

Chương 4. Tầng mạng

Tổng quan
Giao thức IP
Định tuyến



1

1

1. Tổng quan về tầng mạng

Tổng quan
Giao thức IP
Định tuyến



2

2

Tầng mạng trên kiến trúc phân tầng



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 giữa các nút mạng qua môi trường liên mạng
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

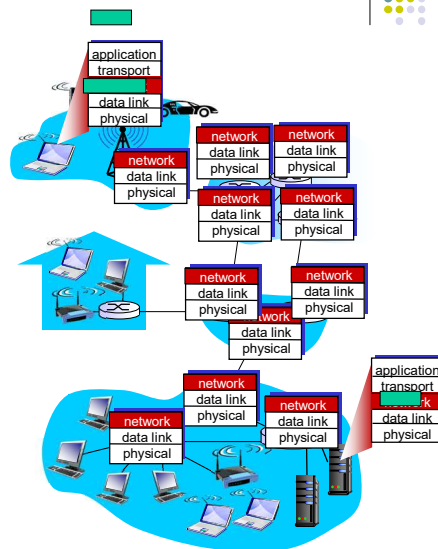
3

3

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



4

4

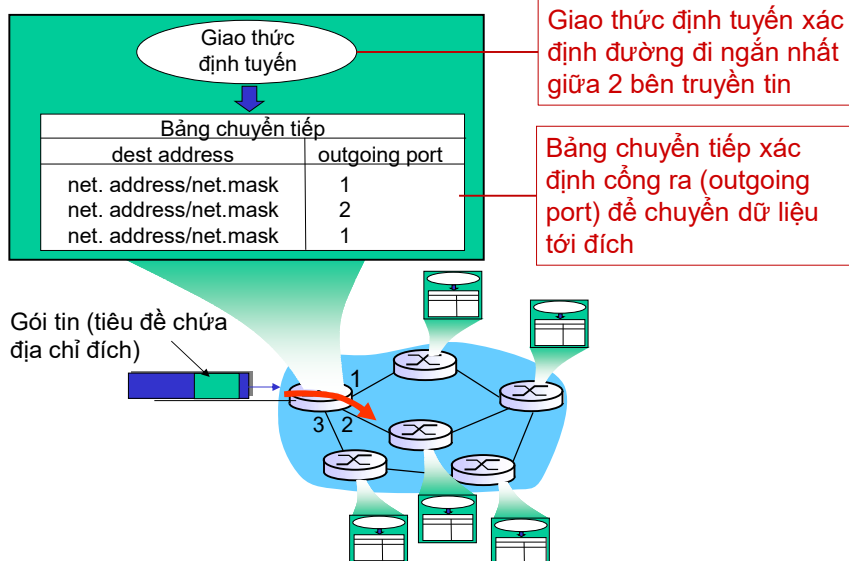
Chức năng chính

- **Định tuyến (Routing)**: Tìm tuyến đường (qua các nút trung gian) để gửi dữ liệu từ nguồn tới đích
- **Chuyển tiếp (Forwarding)**: Chuyển gói tin trên cổng vào tới cổng ra theo tuyến đường
- **Định địa chỉ (Addressing)**: Định danh cho các nút mạng
- **Đóng gói dữ liệu (Encapsulating)**: Nhận dữ liệu từ giao thức ở trên, thêm tiêu đề mang thông tin điều khiển quá trình truyền dữ liệu từ nguồn tới đích
- **Đảm bảo chất lượng dịch vụ (QoS)**: đảm bảo các thông số phù hợp của đường truyền theo từng dịch vụ

5

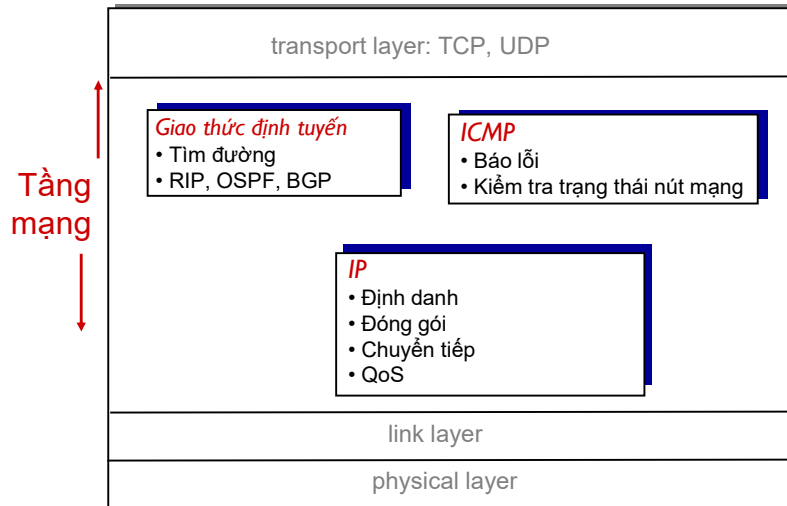
5

Định tuyến và chuyển tiếp



6

Các giao thức tầng mạng



7

7

2. Giao thức IP



8

8

2.1. Đặc điểm của giao thức IP



- Internet Protocol
- 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)
- Là giao thức được định tuyến (**routed protocol**)
 - Đòi hỏi phải có các giao thức định tuyến để xác định trước đường đi cho dữ liệu.
- Hiện nay có 2 phiên bản: IPv4 và IPv6

9

9

Chức năng cơ bản của IP



- Định địa chỉ: địa chỉ IP
- Đóng gói dữ liệu
 - Dồn kênh/Phân kênh
- Chuyển tiếp: theo địa chỉ IP (sẽ đề cập trong phần sau)
- Đảm bảo chất lượng dịch vụ

10

10

2.2. Địa chỉ IPv4

Lớp địa chỉ IP
CIDR – Địa chỉ IP không phân lớp
Mạng con và mặt nạ mạng
Các địa chỉ IP đặc biệt

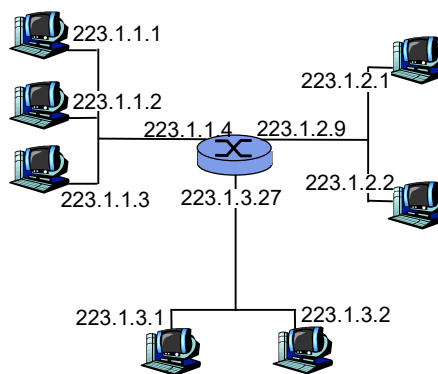


11

11

Địa chỉ IP (IPv4)

- **Địa chỉ IP:** Một số 32-bit để định danh cổng giao tiếp 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



223.1.1.1 = $\underbrace{11011111}_{223} \underbrace{00000001}_1 \underbrace{00000001}_1 \underbrace{00000001}_{12}$

12

Làm thế nào để cấp phát địa chỉ IP?



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

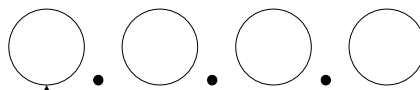
13

13

Biểu diễn địa chỉ IPv4



- Biểu diễn dạng thập phân có chấm



↑
Số nguyên 8 bit

Ví dụ:

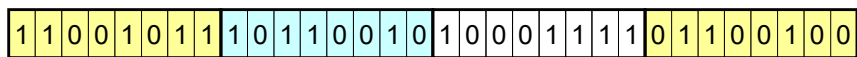
203.178.136.63

259.12.49.192

133.27.4.27

o
x
o

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



203

178

143

100

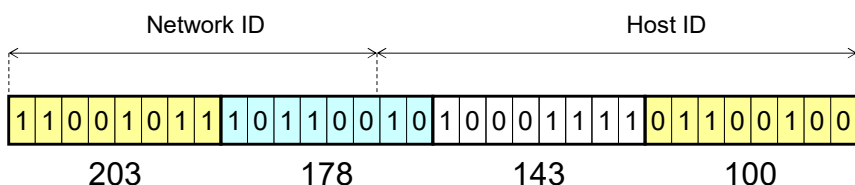
14

14

Địa chỉ IPv4



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



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

15

15

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
- Địa chỉ nhóm (Multicast address): định danh cho nhóm
- Địa chỉ máy trạm (Unicast Address)
 - Gán cho một cổng mạng

16

16

Ví dụ



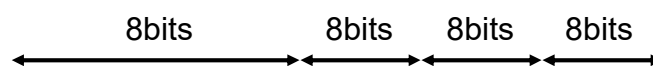
Xác định địa chỉ sau thuộc dạng nào (Địa chỉ mạng/Địa chỉ máy trạm/Địa chỉ quảng bá)?

- (1) 192.168.1.128. NetworkID = 24 bit
- (2) 192.168.1.128. NetworkID = 25 bit
- (3) 10.1.63.255. NetworkID = 18 bit
- (4) 10.1.63.255. NetworkID = 17 bit

17

17

Phân lớp địa chỉ IP (Classful Addressing)



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

	# of network	# of hosts/1 net
Class A	128	$2^{24} - 2$
Class B	16384	65534
Class C	2^{21}	254

18

18

Phân lớp địa chỉ



- Lớp A: 0*
 - NetworkID: 8 bit → HostID: 24 bit
 - Địa chỉ mạng: a.0.0.0
- Lớp B: 10*
 - NetworkID: 16 bit
 - Địa chỉ mạng: a.b.0.0
- Lớp C: 110*
 - NetworkID: 24 bit
 - Địa chỉ mạng: a.b.c.0

19

19

Hạn chế của việc phân lớp địa chỉ



- Lãng phí không gian địa chỉ
 - Việc phân chia cứng thành các lớp (A, B, C, D, E) làm hạn chế việc sử dụng toàn bộ không gian địa chỉ

Cách giải quyết ...

- 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)

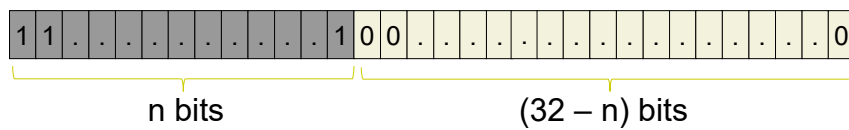
20

20

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:
 - Prefix: /n. Ví dụ: /27
 - n: Số bit của NetworkID
 - Nhị phân:

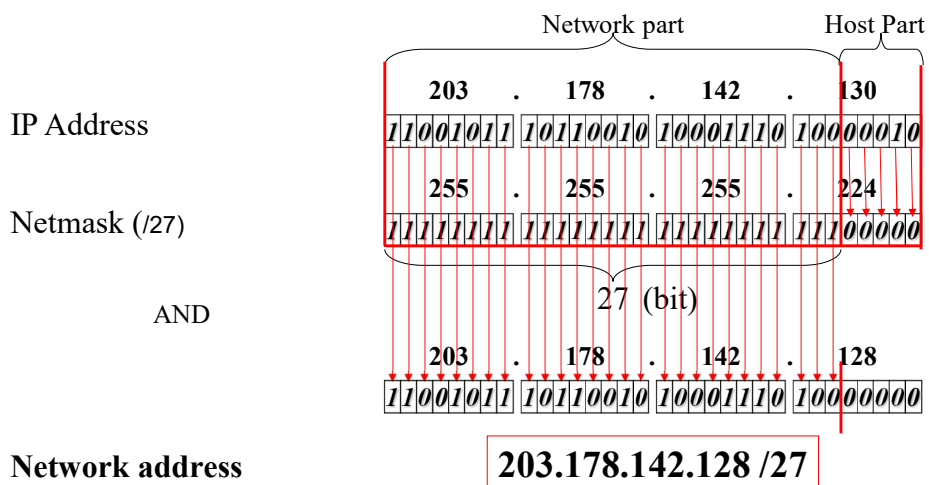


- Ví dụ: /27 = 11111111 11111111 11111111 11100000
- Thập phân có chấm: 255.255.255.224

21

21

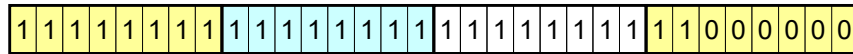
Cách tính địa chỉ mạng



22

22

Mặt nạ mạng và kích thước mạng



255

255

255

192

- Kích thước
 - Theo lũy thừa 2
- [RFC1878](#)
- Trong trường hợp /26
 - Phần máy trạm = 6 bits
 - $2^6 = 64$
 - Octet 4 cho máy trạm:
 - 1 – 62
 - 65 – 126
 - 129 – 190
 - 193 – 254

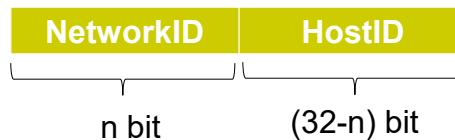
23

23

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):

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

- Mạng có X nút $\rightarrow /n = ?$

$$2^{32-n} - 2 \geq X$$

24

24

Địa chỉ IP và mặt nạ mạng



- 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

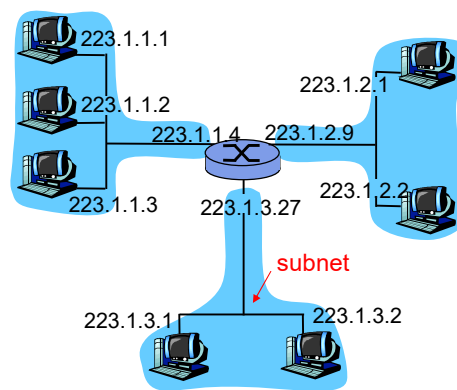
25

25

Mạng con - subnet



- Là một phần của một mạng nào đó
 - ISP thường được gán một khối địa chỉ IP
 - Một vài mạng con sẽ được tạo ra
- Tạo subnet như thế nào
 - Sử dụng một mặt nạ mạng dài hơn



Mạng với 3 mạng con

26

26

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

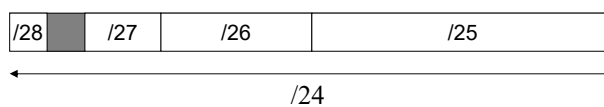
Mặt nạ mạng con = Mặt nạ mạng cũ + Số bit SubnetID
 Chia thành 4 mạng con? → Mặt nạ mạng con: /26

27

27

Ví dụ: Chia làm 4 subnets

- Mạng với mặt nạ /24
- Cần tạo 4 mạng con
 - Mạng với 14 máy tính /28
 - Mạng với 30 máy tính /27
 - Mạng với 31 máy tính /26
 - Mạng với 70 máy tính /25



28

28

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)

29

29

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

30

30

2.3. DHCP



31

31

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

32

32

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

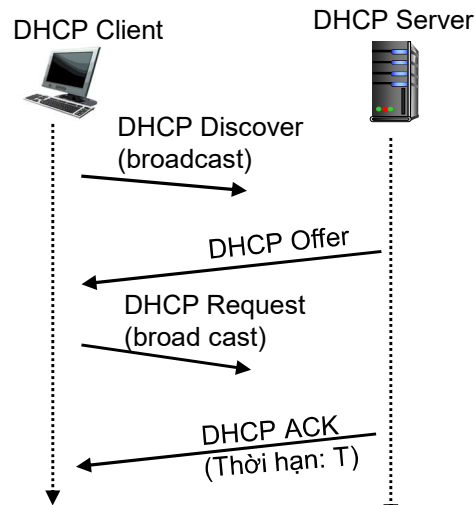
33

33

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

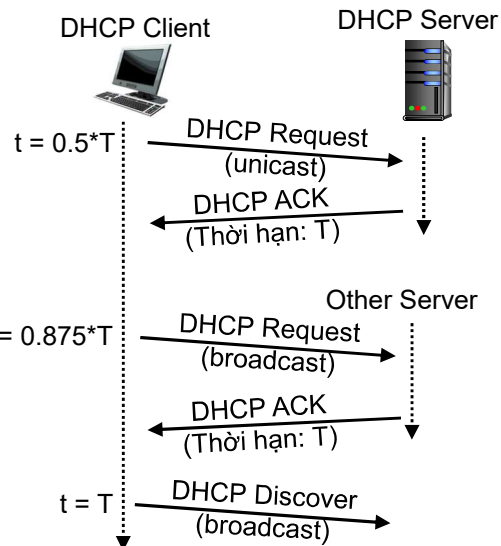


34

34

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

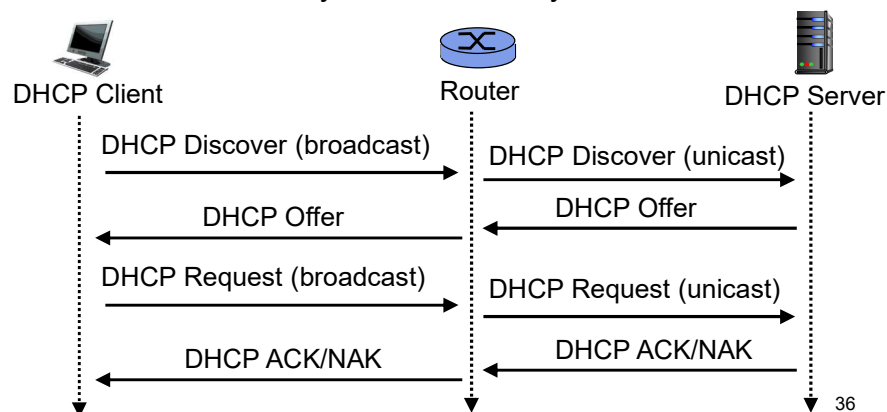


35

35

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



36

36

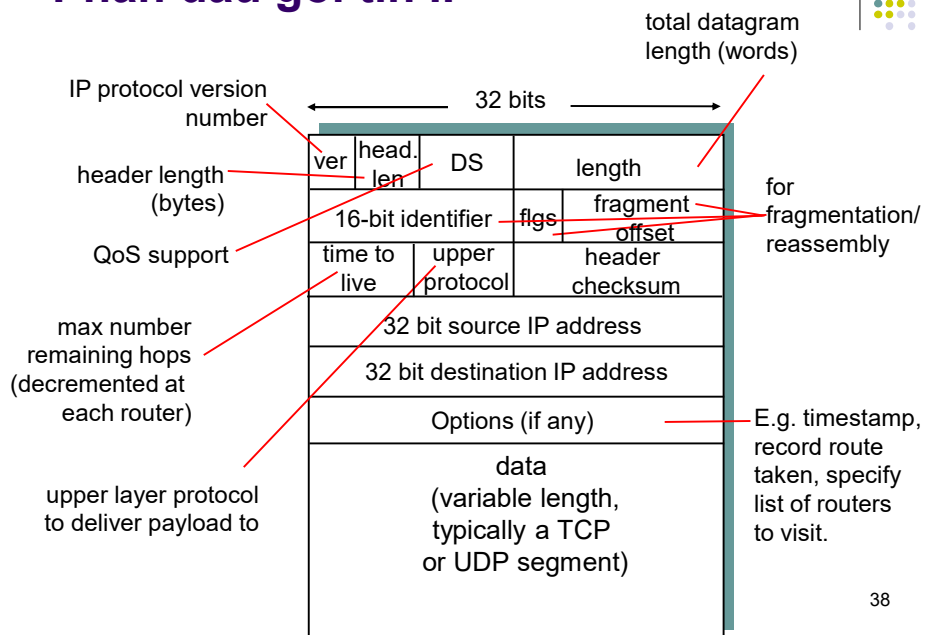
2.4. Khuôn dạng gói tin IP



37

37

Phần đầu gói tin IP



38



IP header (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

39

39



IP header (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)

40

40



IP header (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

41

41



IP header và trường Protocol

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

Protocol:

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

Có thể xem số hiệu giao thức tại

/etc/protocols
C:\WINDOWS\system32\drivers\etc\protocols

42

42

IP header (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

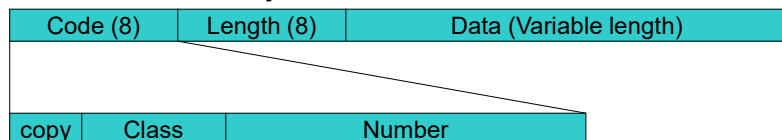
43

43

IP header (5)



- Options: Dùng để thêm vào các chức năng mới.
 - Có thể tới 40 bytes



Copy:

- 0: copy only in first fragment
- 1: copy into all fragment

Class:

- 00: Datagram control
- 01: Reserved
- 10: Debugging and measurement
- 11: Reserved

Number:

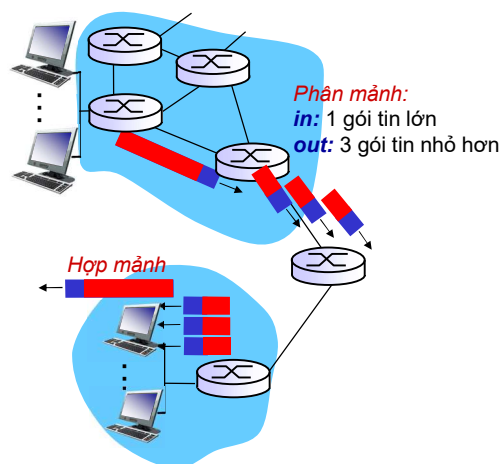
- 00000: End of option
- 00001: No operation
- 00011: Loose source route
- 00100: Timestamp
- 00111: Record route
- 01001: Strict source route

44

44

Phân mảnh gói tin (1)

- Đường truyền có một giá trị MTU (Kích thước đơn vị dữ liệu tối đa)
- 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



45

45

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

46

46

Phân mảnh (3)



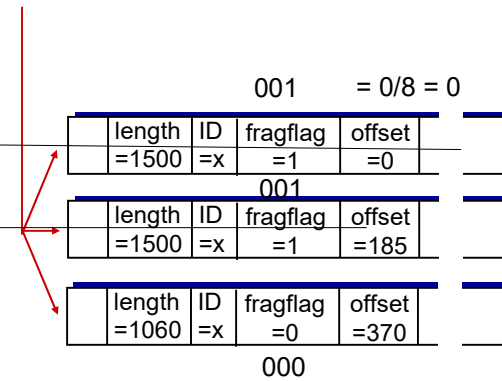
Ví dụ:

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

1480 bytes in data field

offset = 1480/8

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



47

47

2.5. Chuyển tiếp gói tin IP



48

48

Chuyển tiếp gói tin IP

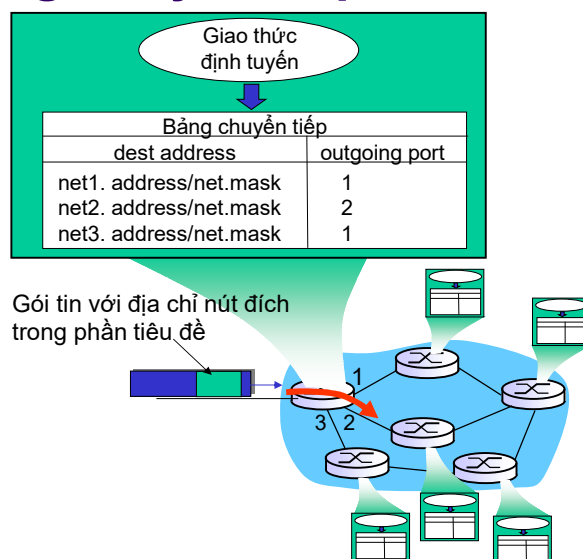


- Mỗi nút mạng sử dụng bảng chuyển tiếp (Forwarding Table)
 - Là một phần của bảng định tuyến (Routing Table)
- Các thông tin:
 - Đích đến (Destination): Địa chỉ mạng/Mặt nạ (/n)
 - Sử dụng địa chỉ 0.0.0.0/0 đại diện cho một đích bất kỳ chưa biết → lỗi ra mặc định
 - Cổng ra (Outgoing port): địa chỉ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

49

49

Bảng chuyển tiếp



50

50

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 ... Outgoing port ... Realtek RTL8109 Family PCI Fast Ethernet NIC -

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

51

51

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



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

52

52

Kết hợp đường đi



- Điều kiện: Đường đi có chung Outgoing Port và k bit đầu tiên (prefix) trong phần NetworkID của địa chỉ đích giống nhau
- Mục đích: giảm kích thước bảng chuyển tiếp, tăng tốc quá trình tìm kiếm đường ra

Destination	Outgoing Port
1.1.0.0 /16	Se0/2
1.2.0.0 /16	Se0/2
1.3.0.0 /16	Se/01
171.8.0.0 /19	Se0/1
171.8.32.0 /19	Se0/1

Destination	Outgoing Port
1.1.0.0 /16	Se0/2
1.2.0.0 /16	Se0/2
1.3.0.0 /16	Se/01
171.8.0.0 /18	Se0/1

53

53

Nguyên tắc tìm kiếm



- 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 (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

54

54

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 DA của gói tin. Mặt nạ hóa địa chỉ đích của gói tin với các mặt nạ của mạng đích(DA **and** Netmask)
- 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 cổng ra 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.

55

55

Ví dụ



- Cho bảng chuyển tiếp của một router trong bảng 1. Hãy cho biết router xử lý như thế nào khi nhận được các gói tin có các thông số trong bảng 2.

Destination	Outgoing Port
45.45.160.0 /19	E3
45.45.64.0 /19	E1
45.45.96.0 /20	E3
45.64.0.0 /10	E2
45.45.96.128 /26	E1
0.0.0.0 /0	E1

Destination Address	TTL
45.45.100.64	0
45.45.100.64	5
45.45.145.144	15

56

56

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
- Network Address Translation
- 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

57

57

Static NAT

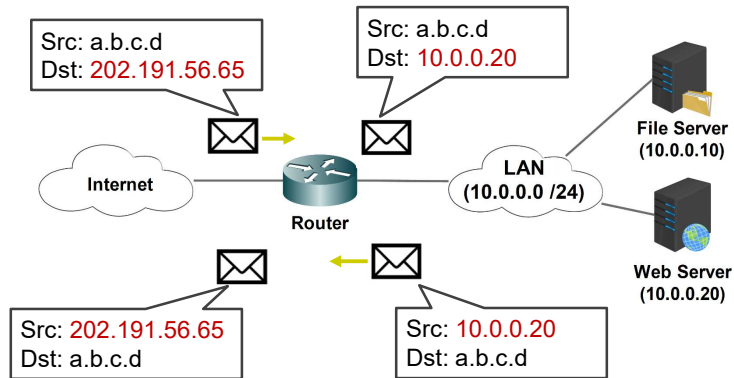


- Là dạng NAT đơn giản nhất
- Một địa chỉ riêng(Private IP) được ánh xạ sang một địa chỉ công cộng(Public IP) xác định
- NAT router sẽ lưu một bảng chuyển đổi trong bộ nhớ
 - Bảng này ánh xạ các địa chỉ nội bộ bên trong với địa chỉ bên ngoài Internet
- Thường sử dụng cho các server nằm trong mạng LAN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

58

58

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

59

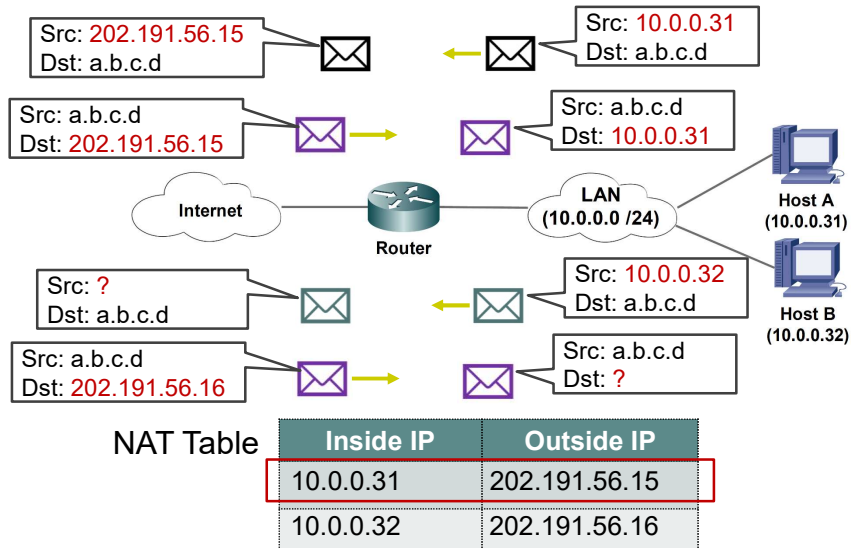
59

Dynamic NAT

- NAT router tự động ánh xạ một dải các địa chỉ riêng với một dải các địa chỉ công cộng
- Không cần cấu hình cụ thể 1 địa chỉ riêng được chuyển đổi thành 1 địa chỉ công cộng nào
- Bất kỳ địa chỉ riêng nào sẽ được chuyển đổi tự động sang 1 trong các địa chỉ công cộng đang sẵn có

60

Dynamic NAT – Ví dụ



61

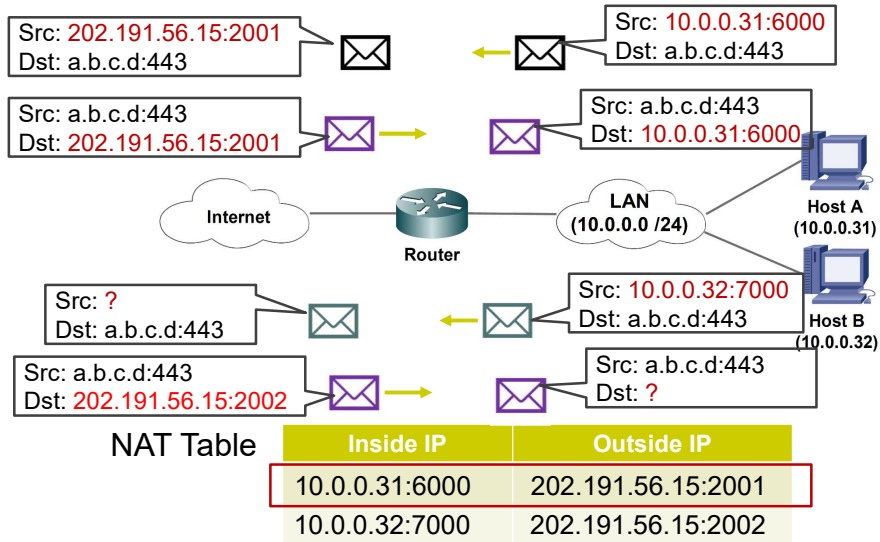
61

Port Address Translation

- Còn được biết là NAT **overloading**
- Là một dạng đặc biệt của Dynamic NAT để tận dụng tổ hợp không gian địa chỉ IPv4
- Cho phép nhiều địa chỉ riêng ánh xạ với một địa chỉ công cộng đã đăng ký kèm các số hiệu cổng khác nhau

62

Dynamic NAT – Ví dụ



63

63

2.6. Giao thức IPv6



64

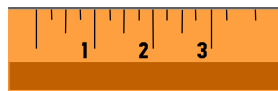
64

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

65

65

Cải tiến trên IPv6(tiếp)



- Tăng tốc độ
 - Khuôn dạng header đơn giản hơn
 - Ít trường hơn
 - Độ dài cố định
 - Bỏ checksum
 - 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

66

66

Địa chỉ IPv6



67

67

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

68

68

Cấu trúc địa chỉ IPv6



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

69

69

Phân vùng địa chỉ

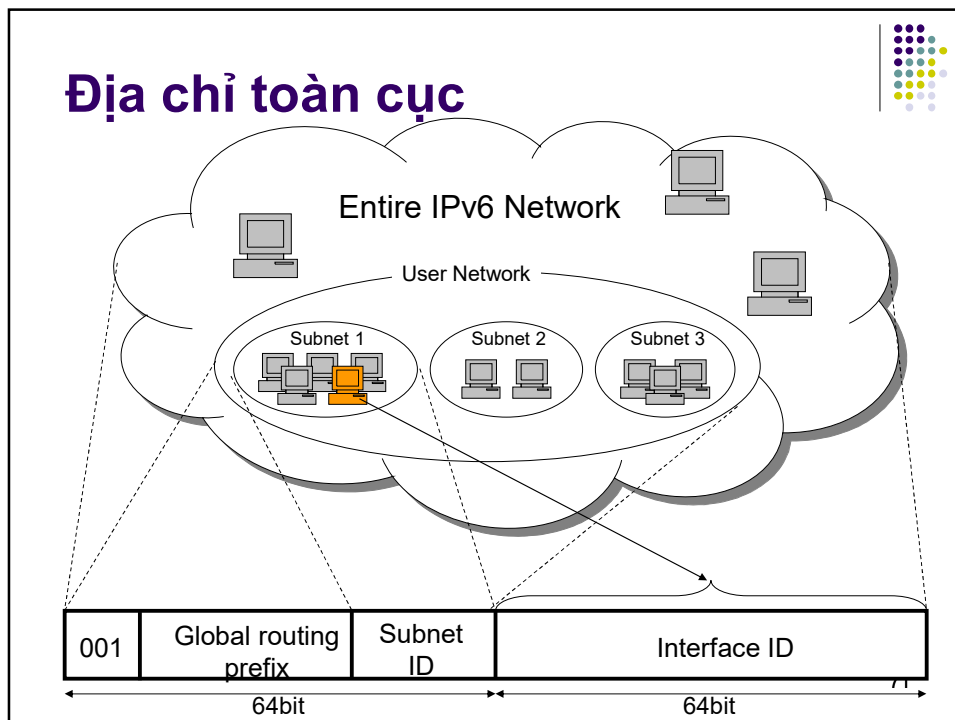


- Địa chỉ toàn cục: Network prefix có
 - Global routing prefix: 48 bit với 3 bit đầu là 001
 - SubnetID: 16 bit
- Địa chỉ liên kết nội bộ: Network prefix có
 - 10 bit đầu là 1111 1110 10
 - 54 bit còn lại mang giá trị 0
- Địa chỉ cục bộ: Network prefix có
 - 10 bit đầu là 1111 1110 11
 - 38 bit kế tiếp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

70

70

Địa chỉ toàn cục



71

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

72

72

Khuôn dạng gói tin IPv6



73

73

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

74

74

Khuôn dạng gói tin IPv6(tiếp)



- 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

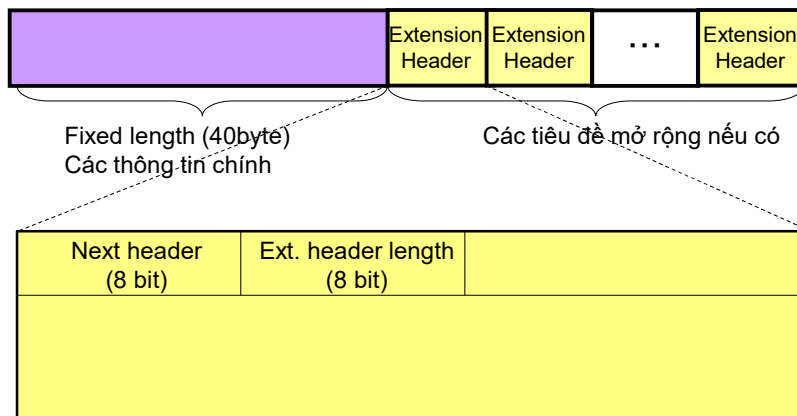
75

75

IPv6 header



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



76

76

3. Internet Control Message Protocol

Tổng quan
Khuôn dạng gói tin
Ping và Traceroute



77

77

3.1. 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:



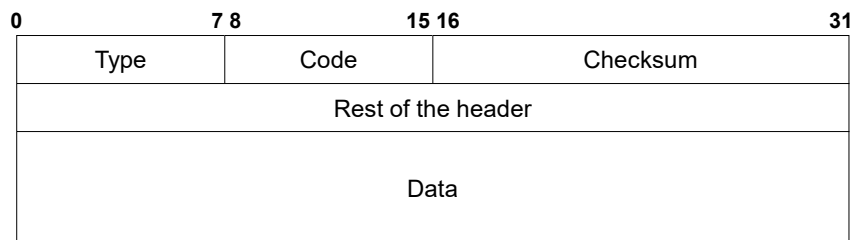
78

78

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



79

79

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	

80

80

3.2. ICMP và các công cụ debug



- ICMP luôn hoạt động song song trong suốt với người sử dụng
- NSD có thể sử dụng ICMP thông qua các công cụ debug
 - ping
 - traceroute

81

81

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)

82

82

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
```

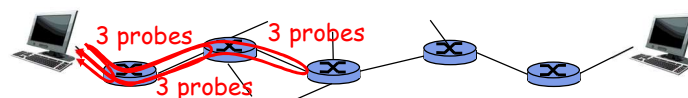
83

83

tracert



- Cú pháp:
 - Windows: `tracert host_address`
 - Linux: `traceroute host_address`
- Bên gửi truyền gói tin UDP 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



84

84

traceroute



Điều kiện kết thúc

- Gói tin đến được đích
- Đích trả về gói tin ICMP “time exceeded” (type 11)
- Khi nguồn nhận được gói tin ICMP này sẽ dừng lại

85

85

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  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]
```

Tại sao trở có đột biến?

```
Trace complete.
```

86

86

4. Giao thức ARP



87

87

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

88

88

Hoạt động của ARP

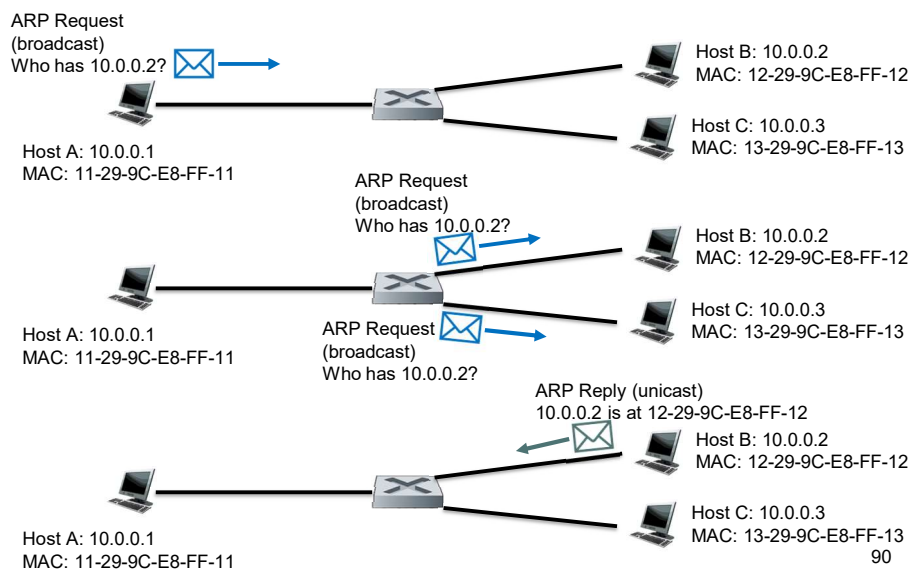


- Mỗi nút trong mạng LAN sử dụng bảng ARP Table:
 - Ánh xạ <Địa chỉ IP, Địa chỉ MAC, TTL>
 - TTL: Thời gian giữ ánh xạ trong bảng
- Khi tìm địa chỉ MAC của một nút khác, nút mạng gửi quảng bá gói tin ARP Request lên trên mạng.
 - Gói tin ARP Request chứa địa chỉ IP của nút cần tìm kiếm
- Nút mạng mang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

89

89

Hoạt động của ARP – Ví dụ

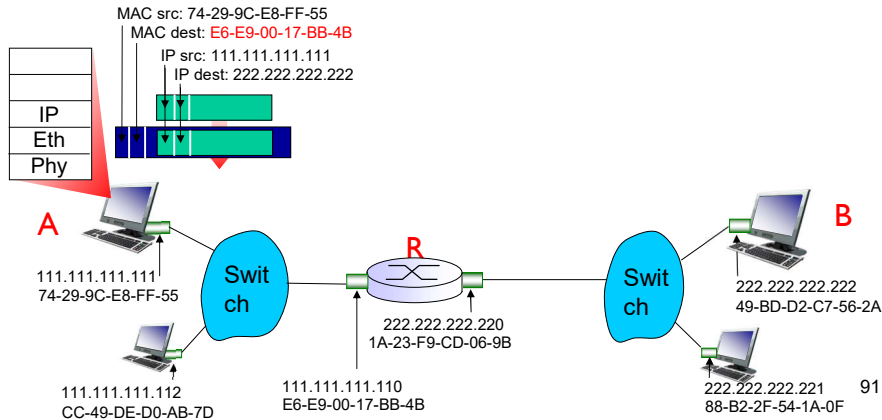


90

Chuyển tiếp dữ liệu tới mạng khác

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

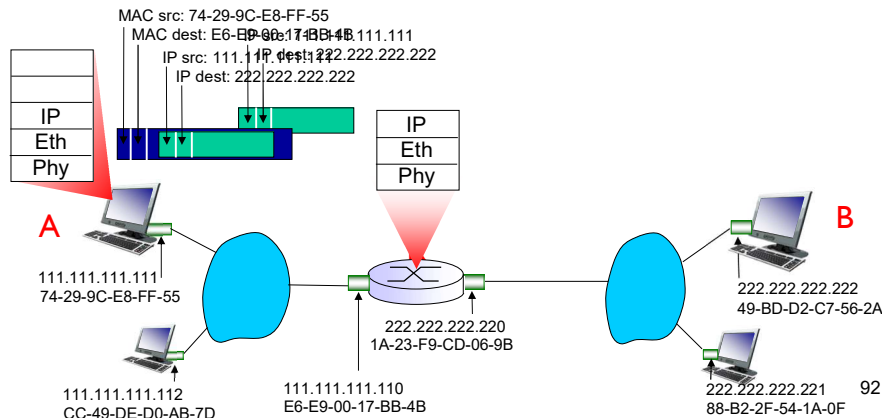
- A soạn một gói tin IP với địa chỉ nguồn là A, địa chỉ đích là B
- Gói tin chuyển xuống tầng liên kết dữ liệu: đóng gói thành khung tin tầng 2 với địa chỉ MAC nguồn là A, địa chỉ MAC đích là R



91

Chuyển tiếp dữ liệu tới mạng 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

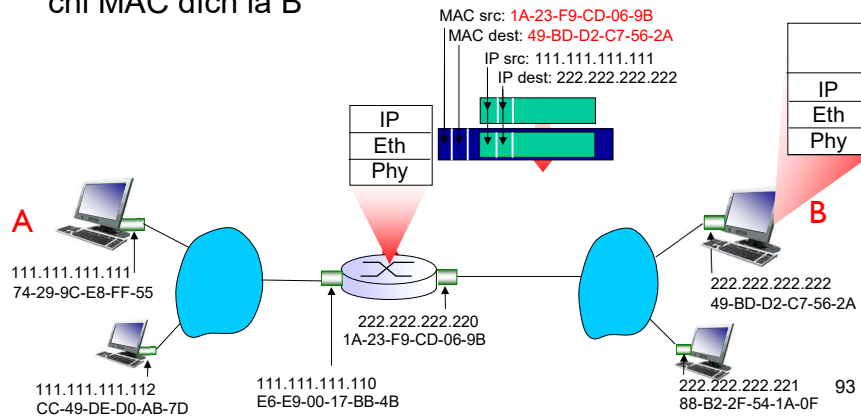


92

Chuyển tiếp dữ liệu tới mạng 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



93

5. Định tuyến



94

94

5.1. Khái niệm cơ bản về định tuyến

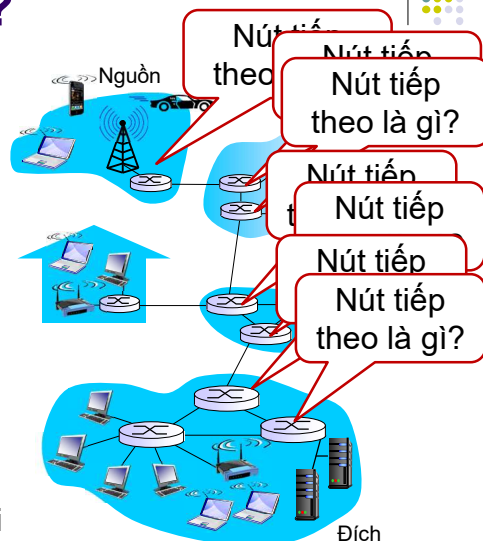


95

95

Định tuyến là gì?

- *Tim đường đi* tốt nhất để *chuyển tiếp* dữ liệu tới nút đích
- Trong mạng IP chuyển mạch gói:
 - Các gói tin được chuyển tiếp độc lập
 - Định tuyến trên từng chặng: Mỗi nút không chỉ ra toàn bộ các chặng trên đường đi tới đích



96

96

Các thành phần của định tuyến



- Thông tin định tuyến: Thông số của đường truyền được sử dụng làm độ đo tính toán chi phí đường đi
- Bảng định tuyến: Lưu thông tin đường đi đã tìm được tới các mạng đích
- Giải thuật, giao thức định tuyến: cách thức tìm đường đi và trao đổi thông tin định tuyến giữa các nút mạng

97

97

Bộ định tuyến (router)



- 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 định tuyến
- Có nhiều cổng kết nối mạng
- Phù hợp với nhiều dạng lưu lượng và phạm vi của mạng

98

98

Một số ví dụ...

Router ngoại vi

- BUFFALO BHR-4RV
- PLANEX GW-AP54SAG
- YAMAHA RTX-1500
- Cisco 2600

Router mạng trực

- Cisco CRS-1

Router cỡ trung

- Hitachi GR2000-1B
- Juniper M10
- Foundry Networks NetIron 800
- Cisco 3700

<http://www.cisco.com.vn>
<http://www.juniper.net/>
<http://www.buffalotech.com>

99

5.2. Bảng định tuyến

Destination	Outgoing Port	Next hop	Cost
2.2.0.0 /16	e1	-	0
4.0.0.0 /8	e0	-	0
1.0.0.0 /8	e0	4.0.0.1	1
3.0.0.0 /14	e0	4.0.0.3	1

- Destination : địa chỉ mạng đích
 - Định tuyến classless: Sử dụng địa chỉ không phân lớp
 - Định tuyến classful: Sử dụng địa chỉ phân lớp
- Outgoing Port : cổng ra cho gói tin để tới mạng đích
- Next hop : địa chỉ cổng nhận gói tin của nút kế tiếp
- Cost : chi phí gửi gói tin từ nút đang xét tới đích

100

100

Vấn đề cập nhật bảng định tuyến



- Sự thay đổi cấu trúc mạng: thêm mạng mới, một nút mạng bị mất kết nối
- Sự cần thiết phải cập nhật bảng định tuyến
 - Cho tất cả các nút mạng (về lý thuyết)
 - Thực tế, chỉ một số nút mạng phải cập nhật
- Làm thế nào để cập nhật ?
 - Định tuyến tĩnh: Các mục trong bảng định tuyến được sửa đổi thủ công bởi người quản trị
 - Định tuyến động: Tự động cập nhật bảng định tuyến các giao thức định tuyến

101

101

Định tuyến tĩnh

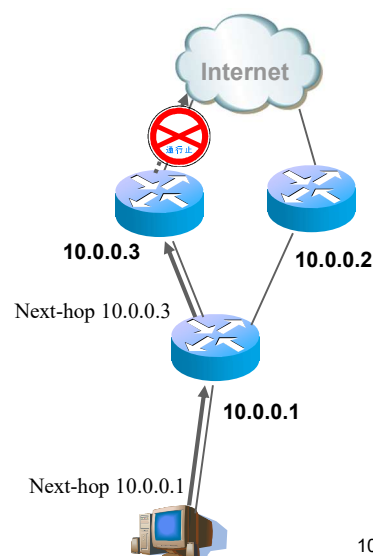


- Khi có sự cố:
 - Không thể nối vào Internet kể cả khi có tồn tại đường đi dự phòng
 - Người quản trị mạng cần thay đổi

Bảng định tuyến của 10.0.0.1 (1 phần)

Prefix	Next-hop
0.0.0.0/0	10.0.0.3

Kết nối bị lỗi



102

102

Định tuyến động

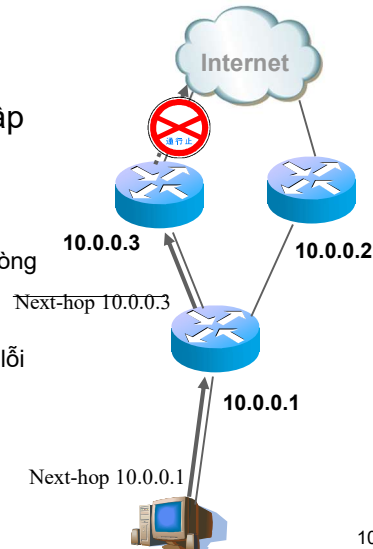
- Khi có sự cố:
 - Đường đi thay thế được cập nhật một cách tự động

Bảng định tuyến 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



103

103

Tập trung hay phân tán?

- Tập trung
 - Thu thập thông tin kết nối của toàn bộ mạng
 - Sử dụng các giải thuật tìm đường đi trên đồ thị (thường sử dụng giải thuật dạng link-state)
 - Phân bổ bảng định tuyến từ nút trung tâm tới các nút
- Phân tán
 - Mỗi nút tự xây dựng bảng chọn đường riêng
 - Giải thuật định tuyến: Link-state hoặc distance-vector
 - Được sử dụng phổ biến trong thực tế

104

104

5.3. Các giải thuật và giao thức định tuyến

Giải thuật Dijkstra và Bellman-Ford
Giao thức dạng link-state và dạng distance-vector



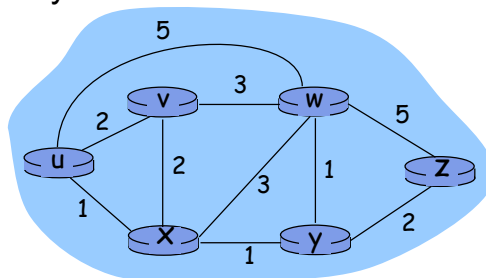
105

105

Biểu diễn mạng bởi đồ thị



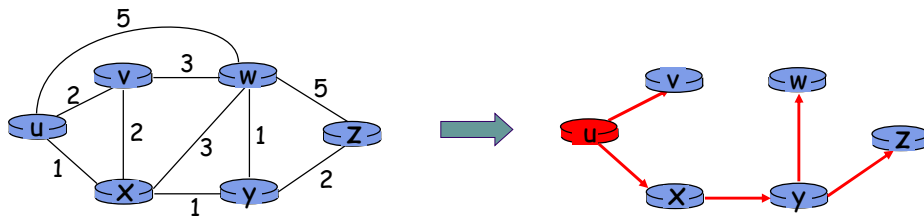
- Đồ thị với các nút (bộ định tuyến) và các cạnh (liên kết)
- Chi phí cho việc sử dụng mỗi liên kết $c(x,y)$
 - Băng thông, độ trễ, chi phí, mức độ tắc nghẽn...
- Giải thuật định tuyến: Xác định đường đi ngắn nhất giữa hai nút bất kỳ



106

106

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

107

107

5.3.1. Giải thuật dạng distance-vector

Phương trình Bellman-Ford (quy hoạch động)

Định nghĩa

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

$c(x,v)$: chi phí từ x tới hàng xóm v

Giải phương trình:

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

cho tất cả các v là hàng xóm của x

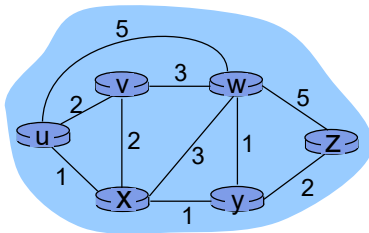
108

108

Minh họa Bellman-Ford Eq.



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



B-F eq. cho ta biết:

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

$$= \min \{ 2 + 5, 1 + 3, 5 + 3 \} = 4$$

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

109

109

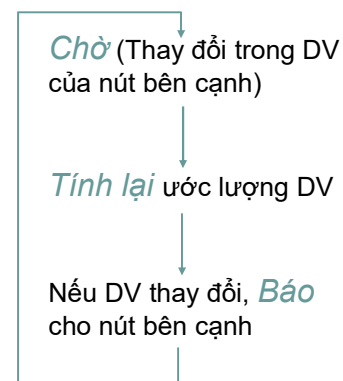
Giải thuật dạng distance-vector (2)



ý tưởng cơ bản:

- DV: Vector khoảng cách, tạm coi là đường đi ngắn nhất của từ một nút tới nút khác
- Mỗi nút tính toán lại DV nếu:
 - Liên kết thay đổi
 - Nhận được DV từ hàng xóm
- Thông báo cho hàng xóm nếu DV thay đổi
 - Hàng xóm sẽ thông báo cho hàng xóm của chúng

Mỗi nút:



110

110

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

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

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

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

thời gian →

111

111

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

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

chi phí tới			
	x	y	z
x	0	2	3
y	2	0	1
z	3	1	0

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

nút y

chi phí tới			
	x	y	z
x	∞	∞	∞
y	2	0	1
z	∞	∞	∞

chi phí tới			
	x	y	z
x	0	2	7
y	2	0	1
z	7	1	0

chi phí tới			
	x	y	z
x	0	2	3
y	2	0	1
z	3	1	0

nút z

chi phí tới			
	x	y	z
x	∞	∞	∞
y	∞	∞	∞
z	7	1	0

chi phí tới			
	x	y	z
x	0	2	7
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

thời gian →

112

112

Vấn đề đếm tới vô cùng



- Count-to-infinity
- Xảy ra khi chi phí một liên kết tăng lên so với ban đầu
- Mỗi router không biết đường đi đến router khác lại đi qua chính nó
- Sau mỗi lần trao đổi và cập nhật bảng định tuyến, khoảng cách ước lượng sẽ tăng dần đến vô cùng

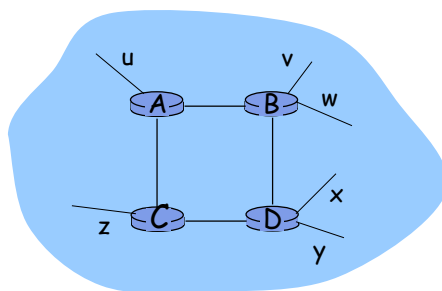
113

113

RIP (Routing Information Protocol)



- IGP
- RIP v.1, phiên bản mới RIP v.2
- Giao thức dạng vector khoảng cách
- Chọn đường đi theo số nút mạng đi qua



Từ nút A:

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

114

114

RIP: Trao đổi thông tin



- Trao đổi bảng chọn đường
- 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

115

115

RIP: Đề tranh lỗi lặp vô hạn



- Giới hạn số hop tối đa
 - v1 : 15
 - v2 : 30
- “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 (hoặc 31)
 - Liên kết chuyển sang trạng thái hold-down

116

116

5.3.2. Giải thuật dạng link-state



- Mỗi nút thu thập thông tin từ các nút khác để xây dựng topo của mạng
- Áp dụng giải thuật tìm đường đi ngắn nhất tới mọi nút trong mạng
- Khi có sự thay đổi bảng định tuyến, mỗi nút gửi thông tin thay đổi cho tất cả các nút khác trong mạng.

117

117

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

118

118

Các thủ tục



- **Init():**
Với mỗi nút v , $d[v] = \infty$, $p[v] = \text{NIL}$
 $d[s] = 0$
- **update(u, v)**, trong đó (u, v) là một cạnh nào đó của G
if $d[v] > d[u] + c(u, v)$ then
 $d[v] = d[u] + c(u, v)$
 $p[v] = u$

119

119

Dijkstra's Algorithm



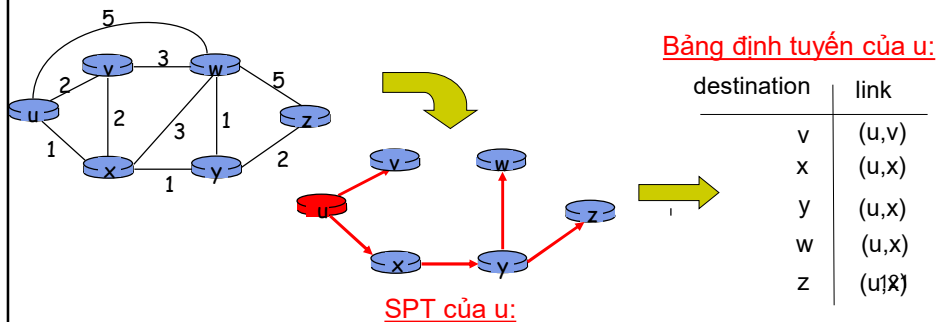
1. **Init()** ;
2. $T = \emptyset$;
3. **Repeat**
4. $u: u \notin T \mid d(u)$ là bé nhất ;
5. $T = T \cup \{u\}$;
6. for all $v \in \text{neighbor}(u)$ và $v \notin T$
7. **update**(u, v) ;
8. **Until** $T = V$

120

120

Dijkstra's algorithm: 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					



121

OSPF: Open Shortest Path First

- Tài liệu mô tả: RFC2740
- Thông tin định tuyến: Thông tin về trạng thái liên kết - LSA (link state advertisement)
- Hoạt động cơ bản:
 - Quảng bá thông tin LSA khi trạng thái liên kết thay đổi
 - Thu thập thông tin LSA và xây dựng topology của mạng.
 - Áp dụng giải thuật Dijkstra tìm đường đi ngắn nhất

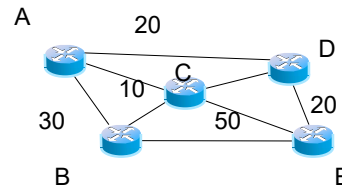
122

122

Thông tin chọn đường



- 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



123

123

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

124

124

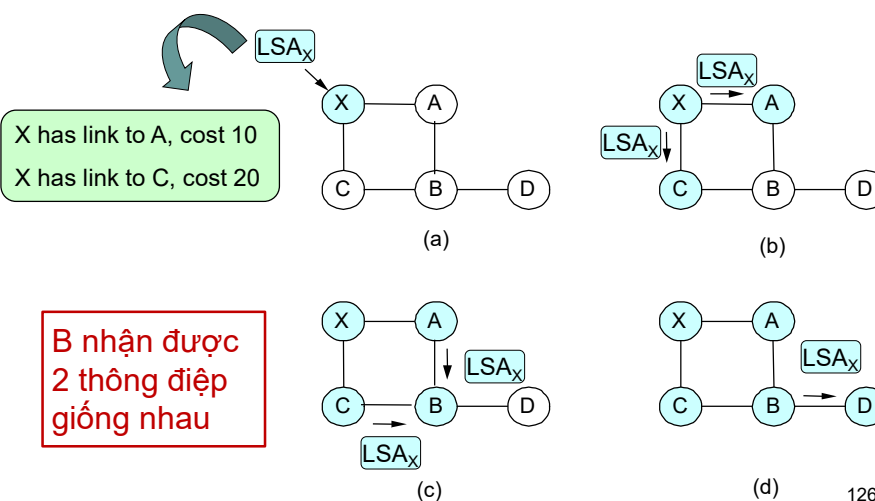
Chi phí mặc định của OSPF

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

125

125

Quảng bá thông tin LSA-Flooding



126

126



So sánh các giải thuật LS và DV

Thông điệp trao đổi

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

127

127



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 chọn đường	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

128

128

5.4. Định tuyến liên vùng



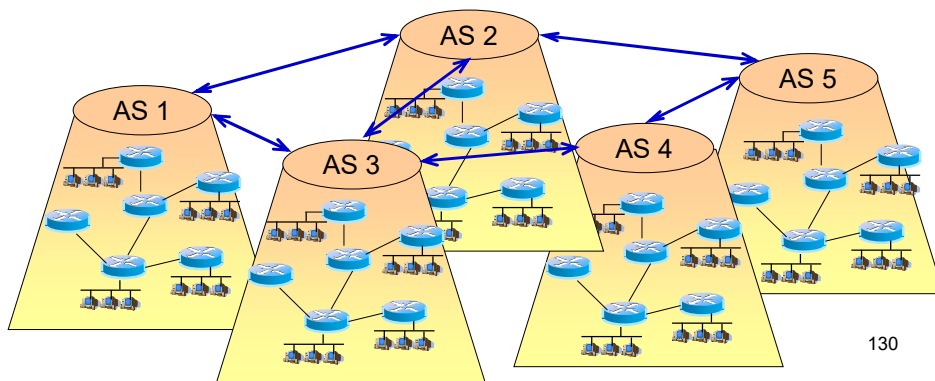
129

129

Phân cấp trong định tuyến



- Internet = Mạng của các mạng
- Các mạng được quản lý bởi các tổ chức khác nhau, có chính sách rất khác nhau, trong đó có định tuyến
- Mỗi mạng như vậy có thể gọi là một hệ tự trị - Autonomous System (AS)



130

130

Khái niệm hệ tự trị - AS



- Góc nhìn định tuyến: 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
.....

131

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

131

Phân cấp giao thức định tuyến

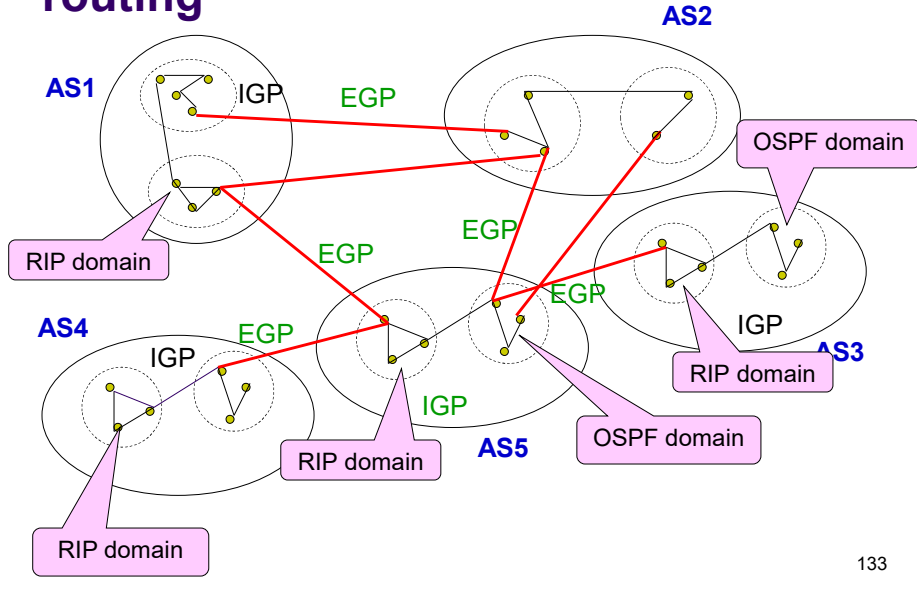


- **Trong một hệ tự trị:** Giao thức định tuyến nội vùng
 - IGP: *Interior Gateway Protocol*
 - RIP: Routing Information Protocol
 - OSPF: Open Shortest Path First
 - IS-IS, IGRP, EIGRP (Cisco)...
- **Giữa các hệ tự trị:** Giao thức định tuyến liên vùng
 - EGP: *Exterior Gateway Protocol*
 - BGP (v4):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

132

132

Extra-domain và Inter-domain routing



133

5.4.1. Giới thiệu chung về BGP



134

134

Yêu cầu mới với các giao thức định tuyến liên vùng



- Mỗi AS có chính sách riêng về định tuyến
 - Nhà cung cấp X không muốn truyền dữ liệu của họ qua mạng nhà cung cấp Y
 - Nhà cung cấp X không muốn dữ liệu của nhà cung cấp Z truyền qua hệ thống mạng của họ
 - Không thể thể hiện những chính sách này trong khái niệm “đường đi ngắn nhất”
- Mỗi AS muốn “tự trị”: tự chọn các giao thức định tuyến nội vùng và các chính sách định tuyến đi kèm
- Mỗi AS muốn tính “riêng tư”: tự chọn topology, các chính sách điều phối hoạt động

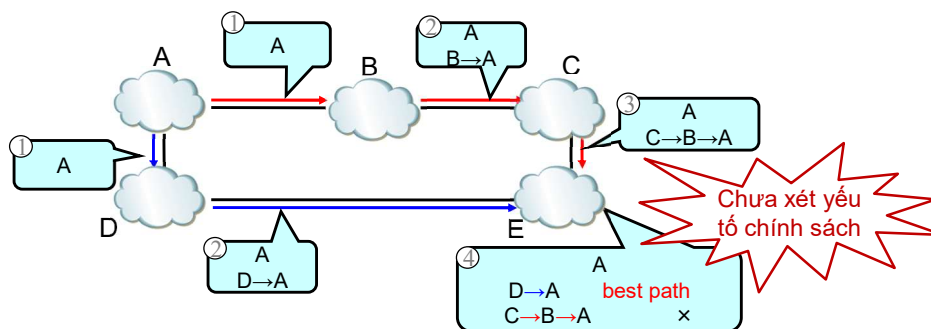
135

135

Chọn giải thuật định tuyến nào?



- Giữa các AS nên dùng giao thức nào?
 - Khó có một chính sách và đơn vị chi phí chung
 - LS: Chi phí không đồng nhất, CSDL quá lớn
 - DV: Mạng quá rộng, khó hội tụ
- Giải pháp: BGP chọn đường theo path-vector



136

136

BGP – Border Gateway Protocol

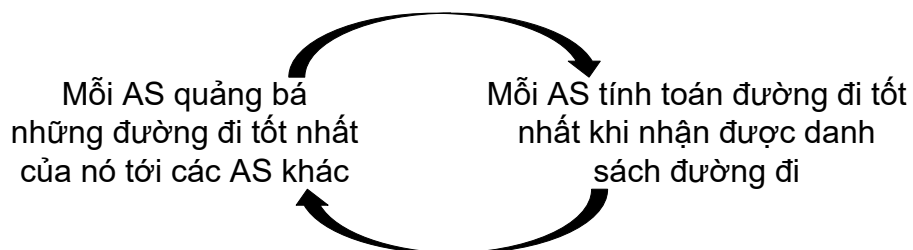


- Yếu tố gắn kết của Internet, kết nối các hệ tự trị
- Trao đổi thông tin đường đi NLRI (Network Layer Reachability Information)
 - Cho phép một AS biết được thông tin đi đến AS khác
 - Gửi thông tin này vào bên trong AS đó
 - Xác định đường đi tốt nhất dựa trên thông tin đó và các chính sách chọn đường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

137

137

BGP: Ý tưởng tìm đường



Có vẻ như BGP sử dụng lại ý tưởng nào đó!

138

138

BGP sử dụng ý tưởng của DV



- Quảng bá thông tin về đường đi tới đích
- Không chia sẻ thông tin về topology của mạng
- Lặp liên tục cho đến khi hội tụ (tìm thấy đường đi tốt nhất)
- Với 4 điểm khác biệt

139

139

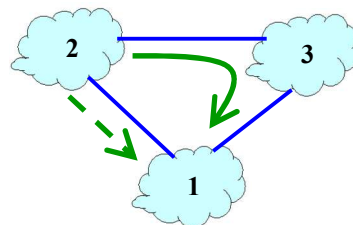
BGP khác DV (1)

Không chọn đường đi ngắn nhất



- Đường đi ngắn nhất (qua ít AS nhất) là mục tiêu rất quan trọng trong định tuyến liên vùng (vì sao?) nhưng không phải là ưu tiên hàng đầu.
- BGP ưu tiên chọn đường theo chính sách trước

Nút 2 có thể chọn
"2, 3, 1" thay vì "2, 1"



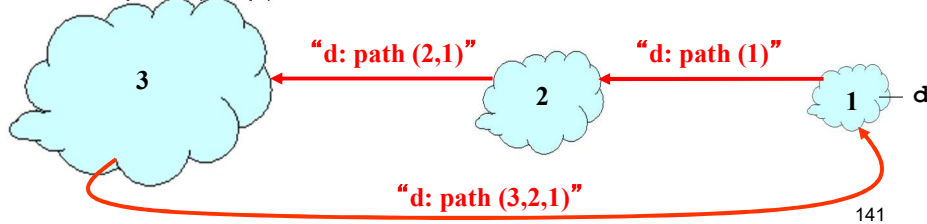
140

140

BGP khác DV (2)

Định tuyến theo vector đường đi

- Ý tưởng chính: quảng bá toàn bộ các chặng trên đường đi
 - Distance vector: chọn đường dựa trên chi phí của đường đi tới các đích → có thể xuất hiện đường đi quẩn do vấn đề đếm tới vô cùng (count-to-infinity)
 - Path vector: chọn đường dựa trên các chặng của đường đi tới đích → dễ dàng phát hiện các đường đi quẩn (loop)

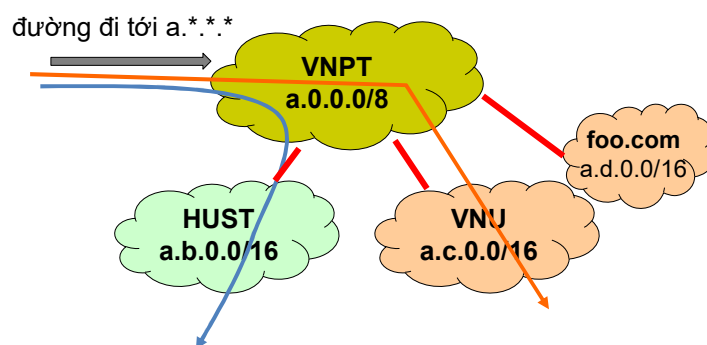


141

BGP khác DV (3)

Kết hợp đường đi

- BGP có khả năng kết hợp các đường đi tới các mạng con không đầy đủ



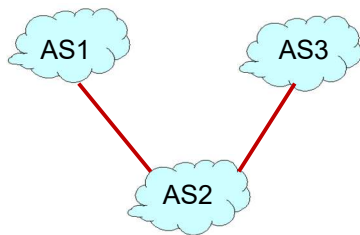
142

BGP khác DV (4)

Quảng bá có chọn lựa



- Một AS có thể chọn để không quảng bá đường đi tới một đích nào đó
- Nói một cách khác, một AS có đường đi tới AS đích nhưng không đảm bảo sẽ vận chuyển mọi lưu lượng mạng tới đó

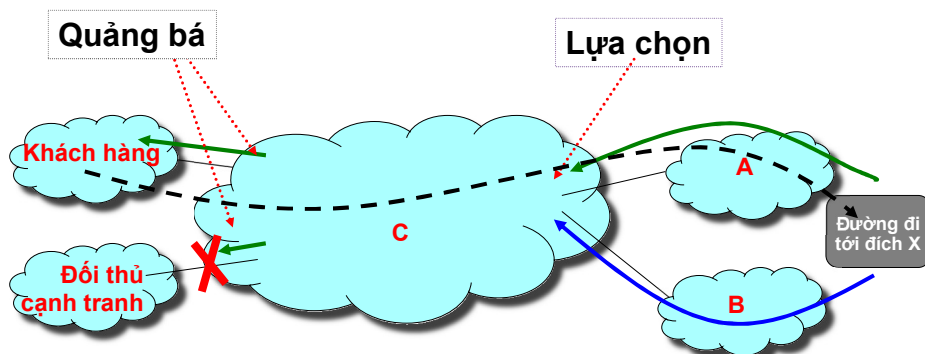


Ví dụ: AS2 không muốn vận chuyển thông tin từ AS1 tới AS3

143

143

Quảng bá có chọn lựa



- **Lựa chọn:** Đường đi nào được dùng để chuyển dữ liệu tới AS đích?
 - Kiểm soát thông tin ra khỏi AS
- **Quảng bá:** Quảng bá cho AS khác đường đi nào?
 - Kiểm soát thông tin vào AS

144

144

Quảng bá có chọn lựa: Một số chính sách điển hình



- Thay đổi mục tiêu khi định tuyến:
 - Tối thiểu hóa chi phí, tối đa hóa lợi nhuận
 - Tối đa hóa hiệu năng (đường đi qua ít AS nhất)
 - Tối thiểu hóa lưu lượng qua AS (định tuyến kiểu “hot potato”)
 - ...
- BGP có cơ chế gán thuộc tính cho các tuyến đường để thực hiện các mục tiêu trên

145

145

5.4.2. Hoạt động của BGP



146

146

eBGP, iBGP và IGP



- BGP cài đặt trên các router biên của AS(kết nối tới các AS khác): 2 phiên hoạt động
 - External BGP (eBGP): thực hiện trao đổi thông điệp với các router biên trên AS khác để tìm đường đi tới đích nằm ngoài AS của nó
 - Internal BGP (iBGP): trao đổi thông điệp với các router biên và router nội vùng cùng AS để quảng bá đường đi tới đích nằm ngoài AS của nó
- IGP : Interior Gateway Protocol = Intra-domain Routing Protocol
 - Cài đặt trên router nội vùng
 - Tìm đường đi tới đích nằm trong vùng AS
 - Dữ liệu tới đích ngoài AS sẽ được chuyển tới router biên

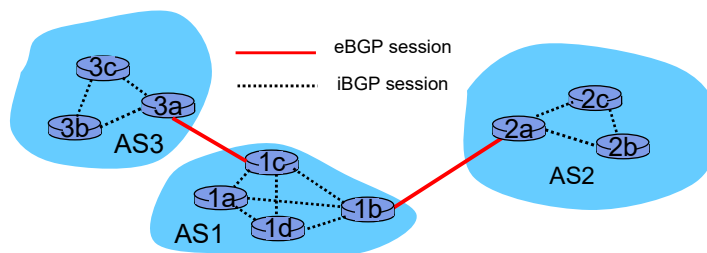
147

147

eBGP và iBGP – Ví dụ



- Quảng bá thông tin đường đi
 1. 3a gửi tới 1c bằng eBGP
 2. 1c gửi thông tin nội bộ tới (1b, 1d, ...) trong AS1 bằng iBGP
 - 1b: Router biên cài BGP
 - 1a,1d: Router nội vùng cài IGP
 3. 2a nhận thông tin từ 1b bằng eBGP



148

148

Các thông điệp BGP cơ bản

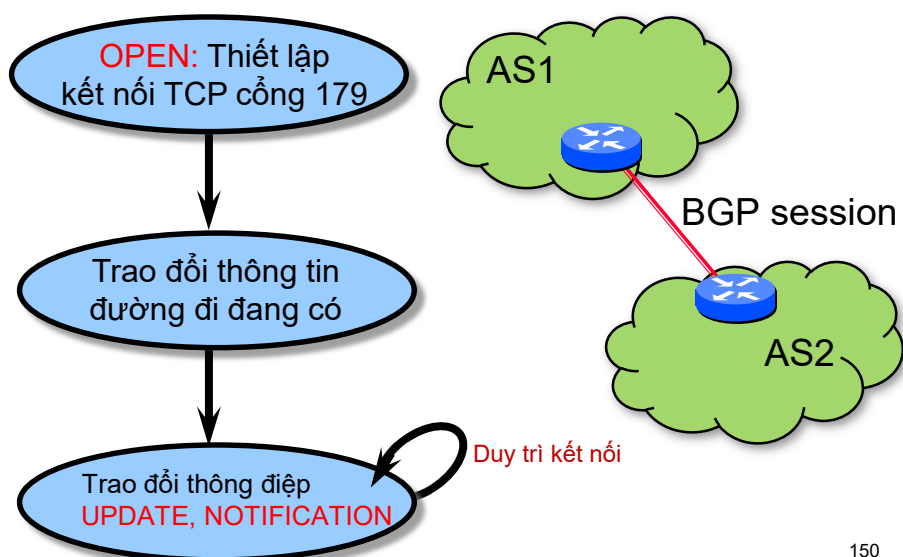


- OPEN:
 - Thiết lập phiên trao đổi thông tin đường đi
 - Sử dụng TCP, cổng 179
- NOTIFICATION: thông báo các sự kiện bất thường
- UPDATE:
 - Thông báo về đường đi mới
 - Thông báo về đường đi không còn khả dụng
- KEEPALIVE: thông báo duy trì kết nối TCP
 - Kết nối của TCP cung cấp cho BGP là nửa duy trì (semi-persistent)

149

149

Phiên trao đổi thông tin của BGP



150

150

Còn rất nhiều vấn đề về tầng mạng



- Giao thức IPv6
- Mobile IP
- Định tuyến trên mạng cáp quang
- Định tuyến trên mạng không dây
- Broadcast và Multicast
- An toàn bảo mật thông tin tầng mạng
- An toàn bảo mật cho các giao thức định tuyến...

151

151

Tài liệu tham khảo



- Keio University
- “Computer Networking: A Top Down Approach”, J.Kurose
- “Computer Network”, Berkeley University

152

152