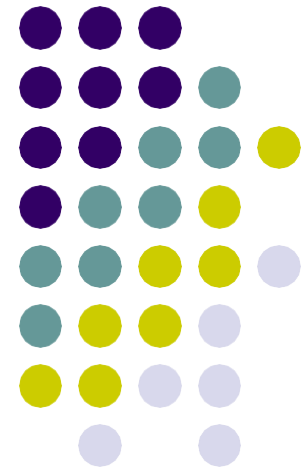


Chương 7:

Các giao thức chọn đường

Giảng viên: Nguyễn Đức Toàn

Bộ môn Truyền thông và Mạng máy tính
Viện CNTT&TT - ĐHBK Hà Nội



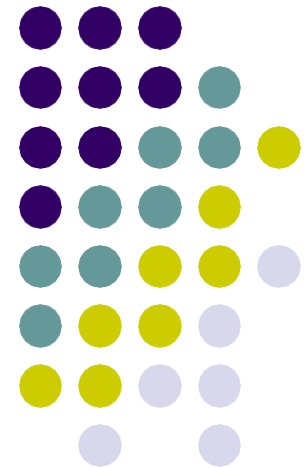


Tổng quan

- Tuần trước: Khái niệm chung về chọn đường.
 - Chọn đường tĩnh và chọn đường động
 - Các giao thức chọn đường
- Tuần này: Các giao thức chọn đường cụ thể trên Internet
 - Vấn đề phân cấp trong chọn đường
 - Chọn đường nội vùng (Intra-domain): RIP, OSPF
 - Chọn đường liên vùng (Inter-domain): BGP

Phân cấp trong chọn đường

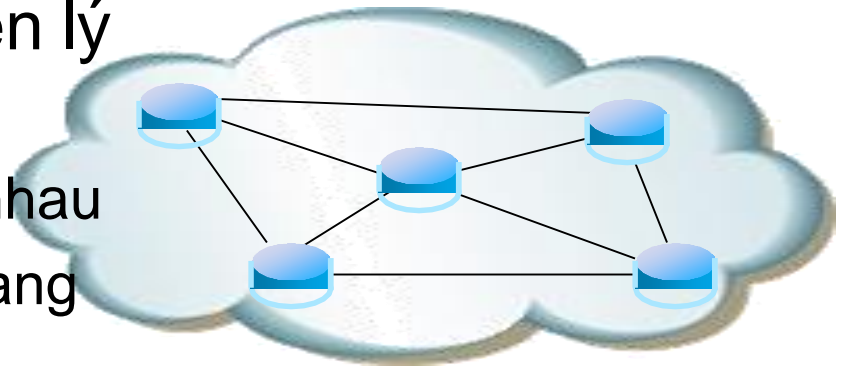
Các hệ tự trị
Chọn đường nội vùng
Chọn đường liên vùng





Tổng quan

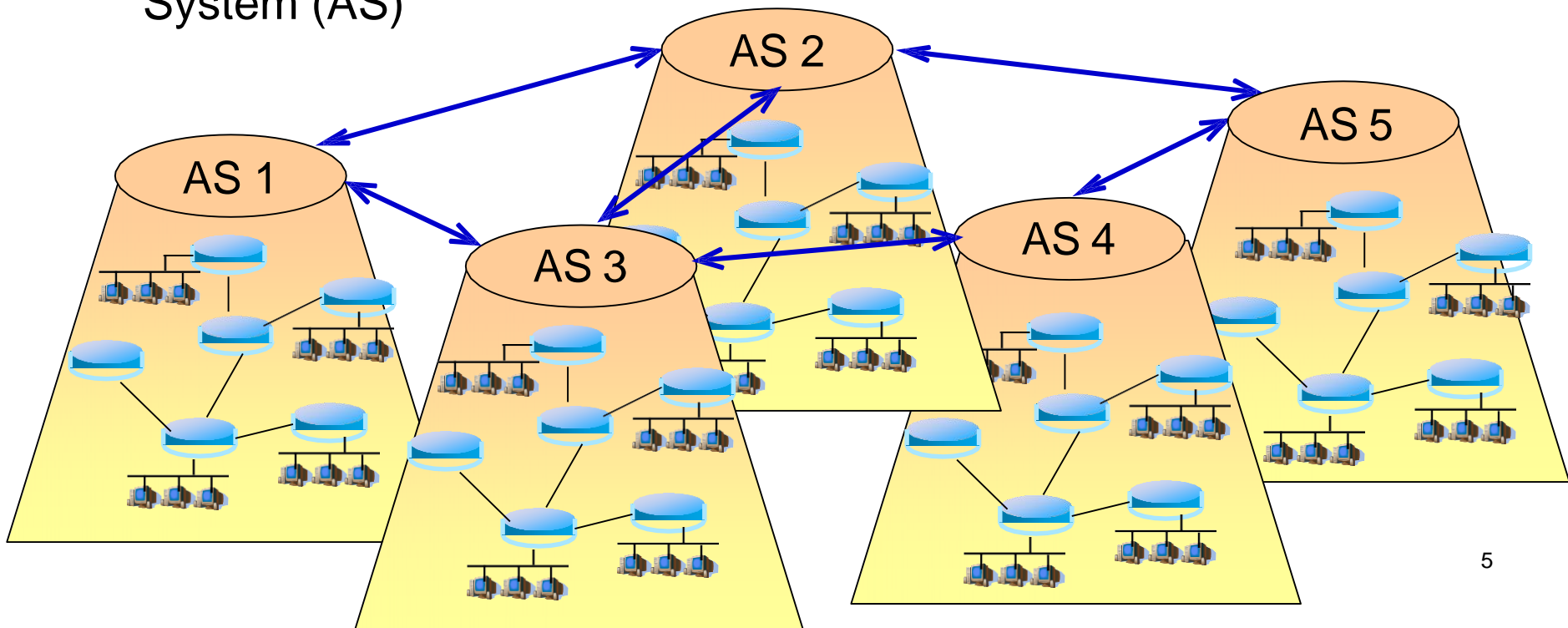
- Vấn đề chọn đường đã học được xem xét trong điều kiện lý tưởng
 - Các nút mạng có vai trò như nhau
 - Chỉ có một mạng duy nhất, mạng “phẳng”
- Thực tế không giống như vậy
- Tính mở rộng: Internet có hàng triệu (tỷ) máy trạm, chọn đường bằng LS hay DV?
 - LS: Quá tải thông tin chọn đường
 - DV: Có hội tụ được không?





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

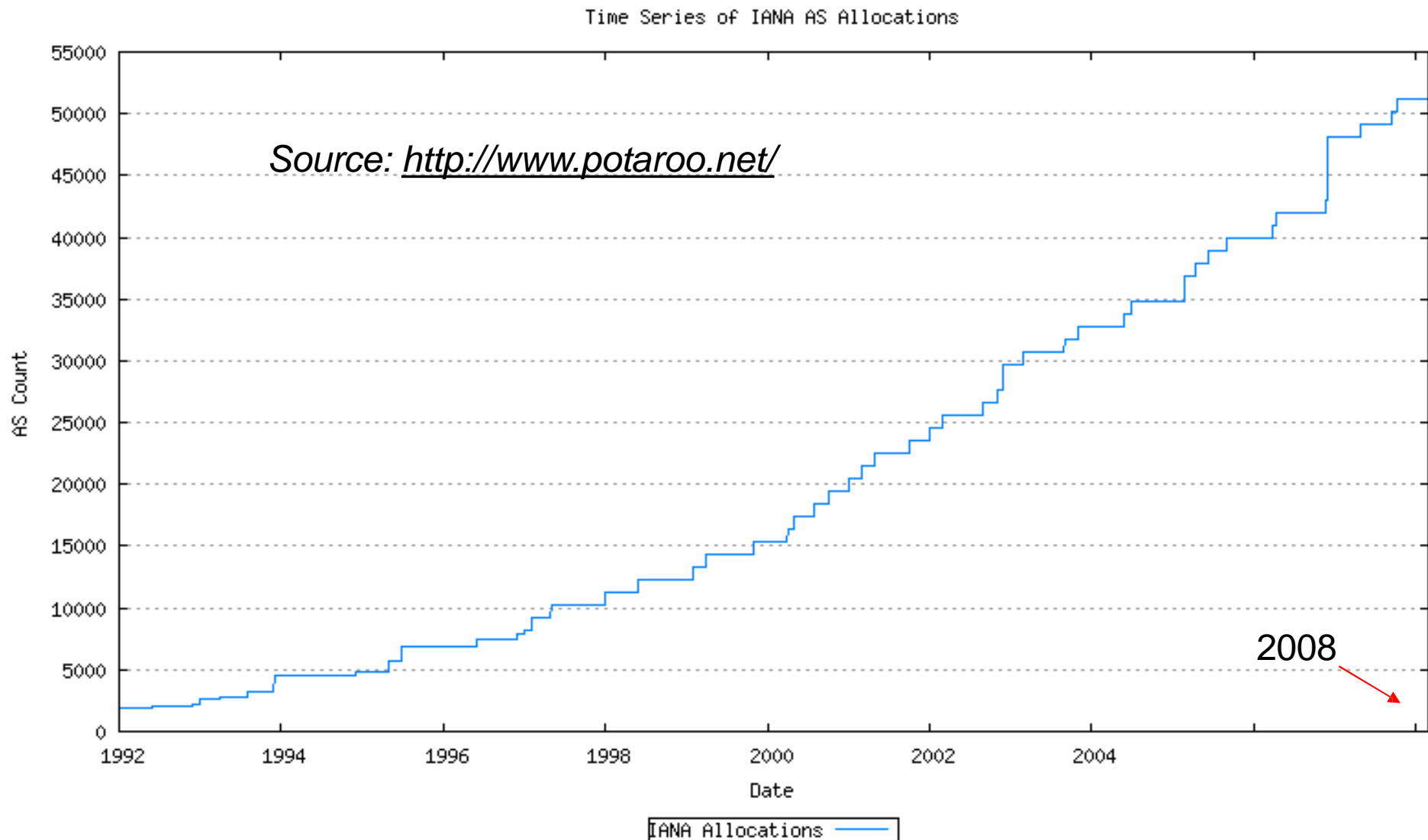
[24066](#) VNNIC-AS-VN Vietnam Internet Network Information Center

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

.....

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

Số lượng ASN cấp phát bởi IANA

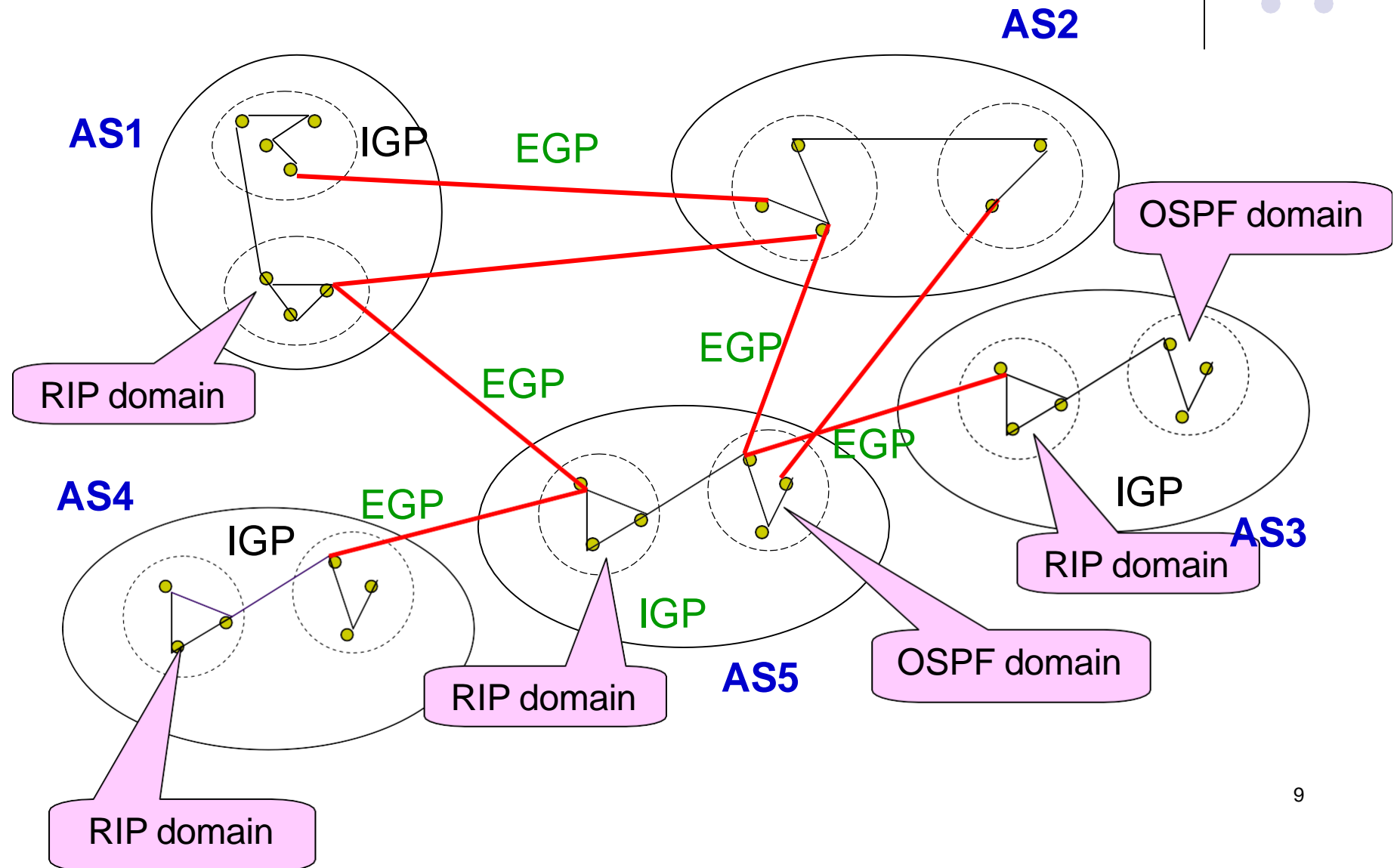


Phân cấp giao thức chọn đường



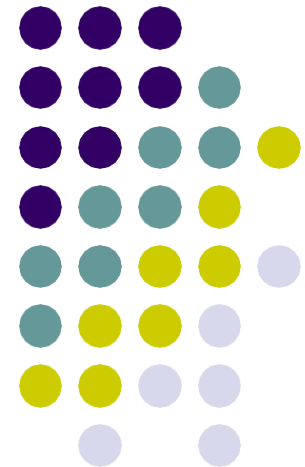
- **Trong một hệ tự trị:** Giao thức chọn đường nội vùng
 - **IGP:** *Interior Gateway Protocol*
 - RIP: Routing Information Protocol
 - OSPF: Open Shortest Path First
 - IS-IS, IGRP, EIGRP (Cisco)...
- **Giữa các hệ tự trị:** Giao thức chọn đường liên vùng
 - **EGP:** *Exterior Gateway Protocol*
 -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

Intra-domain và Inter-domain routing



Chọn đường nội vùng

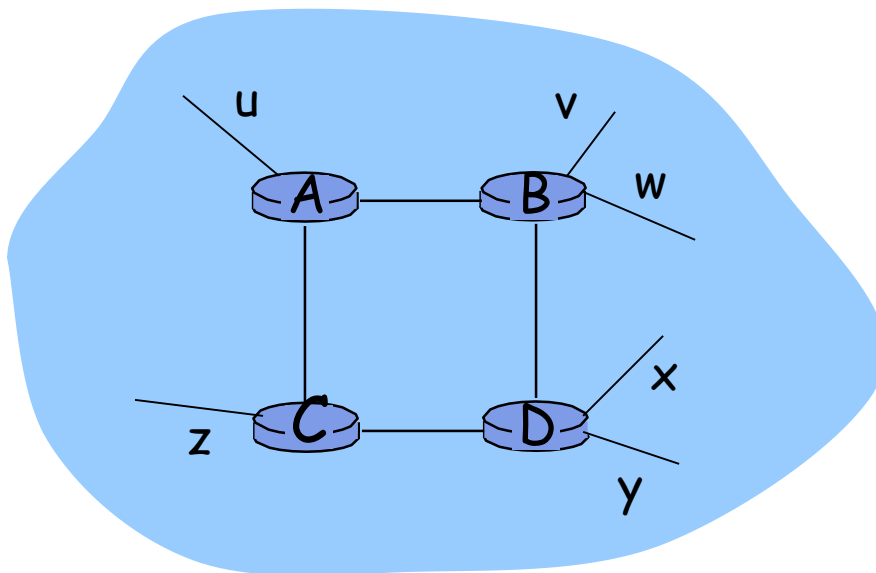
RIP
OSPF



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 (# of hops, max = 15 hops)



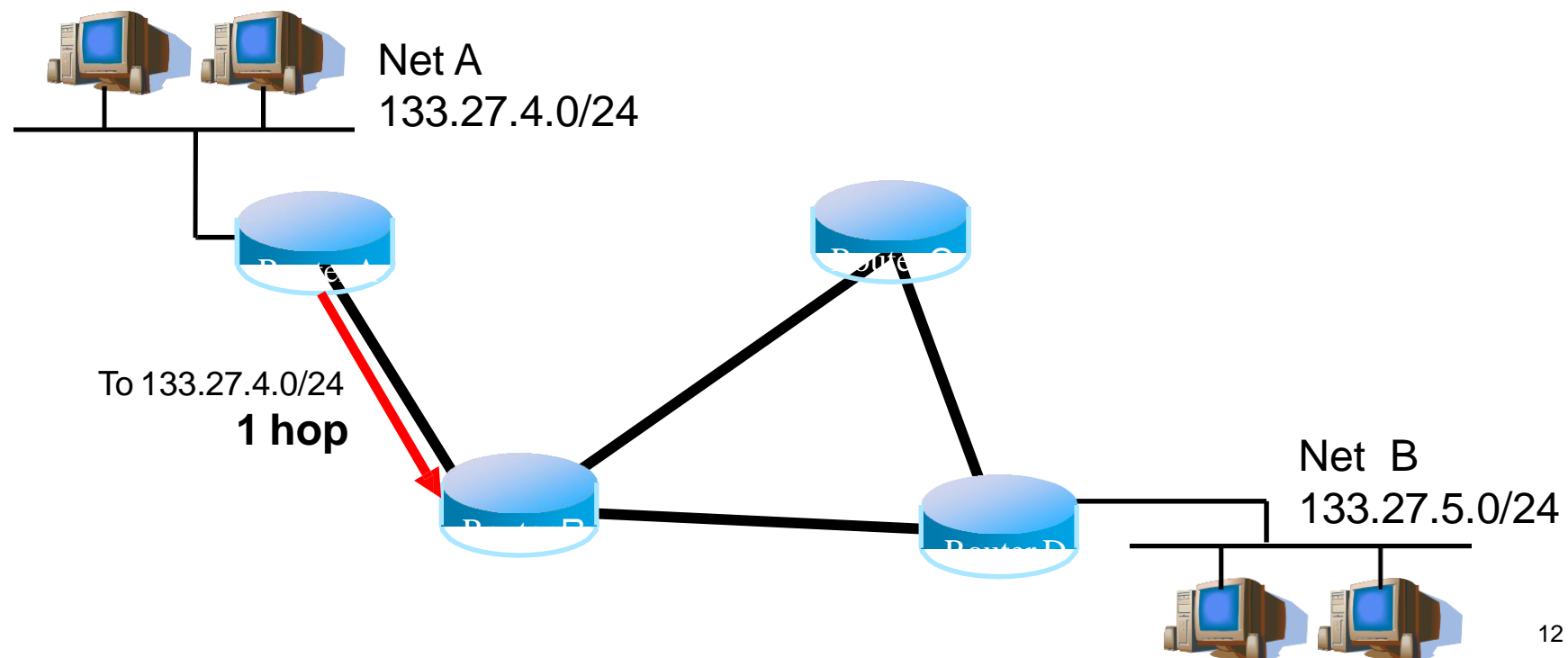
Từ nút A:

<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)



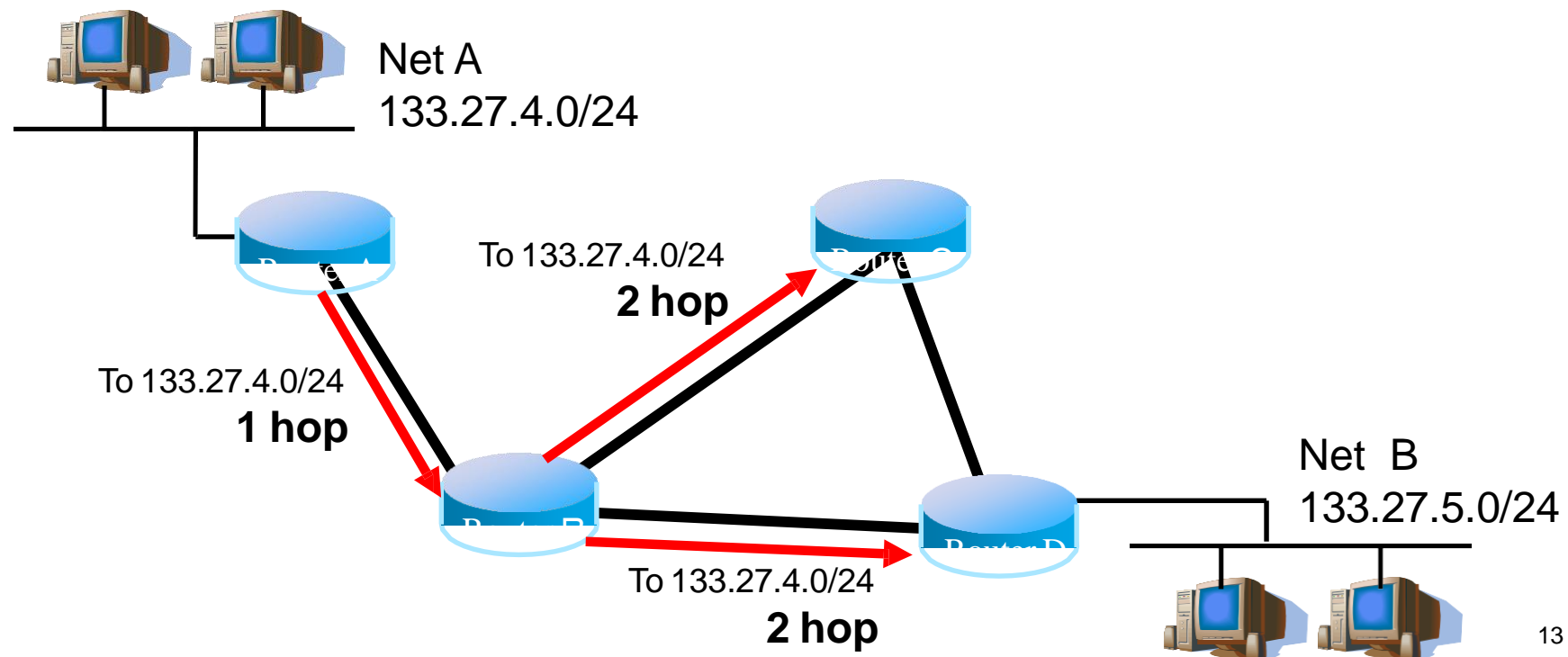
- Bạn của bạn là bạn



Nhắc lại: Chọn đường dạng DV (2)



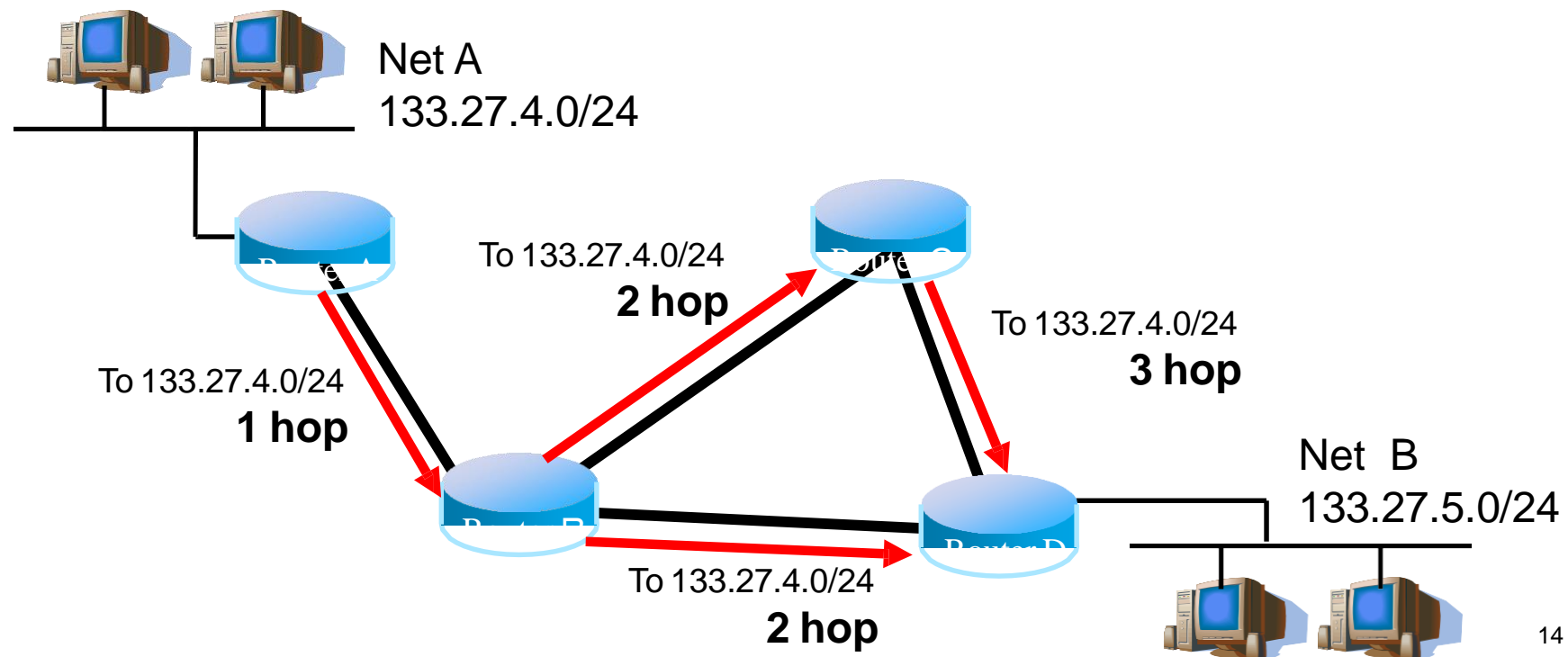
- Bạn của bạn là bạn



Nhắc lại: Chọn đường dạng DV (3)



- Bạn của bạn là bạn

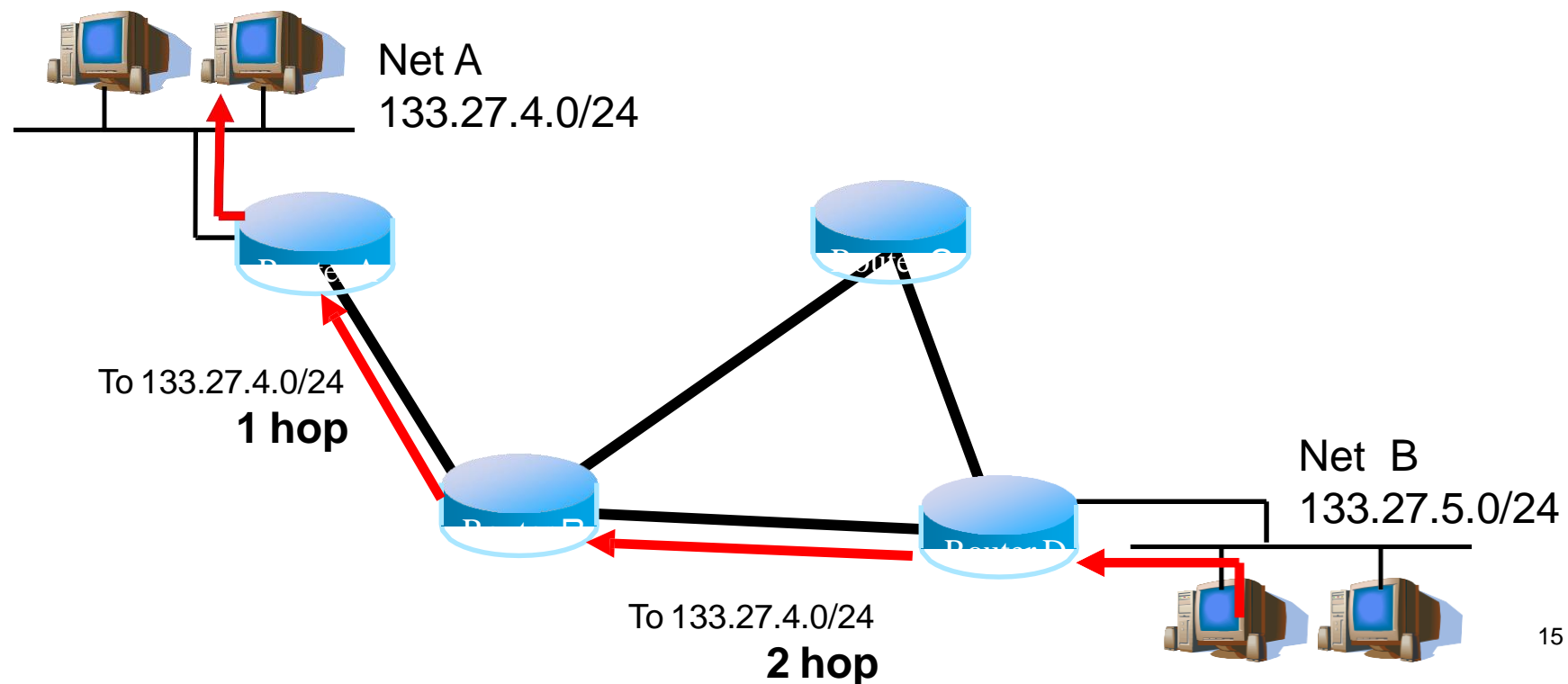




Nhắc lại: Chọn đường dạng DV (4)

- Bạn của bạn là bạn

Lưu ý: Tên của router





RIP: Trao đổi thông tin

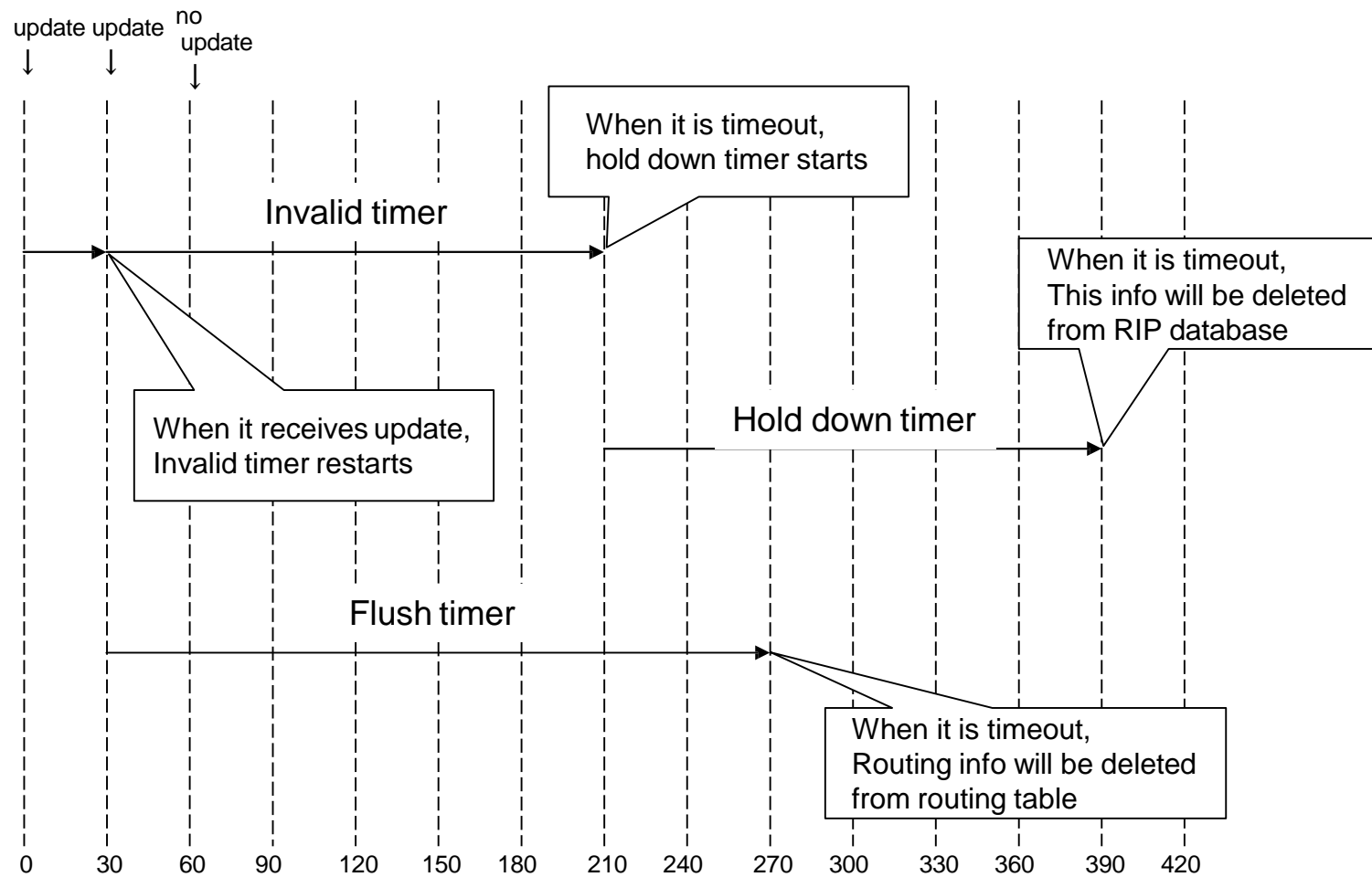
- Trao đổi bảng chọn đường
- Định kỳ
 - Các vector khoảng cách được trao đổi định kỳ - 30s
 - Mỗi thông điệp chứa tối đa 25 mục
 - Trong thực tế, nhiều thông điệp được sử dụng
- Sự kiện
 - Gửi thông điệp cho nút hàng xóm mỗi khi có thay đổi
 - Nút hàng xóm sẽ cập nhật bảng chọn đường của nó

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

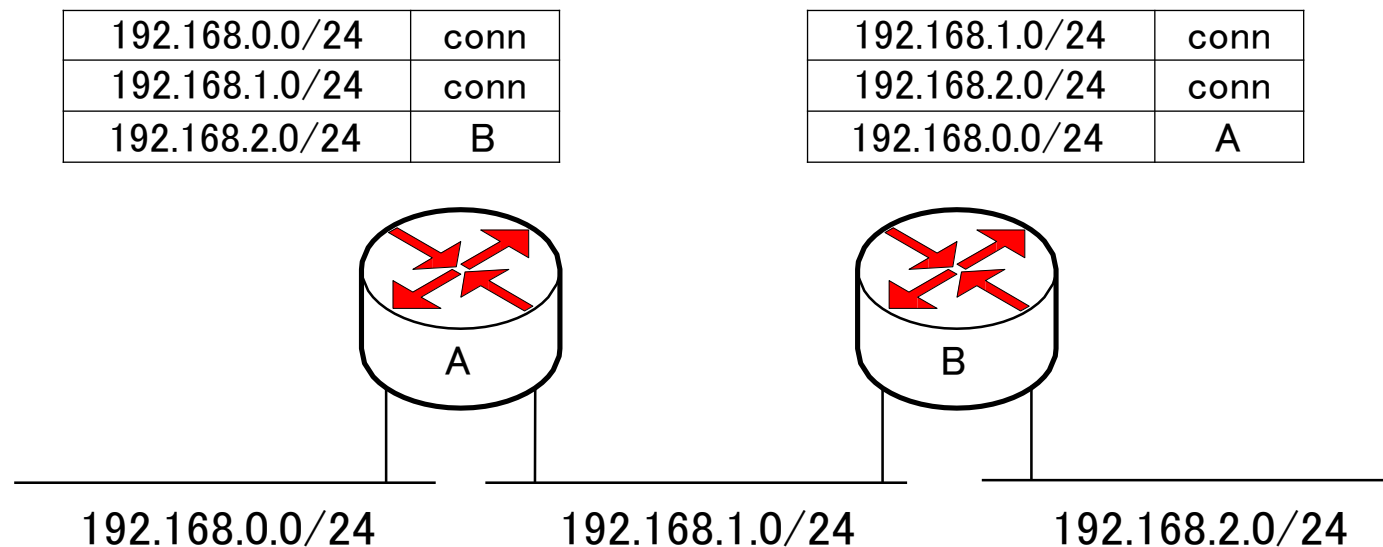
RIP timer (2)



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





OSPF: đề 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

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



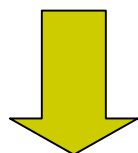
Một số đặc điểm của OSPF

- **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

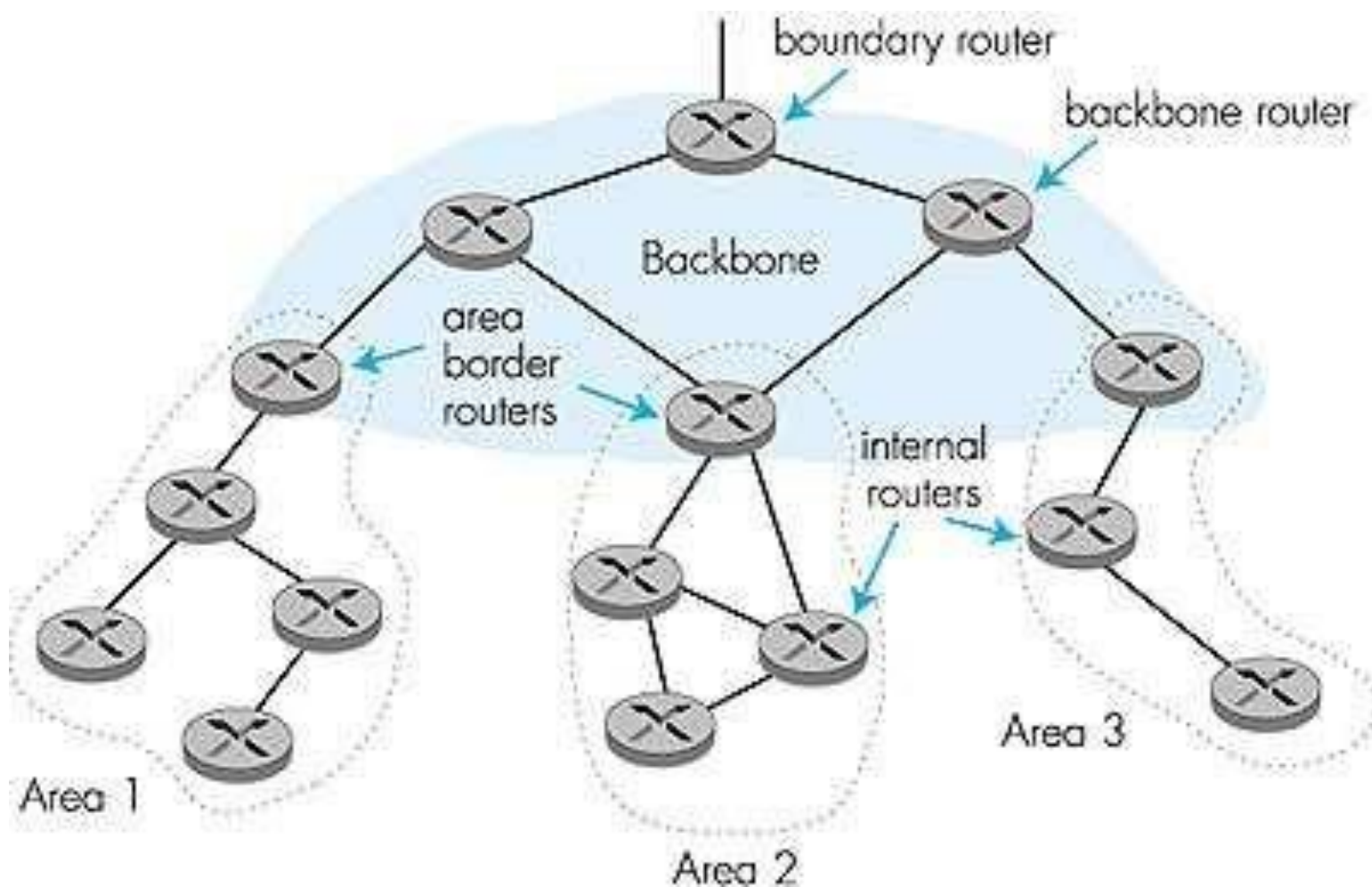
- Trong việc chọn đường, tại sao phải chia mạng thành các vùng nhỏ hơn?
- Nếu có quá nhiều router
 - Thông tin trạng thái liên kết được truyền nhiều lần hơn
 - Phải liên tục tính toán lại
 - Cần nhiều bộ nhớ hơn, nhiều tài nguyên CPU hơn
 - Lượng thông tin phải trao đổi tăng lên
 - Bảng chọn đường lớn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



- Vùng
 - Nhóm các router có cùng thông tin LSA



Phân cấp OSPF





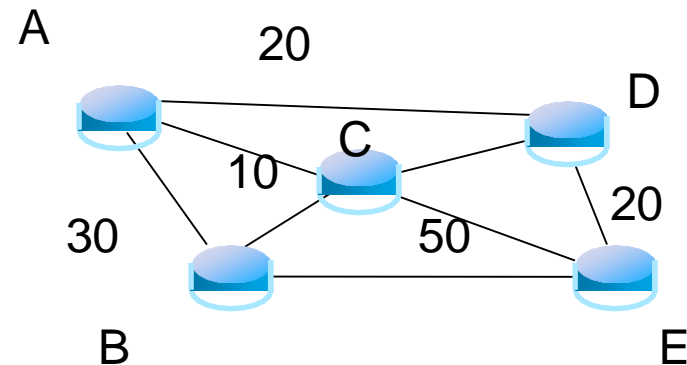
Các dạng router

- **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 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



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

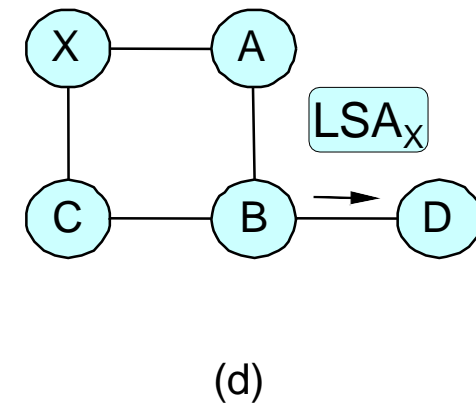
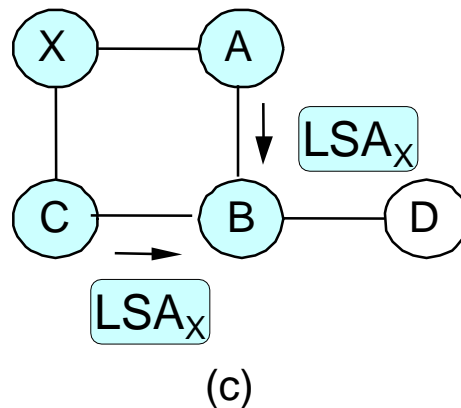
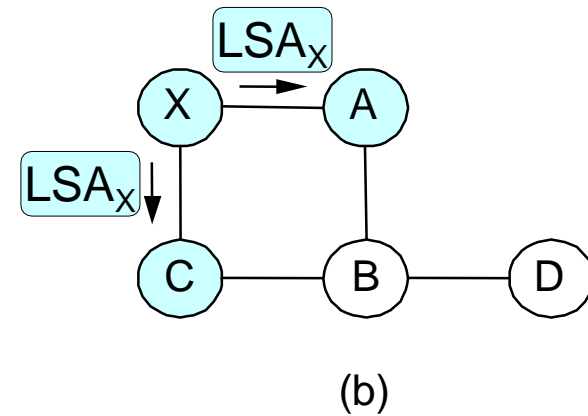
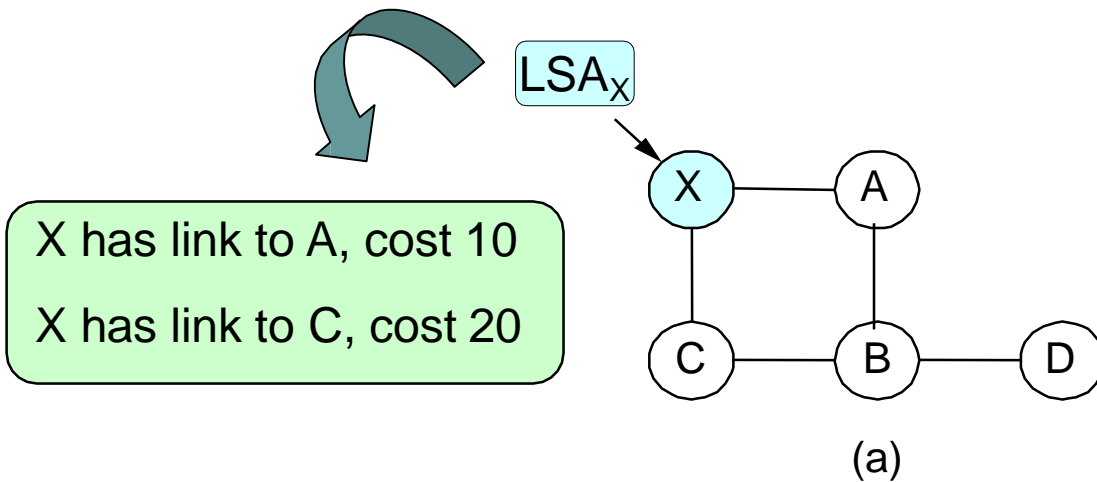


Chi phí mặc định của OSPF

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



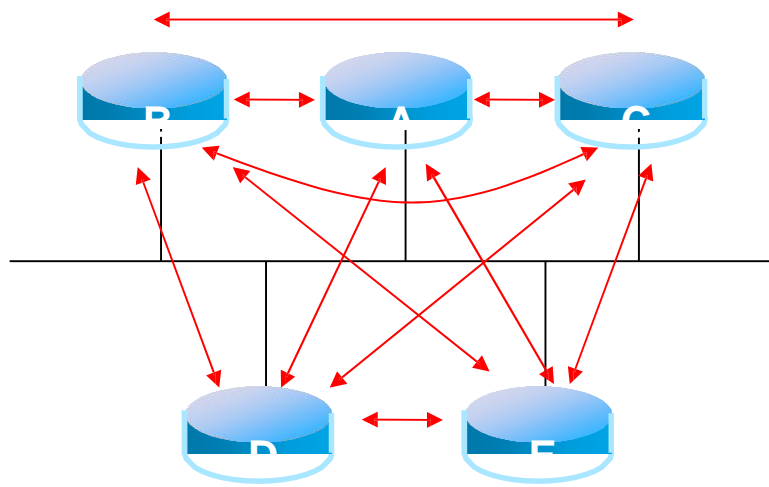
Quảng bá thông tin LSA



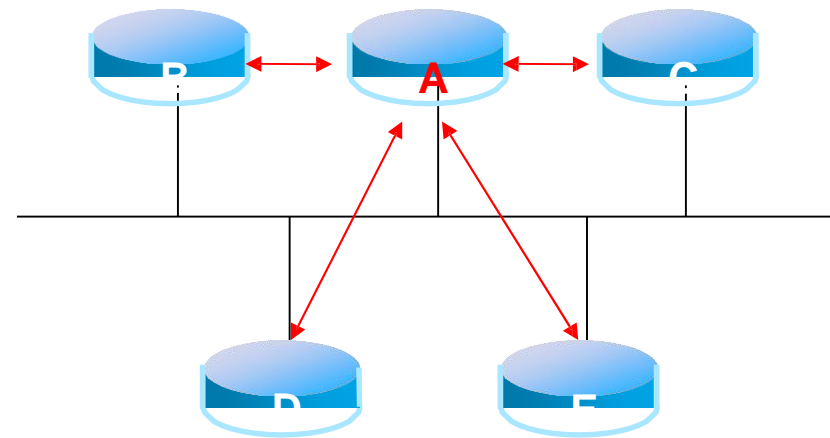


Router đại diện - DR

- 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



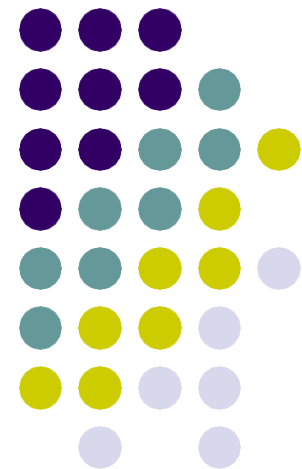
- “Neighbor” và “adjacency” là các k/n khác nhau!
 - Adjacency: có trao đổi thông tin
 - Neighbor: có đường nối trực tiếp
- Mạng quảng bá đa truy cập (e.g Ethernet)
 - Neighbor != Adjacency
- Mạng điểm-nối-điểm
 - Neighbor == Adjacency



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 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

Giao thức chọn đường liên vùng



BGP – Border Gateway Protocol

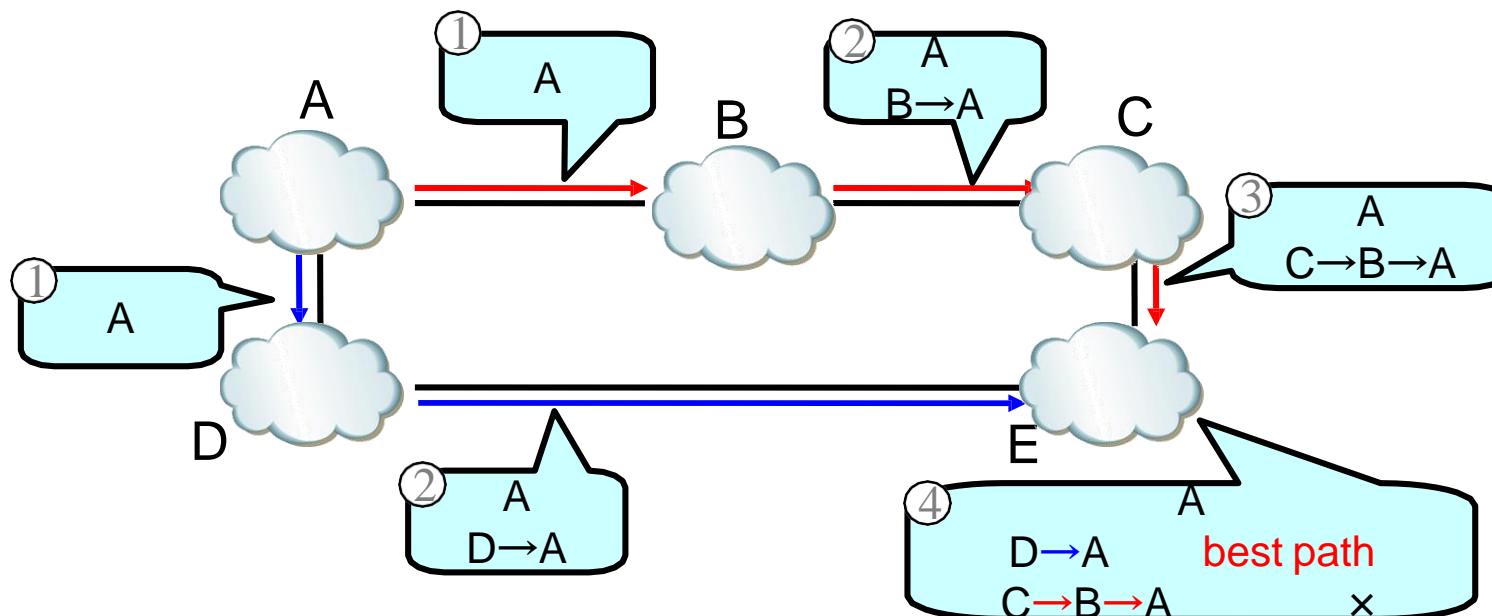


- Yếu tố gắn kết của Internet, kết nối các hệ tự trị
- Trao đổi thông tin NLRI (Network Layer Reachability Information)
 - Cho phép một AS biết được thông tin đi đến AS khác
 - Gửi thông tin này vào bên trong AS đó
 - Xác định đường đi tốt nhất dựa trên thông tin đó và các chính sách chọn đường
- Cho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



BGP: Path vector routing

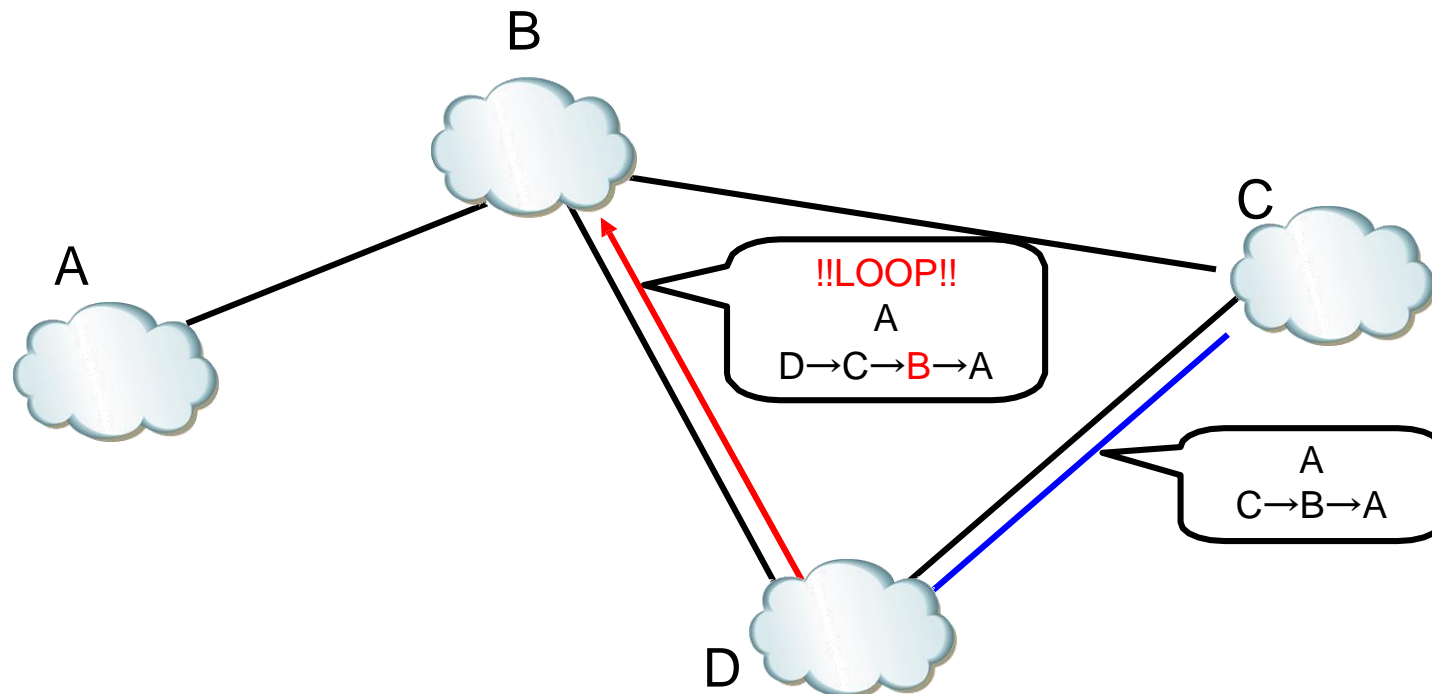
- Giữa các AS nên dùng giao thức nào?
 - Khó có một chính sách và đơn vị chi phí chung
 - LS: Chi phí không đồng nhất, CSDL quá lớn
 - DV: Mạng quá rộng, khó hội tụ
- Giải pháp: Chọn đường theo path-vector



Cơ chế tránh vòng lặp



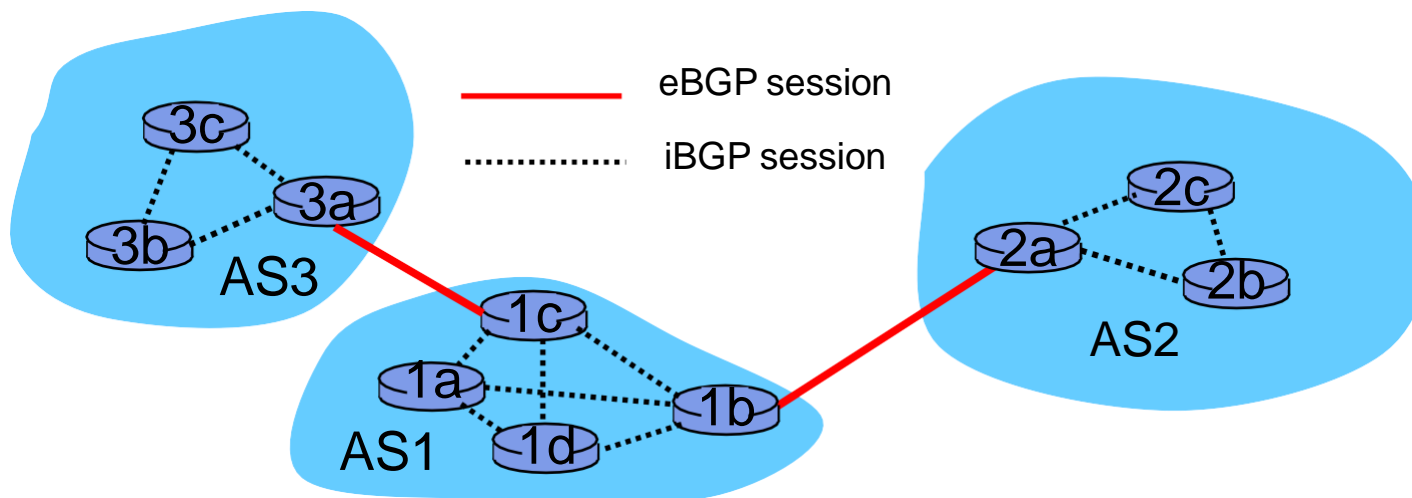
- Dò lại xem router đã có trên path-vector hay chưa
 - B hủy đường đi tới A





eBGP và iBGP

- External BGP vs. Internal BGP
- Phân tán thông tin chọn đường
 1. 3a gửi tới 1c bằng
 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
 - BGP có thể đặt các chính sách
 - Cho đường vào
 - Cho đường ra



Các thuộc tính của đường đi

- ORIGIN
 - Nguồn của thông tin (IGP/EGP/incomplete)
- AS_PATH
- NEXT_HOP
- MED (MULTI_EXIT_DISCRIMINATOR)
- LOCAL_PREF
- ATOMIC_AGGREGATE
- AGGREGATOR
- COMMUNITY



Các bước chọn đường đi

- Bước 1: NEXT_HOP?
- Bước 2: So sánh LOCAL_PREF
- Bước 3: So sánh độ dài AS_PATH
- Bước 4: So sánh ORIGIN
- Bước 5: So sánh MED
- Bước 6: So sánh EBGP/IBGP
- Bước 7: So sánh chi phí tới NEXT_HOP
- Bước 8: So sánh Router ID



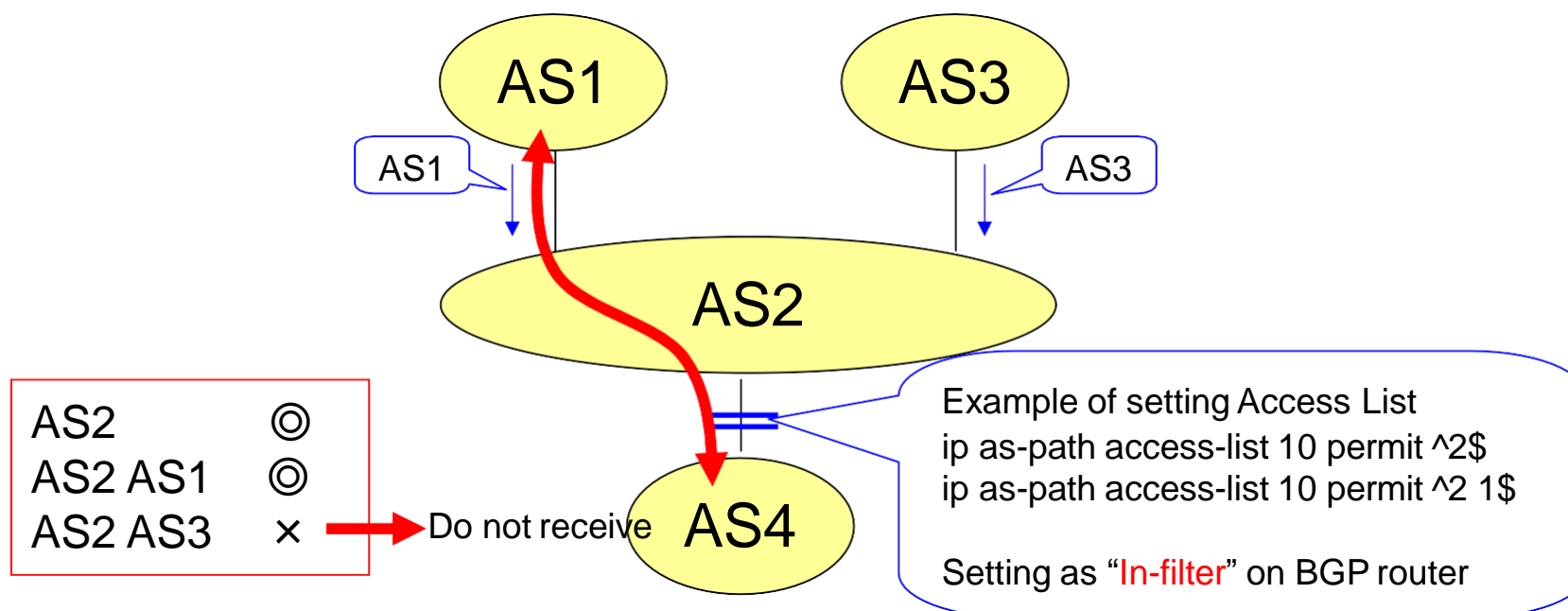
Cài đặt các chính sách ntn?

- Bộ lọc
- Các thuộc tính
 - AS_PATH PREPEND
 - MED
 - LOCAL_PREF



Bộ lọc

- Chỉ trao đổi đường đi nào đã được đăng ký
 - Hạn chế thông qua bộ lọc vào - In-filter
 - Thông qua bộ lọc ra - Out-filter

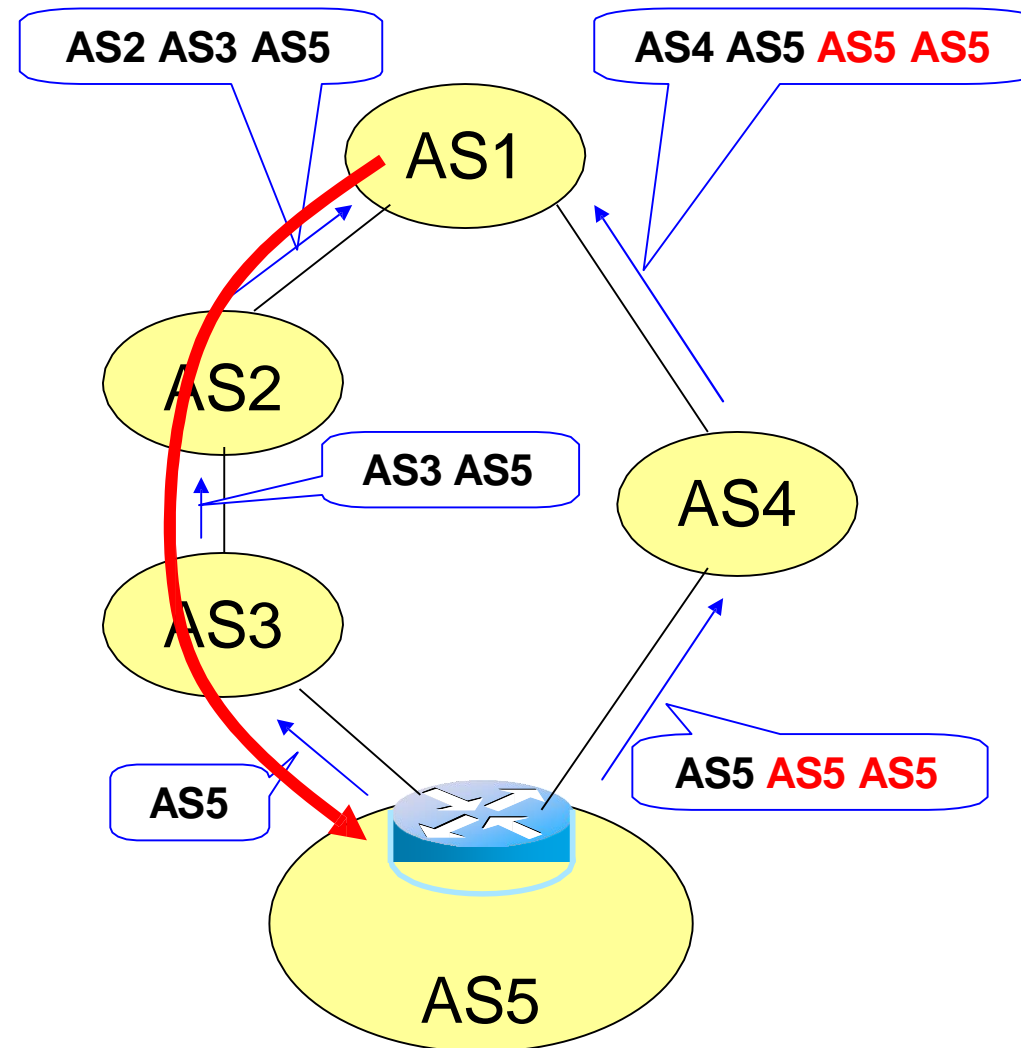


Ví dụ về AS PATH



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

Chọn đường với AS_PATH Prepend



Ví dụ về AS PATH prepend

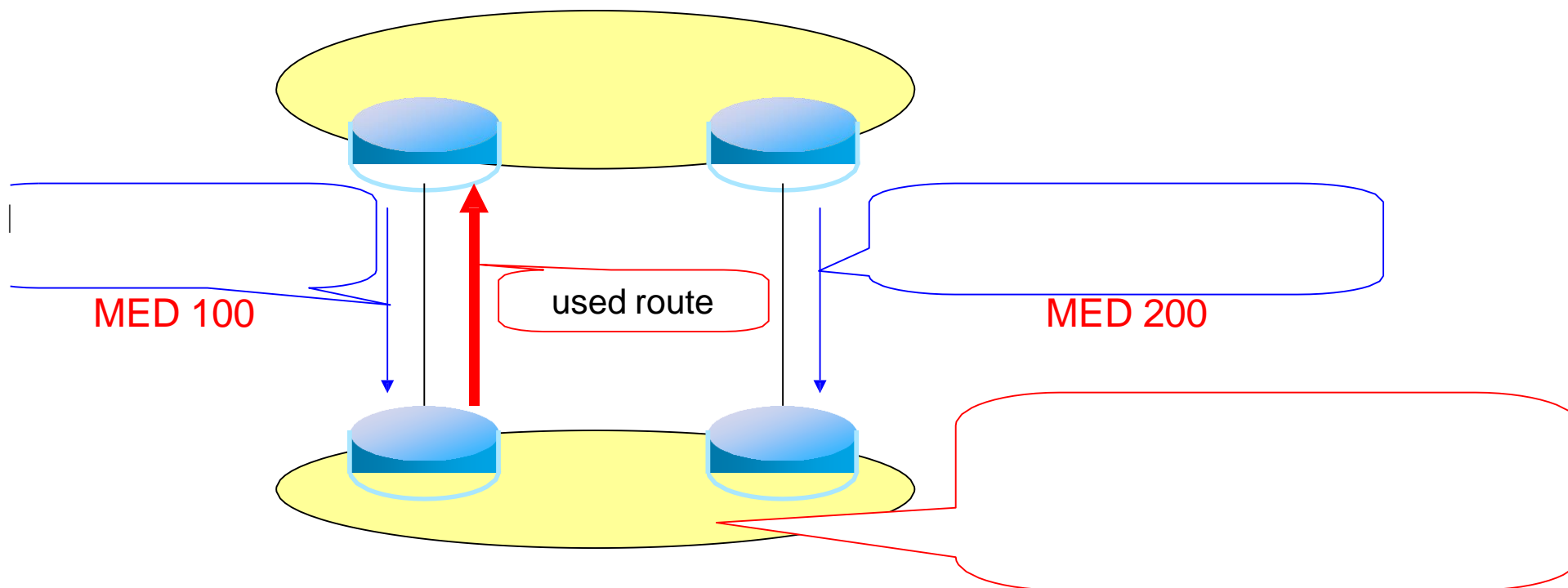


Network	Next Hop	Metric	LocPrf	Weight	Path
8.5.192.0/22	203.178.136.14	100	0	2516	209 13989 13989 13989 13989
	203.178.136.14	100	0	2516	209 13989 13989 13989 13989
	203.178.136.14	100	0	2516	209 13989 13989 13989 13989
8.5.196.0/24	203.178.136.14	100	0	2516	209 13989 13989 13989 13989
	203.178.136.14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



Chọn đường với MED

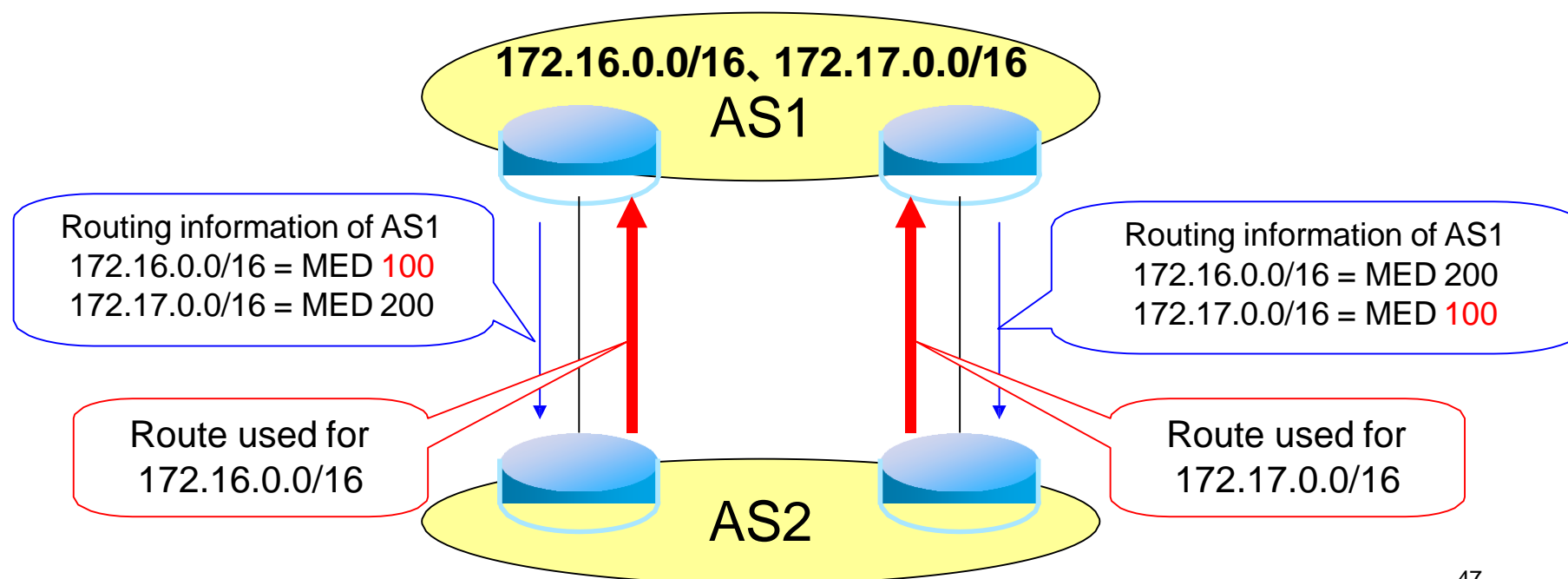
- trong trường hợp 2 AS với nhiều link
- Chọn MED nhỏ hơn
- Áp dụng trong điều khiển lưu lượng





Phân tải với MED

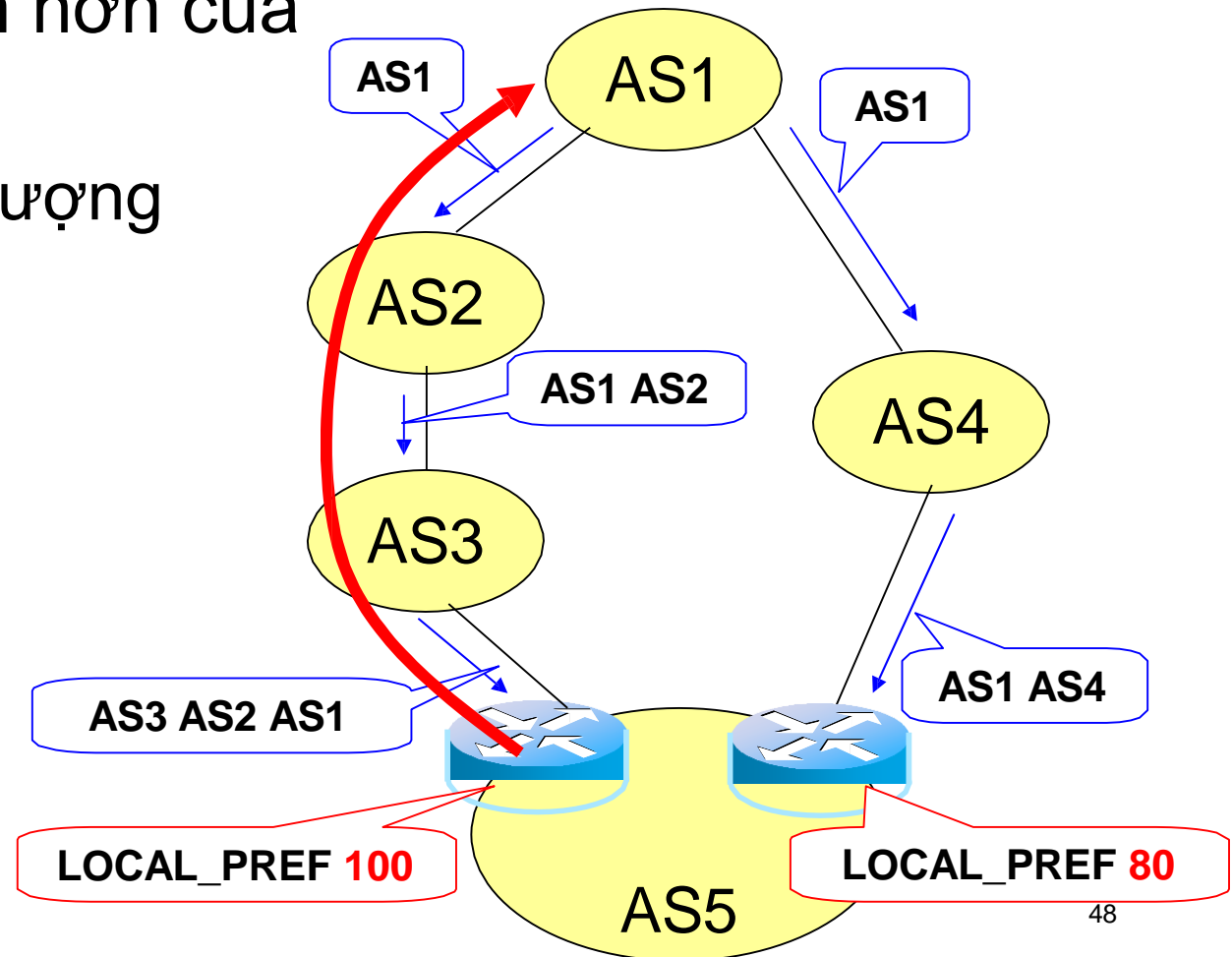
- Đặt giá trị MED khác nhau cho mỗi đường
- Cũng điều khiển lưu lượng





Sử dụng LOCAL_PREF

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



Tóm tắt

- Hierarchical routing
- RIP
- OSPF
- BGP



Tuần tới: Transport Layer



- Nguyên lý tầng giao vận
- UDP và TCP
- điều khiển luồng
- điều khiển tắc nghẽn