mạng với nhau
 - Chuyển tiếp gói tin dựa trên bảng chọn đường
- Có nhiều giao diện
- Phù hợp với nhiều dạng lưu lượng và phạm vi của mạng

Một số ví dụ...



BUFFALO
BHR-4RV



PLANEX
GW-AP54SAG



YAMAHA
RTX-1500



Cisco 2600

Router ngoại vi



Cisco CRS-1

Router mạng trục



Hitachi
GR2000-1B

Juniper M10



Foundry Networks
NetIron 800



Cisco 3700

Router cỡ trung

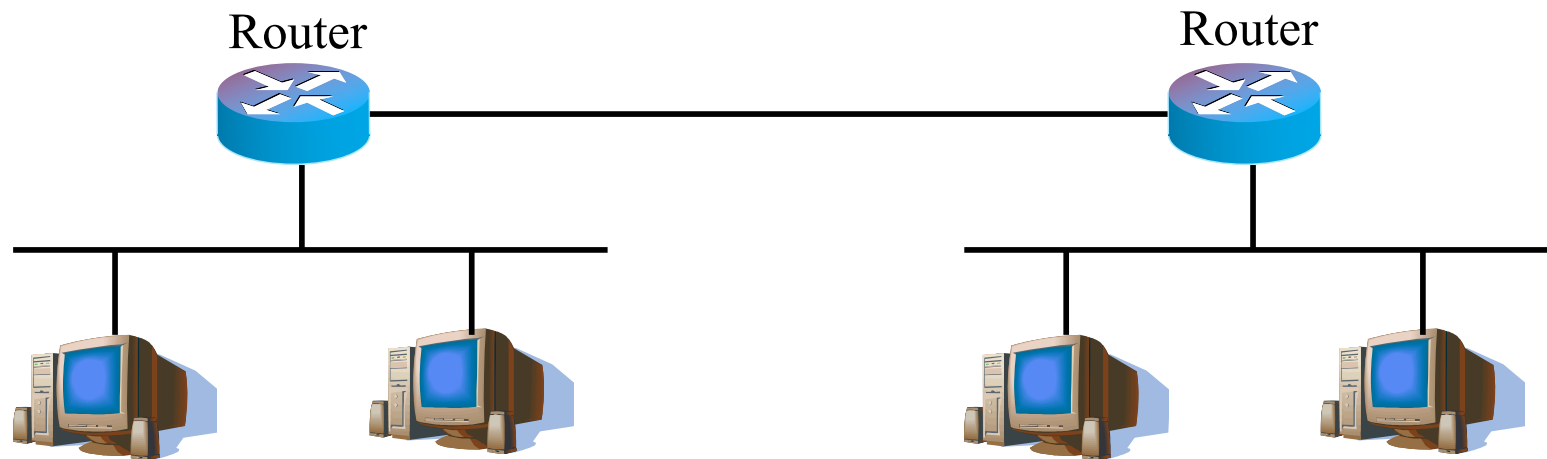
<http://www.cisco.com.vn>

<http://www.juniper.net/>

<http://www.buffalotech.com>

Nguyên tắc chuyển tiếp gói tin IP

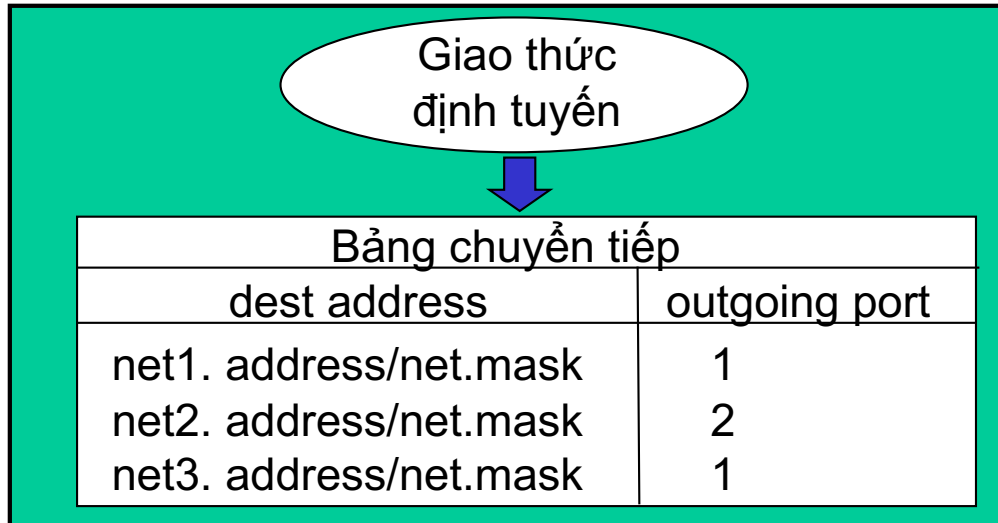
- Khi một máy trạm gửi một gói tin IP tới một máy khác
 - Nếu địa chỉ đích và địa chỉ nguồn thuộc cùng một mạng: Chuyển trực tiếp bằng giao thức tầng 2
 - Nếu địa chỉ đích nằm trên một mạng khác: Chuyển gián tiếp qua bộ định tuyến (chọn đường)



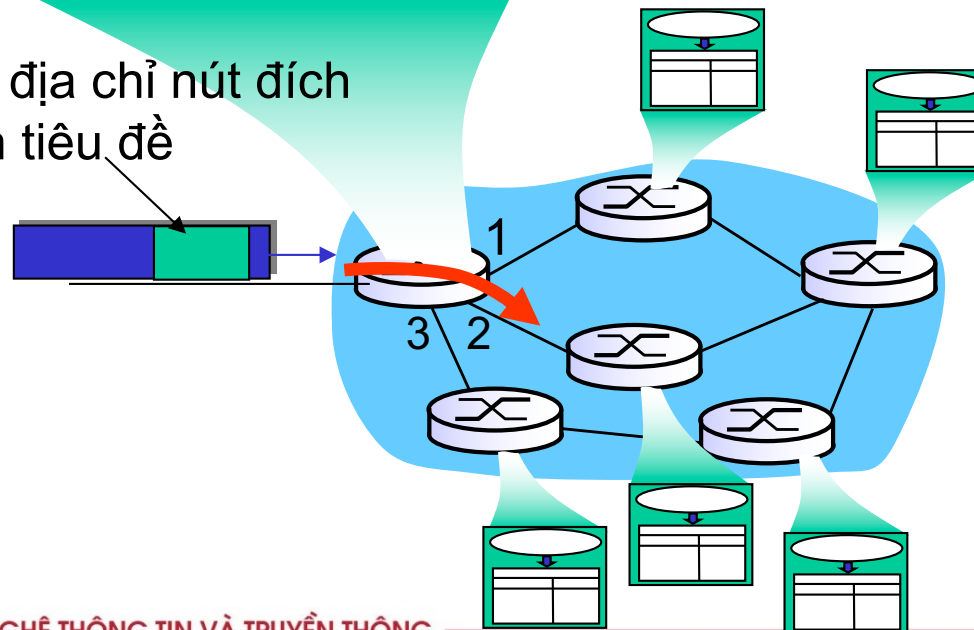
Chuyển tiếp gói tin IP

- Mỗi nút mạng sử dụng bảng chuyển tiếp (Forwarding Table)
 - Là một phần của bảng định tuyến (Routing Table)
- Các thông tin:
 - Đích đến (Destination): Địa chỉ mạng/Mặt nạ (/n)
 - Cổng ra (Outgoing port): địa chỉ của cổng ra trên router để chuyển tới nút kế tiếp trong đường đi
 - Đích đến mặc định: Sử dụng địa chỉ 0.0.0.0/0 đại diện cho một đích bất kỳ chưa biết

Bảng chuyển tiếp



Gói tin với địa chỉ nút đích trong phần tiêu đề



Ví dụ - Bảng chuyển tiếp trên máy trạm

C:\Documents and Settings\tungbt>netstat -r
Route Table

=====

Interface List

0x1 ... **Destination**MS TCP Loopback interface

0x2 ...06 00 11 b2 a1 a5 Realtek RTL8159 Family PCI Fast Ethernet NIC -

=====

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

Ví dụ - Bảng chuyển tiếp trên router (rút gọn)

Destination

Outgoing port

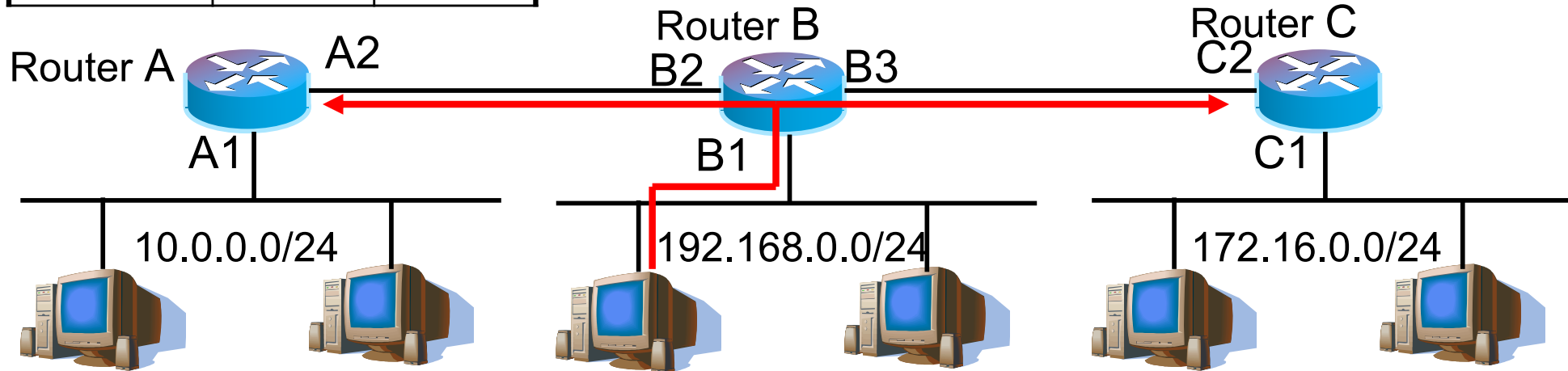
Router# show ip route

```
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
```

Bảng chọn đường và cơ chế chuyển tiếp (1)

Network	Next-hop	Interface
10.0.0.0/24	Direct	A1
172.16.0.0/24	B2	A2
192.168.0.0/24	B2	A2

Network	Next-hop	Interface
10.0.0.0/24	B3	C2
172.16.0.0/24	Direct	C1
192.168.0.0/24	B3	C2



Network	Next-hop	Inf
10.0.0.0/24	A2	B2
172.16.0.0/24	C2	B3
192.168.0.0/24	Direct	B1

Lưu ý quy tắc: **No routes, no reachability!**

Nguyên tắc chọn cổng ra

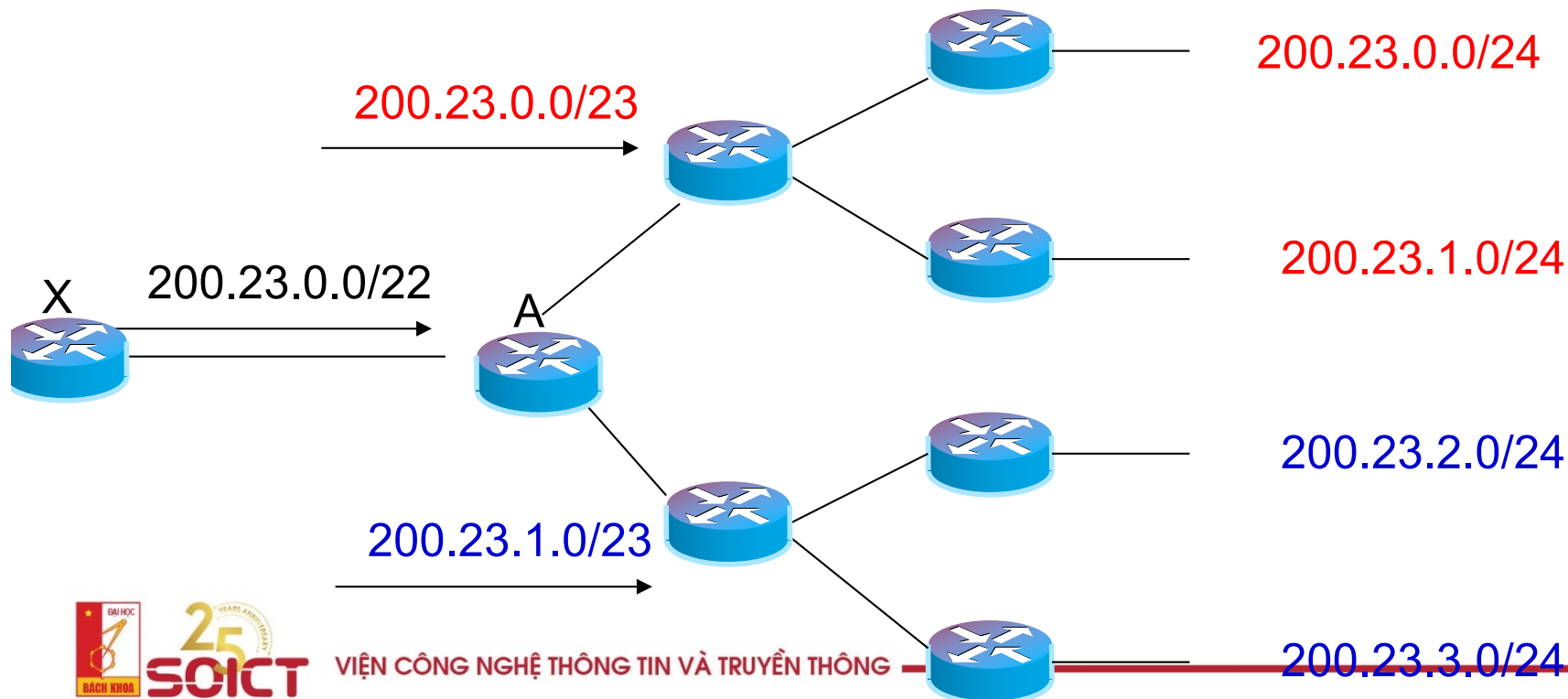
- Khi router nhận được 1 gói tin: ...
- So sánh n bit đầu tiên trên địa chỉ đích gói tin với các bit tương ứng trên địa chỉ mạng đích
 - /n: Mặt nạ mạng đích
 - Nếu có mạng đích khớp chuyển ra cổng tương ứng
 - Nếu không có mạng đích nào khớp, chuyển ra cổng mặc định tương ứng với địa chỉ 0.0.0.0/0 (nếu có)
- Quy tắc “longest matching”: nếu có nhiều mạng đích thỏa mãn, chuyển tiếp tới mạng đích có mặt nạ lớn nhất

Địa chỉ đích của gói tin:
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

Kết hợp đường đi (Routing aggregation)

- Có bao nhiêu mạng con trên mạng Internet?
- Sẽ có rất nhiều mục trong bảng chọn đường?
- Các mạng con kế tiếp với cùng địa chỉ đích có thể được tổng hợp lại để làm giảm số mục trong bảng chọn đường.



Quy tắc “Longest matching”

Địa chỉ đích:

11.1.2.5 = 00001011.00000001.00000010.00000101

11.1.3.6 = 00001011.00000001.00000011.00000110

Đường đi 1:

11.1.2.0/24 = 00001011.00000001.00000010.00000000

Đường đi 2:

11.1.0.0/16 = 00001011.00000001.00000000.00000000

Đường đi 3:

11.0.0.0/8 = 00001011.00000000.00000000.00000000

Tại sao phải cần quy tắc này?

Chuyển tiếp gói tin trên router

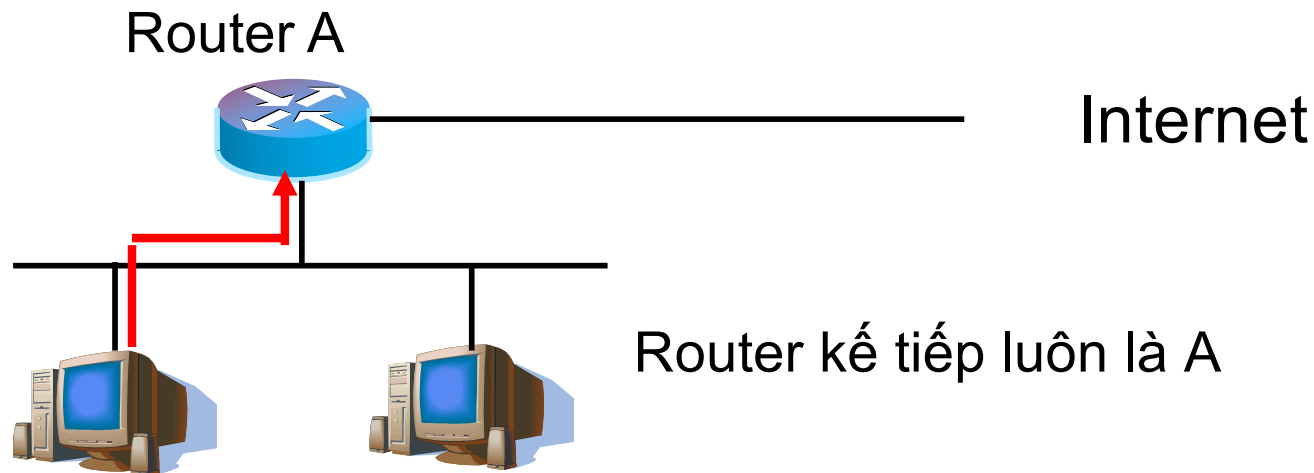
- B1 : Nếu TTL = 1 (hoặc TTL = 0), hủy gói tin và báo lỗi. Kết thúc.
- B2 : Nếu TTL > 1, lấy địa chỉ đích của gói tin. Mặt nạ hóa địa chỉ đích của gói tin với các mặt nạ của mạng đích trong bảng chuyển tiếp.
- B3 : So sánh kết quả mặt nạ hóa với địa chỉ mạng đích tương ứng. Nếu có mạng đích khớp chuyển tới cổng ra tương ứng, giảm TTL.
- B4 : Nếu không có mạng đích khớp, kiểm tra đường đi mặc định (tương ứng với đích 0.0.0.0 /0)
 - Có cổng mặc định : chuyển gói tin tới cổng mặc định, giảm TTL
 - Không có : hủy gói tin, báo lỗi.

Kết hợp đường đi (2)

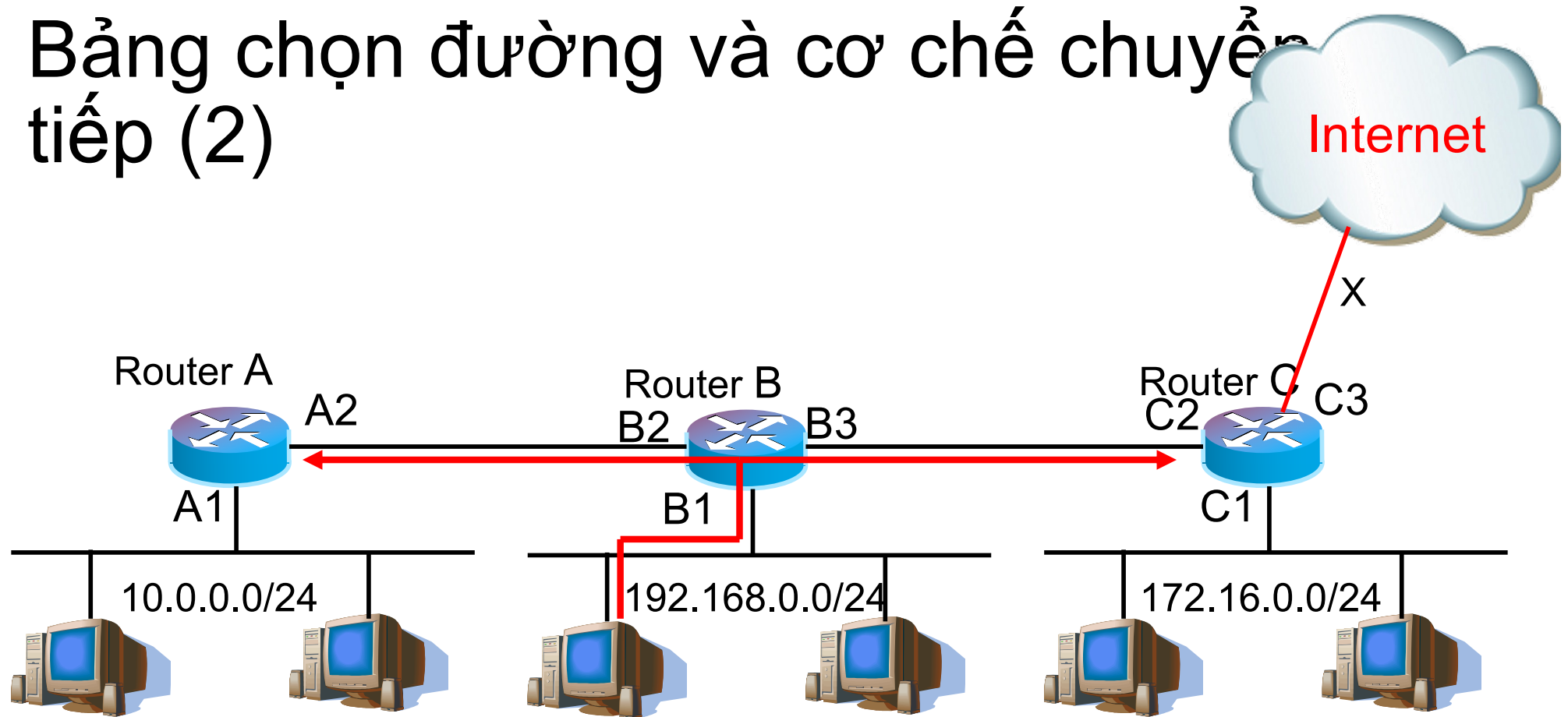
- Ví dụ về Viettel
 - Không gian địa chỉ IP: khá lớn
 - 203.113.128.0 - 203.113.191.255
 - Để kết nối đến một mạng con của Vietel (khách hàng):
Chỉ cần chỉ ra đường đi đến mạng Viettel
- Đường đi mặc định chính là một dạng của việc kết hợp đường
 - 0.0.0.0/0

Đường đi mặc định

- Đường đi mặc định được sử dụng khi không tìm thấy đường đi phù hợp trong bảng chọn đường
 - Đường đi mặc định cũng trở đến một router kết tiếp
 - Trong nhiều trường hợp, đây là đường đi duy nhất
- 0.0.0.0/0
 - Là một trường hợp đặc biệt, chỉ tất cả các đường đi



Bảng chọn đường và cơ chế chuyển tiếp (2)



Network	Next-hop	Inf
10.0.0.0/24	A2	B2
172.16.0.0/24	C2	B3
192.168.0.0/24	Direct	B1
0.0.0.0/0	C2	B3

Q. Mô tả bảng chọn đường trên C

Nếu C nối vào Internet?



ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

4.6. Giao thức ICMP

Giới thiệu chung về ICMP

- Internet Control Message Protocol
- IP là giao thức không tin cậy, hướng không liên kết
 - Thiếu các cơ chế hỗ trợ và kiểm soát lỗi → Phía gửi không biết gói tin có được truyền thành công
- ICMP được sử dụng ở tầng mạng để trao đổi thông tin:
2 chức năng chính
 - Báo lỗi khi chuyển tiếp một gói tin IP không thành công
 - Truy vấn thông tin trạng thái
- Gói tin ICMP được đóng gói vào trong gói tin IP:



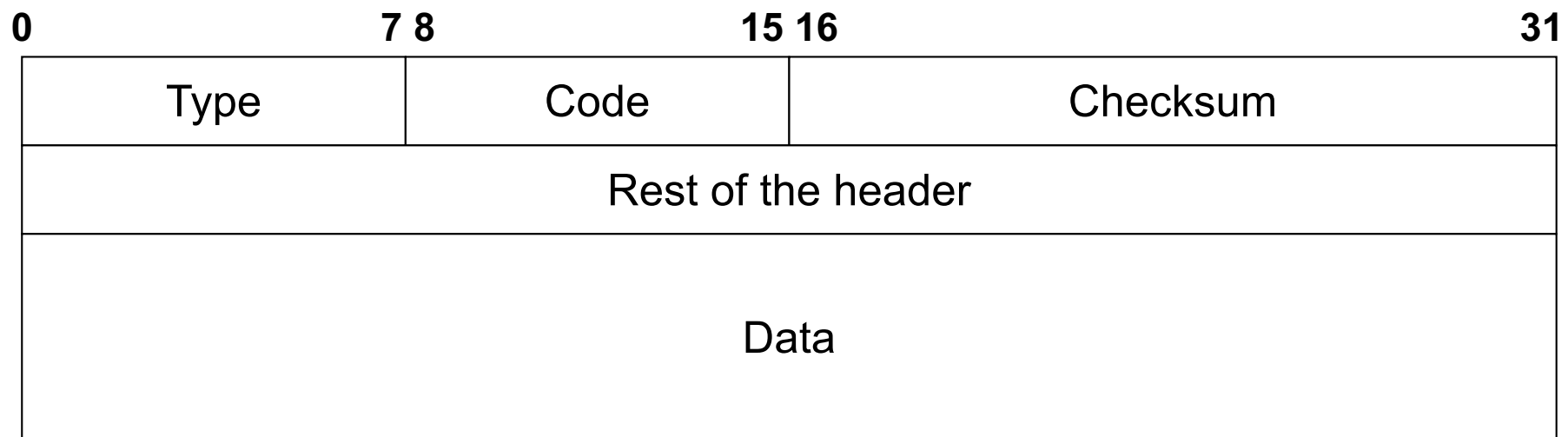
IP header and Protocol field

Ver	HLEN	DS	Total Length	
Identification			Flags	Fragmentation offset
TTL	Protocol		Header Checksum	
Source IP address				
Destination IP address				
Option				
Data gói tin ICMP				

Protocol:
1: ICMP
2: IGMP
6: TCP
17: UDP
89: OSPF

Khuôn dạng gói tin ICMP

- Type: dạng gói tin ICMP
- Code: Nguyên nhân gây lỗi
- Checksum
- Rest of header: Mỗi dạng có phần còn lại tương ứng, phụ thuộc vào Type và Code ()



Một số dạng gói tin ICMP

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

ICMP và các công cụ debug

- ICMP luôn hoạt động song song trong suốt với người dùng
- Có thể sử dụng ICMP thông qua các công cụ debug. Ví dụ:
 - ping
 - traceroute

ping

- Tác dụng: Kiểm tra trạng thái kết nối với một nút mạng
- Cú pháp: ping *host_address*
 - *host_address* là địa chỉ IP hoặc tên miền
- Hoạt động
 - Bên gửi: Gửi gói tin ICMP Echo Request (Type = 8)
 - Bên nhận: Trả gói tin ICMP Echo Reply (Type = 0)
- Trường dữ liệu chứa thời gian gửi gói tin
 - Khi nhận được phản hồi thì tính được trễ khứ hồi (RTT - Round trip time)

Ping: Ví dụ

```
C:\Documents and Settings\admin>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:
```

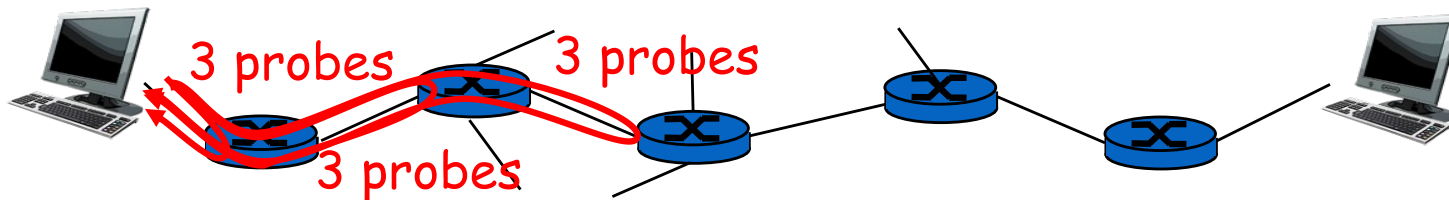
```
    Packets: Sent = 4, Received = 4, Lost = 0 (0% loss),
```

```
Approximate round trip times in milli-seconds:
```

```
    Minimum = 529ms, Maximum = 600ms, Average = 556ms
```

traceroute

- Cú pháp:
 - Windows: `tracert host_address`
 - Linux: `traceroute host_address`
- Bên gửi truyền gói tin cho bên nhận
 - Lượt thứ n có TTL = n
- Khi gói tin lượt gửi thứ n đến router thứ n :
 - Router hủy gói tin
 - Gửi một gói tin ICMP báo lỗi “Time exceeded” (Type = 11)
 - Dữ liệu có chứa tên và địa chỉ IP của router
- Khi nhận được gói tin trả lời, bên gửi sẽ tính ra RTT



traceroute

Điều kiện kết thúc

- Gói tin đến được đích
- Đích trả về gói tin ICMP báo lỗi “Time exceeded”
- Khi nguồn nhận được gói tin ICMP này sẽ dừng lại

Traceroute: Ví dụ

```
C:\Documents and Settings\admin>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  395 ms  399 ms  417 ms  ve-4.foundry3.nezu.wide.ad.jp [203.178.138.244]
11  355 ms  356 ms  378 ms  ve-3705.cisco2.komatsu.wide.ad.jp [203.178.136.193]
12  388 ms  398 ms  414 ms  c76.jaist.ac.jp [203.178.138.174]
13  438 ms  377 ms  435 ms  www.jaist.ac.jp [150.65.5.208]
```

Trace complete.



ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

4.7. Giao thức ARP

Địa chỉ MAC và ARP

- Address Resolution Protocol
- Tìm địa chỉ MAC (định danh tầng liên kết dữ liệu) của một nút mạng khi đã biết địa chỉ IP
- Tại sao cần ARP?
 - Truyền tin trên tầng mạng dùng địa chỉ IP
 - Truyền tin trên tầng liên kết dữ liệu dùng địa chỉ MAC
 - Khi gửi: dữ liệu chuyển từ tầng mạng xuống tầng liên kết dữ liệu.
 - Dữ liệu gửi trong mạng LAN: Máy nguồn cần phải biết địa chỉ MAC của máy đích
 - Dữ liệu gửi ra ngoài mạng LAN: Máy nguồn phải biết địa chỉ MAC của bộ định tuyến mặc định

Hoạt động của ARP

- Mỗi nút trong sử dụng bảng ARP Table:
 - Ánh xạ <Địa chỉ IP, Địa chỉ MAC, TTL>
 - TTL: Thời gian giữ ánh xạ trong bảng
- Khi tìm địa chỉ MAC của một nút khác, nút mạng gửi quảng bá gói tin ARP Request lên trên mạng.
 - Gói tin ARP Request chứa địa chỉ IP của nút cần tìm kiếm
- Nút mạng mang địa chỉ IP được hỏi sẽ gửi ARP Reply trả lời

Hoạt động của ARP – Ví dụ

ARP Request

(broadcast)

Who has 10.0.0.2?



Host B: 10.0.0.2
MAC: 12-29-9C-E8-FF-12



Host C: 10.0.0.3
MAC: 13-29-9C-E8-FF-13

Host A: 10.0.0.1

MAC: 11-29-9C-E8-FF-11

ARP Request

(broadcast)

Who has 10.0.0.2?



Host B: 10.0.0.2
MAC: 12-29-9C-E8-FF-12



Host C: 10.0.0.3
MAC: 13-29-9C-E8-FF-13

Host A: 10.0.0.1

MAC: 11-29-9C-E8-FF-11

ARP Request

(broadcast)

Who has 10.0.0.2?



Host B: 10.0.0.2
MAC: 12-29-9C-E8-FF-12



Host C: 10.0.0.3
MAC: 13-29-9C-E8-FF-13

Host A: 10.0.0.1

MAC: 11-29-9C-E8-FF-11

ARP Reply (unicast)

10.0.0.2 is at 12-29-9C-E8-FF-12



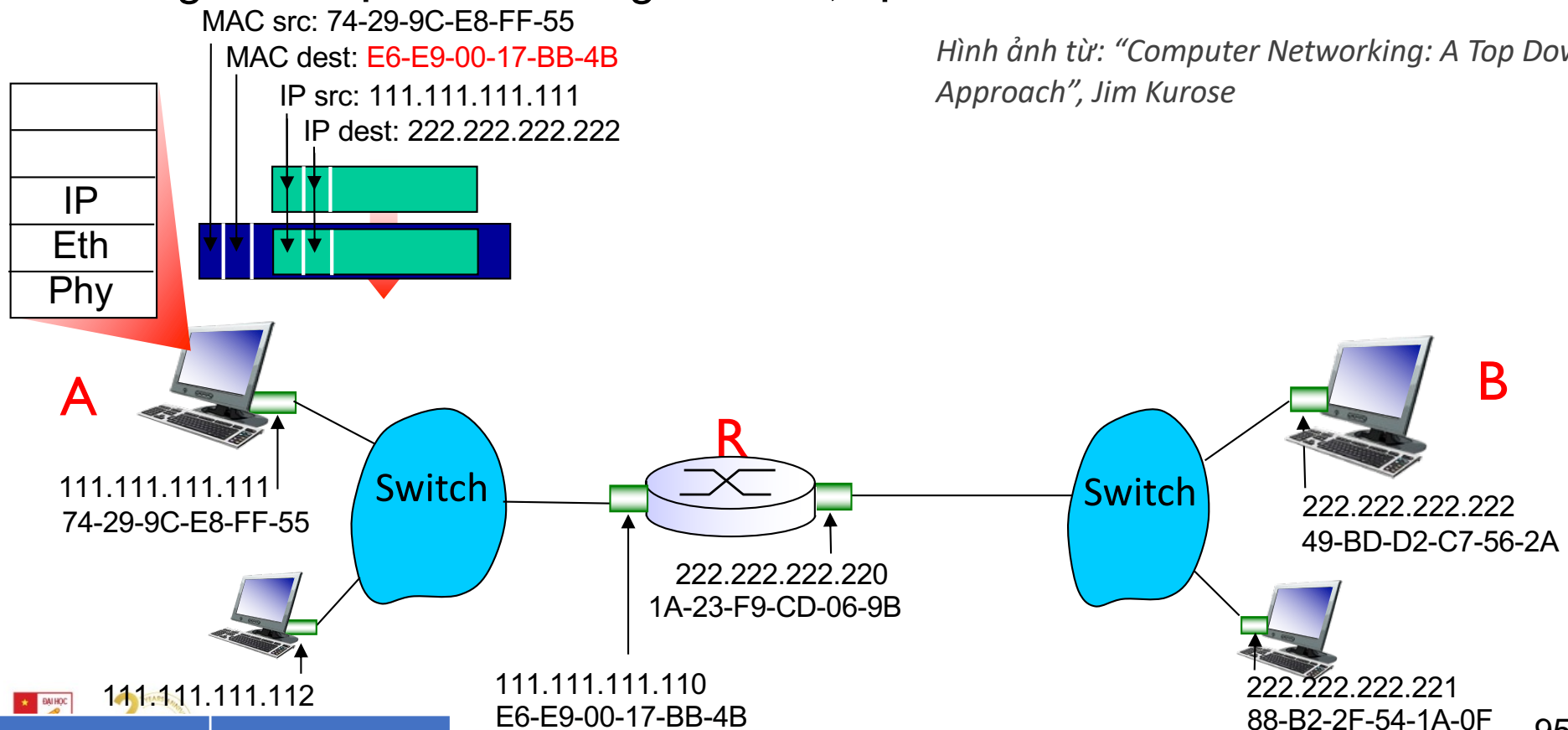
SOICT

Chuyển tiếp dữ liệu tới LAN khác

Ví dụ: **Gửi dữ liệu từ A tới B qua router R**

- A soạn một gói tin IP với địa chỉ nguồn là A, địa chỉ đích là B
- Gói tin chuyển xuống tầng liên kết dữ liệu: đóng gói thành khung tin tầng 2 với địa chỉ MAC nguồn là A, địa chỉ MAC đích là R

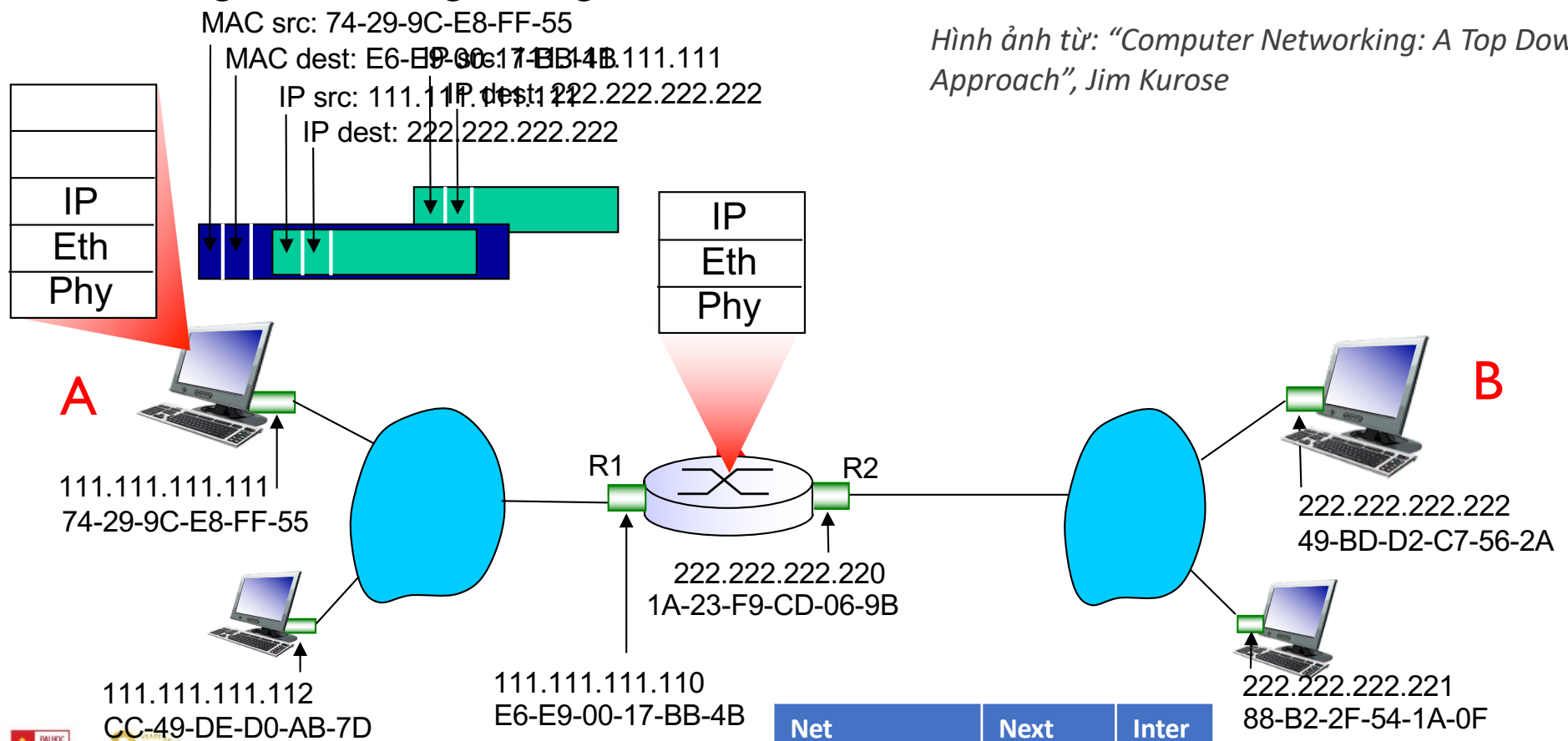
Hình ảnh từ: "Computer Networking: A Top Down Approach", Jim Kurose



Net	Next hop
222.222.222.0	111.111.111.110

Chuyển tiếp dữ liệu tới LAN khác

- ❖ Khung tin được chuyển từ A tới R
- ❖ Tại R: khung tin được bóc bỏ header, chuyển lên cho tầng mạng dưới dạng một gói tin IP



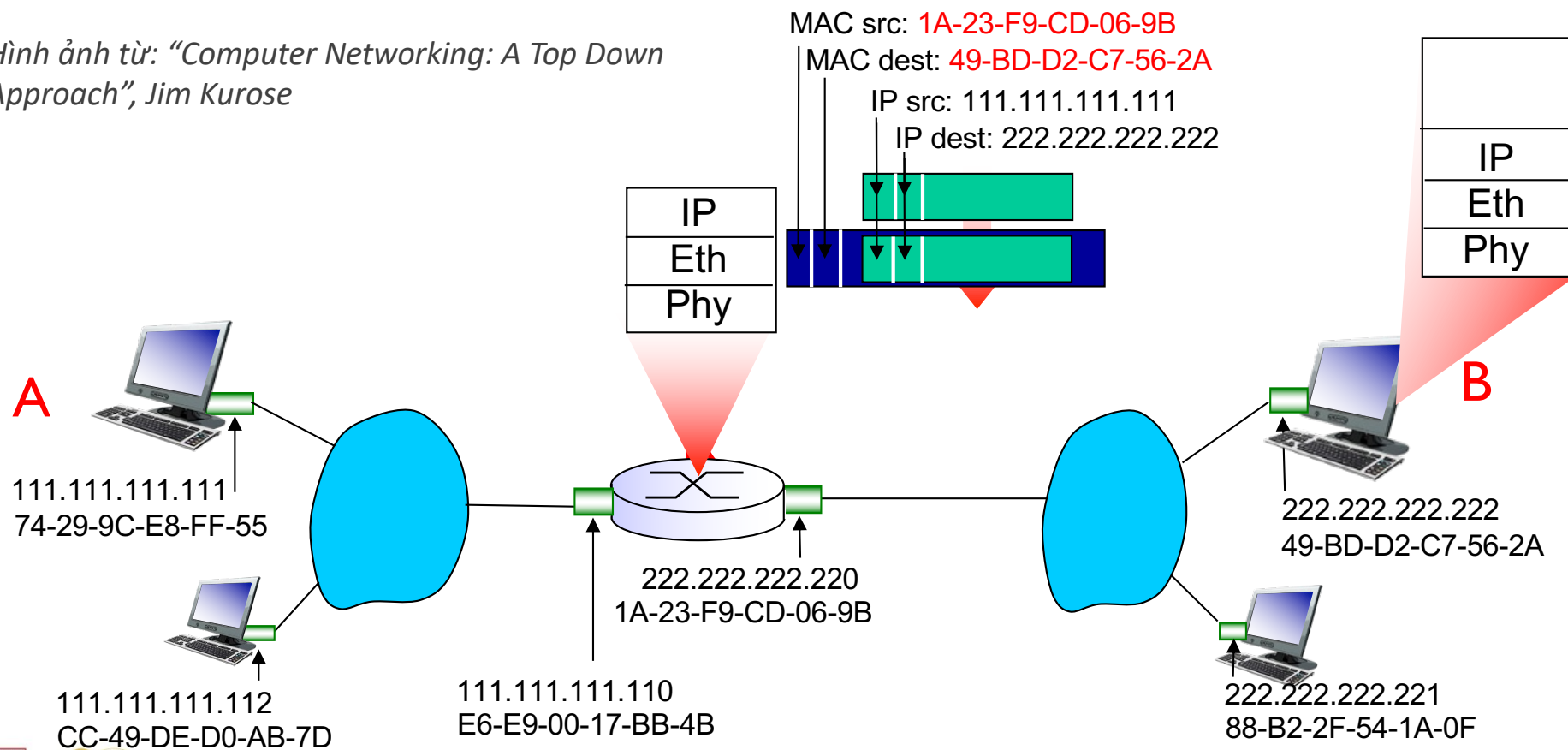
Hình ảnh từ: "Computer Networking: A Top Down Approach", Jim Kurose

Net	Next hop	Inter face
222.222.222.0	Diirect	R2

Chuyển tiếp dữ liệu tới LAN khác

- ❖ R chuyển tiếp gói tin với địa chỉ IP nguồn là A, IP đích là B
- ❖ Gói tin chuyển xuống tầng liên kết dữ liệu: đóng gói thành khung tin tầng 2 với địa chỉ MAC nguồn là R, địa chỉ MAC đích là B

Hình ảnh từ: "Computer Networking: A Top Down Approach", Jim Kurose





ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

Chọn đường tĩnh và chọn đường động

Chọn đường tĩnh

Chọn đường động

Ưu điểm – nhược điểm

Vấn đề cập nhật bảng chọn đường

- Sự thay đổi cấu trúc mạng: thêm mạng mới, một nút mạng bị mất điện
- Sự cần thiết phải cập nhật bảng chọn đường
 - Cho tất cả các nút mạng (về lý thuyết)
 - Thực tế, chỉ một số nút mạng phải cập nhật

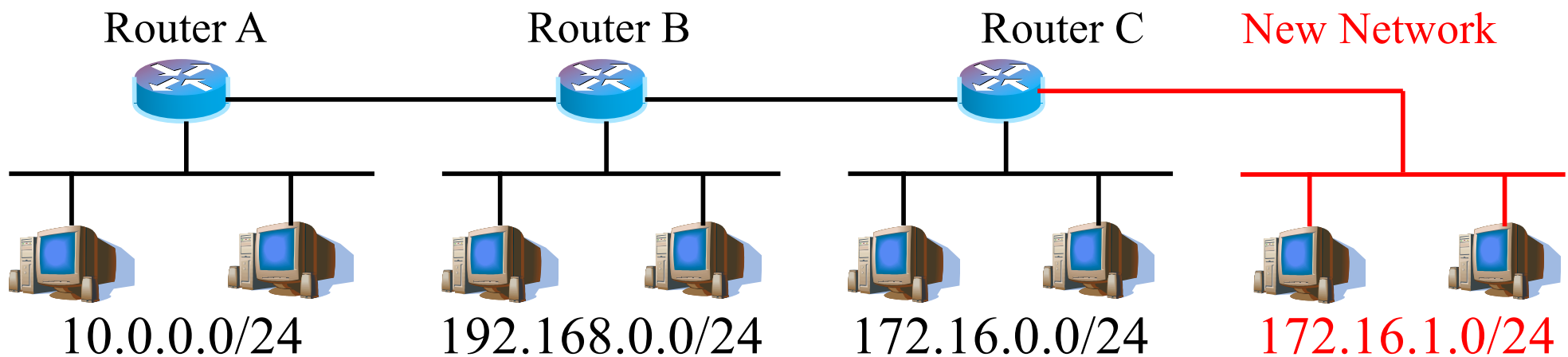
Network	Next-hop
192.168.0.0/24	B
172.16.0.0/24	B

172.16.1.0/24 B

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

172.16.1.0/24 C

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



Làm thế nào để cập nhật?

- Chọn đường tĩnh
 - Các mục trong bảng chọn đường được sửa đổi thủ công bởi người quản trị
- Chọn đường động
 - Tự động cập nhật bảng chọn đường
 - Bảng các giao thức chọn đường

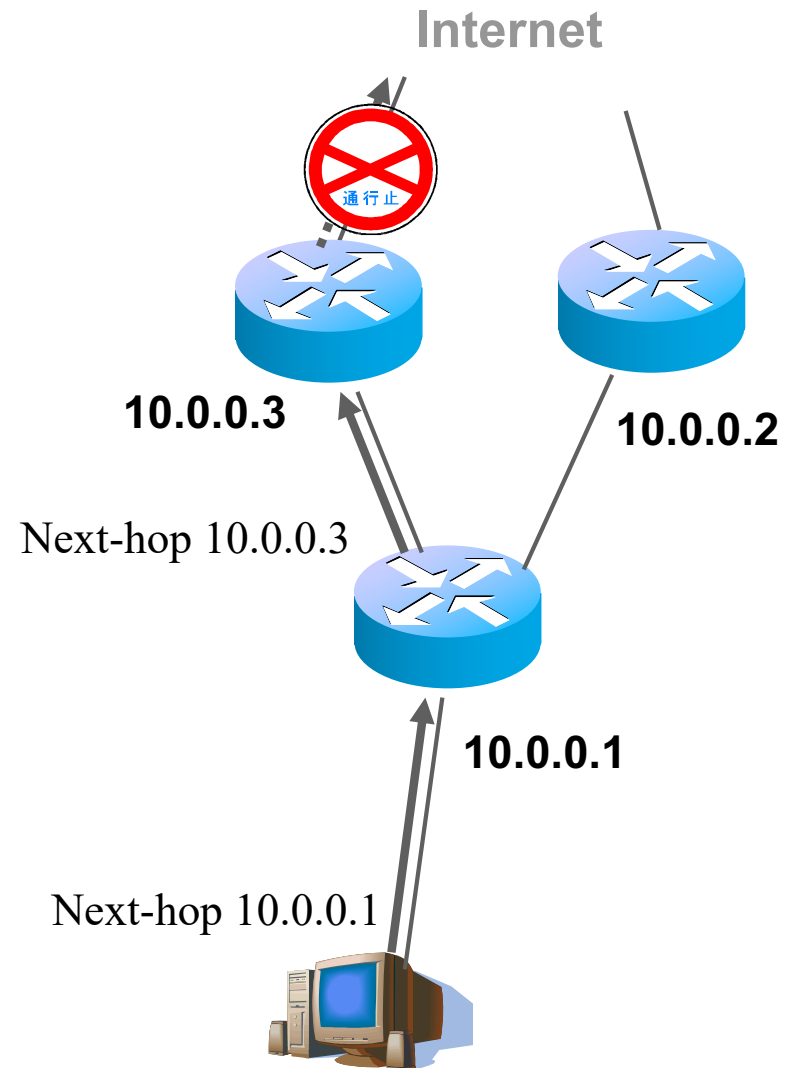
Chọn đường tĩnh

- Khi có sự cố:
 - Không thể nối vào Internet kể cả khi có tồn tại đường đi dự phòng
 - Người quản trị mạng cần thay đổi

Bảng chọn đường của 10.0.0.1 (1 phần)

Prefix	Next-hop
0.0.0.0/0	10.0.0.3

Kết nối bị lỗi



Chọn đường động

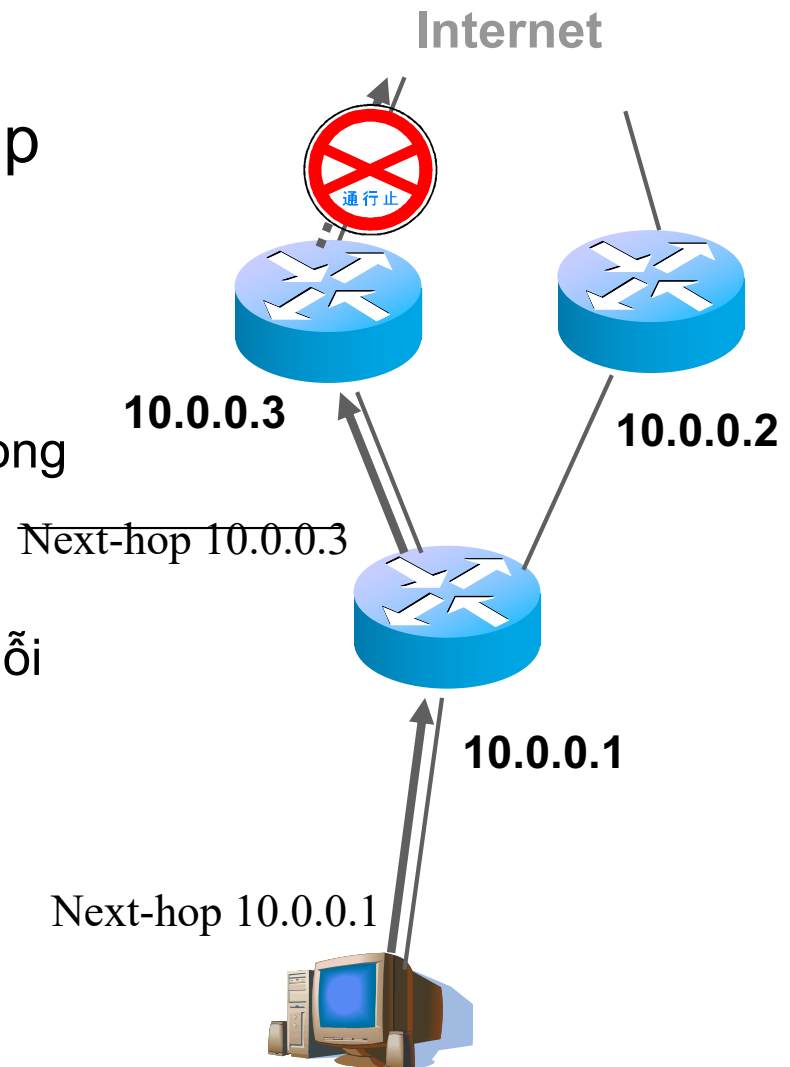
- Khi có sự cố:
 - Đường đi thay thế được cập nhật một cách tự động

Bảng chọn đường của 10.0.0.1 (1 phần)

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

Kết nối dự phòng

Kết nối bị lỗi



Đặc điểm của chọn đường tĩnh

- Ưu điểm
 - Ổn định
 - An toàn
 - Không bị ảnh hưởng bởi các yếu tố tác động
- Nhược điểm
 - Cứng nhắc
 - Không thể sử dụng tự động kết nối dự phòng
 - Khó quản lý

Chọn đường động

- Ưu
 - Dễ quản lý
 - Tự động sử dụng kết nối dự phòng
- Nhược
 - Tính an toàn
 - Các giao thức chọn đường phức tạp và khó hiểu
 - Khó quản lý



ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

Các giải thuật và giao thức chọn đường

Giải thuật Dijkstra và Bellman-Ford

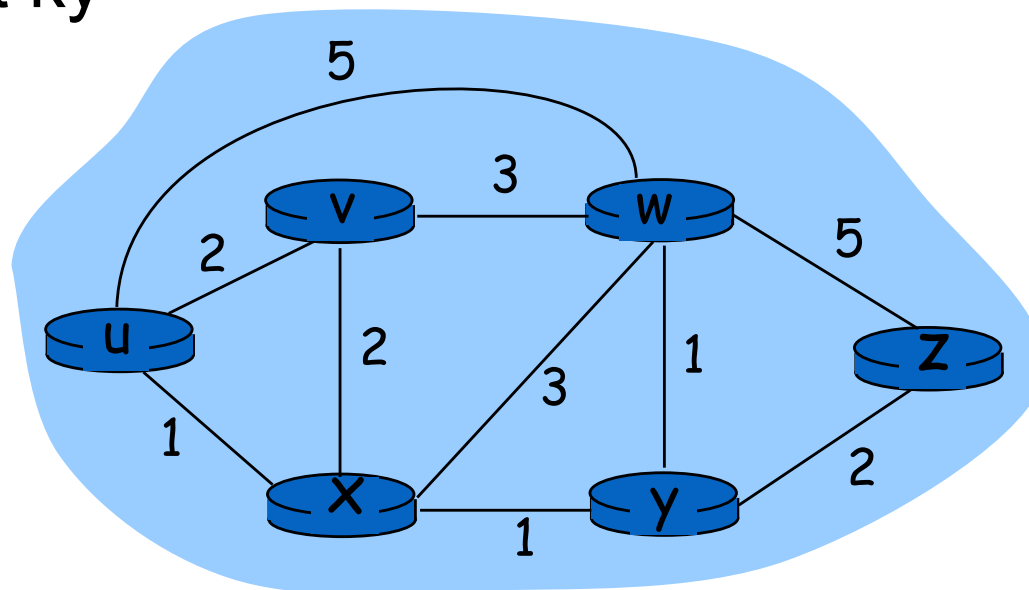
Giao thức dạng link-state và dạng
distance-vector

Vai trò của giải thuật và giao thức chọn đường

- Cần thiết phải xây dựng bảng chọn đường một cách tự động tại các router
- Giải thuật chọn đường thực hiện tính toán đường đi
- Trong hoặc trước quá trình tính toán cần thu thập topo hoặc trao đổi thông tin giữa các nút mạng
 - → giao thức chọn đường thực hiện việc trao đổi này.
 - Giao thức chọn đường còn quy định các quy trình duy trì bảng chọn đường.

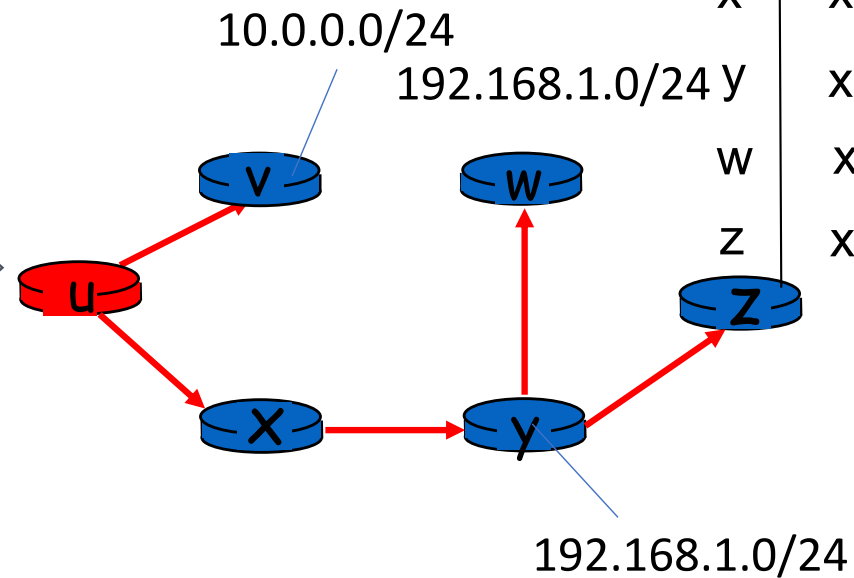
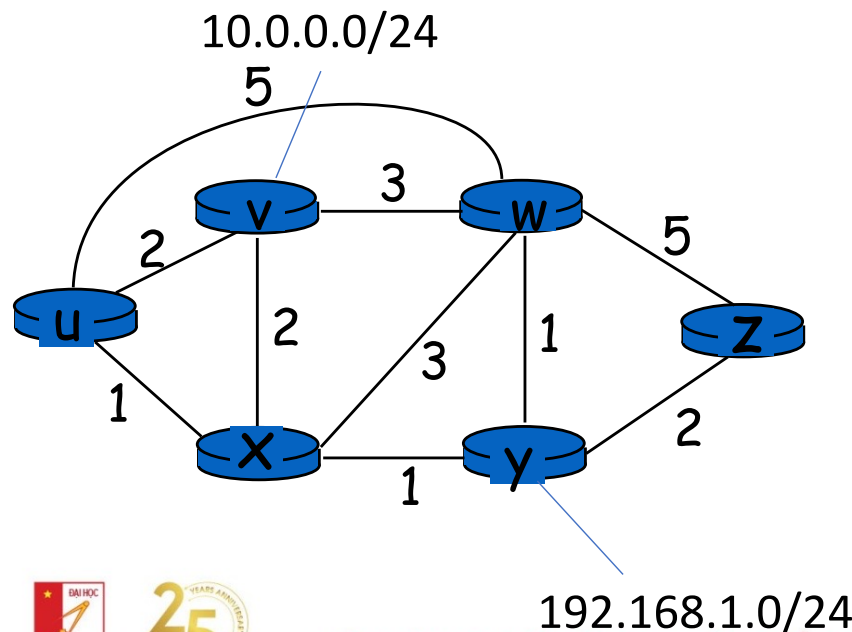
Biểu diễn mạng bởi đồ thị

- Đồ thị với các nút (bộ định tuyến) và các cạnh (liên kết)
- Chi phí cho việc sử dụng mỗi liên kết $c(x,y)$
 - Băng thông, độ trễ, chi phí, mức độ tắc nghẽn...
- Giải thuật chọn đường: Xác định đường đi ngắn nhất giữa hai nút bất kỳ



Cây đường đi ngắn nhất - SPT

- SPT – Shortest Path Tree
- Các cạnh xuất phát từ nút gốc và tới các lá
- Đường đi duy nhất từ nút gốc tới nút v, là đường đi ngắn nhất giữa nút gốc và nút v
- Mỗi nút sẽ có một SPT của riêng nút đó



destination	Next hc
v	v
x	x
192.168.1.0/24 y	x
w	x
z	x

Giải thuật Dijkstra: Pseudo code

Ký hiệu

- $G = (V, E)$: Đồ thị với tập đỉnh V và tập cạnh E
- $c(x, y)$: chi phí của liên kết x tới y ; $= \infty$ nếu không phải 2 nút kề nhau
- $d(v)$: chi phí hiện thời của đường đi từ nút nguồn tới nút đích. v
- $p(v)$: nút ngay trước nút v trên đường đi từ nguồn tới đích
- T : Tập các nút mà đường đi ngắn nhất đã được xác định

Giải thuật Dijkstra: Pseudo code (1)

- **Init():**

Với mỗi nút v , $d[v] = \infty$, $p[v] = \text{NIL}$

$d[u] = 0$

- **Improve(x, v)**, trong đó (x, v) là một cạnh nào đó của G

if $d[v] > d[x] + c(x, v)$ then

$d[v] = d[x] + c(x, v)$

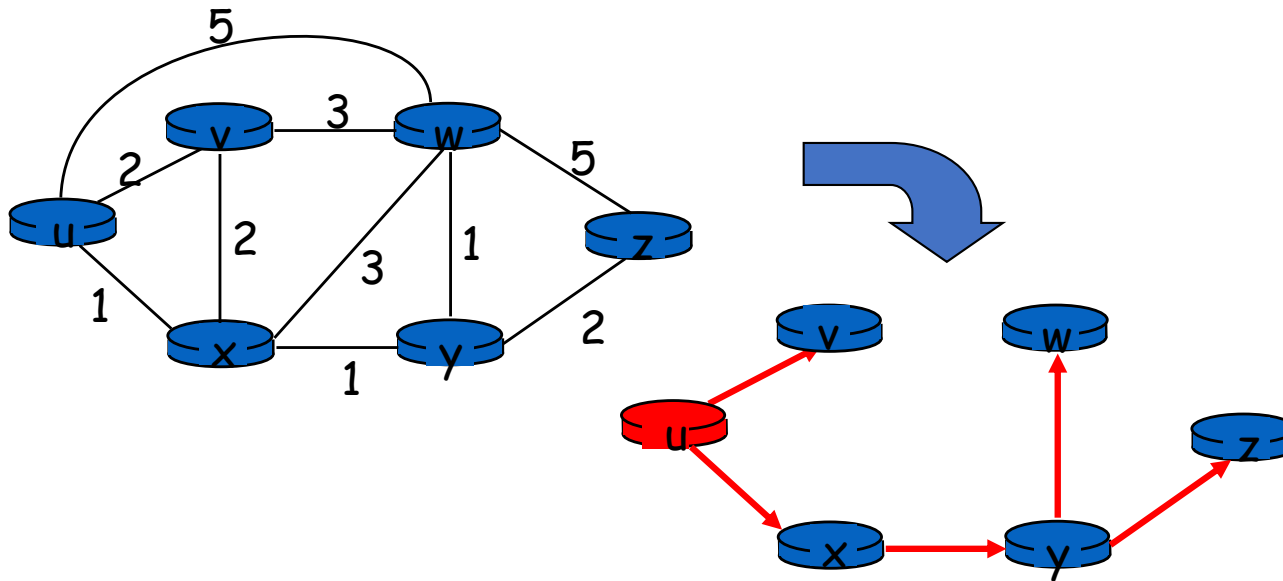
$p[v] = x$

Giải thuật Dijkstra: Pseudo code (2)

1. **Init()** ;
2. $T = \Phi$;
3. **Repeat**
4. $x: x \notin T \mid d(x)$ là bé nhất ;
5. $T = T \cup \{x\}$;
6. for all $v \in \text{neighbor}(x)$ và $v \notin T$
7. *update*(x, v) ;
8. **Until** $T = V$

Giải thuật Dijkstra: Ví dụ

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					



Bảng chọn đường của u:

destination	link
v	(u,v)
x	(u,x)
y	(u,x)
w	(u,x)
z	(u,x)

SPT của u:

Sử dụng giải thuật Dijkstra trong link-state protocol

Ý tưởng

- Mỗi nút đều có sơ đồ và chi phí mỗi link kề với nó
 - Quảng bá “Link-state”
 - Mỗi nút có cùng thông tin
- Tìm đường đi chi phí nhỏ nhất từ một nút (‘nguồn’) tới tất cả các nút khác
 - dùng để xây dựng **bảng chọn đường**
- Còn được gọi là giải thuật dạng link-state.

Giải thuật Bellman-ford

- Mỗi nút lưu trữ một vector khoảng cách tới mọi đích \rightarrow distance vector
- Tính toán (cải thiện) đường đi dựa trên việc thử đi qua hàng xóm \rightarrow giải thuật Distance-vector.

Phương trình Bellman-Ford (quy hoạch động)

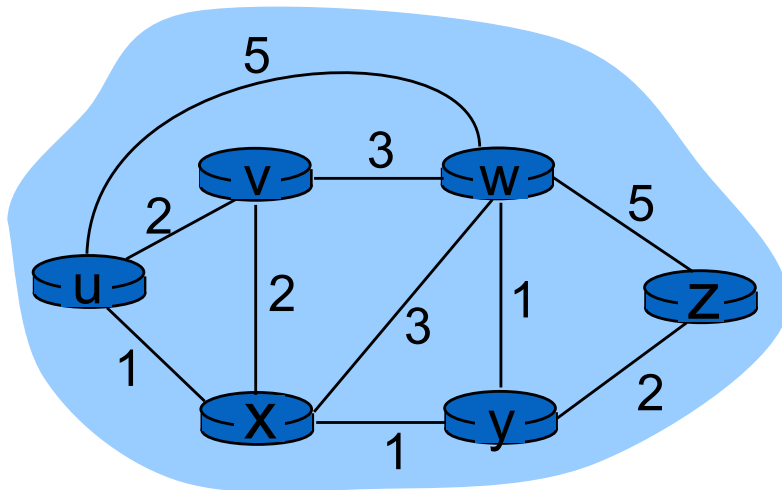
$d_x(y)$:= chi phí của đường đi ngắn nhất từ x tới y

Ta có
$$d_x(y) = \min_v \{c(x,v) + d_v(y)\}$$

cho tất cả các v là hàng xóm của x

Giải thuật Bellman-ford: Ví dụ

Dễ thấy, $d_v(z) = 5$, $d_x(z) = 3$, $d_w(z) = 3$



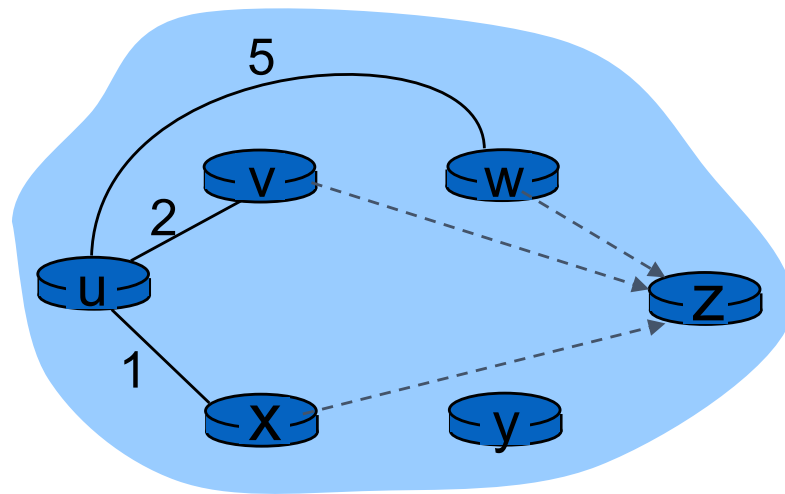
B-F eq. cho ta biết:

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

Nút nào làm giá trị trên nhỏ nhất → Lựa chọn là nút kế tiếp trong bảng chọn đường

Distance-vector algorithm Bellman-Ford (2)

- Vision of u:
 - Know only neighbour 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:

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

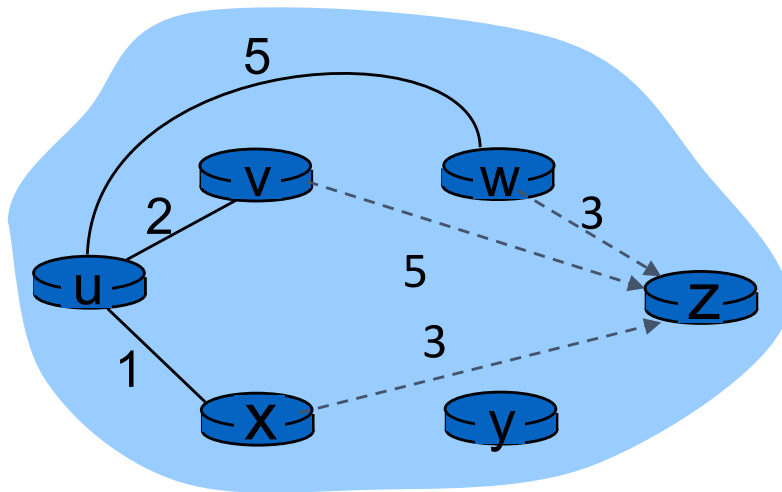
Amongst all paths from $u \rightarrow z$, u chooses to go via the neighbor that makes the path shortest

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. :

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

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

Giải thuật Bellman-ford

ý tưởng cơ bản:

- DV: Vector khoảng cách, tạm coi là đường đi ngắn nhất của từ một nút tới những nút khác
- Mỗi nút định kỳ gửi DV của nó tới các nút bên cạnh
- Khi nút x nhận được 1 DV, nó sẽ cập nhật DV của nó qua pt Bellman-ford
- Với một số điều kiện, ước lượng $D_x(y)$ sẽ hội tụ dần đến giá trị nhỏ nhất $d_x(y)$

Mỗi nút:

Chờ (Thay đổi trong DV của nút bên cạnh)

Tính lại ước lượng DV

Nếu DV thay đổi, *Báo* cho nút bên cạnh

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

$$= \min\{2+0, 7+1\} = 2$$

nút x

chi phí tới

	x	y	z
x	0	2	7
y	∞	∞	∞
z	∞	∞	∞

chi phí tới

	x	y	z
x	0	2	3
y	2	0	1
z	7	1	0

nút y

chi phí tới

	x	y	z
x	∞	∞	∞
y	2	0	1
z	∞	∞	∞

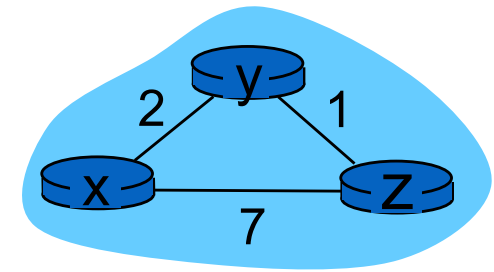
nút z

chi phí tới

	x	y	z
x	∞	∞	∞
y	∞	∞	∞
z	7	1	0

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

$$= \min\{2+1, 7+0\} = 3$$



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

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

nút x

chi phí tới

	x	y	z
x	0	2	7
y	∞	0	1
z	∞	∞	0

nút y

từ

	x	y	z
x	∞	∞	∞
y	2	0	1
z	∞	∞	∞

nút z

từ

	x	y	z
x	∞	∞	∞
y	∞	∞	∞
z	7	1	0

chi phí tới

	x	y	z
x	0	2	3
y	2	0	1
z	7	1	0

từ

từ

từ

	x	y	z
x	0	2	7
y	2	0	1
z	7	1	0

	x	y	z
x	0	2	7
y	2	0	1
z	3	1	0

từ

từ

từ

chi phí tới

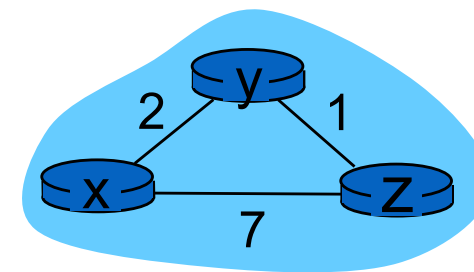
	x	y	z
x	0	2	3
y	2	0	1
z	3	1	0

chi phí tới

	x	y	z
x	0	2	3
y	2	0	1
z	3	1	0

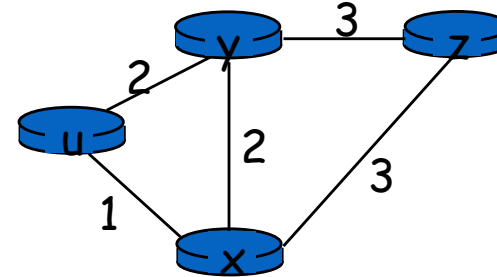
chi phí tới

	x	y	z
x	0	2	3
y	2	0	1
z	3	1	0



Bài tập

- Thực hiện giải thuật Bellman-Ford tính đường đi ngắn nhất cho đồ thị sau
- Yêu cầu, minh họa bằng bảng vector khoảng cách như trong ví dụ trước
- Ví dụ vector khoảng cách tại nút u chi phí tới



	u	x	y	z
u	0	1	2	∞
x				
y				
z				

Tập trung hay phân tán

- Thông tin chọn đường là cần thiết để xây dựng bảng chọn đường
- Tập trung hay phân tán?
 - Tập trung:
 - Mỗi router có thông tin đầy đủ về trạng thái của mạng
 - Giao thức dạng “link state”
 - Phân tán:
 - Các nút chỉ biết được trạng thái của liên kết vật lý tới nút kế bên
 - Liên tục lặp lại việc tính toán và trao đổi thông tin với nút kế bên
 - Giao thức dạng “distance vector”
 - “Bạn của bạn cũng là bạn”

So sánh các giải thuật

Link-state:LS, Sử dụng Dijkstra

Distance vector: DV, Sử dụng Bellman-Ford

Thông điệp trao đổi

- **LS:** n nút, E cạnh, $O(nE)$ thông điệp
- **DV:** Chỉ trao đổi giữa các hàng xóm
 - thay đổi

Tốc độ hội tụ

- **LS:** Thuật toán: $O(n^2)$ cần $O(nE)$ thông điệp
- **DV:** Thay đổi

Sự chắc chắn: Giải sử một router hoạt động sai

LS:

- nút gửi các chi phí sai
- Mỗi nút tính riêng bảng chọn đường -> có vẻ chắc chắn hơn

DV:

- DV có thể bị gửi sai
- Mỗi nút tính toán dựa trên các nút khác
 - Lỗi bị lan truyền trong mạng



ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

Các giao thức chọn đường

Đọc trước: Chapter 5- Computer
Networks, Tanenbaum



ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

Phân cấp trong chọn đường

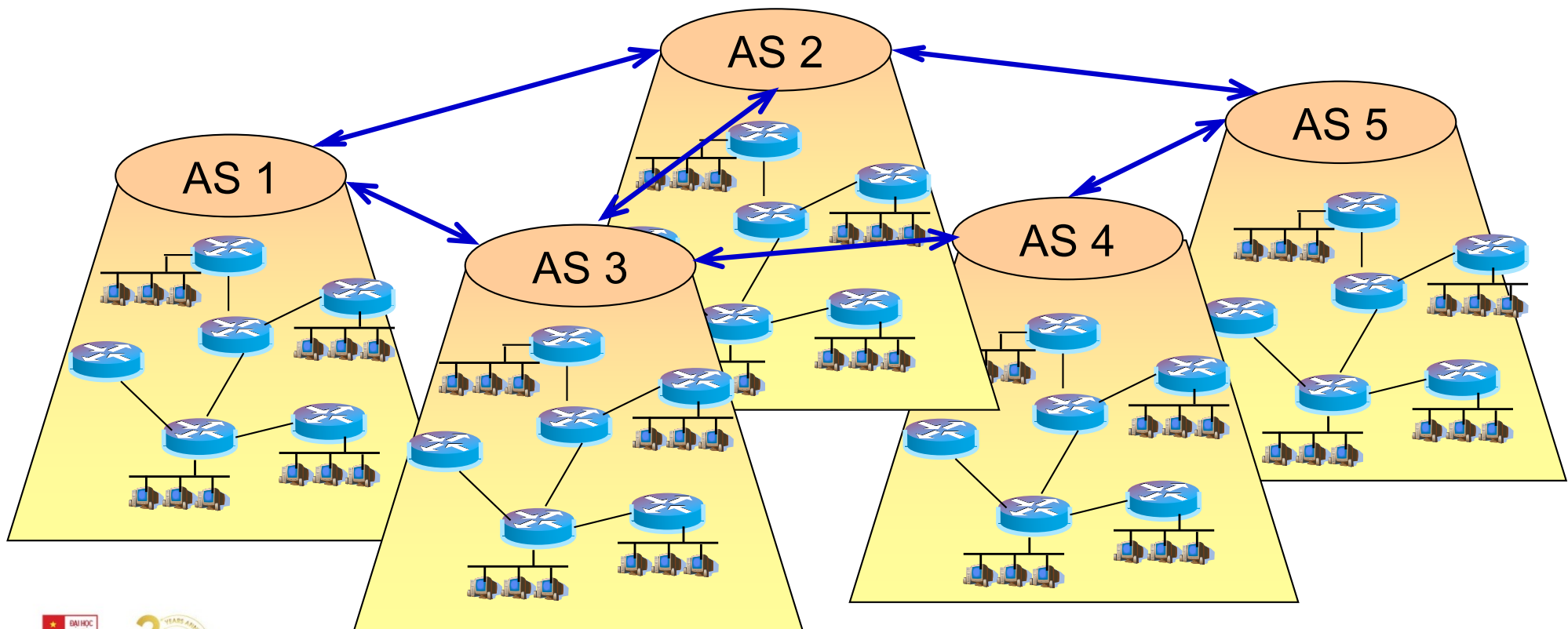
Các hệ tự trị

Chọn đường nội vùng

Chọn đường liên vùng

Kiến trúc phân cấp của Internet

- Internet = Mạng của các mạng
- Mỗi mạng có thể lựa chọn riêng cho mình một chiến lược chọn đường riêng.
- Mỗi mạng như vậy có thể gọi là một hệ tự trị - Autonomous System (AS)



Khái niệm hệ tự trị - AS

- Tập hợp các nút mạng có cùng chính sách chọn đường (Giao thức, quy ước chi phí...)
- Các ASes được nối kết thông qua các router hay gateway
- Mỗi hệ tự trị có một số hiệu riêng – AS number (ASN - 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

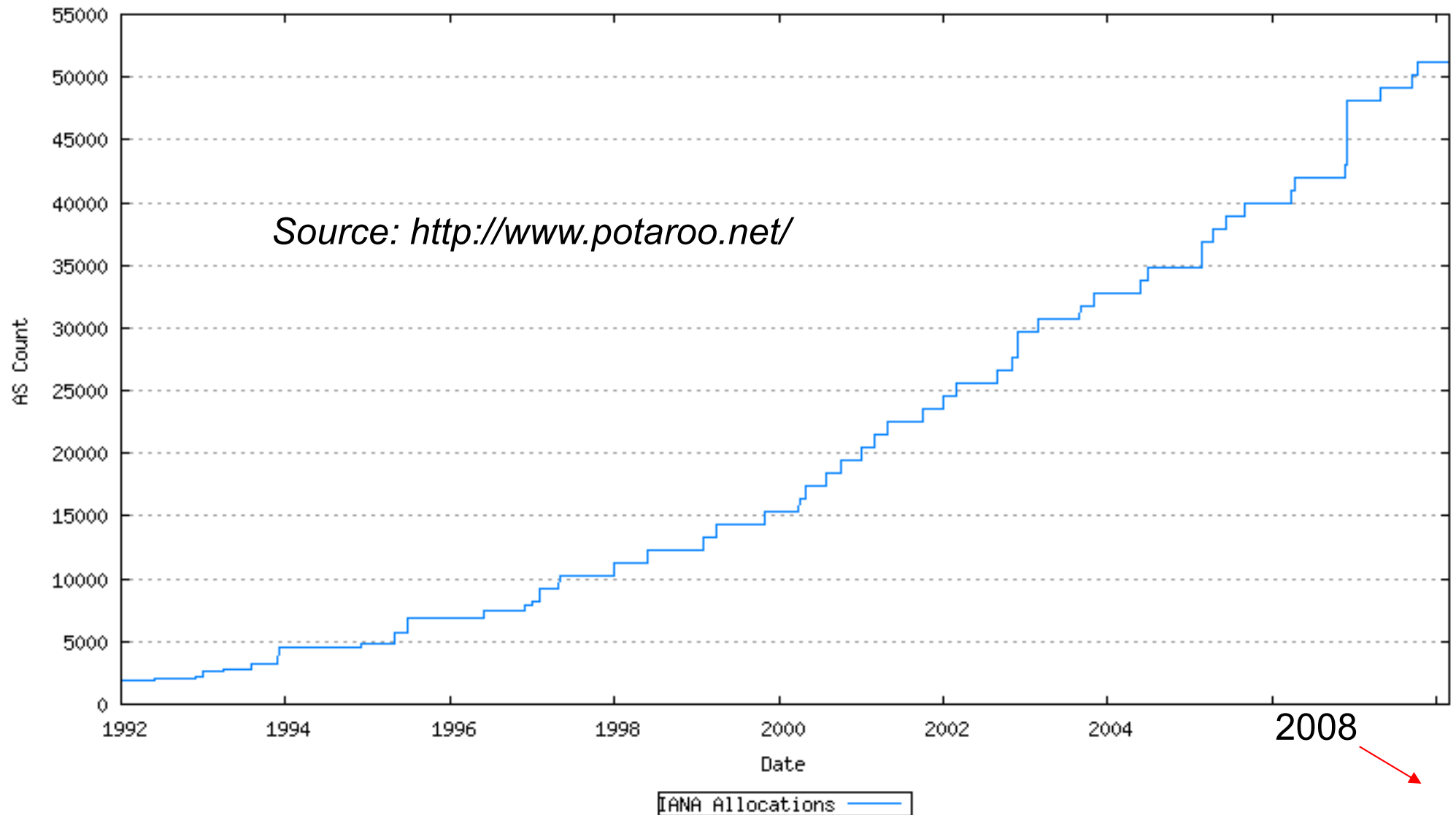
[17981](#) CAMBOTECH-KH-AS ISP Cambodia

.....

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

Số lượng ASN cấp phát bởi IANA

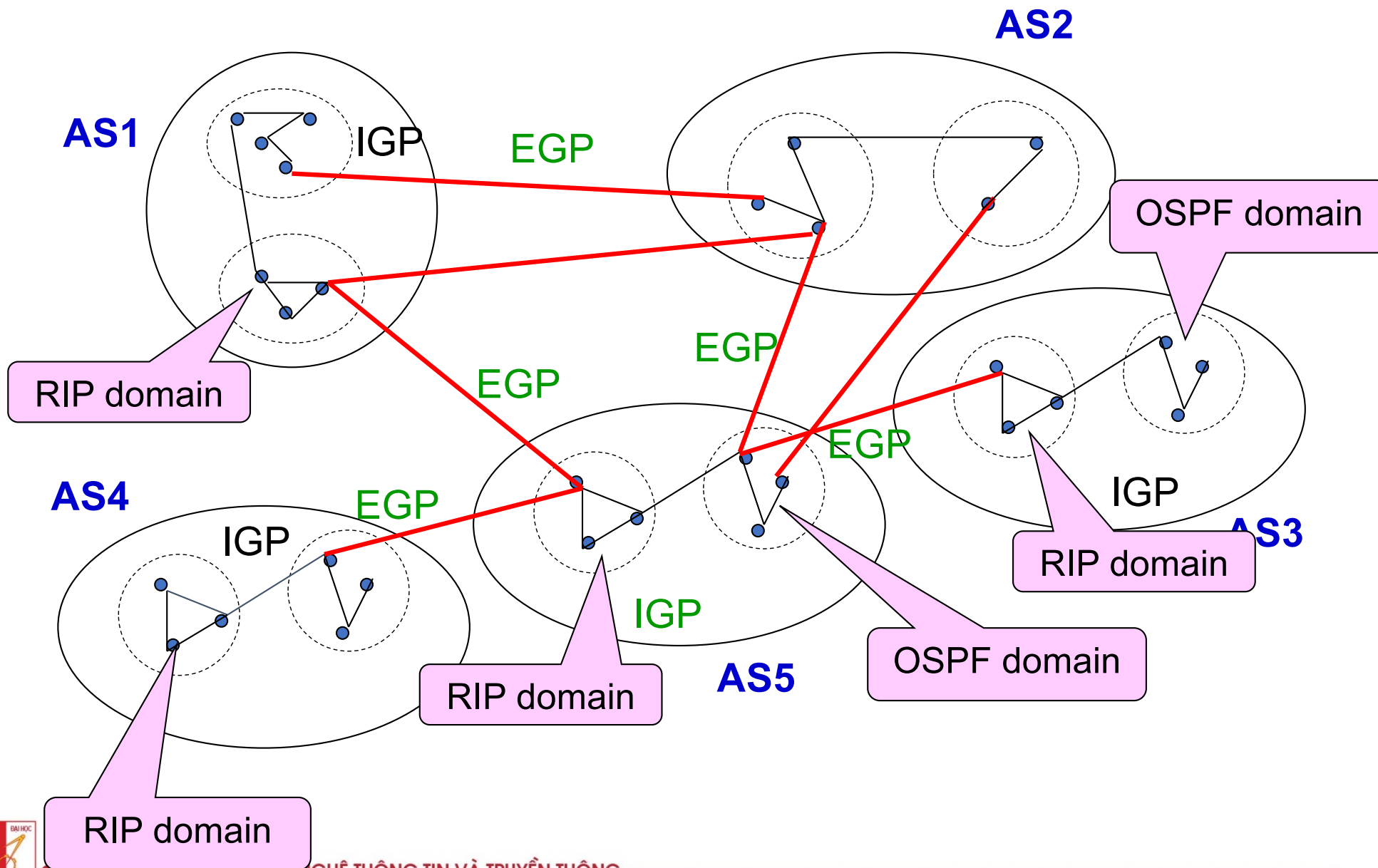
Time Series of IANA AS Allocations



Phân cấp giao thức chọn đường

- **Trong một hệ tự trị:** Giao thức chọn đường nội vùng
 - IGP: *Interior Gateway Protocol*
 - RIP: Routing Information Protocol
 - OSPF: Open Shortest Path First
 - IS-IS, IGRP, EIGRP (Cisco)...
- **Giữa các hệ tự trị:** Giao thức chọn đường liên vùng
 - EGP: *Exterior Gateway Protocol*
 - BGP (v4): Border Gateway Protocol

Intra-domain và Inter-domain routing





ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

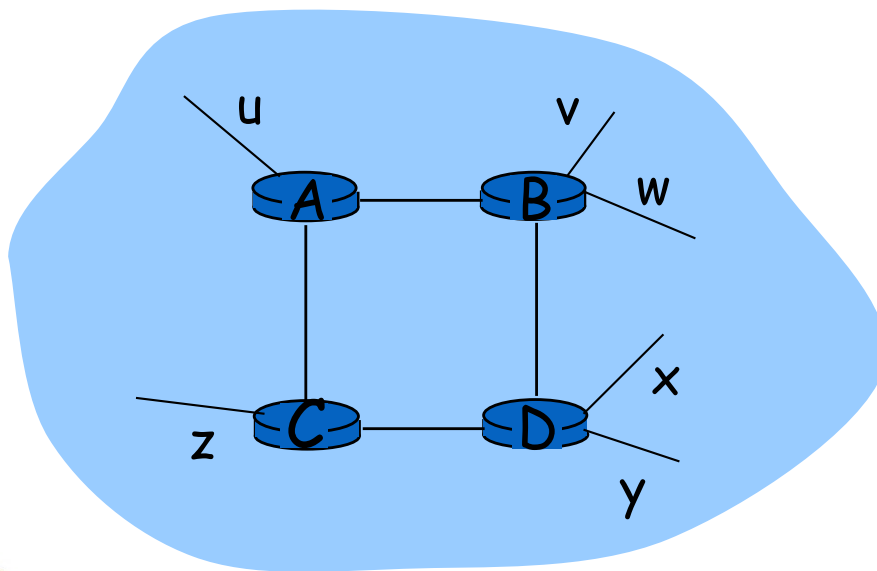
Chọn đường nội vùng

RIP

OSPF

RIP (Routing Information Protocol)

- RIP v.1, phiên bản mới RIP v.2
- Giao thức dạng Distance vector (vector khoảng cách)
- Chọn đường đi theo số nút mạng đi qua (# of hops, max = 15 hops)
- RIP v.1: RFC-1058 (www.ietf.org)

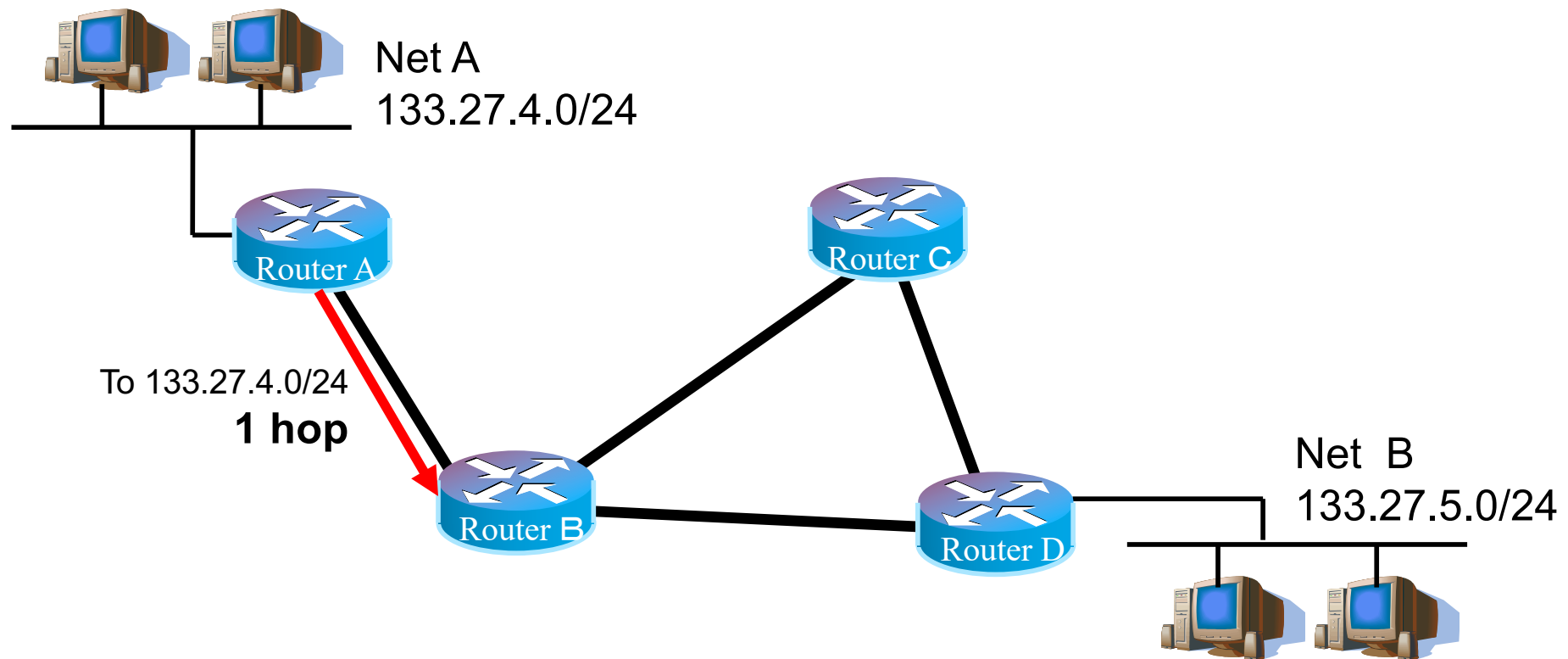


Từ nút A:

<u>Đích đến</u>	<u>số nút</u>
u	1
v	2
w	2
x	3
y	3
z	2

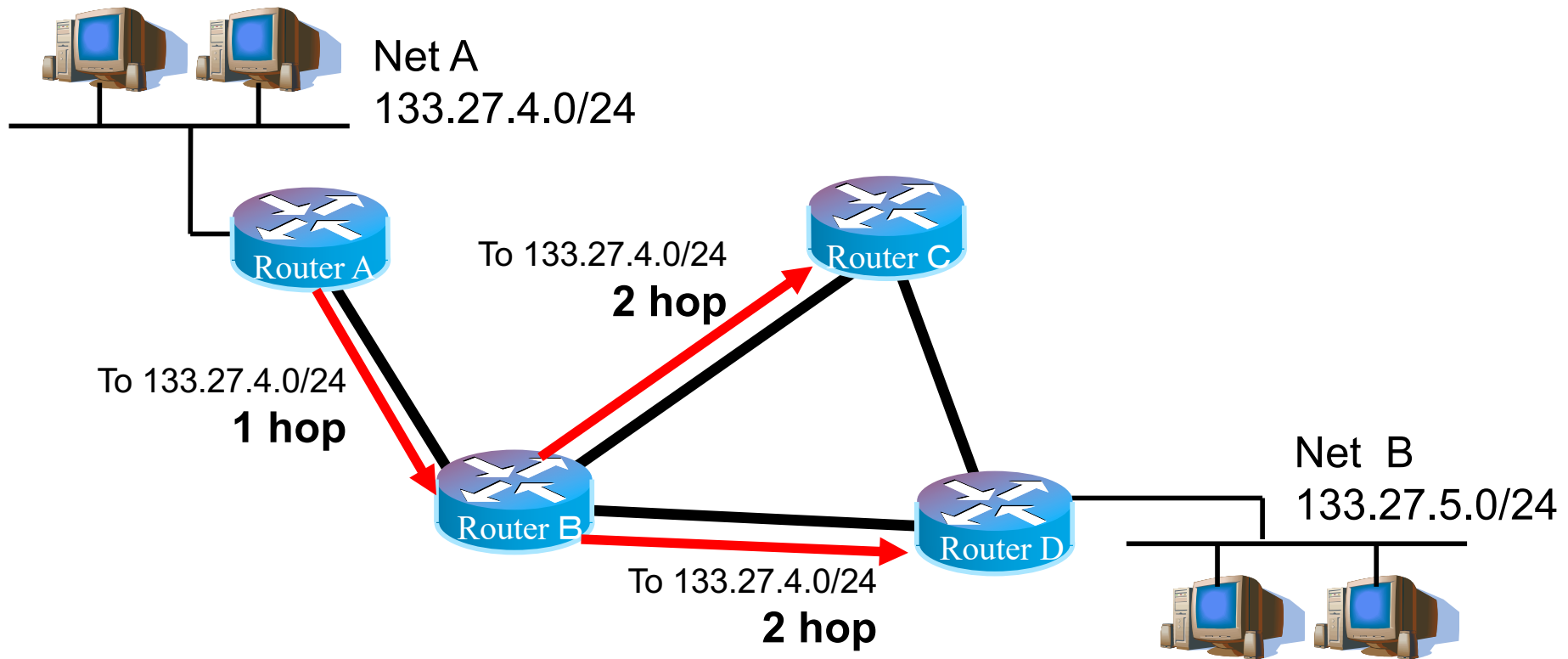
Nhắc lại: Chọn đường dạng DV (1)

- Bạn của bạn là bạn



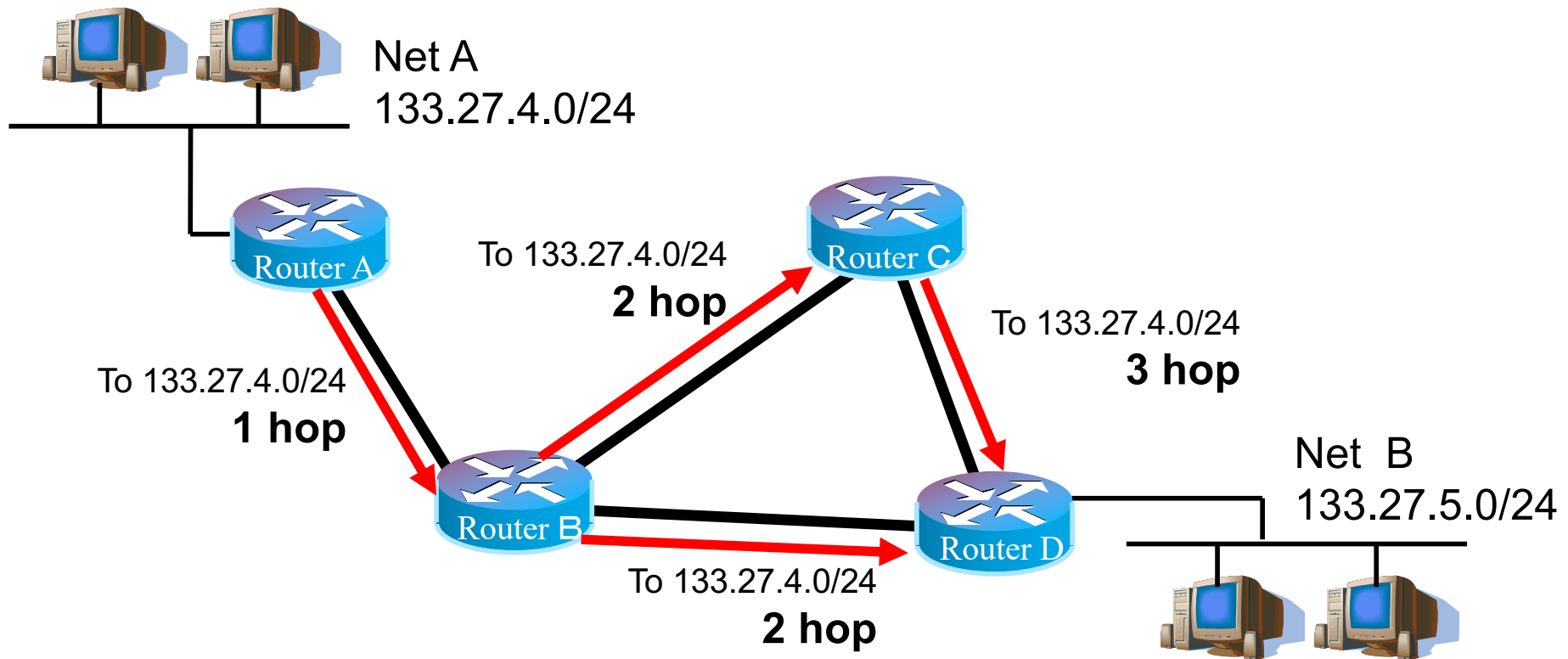
Nhắc lại: Chọn đường dạng DV (2)

- Bạn của bạn là bạn



Nhắc lại: Chọn đường dạng DV (3)

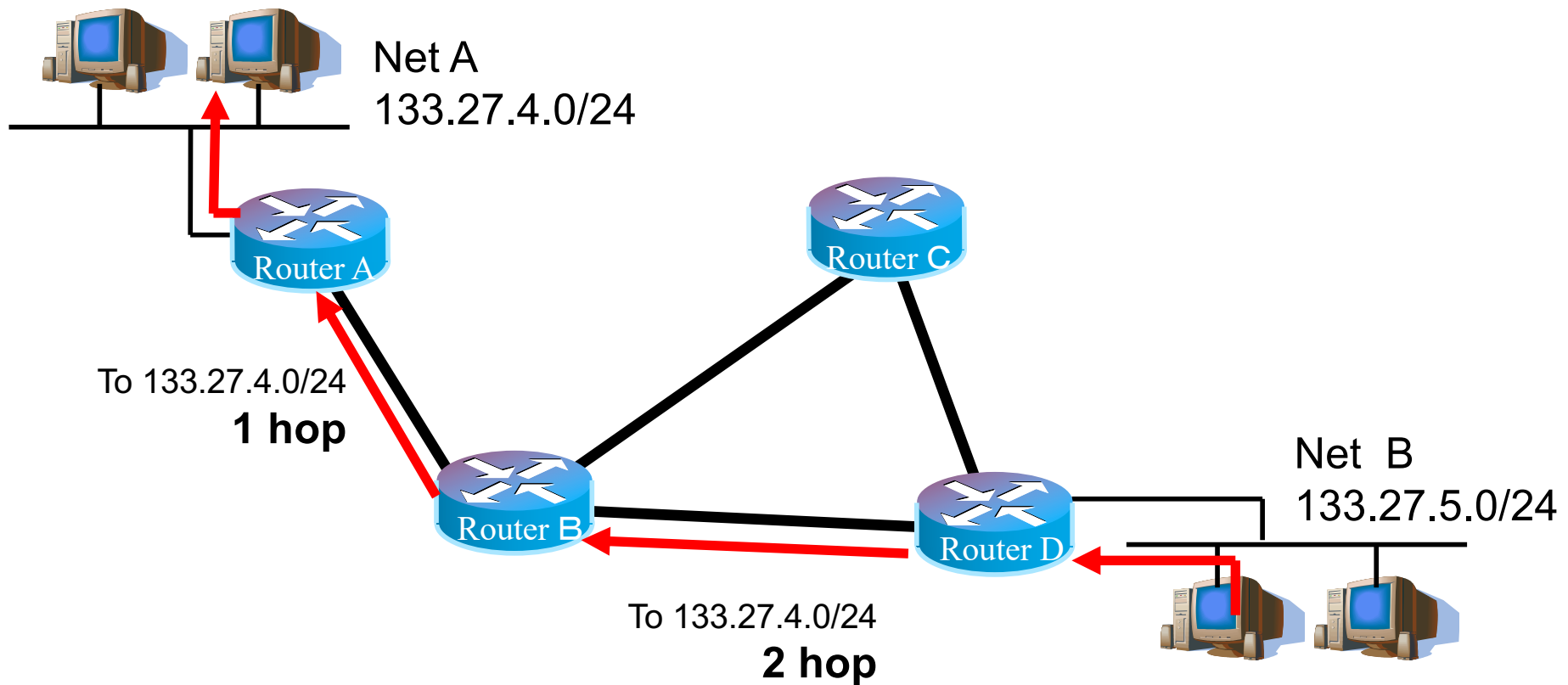
- Bạn của bạn là bạn



Nhắc lại: Chọn đường dạng DV (4)

- Bạn của bạn là bạn

Lưu ý: Tên của router



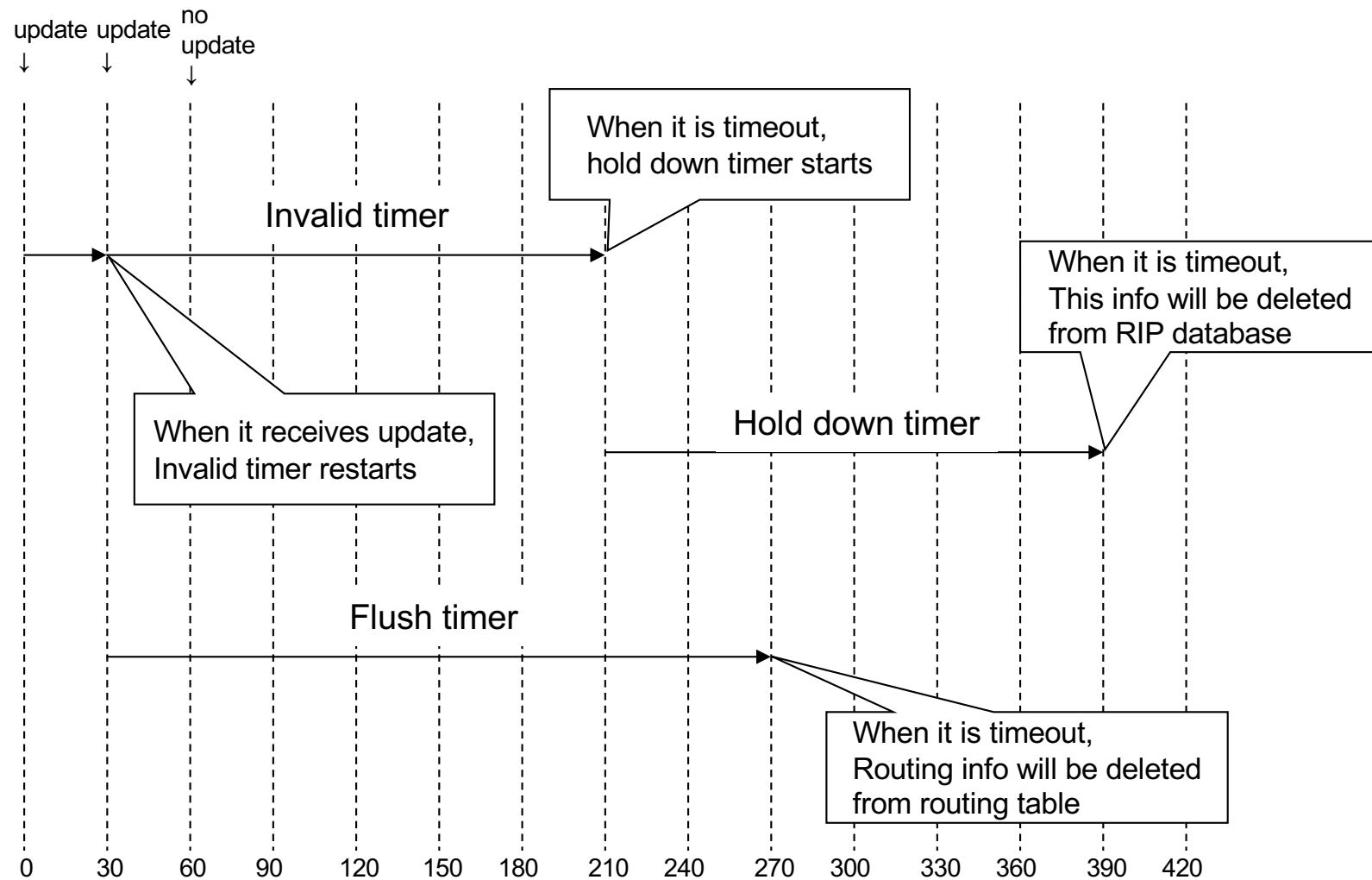
RIP: Trao đổi thông tin

- Trao đổi bảng vector khoảng cách
- Định kỳ
 - Các vector khoảng cách được trao đổi định kỳ - 30s
 - Mỗi thông điệp chứa tối đa 25 mục
 - Trong thực tế, nhiều thông điệp được sử dụng
- Sự kiện
 - Gửi thông điệp cho nút hàng xóm mỗi khi có thay đổi
 - Nút hàng xóm sẽ cập nhật bảng chọn đường của nó

Các bộ đếm thời gian - RIP timer (1)

- Update timer
 - Dùng để trao đổi thông tin cứ 30s
- Invalid timer
 - Khởi tạo lại mỗi khi nhận được thông tin chọn đường
 - Nếu sau 180s không nhận được thông tin -> trạng thái hold-down
- Hold down timer
 - Giữ trạng thái hold-down trong 180s
 - Chuyển sang trạng thái down
- Flush timer
 - Khởi tạo lại mỗi khi nhận được thông tin chọn đường
 - Sau 120s, xóa mục tương ứng trong bảng chọn đường

RIP timer (2)



Lỗi lặp vô hạn (Ping-pong failure)

- Nếu 192.168.0.0/24 bị lỗi...
 - B cập nhật thông tin về 192.168.0.0 cho A
 - Các gói tin đến 192.168.0.0/24 sẽ bị quẩn
- A cập nhật thông tin về 192.168.0.0 cho B
 - Luẩn quẩn, vô hạn

RIP: Để tránh lỗi lặp vô hạn

- Giới hạn số hop tối đa
 - 16
- “Split horizon”
 - Thông tin chọn đường không được quay về nút nguồn
- “Poison reverse”
 - Khi liên kết bị lỗi, gửi giá trị của chi phí là 16
 - Liên kết chuyển sang trạng thái hold-down

OSPF: Open Shortest Path First

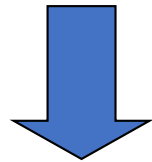
- *Open*: Chuẩn mở của IETF (phiên bản 3, định nghĩa trong [RFC 2740](#))
- *Shortest Path First*: Cài đặt giải thuật Dijkstra.
- Loại Link-state
- Thông tin về trạng thái liên kết - LSA (link state advertisement) được quảng bá “tràn ngập” trên toàn AS

Một số đặc điểm của OSPF

- **An toàn**: thông điệp OSPF được bảo vệ
- Với các AS lớn: OSPF **phân cấp**
- **Địa chỉ không phân lớp** (Variable-Length Subnet Masking -VLSM)
- Mỗi link sẽ có nhiều giá trị về **chi phí** khác nhau dựa trên TOS (tuy nhiên hơi phức tạp và chưa được sử dụng)

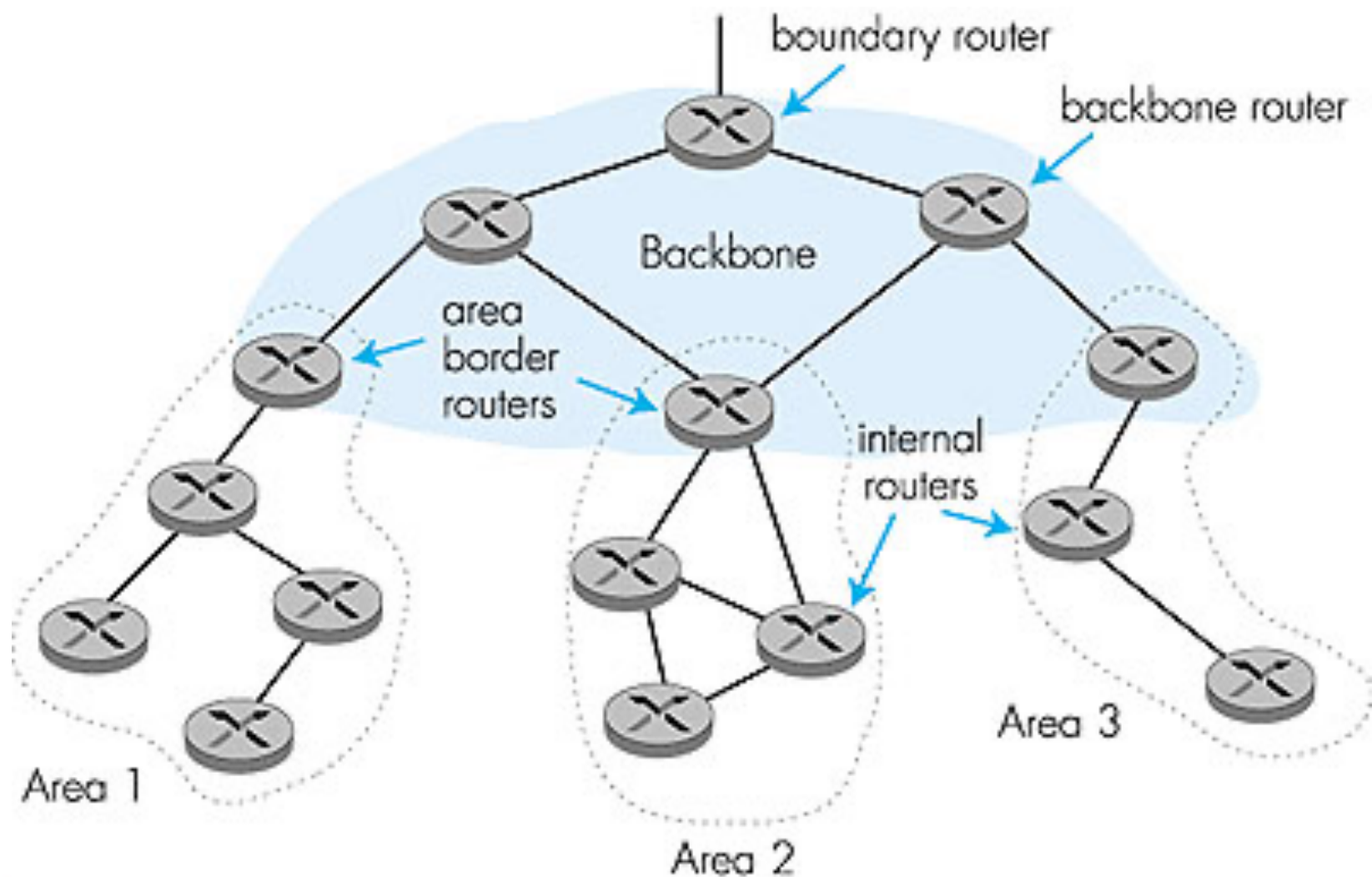
Phân cấp OSPF

- Trong việc chọn đường, tại sao phải chia mạng thành các vùng nhỏ hơn?
- Nếu có quá nhiều router
 - Thông tin trạng thái liên kết được truyền nhiều lần hơn
 - Phải liên tục tính toán lại
 - Cần nhiều bộ nhớ hơn, nhiều tài nguyên CPU hơn
 - Lượng thông tin phải trao đổi tăng lên
 - Bảng chọn đường lớn hơn



- Vùng
 - Nhóm các router có cùng thông tin LSA

Phân cấp OSPF

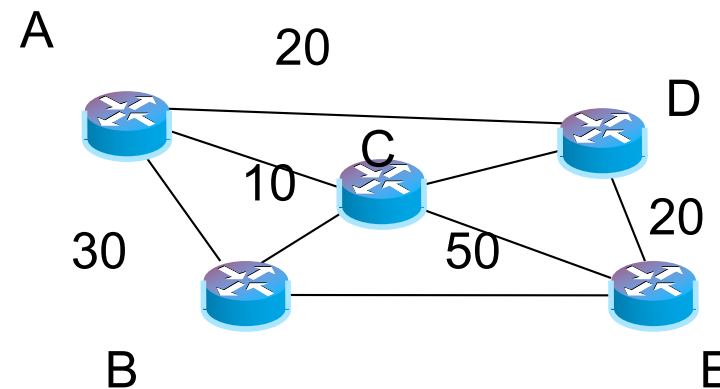


Các dạng router

- **ABR - Area border routers:** Quản lý 1 vùng và kết nối đến các vùng khác
- **ASBR - Autonomous system boundary router:** Nối đến các AS khác
- **BR - backbone routers:** thực hiện OSPF routing trong vùng backbone
- **Internal Router** – Thực hiện OSPF bên trong một vùng

Thông tin liên kết

- Link-State Advertisement (LSA): Chỉ ra một nút được nối tới nút nào (link) và chi phí (cost) tương ứng
- Ví dụ: nút A
 - link to B, cost 30
 - link to D, cost 20
 - link to C, cost 10
- Ví dụ: nút D
 - link to A, cost 20
 - link to E, cost 20
 - link to C, cost 50



Chi phí trong giao thức OSPF - metric

- Giá trị mặc định
100Mbps / bandwidth of interface
 - Hiện nay người quản trị có thể gán giá trị này
- Khi tính toán bảng chọn đường
 - Chọn đường đi chi phí nhỏ nhất
- Chi phí bằng nhau
 - Có thể thực hiện cân bằng tải

Chi phí mặc định của OSPF

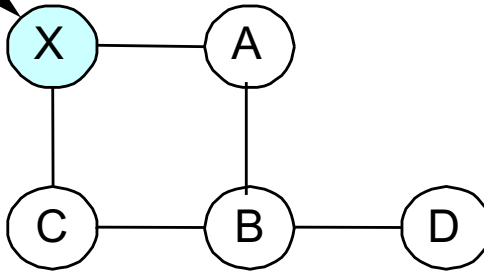
Link Bandwidth	Default OSPF cost
56Kbps serial link	1785
64Kbps serial link	1562
T1 (1.544Mbps) serial link	65
E1 (2.048Mbps) serial link	48
4Mbps Token Ring	25
Ethernet	10
16Mbps Token Ring	6
FDDI or Fast Ethernet	1
Gigabit Ethernet / 10G network	1

Quảng bá thông tin LSA

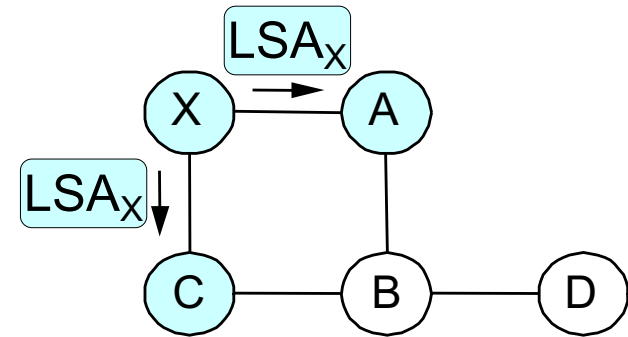
X has link to A, cost 10
X has link to C, cost 20



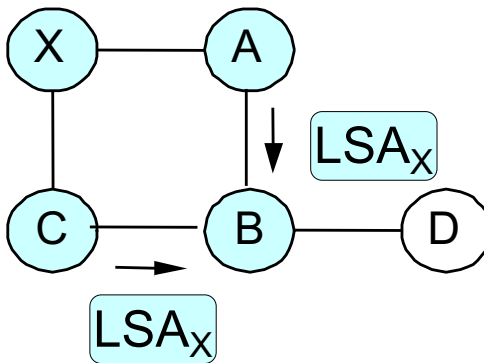
LSA_x



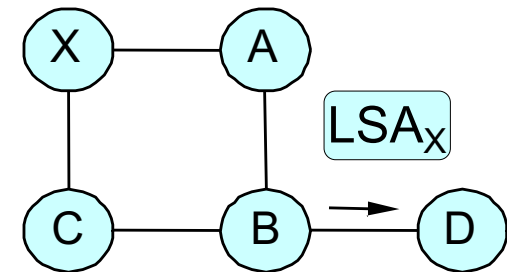
(a)



(b)



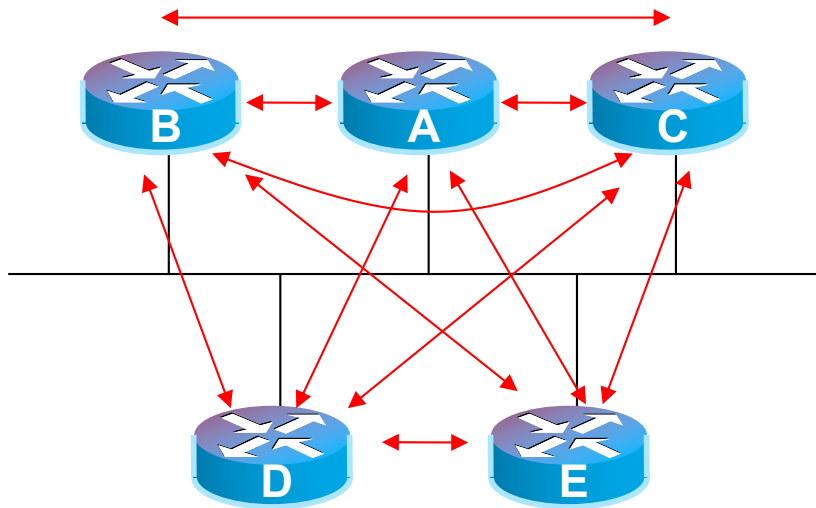
(c)



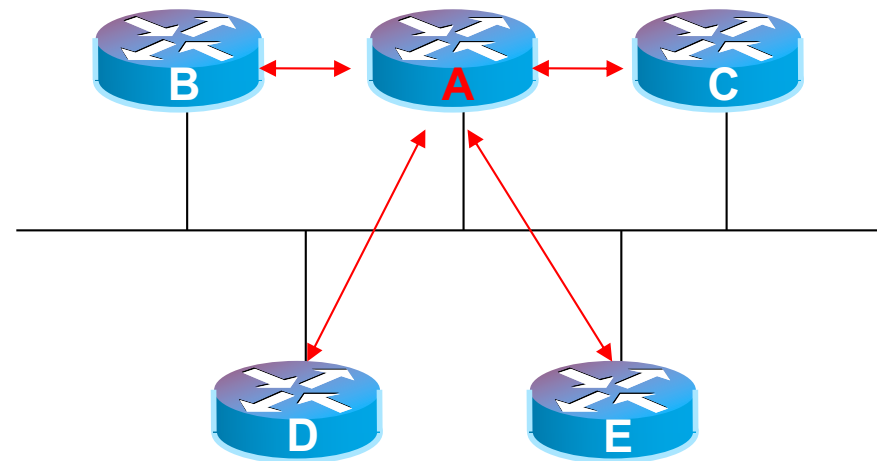
(d)

Router đại diện - DR

- Để tăng hiệu quả của việc quảng bá LSA
- Mỗi router phải lập quan hệ với router đại diện - designated router (DR)
 - Trao đổi thông tin thông qua DR
 - DR dự phòng
- Chọn DR và BDR?



Không có DR



Có DR

Neighbor & Adjacency

- “Neighbor” và “adjacency” là các khái niệm khác nhau!
 - Adjacency: có trao đổi thông tin
 - Neighbor: có đường nối trực tiếp
- Mạng quảng bá đa truy cập (e.g Ethernet)
 - Neighbor != Adjacency
- Mạng điểm-nối-điểm
 - Neighbor == Adjacency

RIP vs. OSPF

	RIP	OSPF
Đặc điểm	<ul style="list-style-type: none">• Router bình đẳng• Cấu hình dễ dàng• Mạng cỡ nhỏ	<ul style="list-style-type: none">• Phân cấp• Cấu hình phức tạp• Mạng cỡ vừa và lớn
Khả năng mở rộng	X	O
Độ phức tạp tính toán	Nhỏ	Lớn
Hội tụ	Chậm	Nhanh
Trao đổi thông tin	Bảng vector khoảng cách	Trạng thái liên kết
Giải thuật	Distant vector	Link-state
Cập nhật hàng xóm	30s	10s (Hello packet)
Đơn vị chi phí	Số nút mạng	Bảng thông



ĐẠI HỌC BÁCH KHOA HÀ NỘI
VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

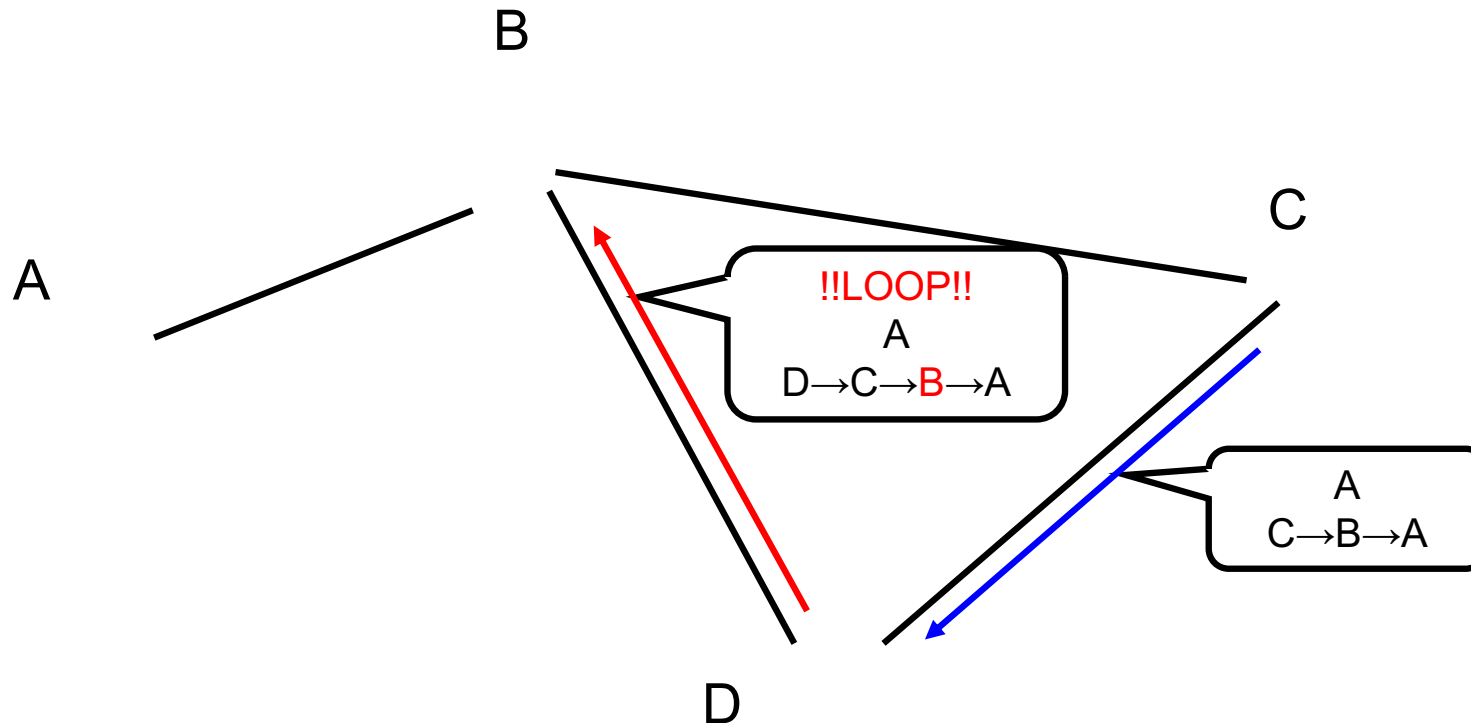
Giao thức chọn đường liên vùng

BGP – Border Gateway Protocol

- Yếu tố gắn kết của Internet, kết nối các hệ tự trị
- Trao đổi thông tin NLRI (Network Layer Reachability Information)
 - Cho phép một AS biết được thông tin đi đến AS khác
 - Gửi thông tin này vào bên trong AS đó
 - Xác định đường đi tốt nhất dựa trên thông tin đó và các chính sách chọn đường
- Cho phép thiết lập các chính sách
 - Chọn đường ra
 - Quảng bá các đường vào

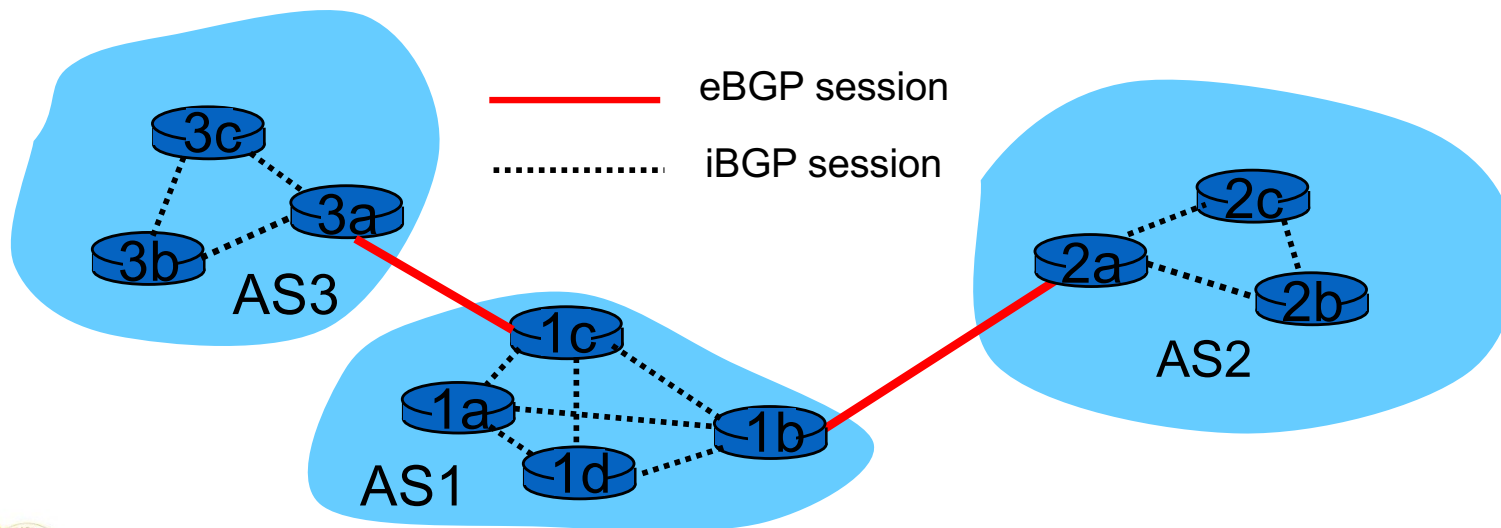
Cơ chế tránh vòng lặp

- Dò lại xem router đã có trên path-vector hay chưa
 - B hủy đường đi tới A



eBGP và iBGP

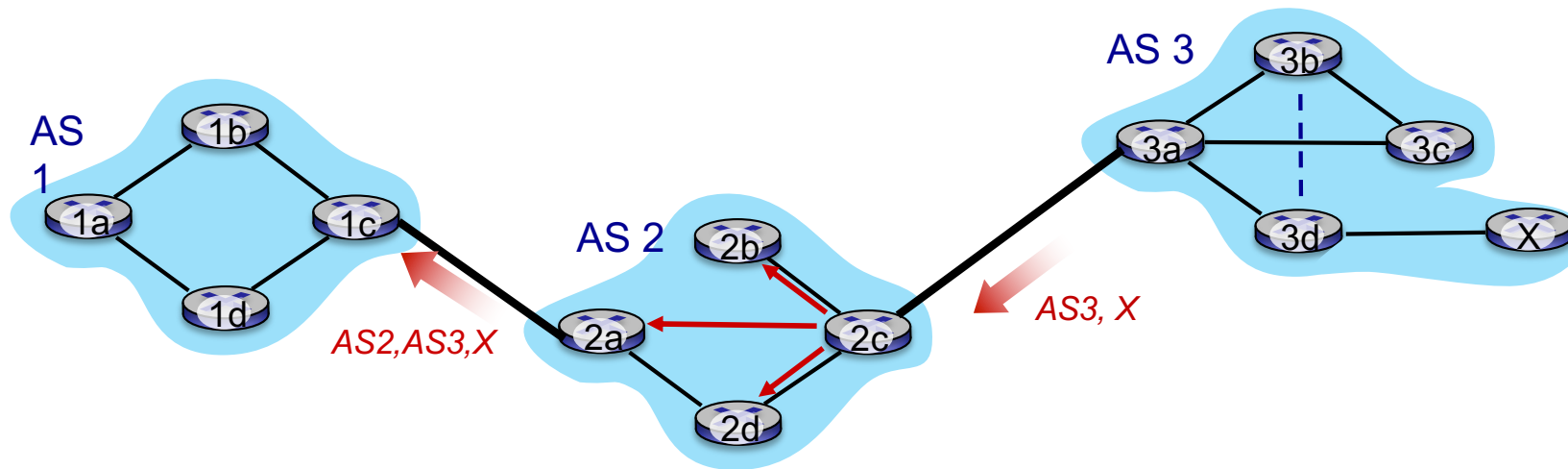
- External BGP vs. Internal BGP
 - External BGP dùng để trao đổi thông tin giữa các router biên thuộc các AS khác nhau
 - Internal BGP dùng để trao đổi thông tin giữa các router biên thuộc cùng AS
- Phân tán thông tin chọn đường
 1. 3a gửi tới 1c bằng eBGP
 2. 1c gửi thông tin nội bộ tới (1b, 1d, ...) trong AS1 bằng iBGP
 3. 2a nhận thông tin từ 1b bằng eBGP



BGP có thể cài đặt các policy

- Khi các router gửi và nhận thông tin chọn đường router BGP có thể áp đặt các chính sách
 - Cho đường vào: chọn lọc đường đi sẽ tiếp nhận
 - Chọn đường đi tốt nhất căn cứ theo chính sách.
 - Cho đường ra: chọn lọc đường đi sẽ quảng bá

BGP path advertisement

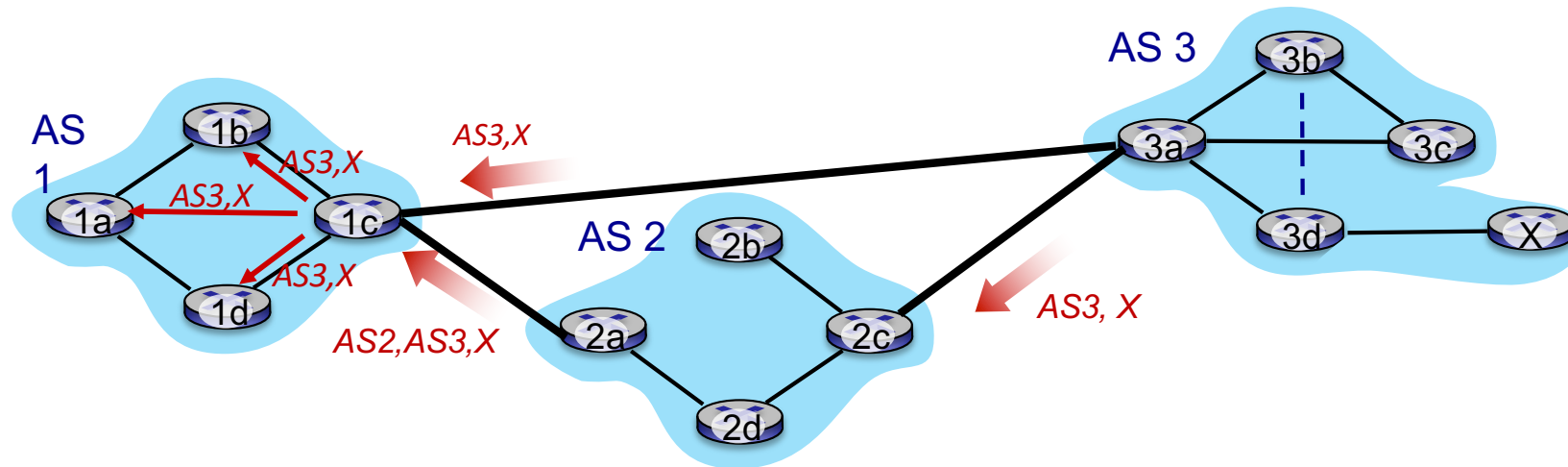


AS2 router 2c receives path advertisement **AS3,X** (via eBGP) from AS3 router 3a

based on AS2 policy, AS2 router 2c accepts path AS3,X, propagates (via iBGP) to all AS2 routers

based on AS2 policy, AS2 router 2a advertises (via eBGP) path **AS2, AS3, X** to AS1 router 1c

BGP path advertisement (more)



border router may learn about **multiple paths to destination:**

AS1 border router 1c learns path **AS2,AS3,X** from 2a

AS1 border router 1c learns path **AS3,X** from 3a

➔ based on *policy*, AS1 border router 1c chooses path **AS3,X** and advertises path within AS1 via iBGP

Các thuộc tính của đường đi

- **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**

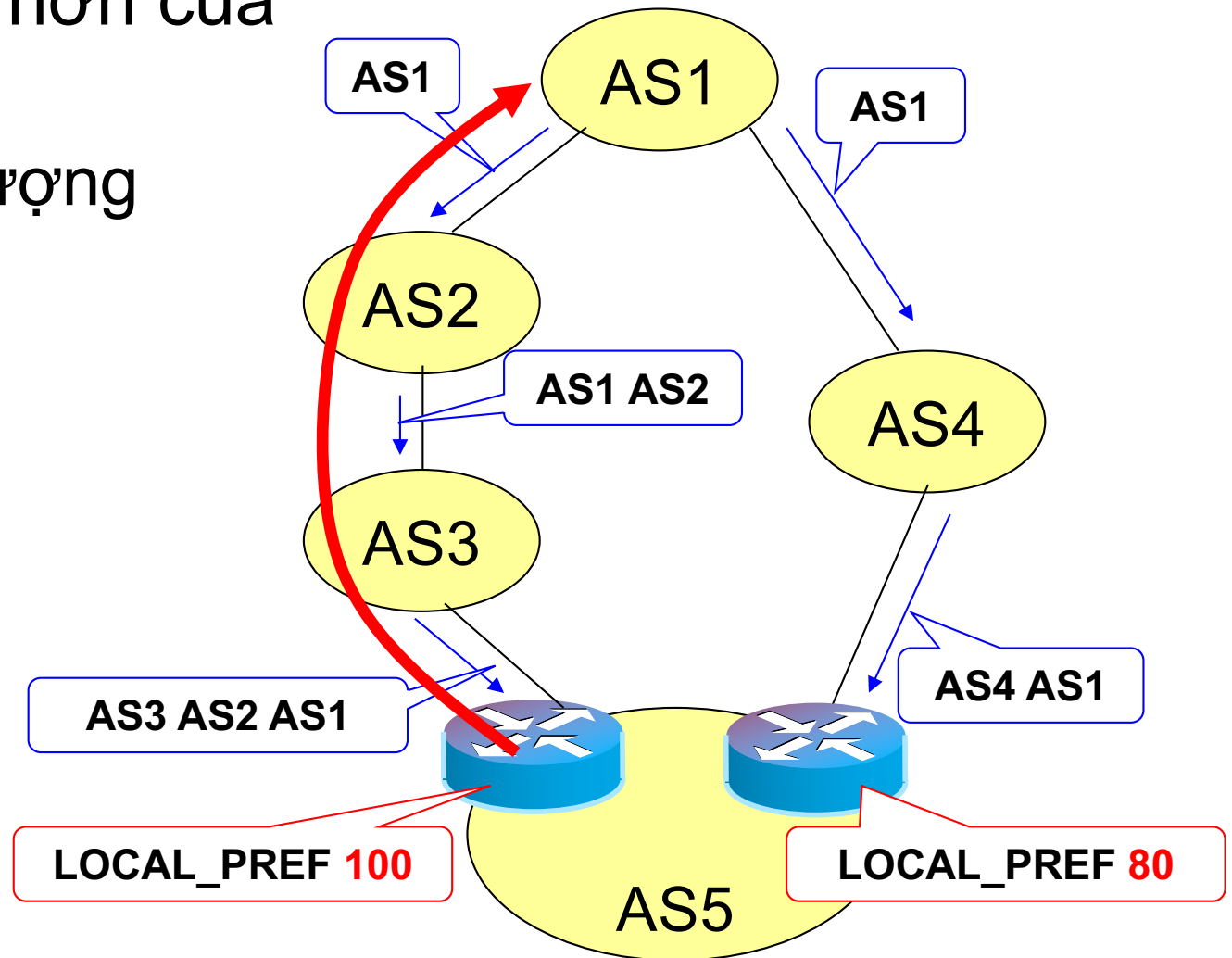
Các bước chọn đường đi

Với các tuyến đường có cùng một đích, chọn tuyến tốt nhất bằng cách so sánh theo thứ tự:

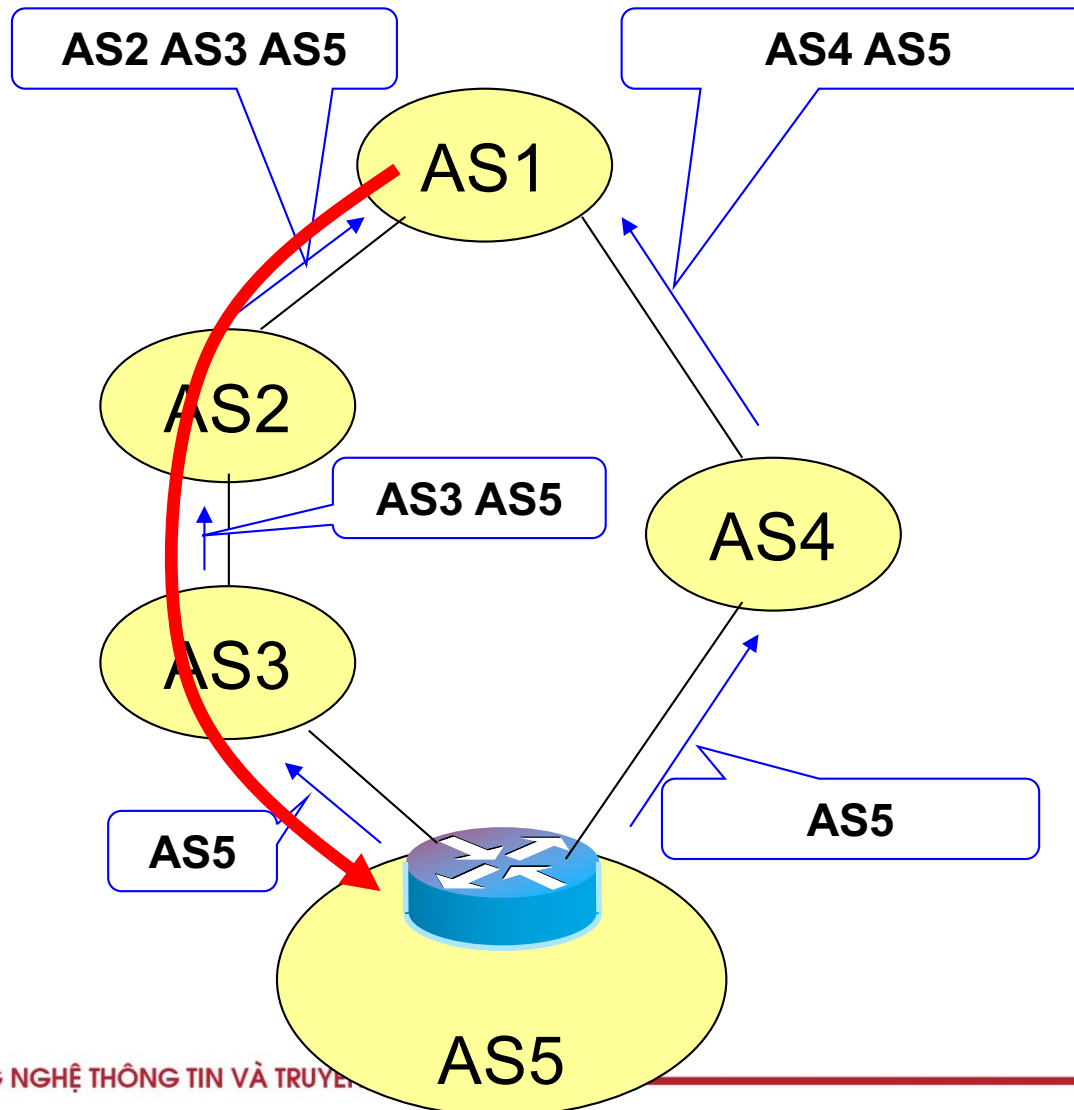
- Bước 1: So sánh LOCAL_PREF
- Bước 2: So sánh độ dài AS_PATH
- Bước 3: So sánh ORIGIN
- Bước 4: So sánh MED
- Bước 5: So sánh EBGP/IBGP
- Bước 6: So sánh chi phí tới NEXT_HOP
- Bước 7: So sánh Router ID

Sử dụng LOCAL_PREF

- Chọn giá trị lớn hơn của LOCAL_PREF
- Điều khiển lưu lượng upbound



Chọn đường với AS_PATH Prepend



Ví dụ về 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

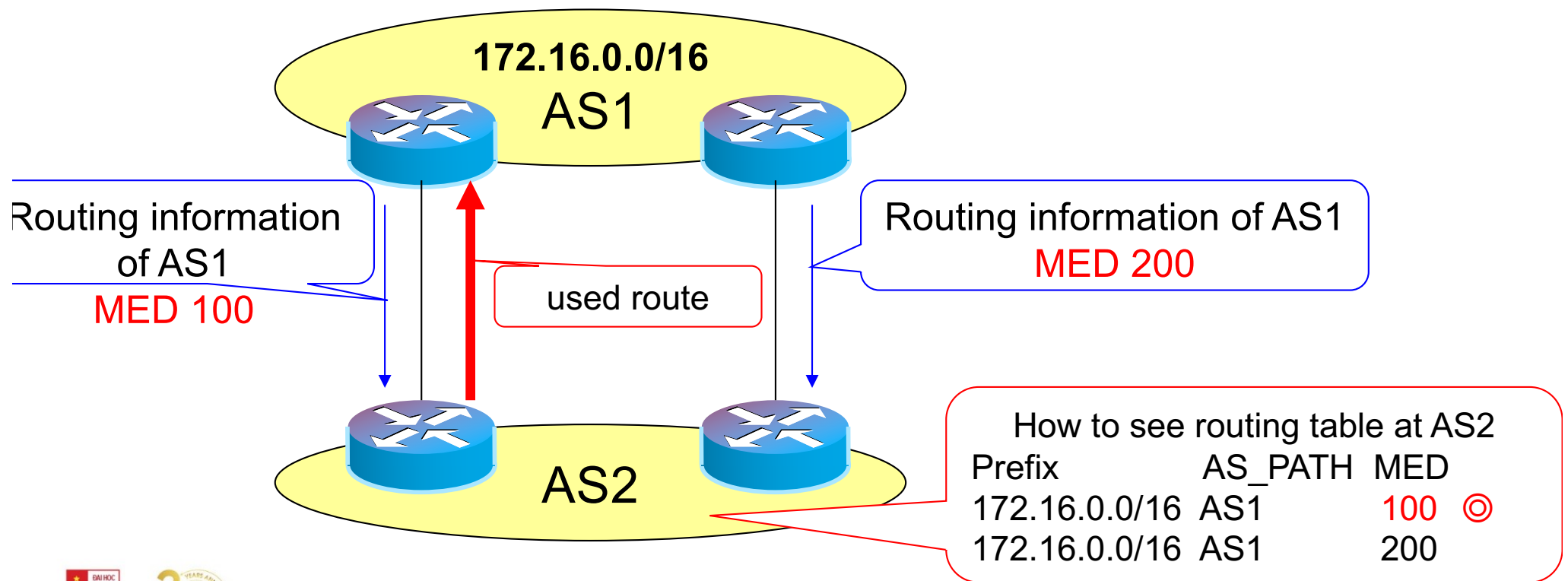
Ví dụ về AS PATH prepend

Network	Next Hop	Metric	LocPrf	Weight	Path
8.5.192.0/22	203.178.136.14	100	0	2516 209	13989 13989 13989 13989
	203.178.136.14	100	0	2516 209	13989 13989 13989 13989
	203.178.136.14	100	0	2516 209	13989 13989 13989 13989
8.5.196.0/24	203.178.136.14	100	0	2516 209	13989 13989 13989 13989
	203.178.136.14	100	0	2516 209	13989 13989 13989 13989
	203.178.136.14	100	0	2516 209	13989 13989 13989 13989
8.5.200.0/22	203.178.136.14	100	0	2516 209	13989 13989 13989 13989
	203.178.136.14	100	0	2516 209	13989 13989 13989 13989

Một số AS được lặp lại trên đường đi để làm đường đi dài hơn và sẽ không được ưu tiên chọn

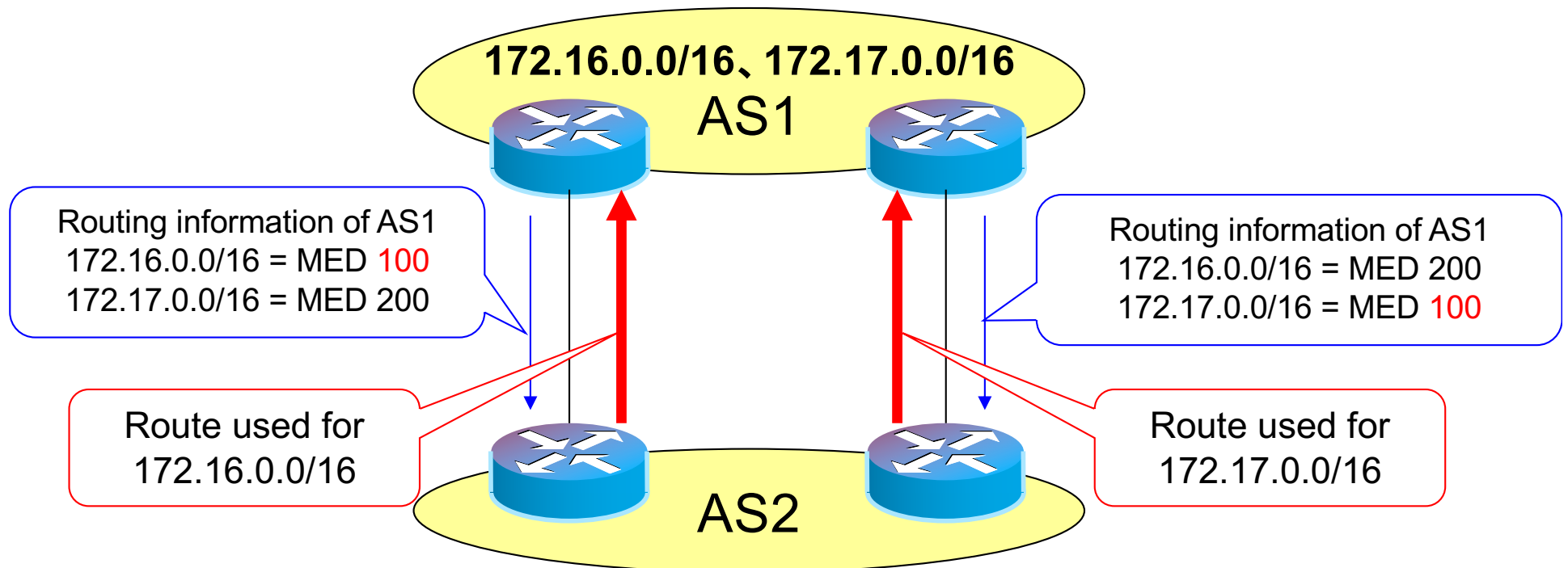
Chọn đường với MED

- Trong trường hợp 2 AS với nhiều link
- Chọn MED nhỏ hơn
- Áp dụng trong điều khiển lưu lượng



Phân tải với MED

- Đặt giá trị MED khác nhau cho mỗi đường
- Cũng điều khiển lưu lượng





25 YEARS ANNIVERSARY
SOICT

VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG
SCHOOL OF INFORMATION AND COMMUNICATION TECHNOLOGY

**Thank you for
your attentions!**

 soict.hust.edu.vn/  fb.com/groups/soict

