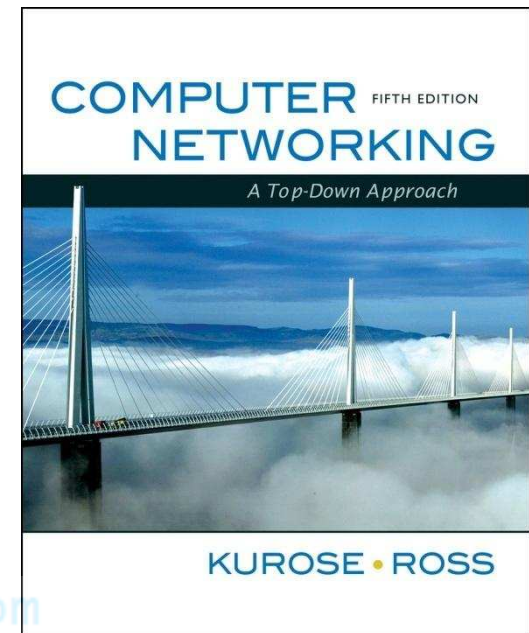


Chapter 4

Network Layer



A note on the use of these ppt slides:

We're making these slides freely available to all (faculty, students, readers). They're in PowerPoint form so you can add, modify, and delete slides (including this one) and slide content to suit your needs. They obviously represent a *lot* of work on our part. In return for use, we only ask the following:

- ☐ If you use these slides (e.g., in a class) in substantially unaltered form, that you mention their source (after all, we'd like people to use our book!)
- ☐ If you post any slides in substantially unaltered form on a www site, that you note that they are adapted from (or perhaps identical to) our slides, and note our copyright of this material.

Thanks and enjoy! JFK/KWR

All material copyright 1996-2009
J.F Kurose and K.W. Ross, All Rights Reserved

*Computer Networking:
A Top Down Approach
5th edition.*

*Jim Kurose, Keith Ross
Addison-Wesley, April
2009.*

Chapter 4: Network Layer

Chapter goals:

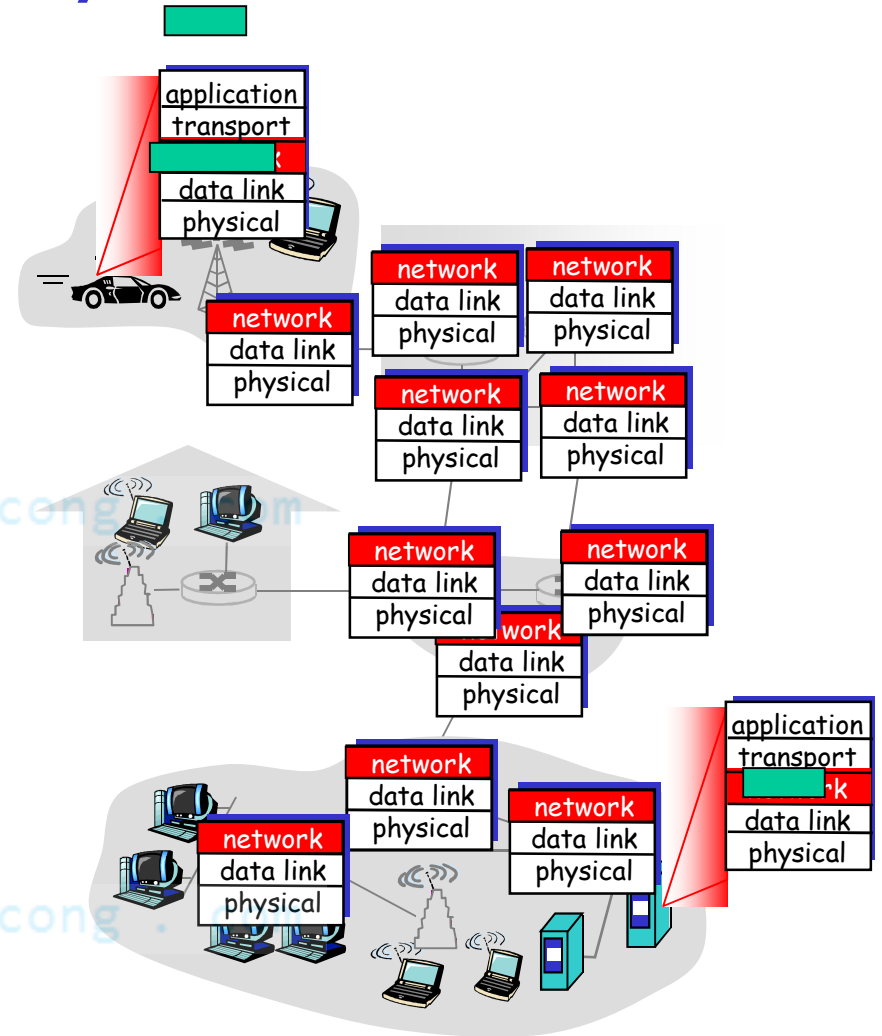
- understand principles behind network layer services:
 - ❖ network layer service models
 - ❖ forwarding versus routing
 - ❖ how a router works
 - ❖ routing (path selection)
 - ❖ dealing with scale
 - ❖ advanced topics: IPv6, mobility
- instantiation, implementation in the Internet

Chapter 4: Network Layer

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- 4.4 IP: Internet Protocol
 - ❖ Datagram format
 - ❖ IPv4 addressing
 - ❖ ICMP
 - ❖ IPv6
- 4.5 Routing algorithms
 - ❖ Link state
 - ❖ Distance Vector
 - ❖ Hierarchical routing
- 4.6 Routing in the Internet
 - ❖ RIP
 - ❖ OSPF
 - ❖ BGP
- 4.7 Broadcast and multicast routing

Network layer

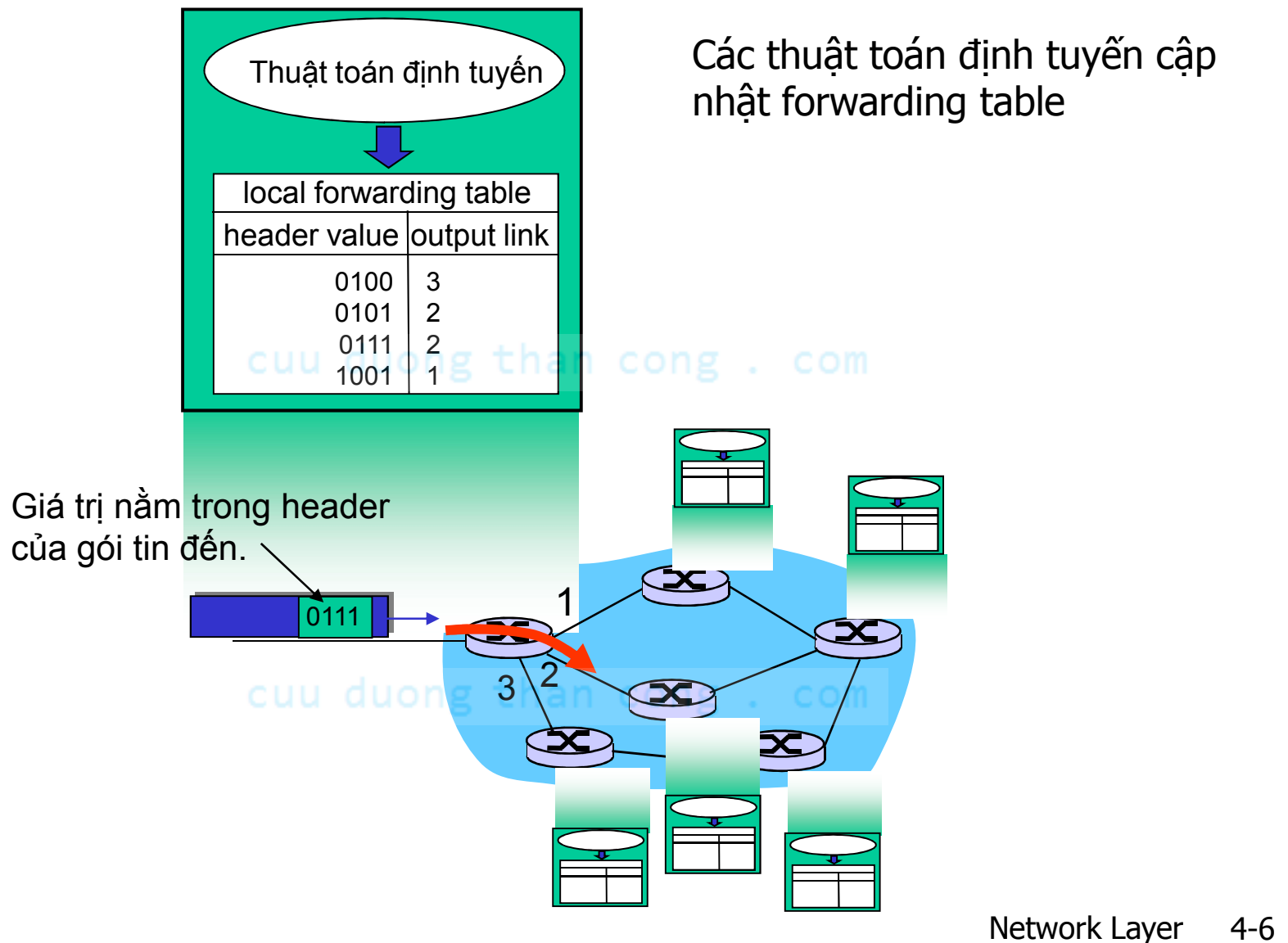
- Vận chuyển segment từ host bên gửi đến host bên nhận.
- Ở bên gửi, đóng gói segment trong datagram
- Ở bên nhận, trao lại segment cho tầng transport.
- Các giao thức của tầng mạng có trong mọi host và router.
- Router kiểm tra các field trong header của tất cả các datagram đi qua nó.



Two Key Network-Layer Functions

- *forwarding*: quá trình chuyển tiếp gói tin từ một ngõ vào đến một ngõ ra phù hợp trong cùng 1 router
- *routing*: quá trình tìm đường đi cho gói tin qua các router để đến đích.
 - ❖ *routing algorithms (thuật toán định tuyến): thuật toán tìm đường đi cho gói tin đến đích.*

Interplay between routing and forwarding



Connection setup

- Ngoài 2 chức năng quan trọng *forwarding* và *routing*, thiết lập kết nối [connection setup] là chức năng quan trọng thứ 3 trong một số kiến trúc mạng:
 - ❖ ATM, frame relay, X.25
- Trước khi các datagram truyền đi, hai host gửi/nhận và các router trung gian thiết lập kết nối ảo.
 - ❖ Các router tham gia vào quá trình này
- So sánh việc thiết lập kết nối giữa tầng network và tầng transport:
 - ❖ **network**: kết nối giữa 2 host truyền nhận (cũng có thể liên quan đến các router trung gian trong trường hợp kết nối ảo (virtual connections))
 - ❖ **transport**: kết nối giữa 2 tiến trình.

Network service model

Q: Tầng network có các mô hình dịch vụ nào?

Ví dụ về dịch vụ cho từng datagram:

- đảm bảo gói tin đến đích [guaranteed delivery].
- đảm bảo gói tin đến đích dưới 40ms [guaranteed delivery with less than 40 msec delay]

Ví dụ về dịch vụ cho dòng datagram:

- Đảm bảo gói tin đến nơi đúng thứ tự như khi gửi [in-order datagram delivery]
- Đảm bảo băng thông tối thiểu [guaranteed minimum bandwidth to flow]
- ...

Network layer service models:

Network Architecture	Service Model	Guarantees ?				Congestion feedback
		Bandwidth	Loss	Order	Timing	
Internet	best effort	none	no	no	no	no (inferred via loss)
ATM	CBR	constant rate	yes	yes	yes	no congestion
ATM	VBR	guaranteed rate	yes	yes	yes	no congestion
ATM	ABR	guaranteed minimum	no	yes	no	yes
ATM	UBR	none	no	yes	no	no

Chapter 4: Network Layer

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- 4.4 IP: Internet Protocol
 - ❖ Datagram format
 - ❖ IPv4 addressing
 - ❖ ICMP
 - ❖ IPv6
- 4.5 Routing algorithms
 - ❖ Link state
 - ❖ Distance Vector
 - ❖ Hierarchical routing
- 4.6 Routing in the Internet
 - ❖ RIP
 - ❖ OSPF
 - ❖ BGP
- 4.7 Broadcast and multicast routing

Network layer connection and connection-less service

- **Mạng datagram** cung cấp các dịch vụ phi-kết-nối ở tầng mạng
- **Mạng mạch ảo** [Virtual-circuit] cung cấp các dịch vụ có-kết-nối ở tầng mạng
- Tương tự như các dịch vụ ở tầng transport, nhưng:
 - ❖ **Dịch vụ:** host-to-host
 - ❖ **Không được lựa chọn:** mạng cung cấp hoặc dịch vụ phi-kết-nối, hoặc có-kết-nối
 - ❖ **Cài đặt:** ở phần network core (i.e. router)

Virtual circuits

“đường truyền từ nơi gửi đến nơi nhận, làm việc tương tự như mạch điện thoại

- ❖ performance-wise
- ❖ network actions along source-to-dest path

- Thiết lập cuộc truyền trước khi truyền dữ liệu, hủy bỏ cuộc truyền.
- Mỗi gói tin mang số hiệu của VC (VC-id), ko phải địa chỉ host bên nhận.
- Mọi router trên đường từ nơi gửi đến nơi nhận, đều duy trì “thông tin trạng thái” về từng kết nối đi qua nó.
- Các tài nguyên đường truyền, router (bandwidth, buffers) có thể được *cấp phát* cho VC (dedicated resources = predictable service)

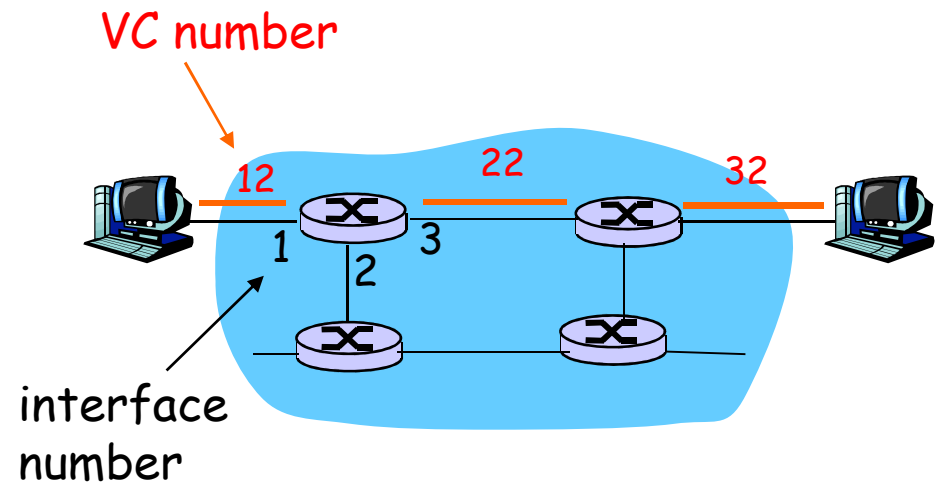
VC implementation

Một VC bao gồm:

1. Một tuyến đi từ nguồn, qua dãy gồm nhiều đường truyền và router, đến đích.
 2. Các *số hiệu VC* [VC numbers], một số cho mỗi đường truyền dọc theo tuyến.
 3. Các dòng trong bảng chỉ đường của router nằm dọc theo tuyến
- Các gói tin trực thuộc VC mang các số hiệu VC (thay vì địa chỉ đích đến)
 - Các số hiệu VC có thể bị thay đổi trên mỗi đường truyền.
 - ❖ Số hiệu VC mới đến từ forwarding table

Forwarding table

Forwarding table của router
phía trên-bên trái:



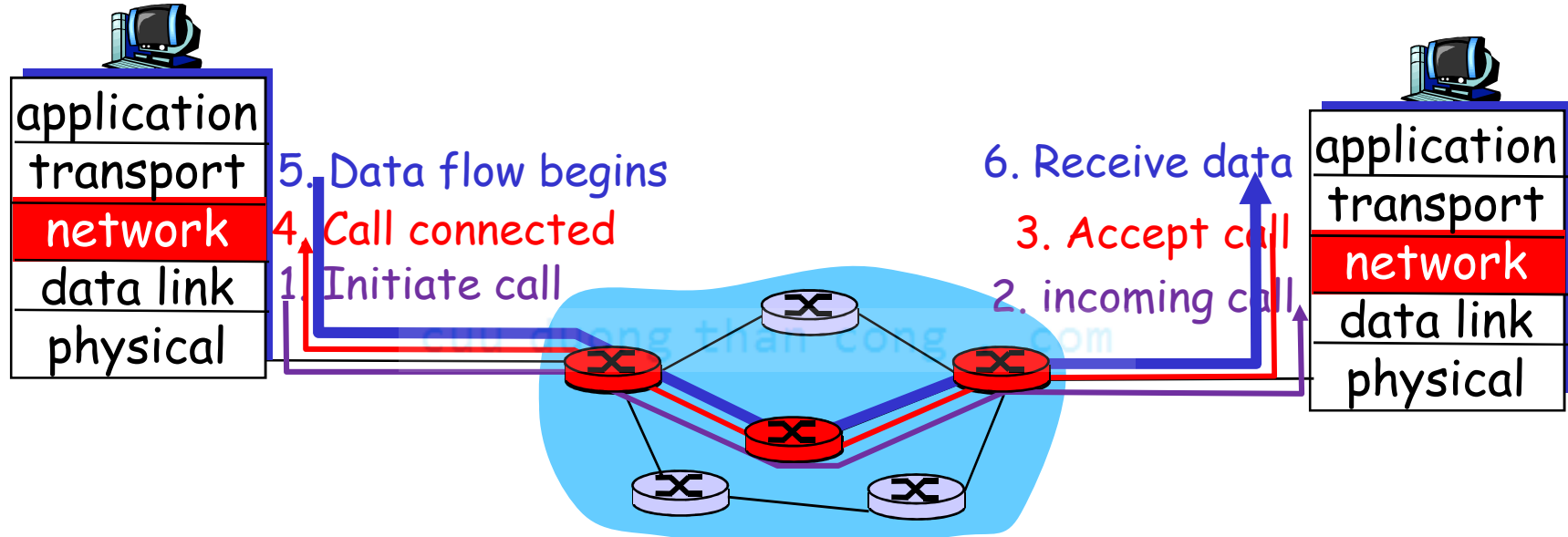
Incoming interface	Incoming VC #	Outgoing interface	Outgoing VC #
1	12	3	22
2	63	1	18
3	7	2	17
1	97	3	87
...

Routers quản lý các thông tin về trạng thái của kết nối!

Virtual circuits: signaling protocols

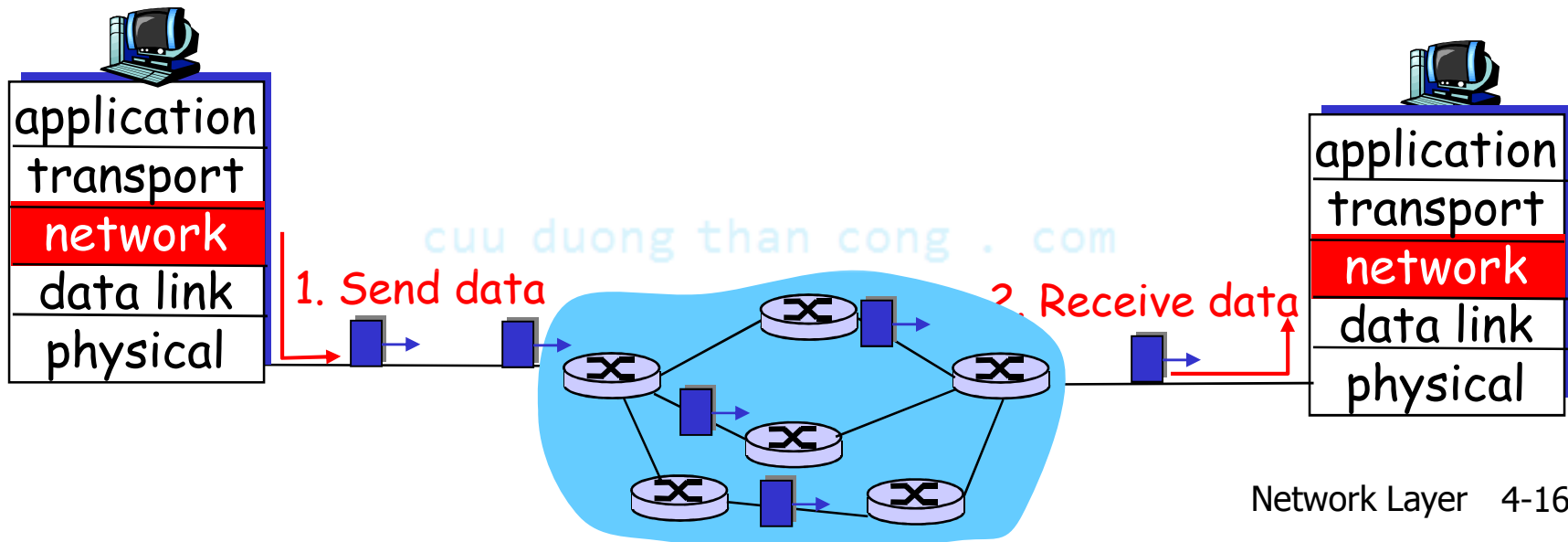
- Được sử dụng để thiết lập, duy trì, hủy bỏ VC
- Được sử dụng trong mạng ATM, frame-relay, X.25
- Không được sử dụng trong mạng Internet hiện nay.

cuu duong than cong . com



Datagram networks

- Không thiết lập cuộc truyền (hoặc kết nối) ở tầng mạng
- routers: Không quản lý thông tin trạng thái kết nối điểm-điểm
 - ❖ Không có khái niệm “kết nối” ở tầng mạng
- Các gói tin được chuyển tiếp giữa các router nhờ sử dụng địa chỉ của host đích.
 - ❖ Các gói tin truyền nhận giữa 2 máy có thể đi theo các tuyến khác nhau.



Forwarding table

4 billion
possible entries

Destination Address Range

Link Interface

11001000 00010111 00010000 00000000

đến

0

11001000 00010111 00010111 11111111

cuu duong than cong . com

11001000 00010111 00011000 00000000

đến

1

11001000 00010111 00011000 11111111

11001000 00010111 00011001 00000000

đến

2

cuu duong than cong . com

11001000 00010111 00011111 11111111

các trường hợp khác

3

Network Layer 4-17

Longest prefix matching

Khi có 1 địa chỉ nguồn khớp được nhiều mục trong forwarding table, ưu tiên khớp cho phần prefix dài nhất.

<u>Prefix Match</u>	<u>Link Interface</u>
11001000 00010111 00010	0
11001000 00010111 00011000	1
11001000 00010111 00011	2
các trường hợp khác	3

Ví dụ

DA: 11001000 00010111 00010110 10100001 Interface nào?

DA: 11001000 00010111 00011000 10101010 Interface nào?

Datagram or VC network: why?

Internet (datagram)

- Dữ liệu được trao đổi giữa các máy tính
 - ❖ “elastic” service, no strict timing requirements.
- Các đầu cuối là máy tính, “thông minh” hơn điện thoại
 - ❖ Có thể thích ứng, kiểm soát và khôi phục khi có lỗi
 - ❖ Bên trong mạng (router) được giữ đơn giản, các xử lý phức tạp được đặt tại “edge” (các host)
- Nhiều loại đường truyền
 - ❖ Có các đặc tính khác biệt
 - ❖ uniform service difficult

ATM (VC)

- Tiến hóa từ mạng điện thoại
- Đàm thoại:
 - ❖ strict timing, reliability requirements
 - ❖ need for guaranteed service
- Các đầu cuối là “không thông minh”
 - ❖ telephones
 - ❖ Các xử lý phức tạp được đặt ở mạng

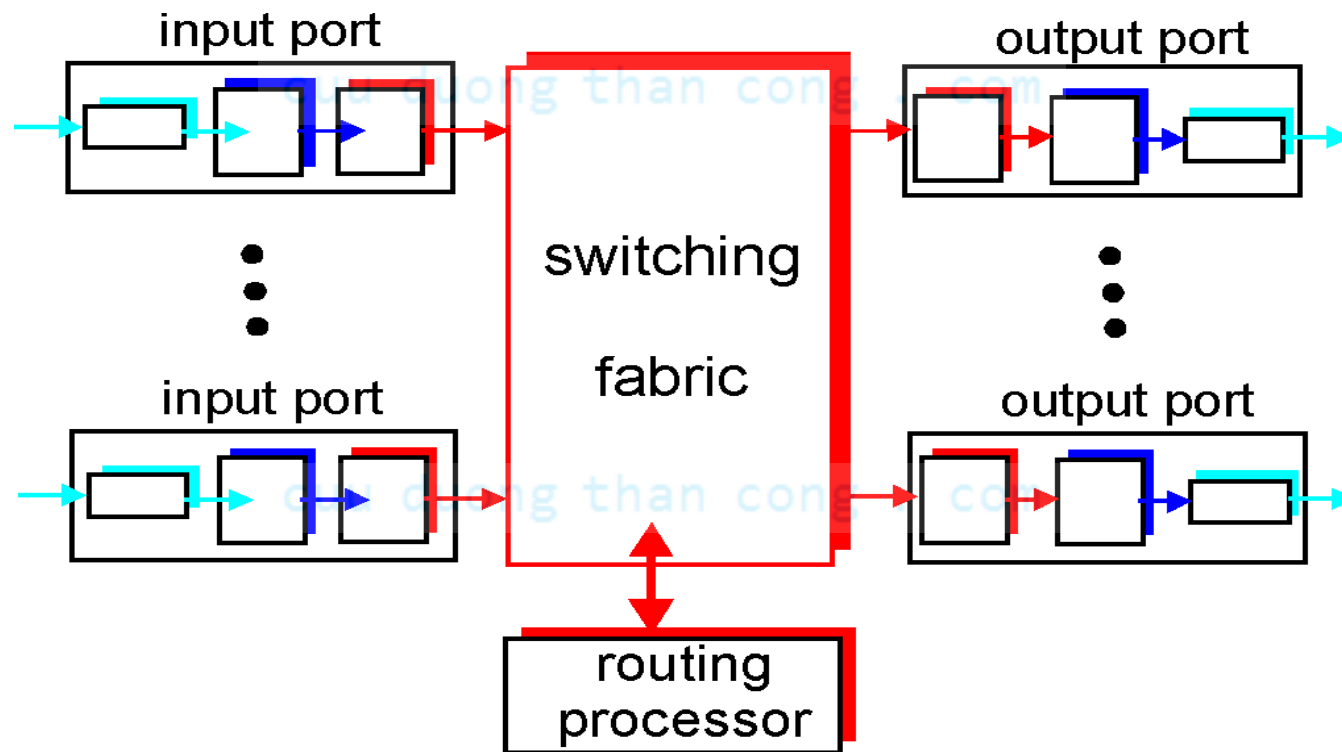
Chapter 4: Network Layer

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- 4.4 IP: Internet Protocol
 - ❖ Datagram format
 - ❖ IPv4 addressing
 - ❖ ICMP
 - ❖ IPv6
- 4.5 Routing algorithms
 - ❖ Link state
 - ❖ Distance Vector
 - ❖ Hierarchical routing
- 4.6 Routing in the Internet
 - ❖ RIP
 - ❖ OSPF
 - ❖ BGP
- 4.7 Broadcast and multicast routing

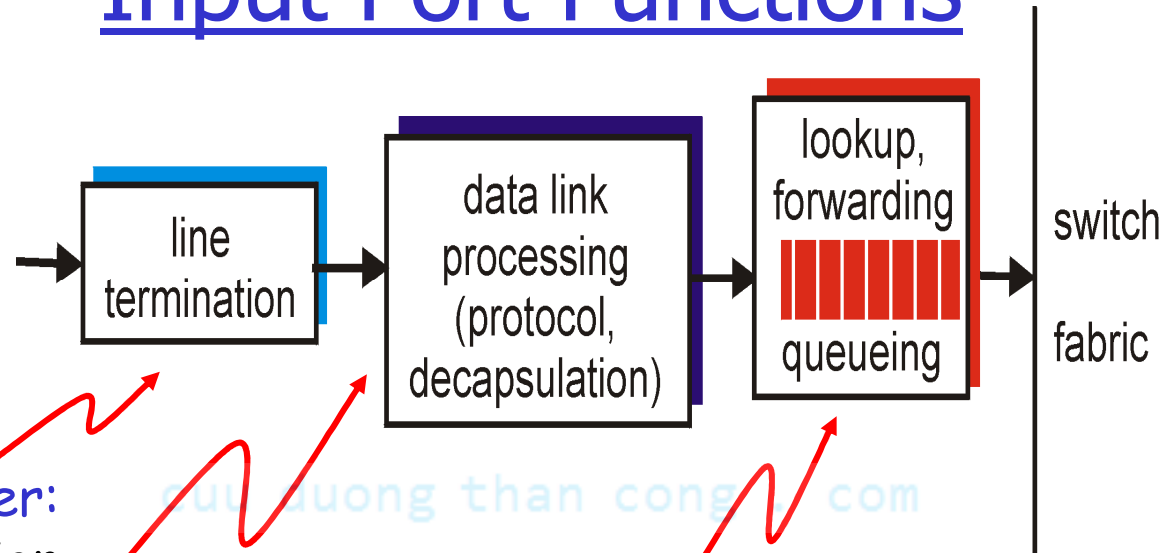
Router Architecture Overview

Hai chức năng chính của router:

- *Routing*: chạy các thuật toán/giao thức định tuyến (RIP, OSPF, BGP)
- *Forwarding*: Chuyển tiếp các datagrams từ ngõ vào đến ngõ ra



Input Port Functions



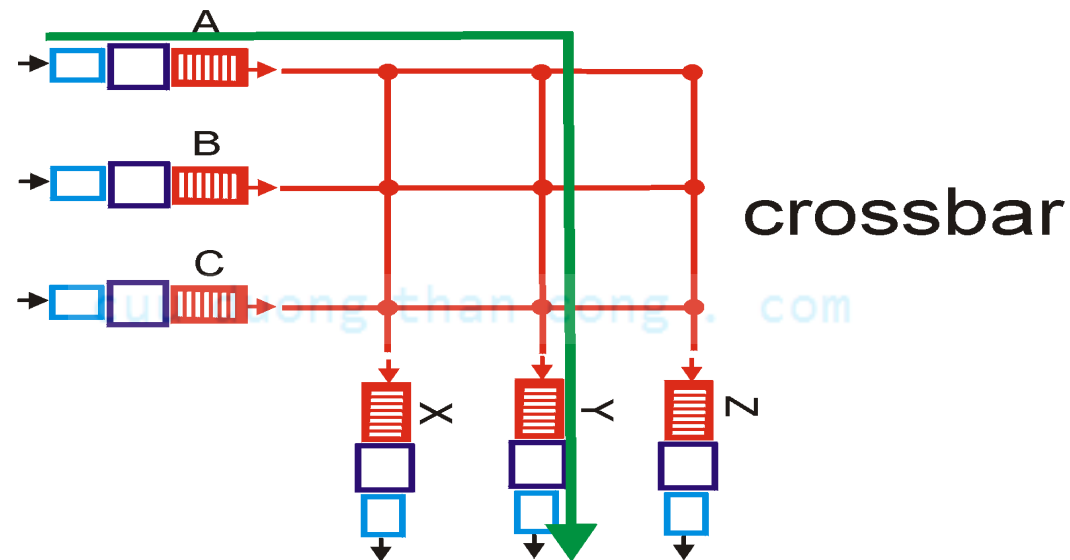
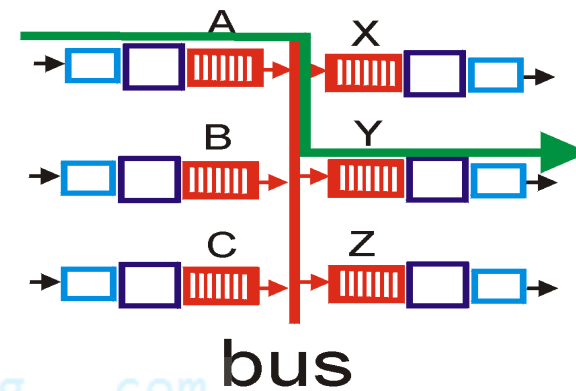
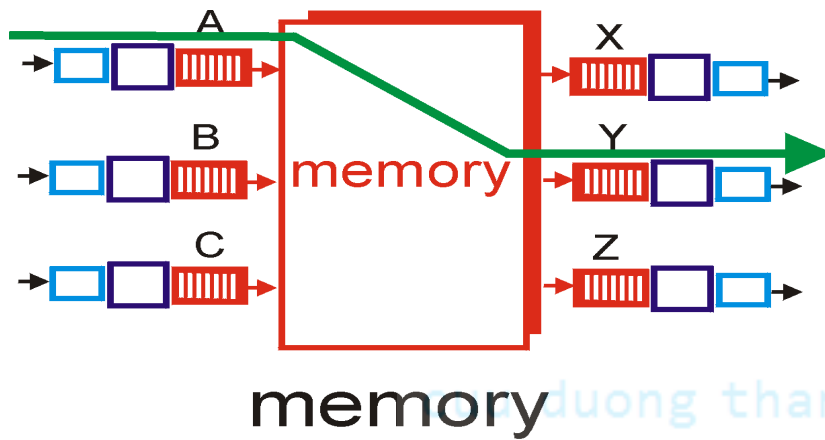
Physical layer:
bit-level reception

Data link layer:
e.g., Ethernet
see chapter 5

Decentralized switching:

- Với 1 datagram có địa chỉ đích, chức năng này tra cứu forwarding table trong bộ nhớ input port để tìm ra output port tương ứng.
- Mục tiêu: hoàn toàn xử lý tại input port để tốc độ đạt tới "line speed"
- Xếp hàng đợi: nếu datagrams đến nhanh hơn tốc độ forwarding vào switch fabric

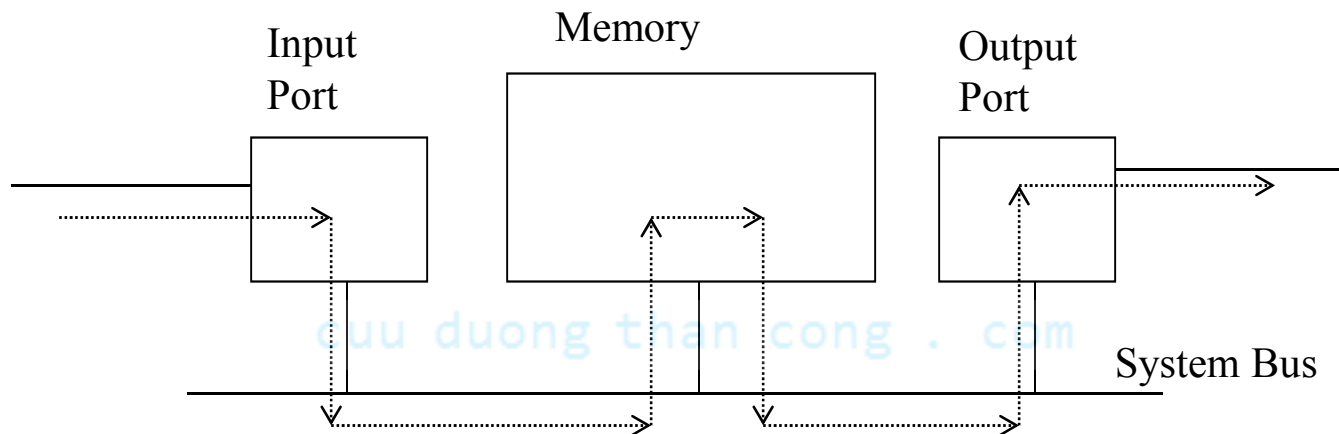
Three types of switching fabrics



Switching Via Memory

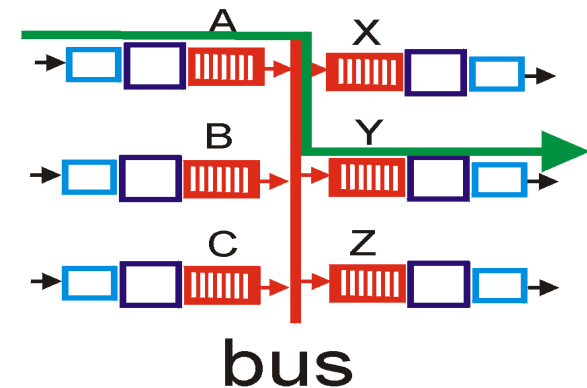
Các router thế hệ đầu:

- Theo cách các máy tính truyền thống xử lý, dưới sự điều khiển trực tiếp của CPU
- Các gói tin được chép trực tiếp vào bộ nhớ hệ thống.
- Tốc độ trao gói tin bị giới hạn bởi tốc độ đọc/ghi bộ nhớ



Switching Via a Bus

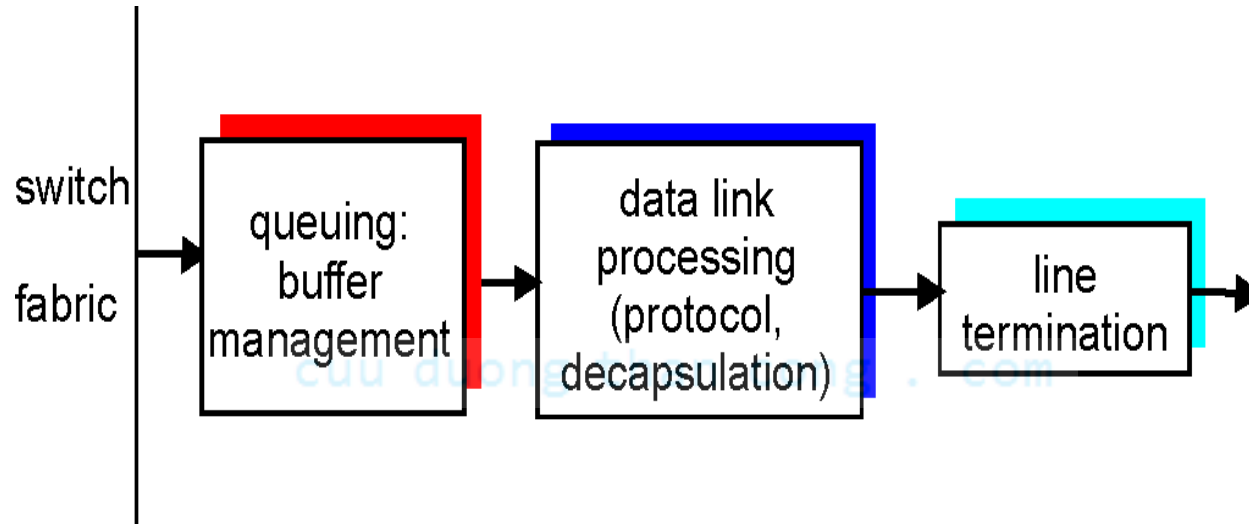
- Các datagram đi từ bộ nhớ ngõ vào đến bộ nhớ ngõ ra thông qua một bus dùng chung.
- **Tranh chấp bus:** mỗi thời điểm chỉ có 1 gói tin qua bus
- Tốc độ trao gói tin bị giới hạn bởi tốc độ bus.
- 32 Gbps bus, Cisco 5600: tốc độ đủ cho các router của mạng truy cập và mạng doanh nghiệp



Switching Via An Interconnection Network

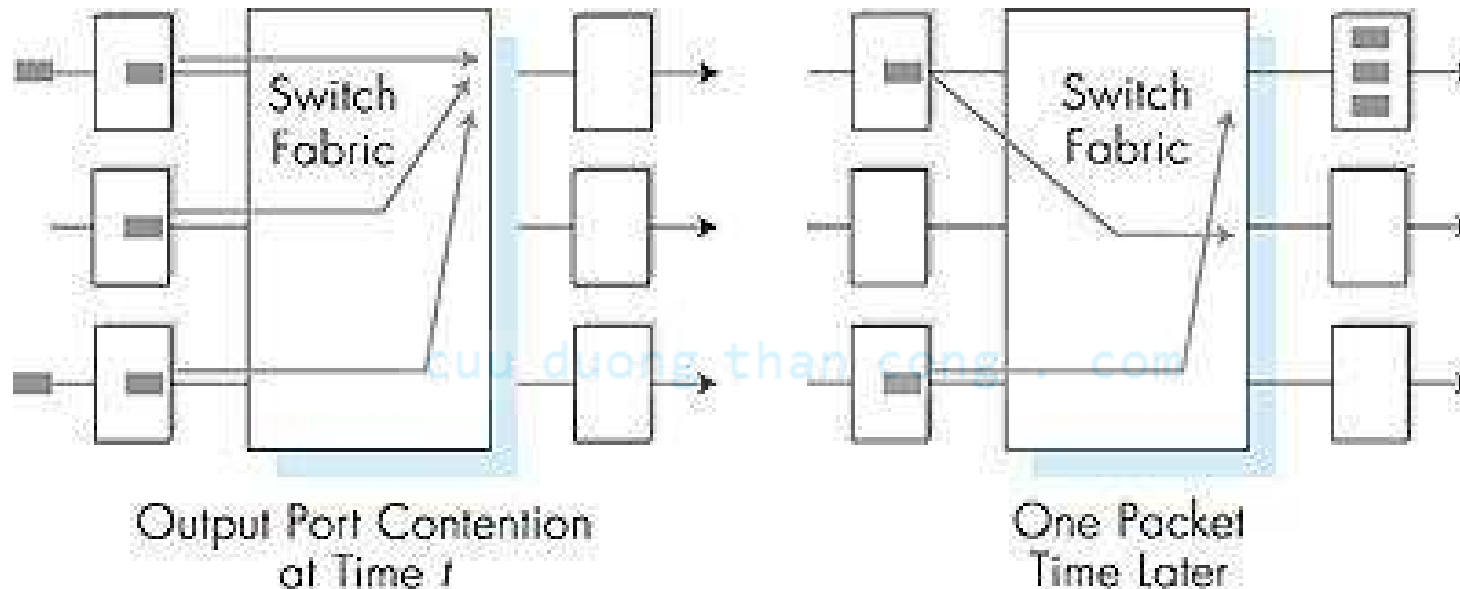
- Giải quyết các hạn chế do cách chuyển mạch bus
- Là cách mà Banyan networks, tiên phong sử dụng để kết nối các bộ xử lý trong hệ thống đa bộ xử lý. Cách này sử dụng hệ thống $2n$ bus ngang-dọc để nối kết n input port với n output port. Gói tin di chuyển dọc theo bus ngang, đến giao điểm với bus dọc dẫn đến outport mà nó muốn đến. Nếu bus dọc ở trạng thái free, thì gói tin được chuyển đến output port. Ngược lại thì gói tin phải xếp hàng đợi ở input port.
- Một khuynh hướng thiết kế khác: chia datagram thành những cell có chiều dài cố định, và chuyển các cell vào fabric.
- Cisco 12000: sử dụng interconnection network, cho phép chuyển ở tốc độ 60 Gbps qua switch fabric

Output Ports



- Lấy datagram từ bộ nhớ, chuyển datagram ra đường truyền.
- **Buffering** cần thiết khi datagram đến từ fabric nhanh hơn tốc độ truyền

Output port queueing



- Các gói tin xếp hàng đợi phía ngõ ra, khi tốc độ gói tin đến 1 ngõ ra vượt quá tốc độ đẩy gói tin ra đường truyền.
- *Khi bộ đệm (buffer) ngõ ra bị tràn, gói tin bị mất!*

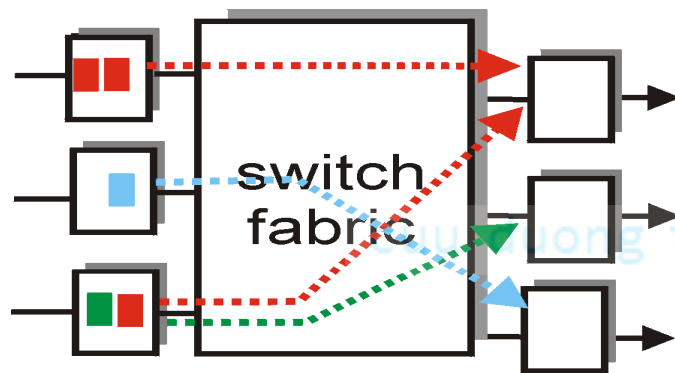
How much buffering?

- RFC 3439: dung lượng bộ đệm trung bình = (“typical” RTT) x (băng thông ngõ ra C)
 - ❖ Thông thường “typical” RTT=250 ms
 - ❖ VD, với C = 10 Gbps thì: dung lượng bộ đệm trung bình = 2.5 Gbits
- Recent recommendation: with N flows, buffering equal to

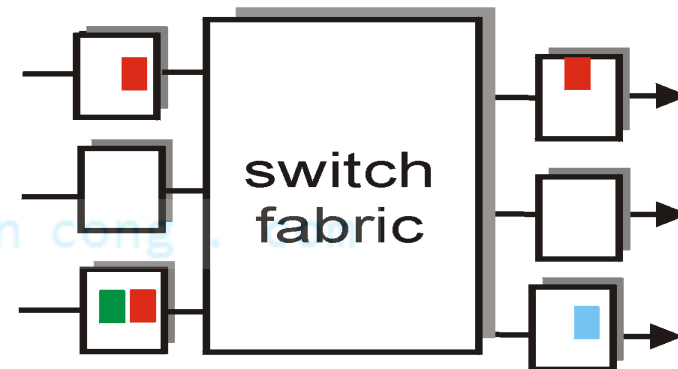
$$\frac{RTT \cdot C}{\sqrt{N}}$$

Input Port Queuing

- Tốc độ của switch fabric thường chậm hơn tốc độ kết hợp của các ngõ vào => các gói tin thường chờ ở các ngõ vào.
- **Hiện tượng "Head-of-the-Line (HOL) blocking"**: các gói tin đứng đầu hàng đợi bên ngõ vào, ngăn các gói tin khác đứng sau trong cùng hàng đợi, ngay cả khi ngõ ra mà gói tin đứng sau hướng đến hoàn toàn trống!
- *=> Nguyên nhân của việc chậm trễ và mất gói tin!*



output port contention
at time t - only one red
packet can be transferred



green packet
experiences HOL blocking

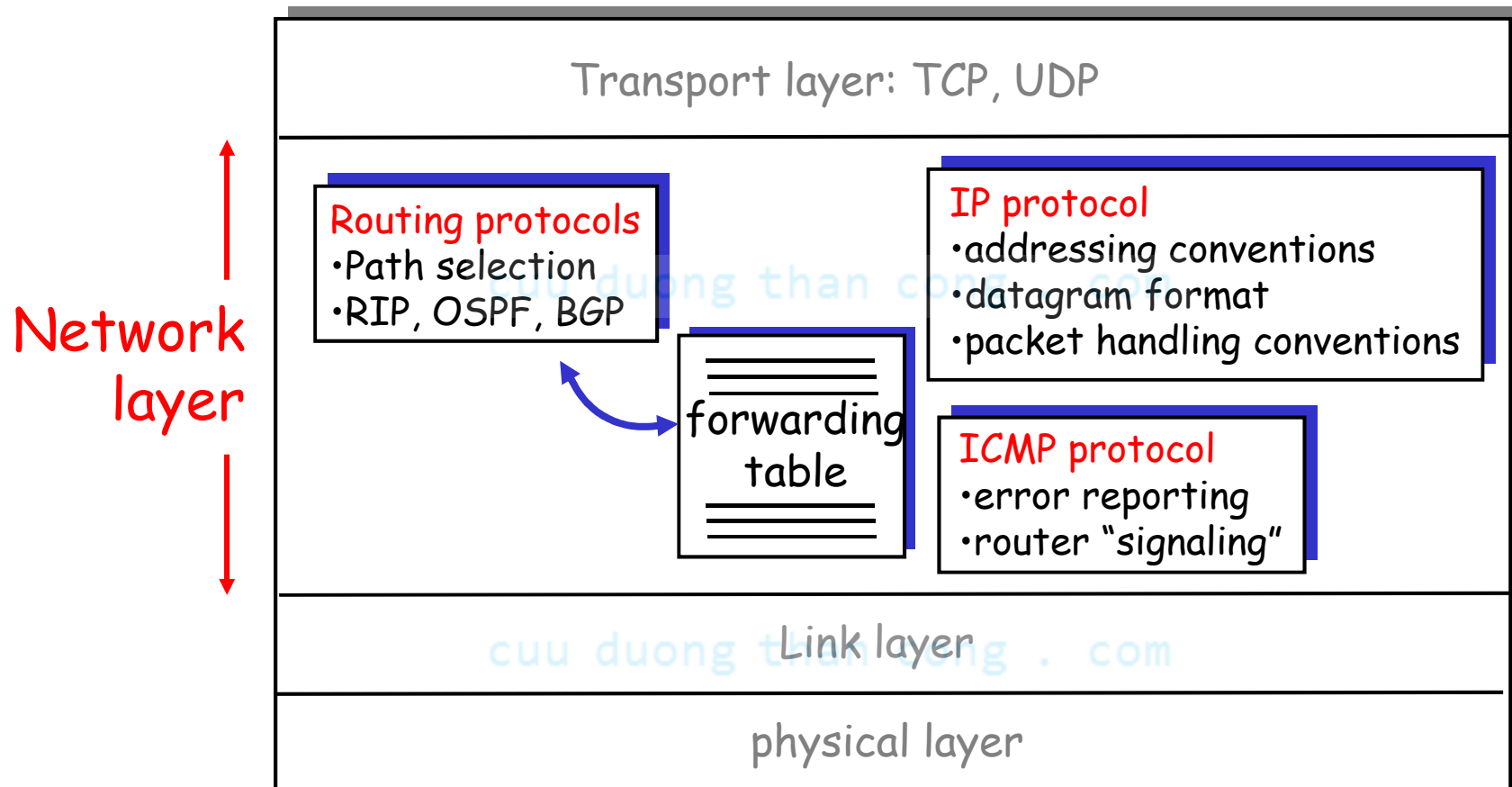
.....yer 4-30

Chapter 4: Network Layer

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- **4.4 IP: Internet Protocol**
 - ❖ Datagram format
 - ❖ IPv4 addressing
 - ❖ ICMP
 - ❖ IPv6
- 4.5 Routing algorithms
 - ❖ Link state
 - ❖ Distance Vector
 - ❖ Hierarchical routing
- 4.6 Routing in the Internet
 - ❖ RIP
 - ❖ OSPF
 - ❖ BGP
- 4.7 Broadcast and multicast routing

The Internet Network layer

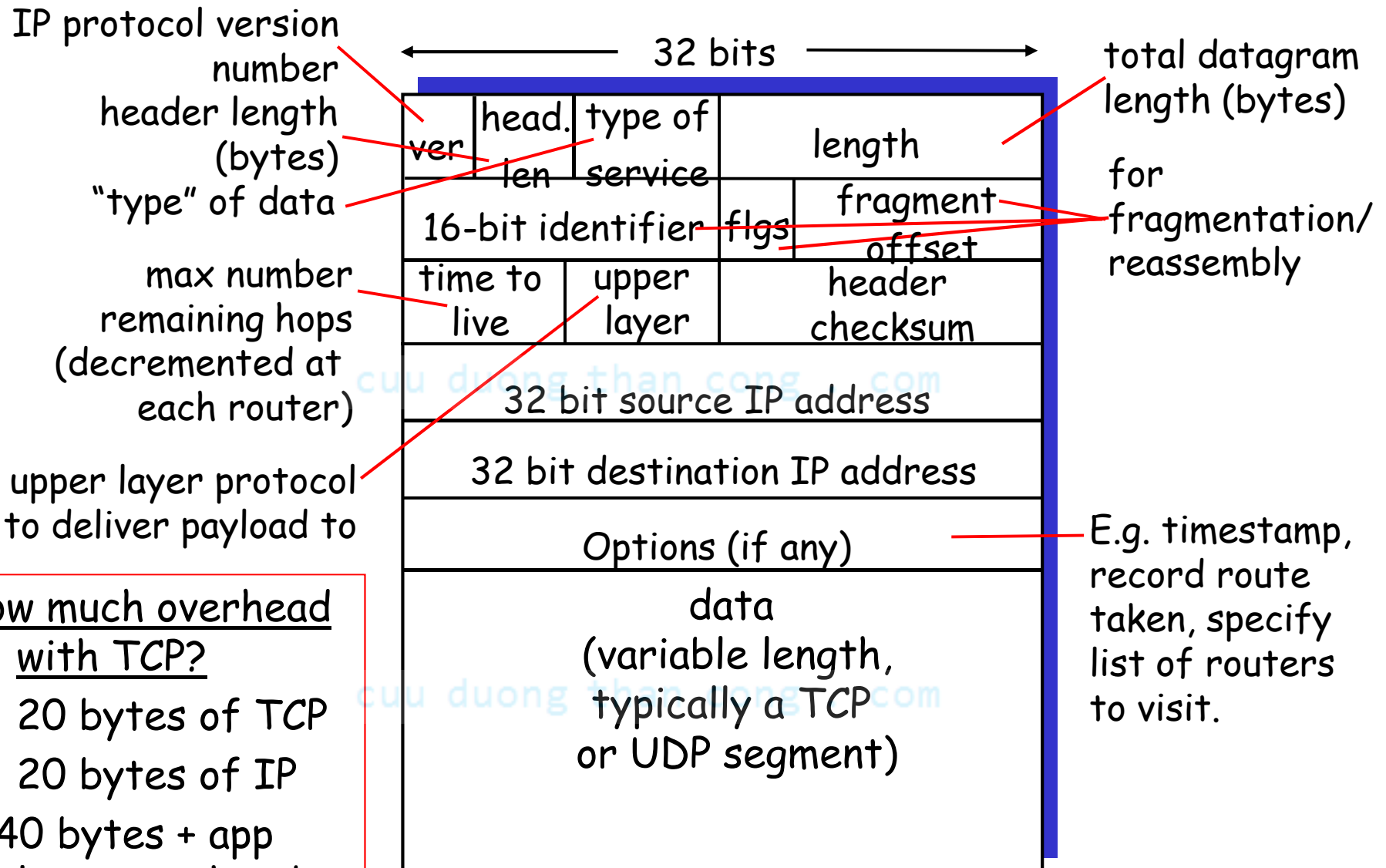
Các chức năng ở tầng Network:



Chapter 4: Network Layer

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- 4.4 IP: Internet Protocol
 - ❖ Datagram format
 - ❖ IPv4 addressing
 - ❖ ICMP
 - ❖ IPv6
- 4.5 Routing algorithms
 - ❖ Link state
 - ❖ Distance Vector
 - ❖ Hierarchical routing
- 4.6 Routing in the Internet
 - ❖ RIP
 - ❖ OSPF
 - ❖ BGP
- 4.7 Broadcast and multicast routing

IP datagram format

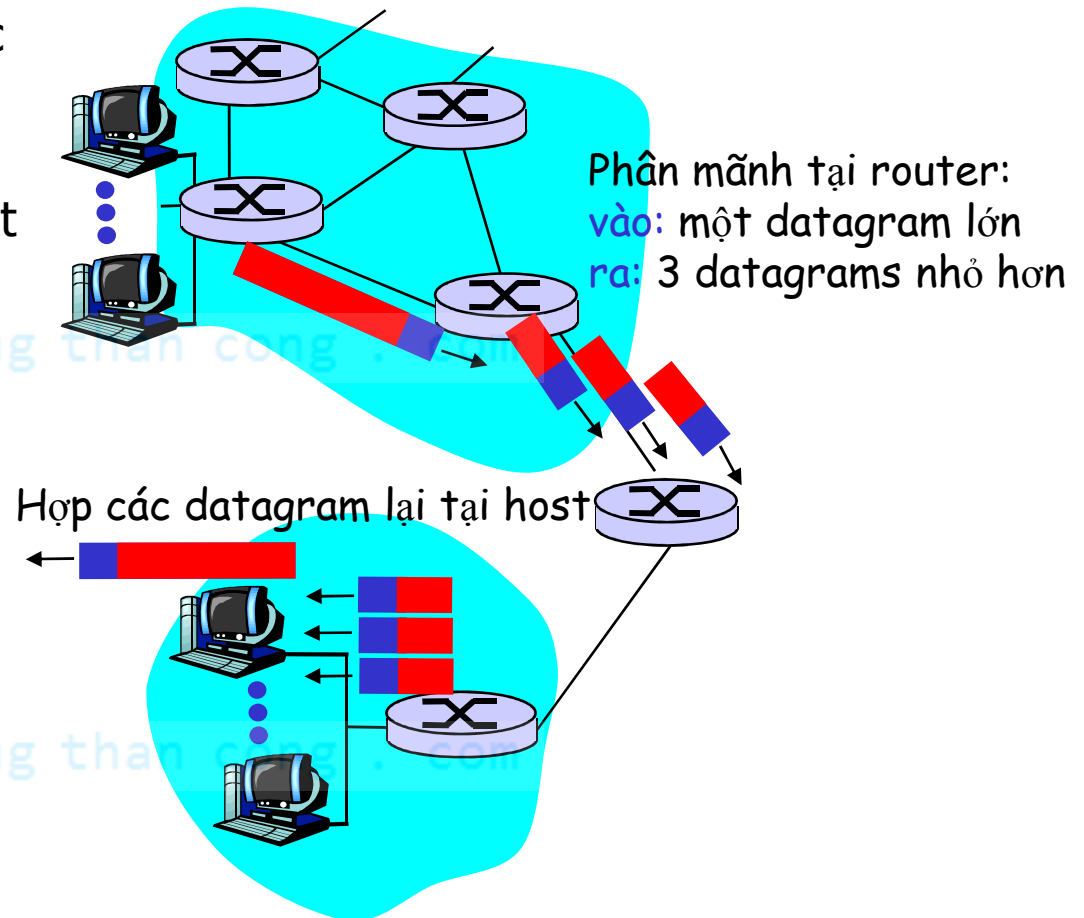


how much overhead with TCP?

- 20 bytes of TCP
 - 20 bytes of IP
- = 40 bytes + app layer overhead

IP Fragmentation & Reassembly

- Mỗi liên kết mạng đều có 1 MTU (max.transfer size) – kích thước khung ở tầng link lớn nhất có thể có.
 - ❖ different link types, different MTUs
- Các IP datagram lớn được chia thành các phân mảnh nhỏ hơn
 - ❖ Một datagram trở thành nhiều datagrams
 - ❖ Việc “hợp lại” các datagram được diễn ra ở máy nhận
 - ❖ Các bit của IP header được sử dụng để xác định, sắp thứ tự các phân mảnh.



IP Fragmentation and Reassembly

VD

- 4000 byte datagram
- MTU = 1500 bytes

1480 bytes in data field

offset =
 $1480/8$

	length	ID	fragflag	offset	
	=4000	=x	=0	=0	

Một datagram lớn được tách thành nhiều datagram nhỏ hơn

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

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

	length	ID	fragflag	offset	
	=1040	=x	=0	=370	

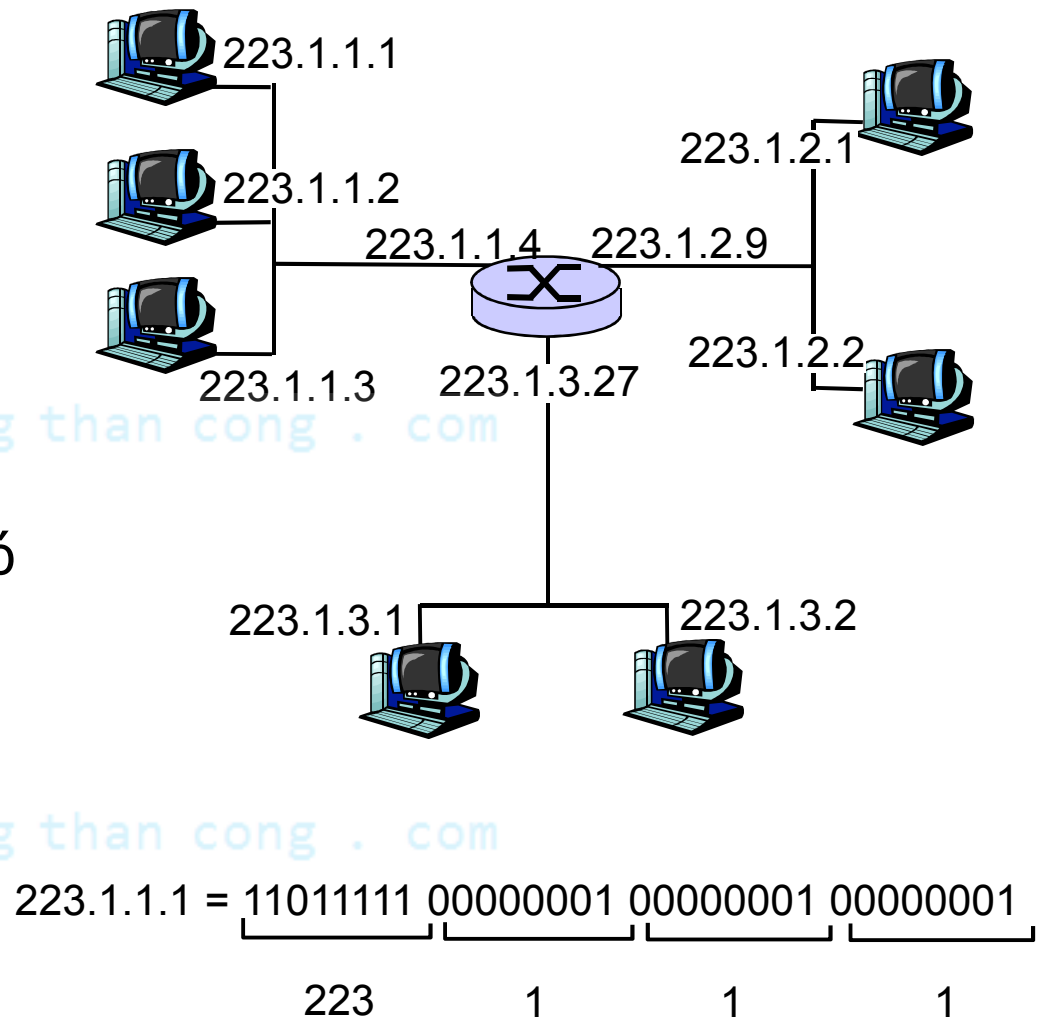
Lưu ý: Mỗi một trị số offset ứng 1 vùng 8 bytes

Chapter 4: Network Layer

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- 4.4 IP: Internet Protocol
 - ❖ Datagram format
 - ❖ IPv4 addressing
 - ❖ ICMP
 - ❖ IPv6
- 4.5 Routing algorithms
 - ❖ Link state
 - ❖ Distance Vector
 - ❖ Hierarchical routing
- 4.6 Routing in the Internet
 - ❖ RIP
 - ❖ OSPF
 - ❖ BGP
- 4.7 Broadcast and multicast routing

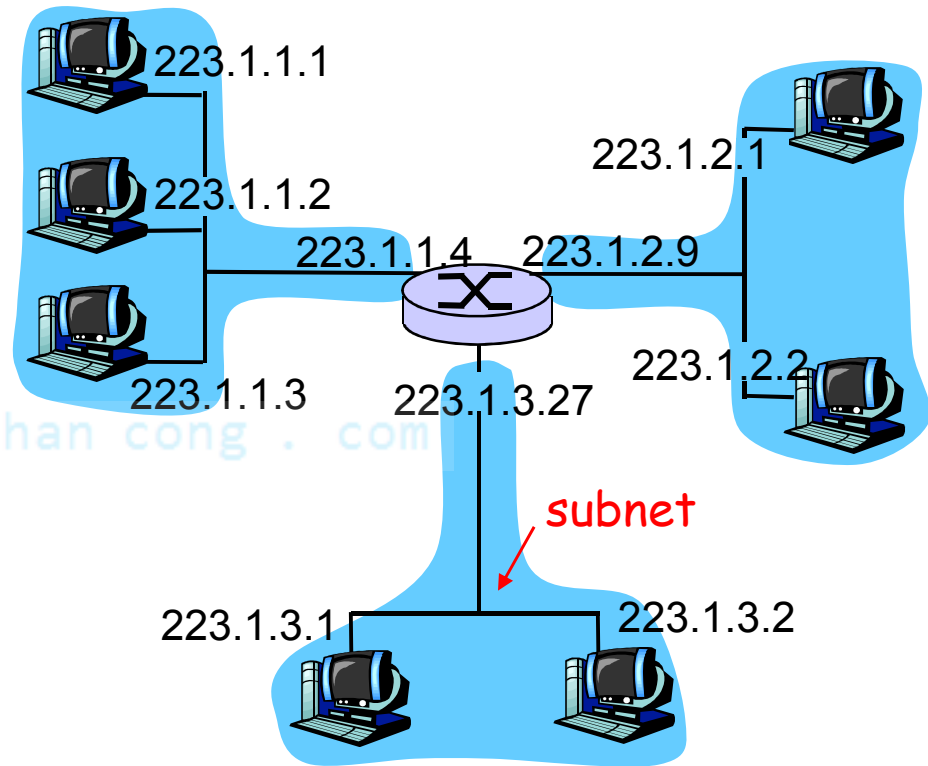
IP Addressing: introduction

- **IP address:** dài 32-bit để xác định *interface* của host/router
- **interface:** nơi giao tiếp giữa host/router và đường truyền vật lý
 - ❖ Thông thường, router có nhiều interfaces
 - ❖ Thông thường host có 1 interface
 - ❖ Địa chỉ IP được liên kết với từng Interface



Subnets

- Địa chỉ IP : 2 phần
 - ❖ Phần subnet (các bit nằm trước)
 - ❖ Phần host (các bit nằm sau)
- *Subnet là gì?*
 - ❖ Là các interface có cùng phần subnet trong địa chỉ IP
 - ❖ Các interface có thể nối trực tiếp nhau mà không qua router trung gian.

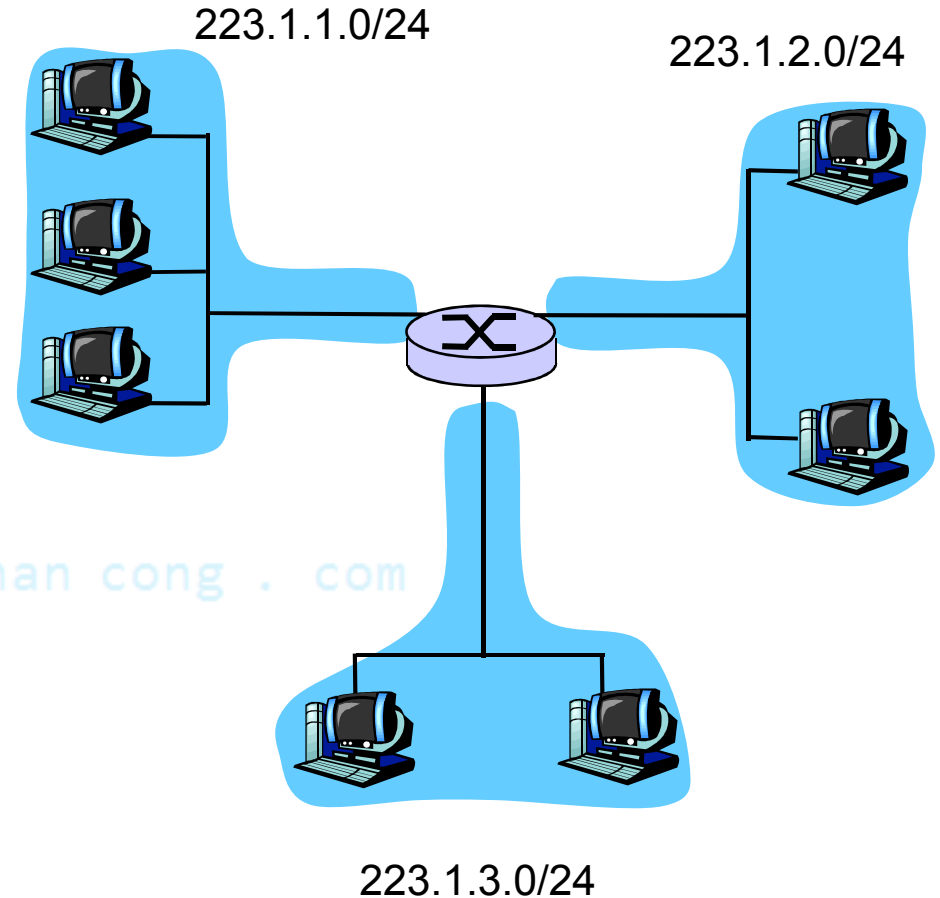


Mạng gồm 3 subnets

Subnets

Quy tắc

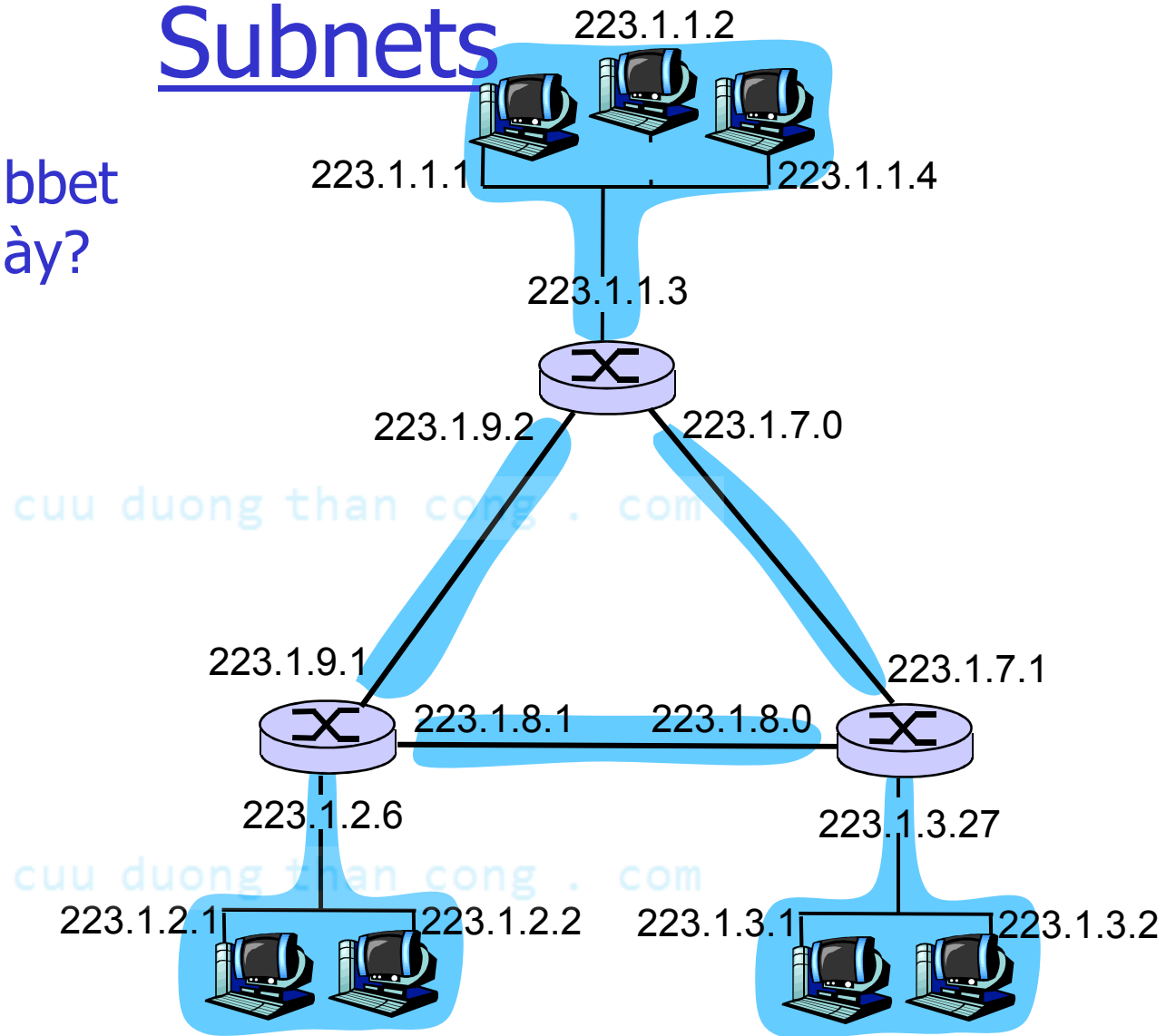
- Để xác định subnet, hãy tách các interface từ các host hoặc router của nó. Với các interface có đường nối trực tiếp, hãy tạo ra một "tiểu đảo" chứa các interface được tách ra. Mạng được tách ra thành nhiều "tiểu đảo" rời nhau. Mỗi "tiểu đảo" chính là một **subnet**.



Subnet mask: /24

Subnets

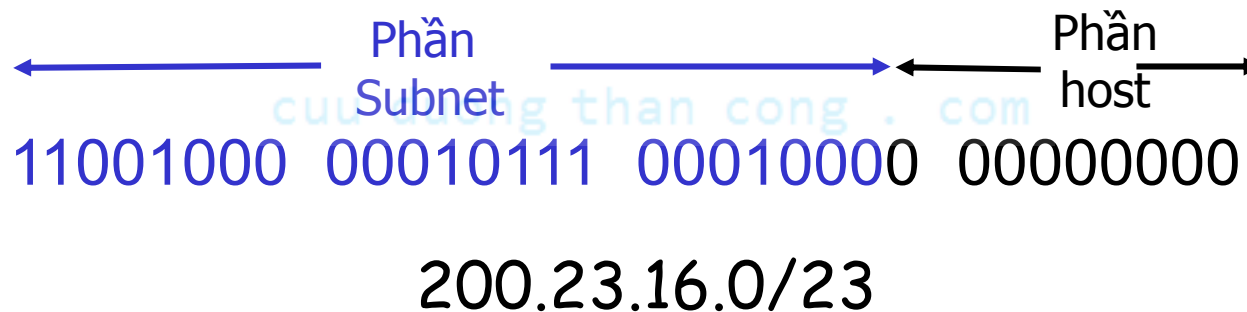
Có bao nhiêu subbet
trong mạng này?



IP addressing: CIDR

CIDR: Classless InterDomain Routing

- ❖ Phần subnet của địa chỉ IP có chiều dài tùy ý.
- ❖ Khuôn dạng địa chỉ IP: **a.b.c.d/x**, trong đó x là số bit của phần subnet trong địa chỉ



IP addresses: how to get one?

Q: Host có địa chỉ IP như thế nào?

- Quản trị mạng ấn định “cứng” trong file
 - ❖ Windows: control-panel->network->configuration->tcp/ip->properties
 - ❖ UNIX: /etc/rc.config
- **DHCP: D**ynamic **H**ost **C**onfiguration **P**rotocol: Host xin cấp phát địa chỉ IP từ 1 server
 - ❖ “plug-and-play”

DHCP: Dynamic Host Configuration Protocol

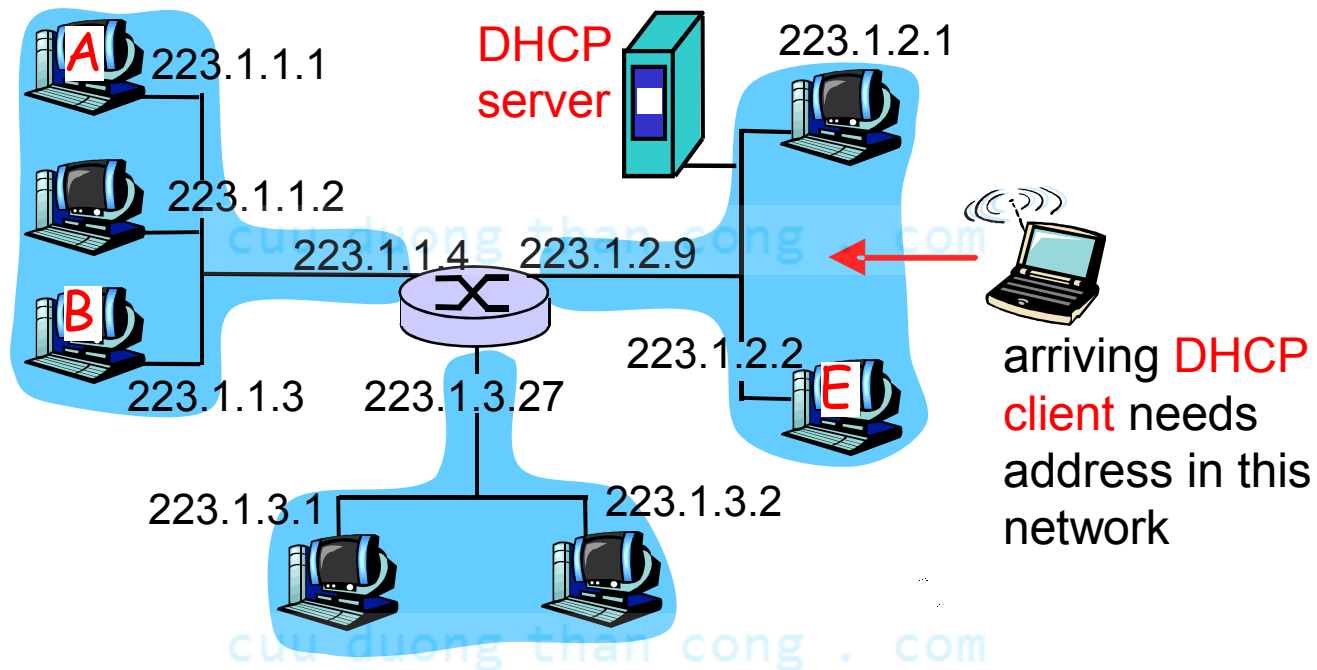
Mục tiêu: Cho phép host xin cấp phát động 1 địa chỉ IP từ server khi host gia nhập mạng.

- ❖ Có thể xin cấp mới địa chỉ IP
- ❖ Cho phép dùng lại địa chỉ IP
- ❖ Hỗ trợ cho người dùng di động, muốn gia nhập mạng.

Giới thiệu chung về DHCP :

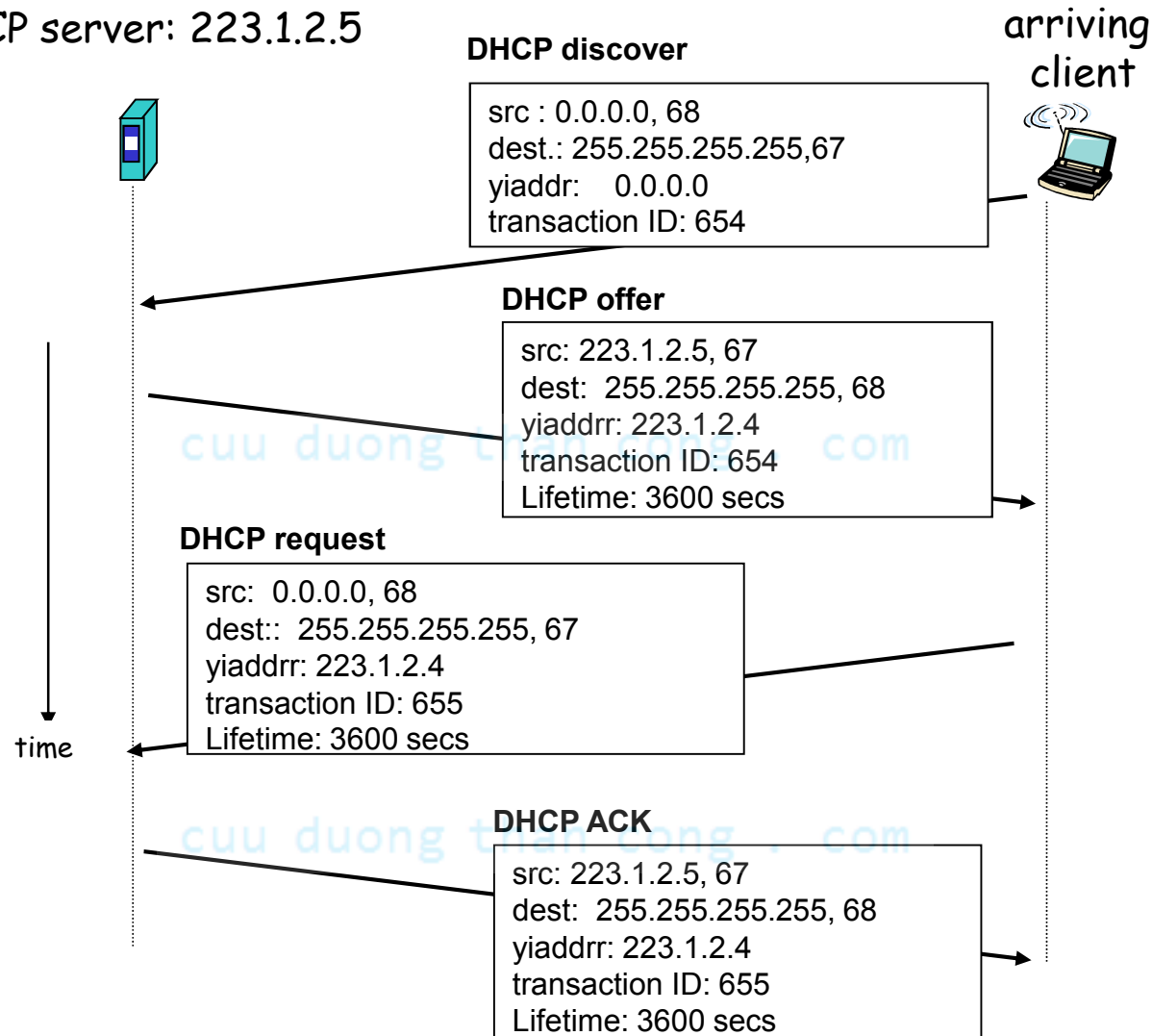
- ❖ host gửi quảng bá thông điệp “DHCP discover”
- ❖ DHCP server trả lời bằng thông điệp “DHCP offer”
- ❖ host yêu cầu địa chỉ IP bằng thông điệp “DHCP request”
- ❖ DHCP server gửi địa chỉ qua thông điệp “DHCP ack”

DHCP client-server scenario



DHCP client-server scenario

DHCP server: 223.1.2.5



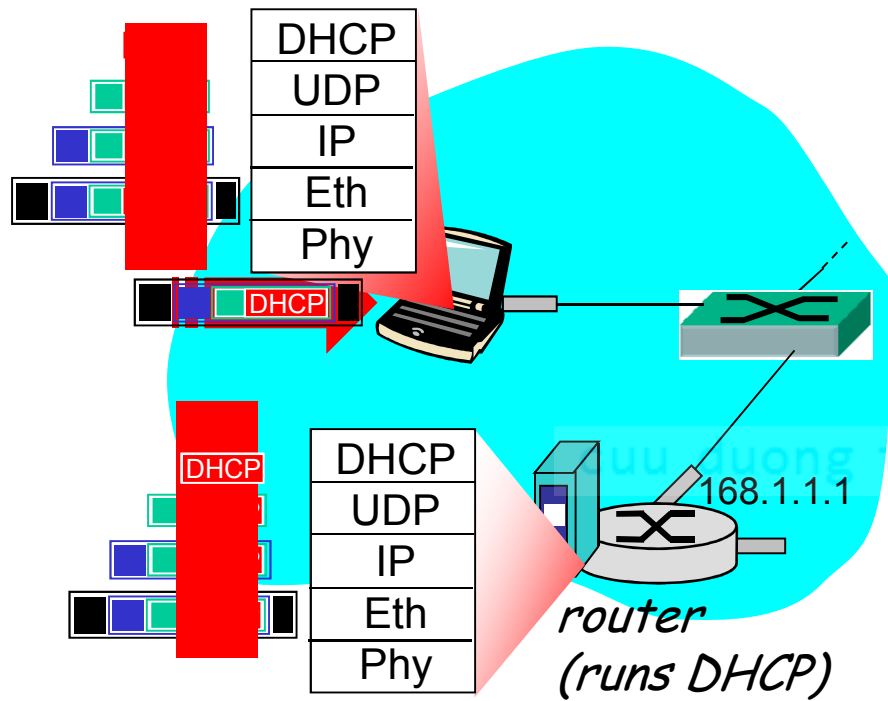
DHCP: more than IP address

DHCP có thể trả về cho host thêm nhiều thông tin khác, ngoài địa chỉ IP cấp phát.

- ❖ Địa chỉ của router đầu chặng (i.e. default gateway)
- ❖ Tên và địa chỉ IP của DNS sever.
- ❖ network mask

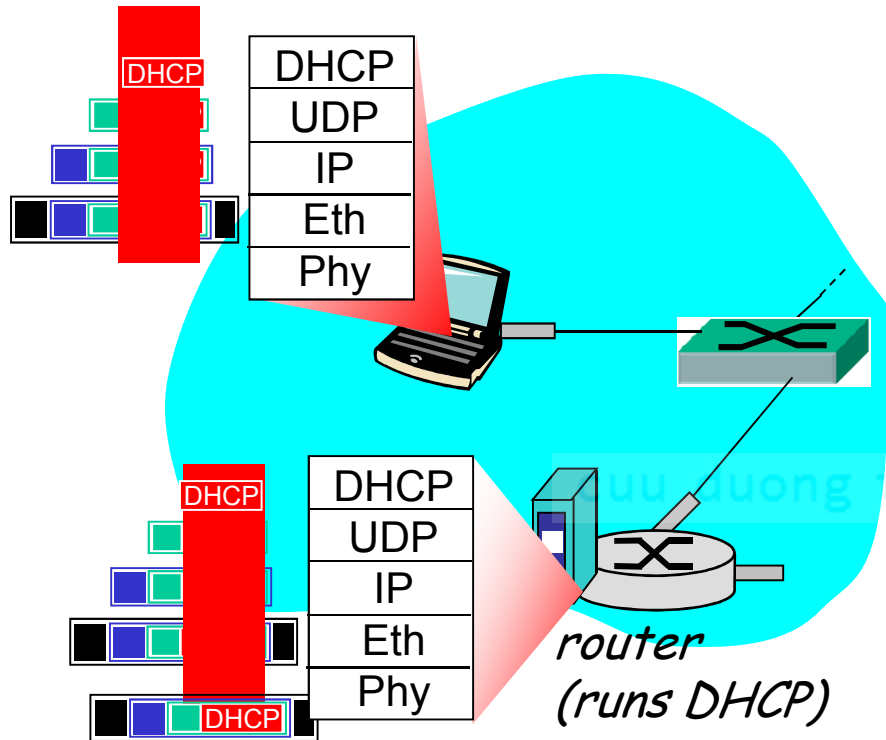
cuu duong than cong . com

DHCP: example



- Máy laptop cần địa chỉ IP, địa chỉ IP của default gateway, địa chỉ IP của DNS Server: Sử dụng DHCP
- Thông điệp "DHCP request", được gói trong UDP segment, rồi được gói tiếp trong IP datagram, rồi được gói tiếp trong Ethernet 802.1 frame
- Ethernet frame được gửi quảng bá (gửi đến mọi máy trên LAN, dest FFFFFFFFFFFFFFFF). DHCP server (đang chạy ở router) nhận được.
- Tại router, Ethernet tách lấy IP datagram, IP tách lấy UDP segment, UDP tách lấy DHCP request.

DHCP: example



- DHCP tạo ra thông điệp "DHCP ACK" chứa địa chỉ IP cấp cho laptop, địa chỉ IP của default gateway, địa chỉ IP của DNS Server.
- Đóng gói thông điệp tại DHCP server, frame được chuyển đến laptop, và tại đây frame được xử lý để lấy thông điệp DHCP ACK
- Giờ đây laptop biết được địa chỉ IP của nó, tên và địa chỉ IP của DNS server, địa chỉ IP của default gateway.

DHCP: wireshark output (home LAN)

Message type: **Boot Request (1)**

Hardware type: Ethernet

Hardware address length: 6

Hops: 0

Transaction ID: 0x6b3a11b7

Seconds elapsed: 0

Bootp flags: 0x0000 (Unicast)

Client IP address: 0.0.0.0 (0.0.0.0)

Your (client) IP address: 0.0.0.0 (0.0.0.0)

Next server IP address: 0.0.0.0 (0.0.0.0)

Relay agent IP address: 0.0.0.0 (0.0.0.0)

Client MAC address: Wistron_23:68:8a (00:16:d3:23:68:8a)

Server host name not given

Boot file name not given

Magic cookie: (OK)

Option: (t=53,l=1) **DHCP Message Type = DHCP Request**

Option: (61) Client identifier

Length: 7; Value: 010016D323688A;

Hardware type: Ethernet

Client MAC address: Wistron_23:68:8a (00:16:d3:23:68:8a)

Option: (t=50,l=4) Requested IP Address = 192.168.1.101

Option: (t=12,l=5) Host Name = "nomad"

Option: (55) Parameter Request List

Length: 11; Value: 010F03062C2E2F1F21F92B

1 = Subnet Mask; 15 = Domain Name

3 = Router; 6 = Domain Name Server

44 = NetBIOS over TCP/IP Name Server

.....

request

Message type: **Boot Reply (2)**

Hardware type: Ethernet

Hardware address length: 6

Hops: 0

Transaction ID: 0x6b3a11b7

Seconds elapsed: 0

Bootp flags: 0x0000 (Unicast)

Client IP address: 192.168.1.101 (192.168.1.101)

Your (client) IP address: 0.0.0.0 (0.0.0.0)

Next server IP address: 192.168.1.1 (192.168.1.1)

Relay agent IP address: 0.0.0.0 (0.0.0.0)

Client MAC address: Wistron_23:68:8a (00:16:d3:23:68:8a)

Server host name not given

Boot file name not given

Magic cookie: (OK)

Option: (t=53,l=1) DHCP Message Type = DHCP ACK

Option: (t=54,l=4) Server Identifier = 192.168.1.1

Option: (t=1,l=4) Subnet Mask = 255.255.255.0

Option: (t=3,l=4) Router = 192.168.1.1

Option: (6) Domain Name Server

Length: 12; Value: 445747E2445749F244574092;

IP Address: 68.87.71.226;

IP Address: 68.87.73.242;

IP Address: 68.87.64.146

Option: (t=15,l=20) Domain Name = "hsd1.ma.comcast.net."

reply

IP addresses: how to get one?

Q: Một tổ chức có được vùng địa chỉ IP như thế nào?

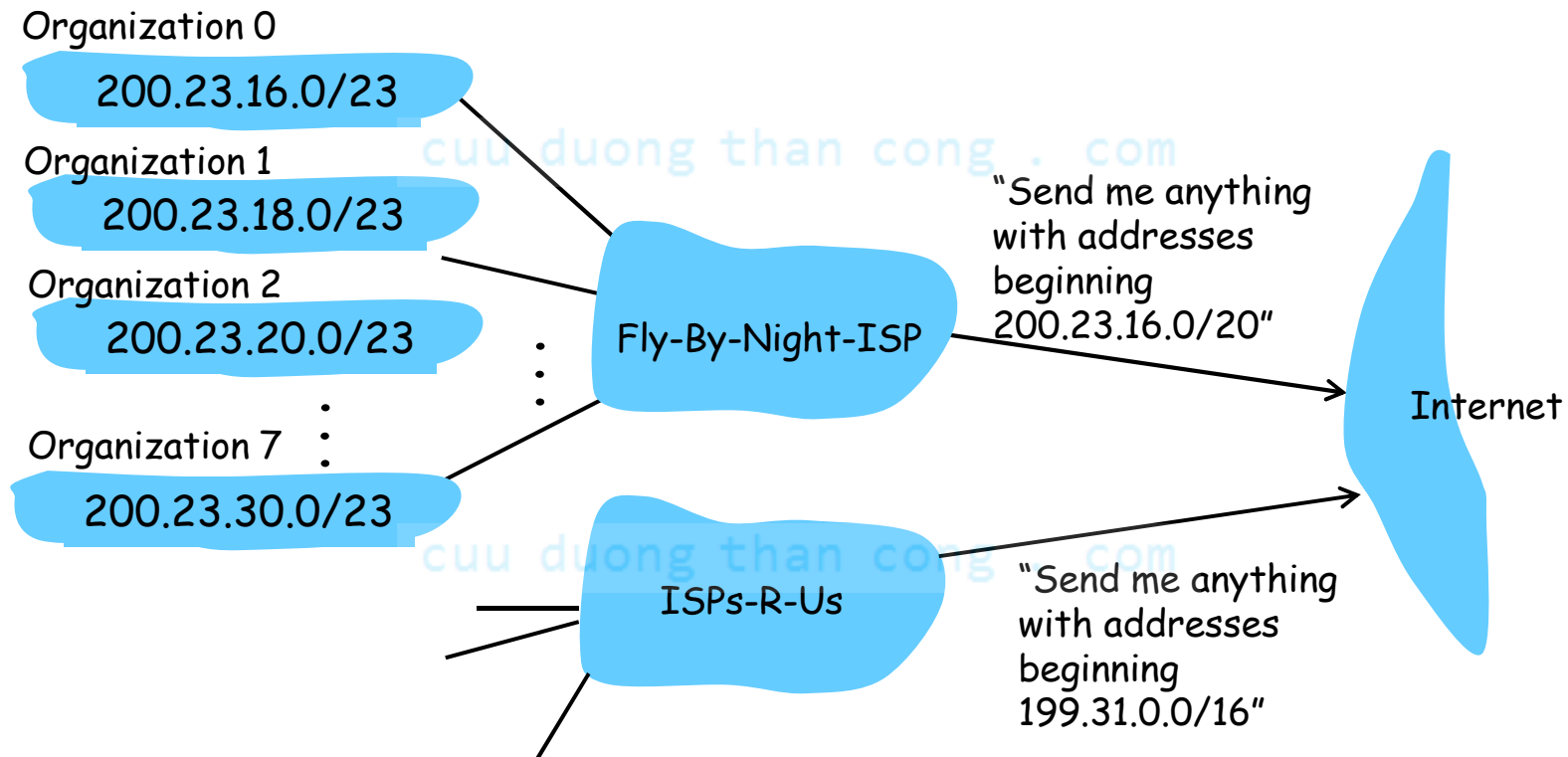
A: Xin ISP cấp phát. ISP sẽ cấp lại cho một subnet lấy từ không gian địa chỉ mà ISP đã được cấp phát trước đó.

cuu duong than cong . com

ISP's block	<u>11001000 00010111 00010000</u> 00000000	200.23.16.0/20
Organization 0	<u>11001000 00010111 00010000</u> 00000000	200.23.16.0/23
Organization 1	<u>11001000 00010111 00010010</u> 00000000	200.23.18.0/23
Organization 2	<u>11001000 00010111 00010100</u> 00000000	200.23.20.0/23
...	cuu duong than cong . com
Organization 7	<u>11001000 00010111 00011110</u> 00000000	200.23.30.0/23

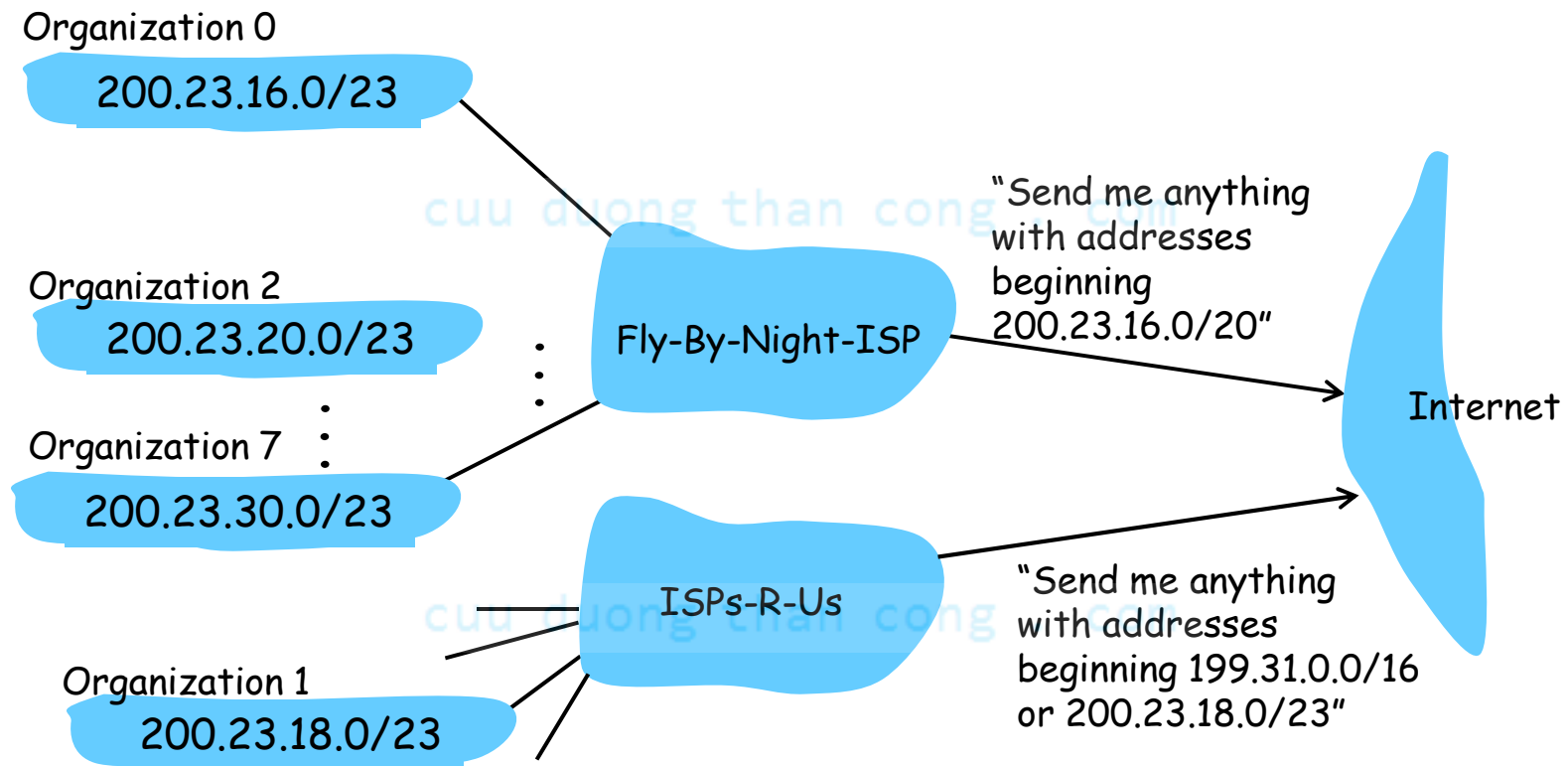
Hierarchical addressing: route aggregation

- Route aggregation: Khả năng sử dụng 1 network prefix để quảng cáo các thông tin định tuyến cho nhiều mạng:



Hierarchical addressing: more specific routes

Trường hợp ISPs-R-Us có thêm 1 tuyến đến Organization 1



IP addressing: the last word...

Q: Một ISP xin một dãy địa chỉ như thế nào?

A: **ICANN**: **I**nternet **C**orporation for **A**ssigned
Names and **N**umbers

- ❖ Cấp phát địa chỉ
- ❖ Quản lý DNS
- ❖ Ấn định tên miền, giải quyết các tranh chấp

IPv4: Tình trạng cạn kiệt và giải pháp

- Tình trạng thiếu địa chỉ IPv4
 - ❖ 1/10/2010: Chỉ còn 190 khoảng triệu địa chỉ (5%)
 - ❖ Dự kiến 5/2011: Hết địa chỉ IPv4!
- Giải pháp: [cuu duong than cong . com](http://cuuduongthancong.com)
 - ❖ NAT (Network Address Translation) + Private Network
 - ❖ Virtual hosting
 - ❖ IPv6
 - ❖ ... [cuu duong than cong . com](http://cuuduongthancong.com)

IPv4: Tình trạng cạn kiệt và giải pháp

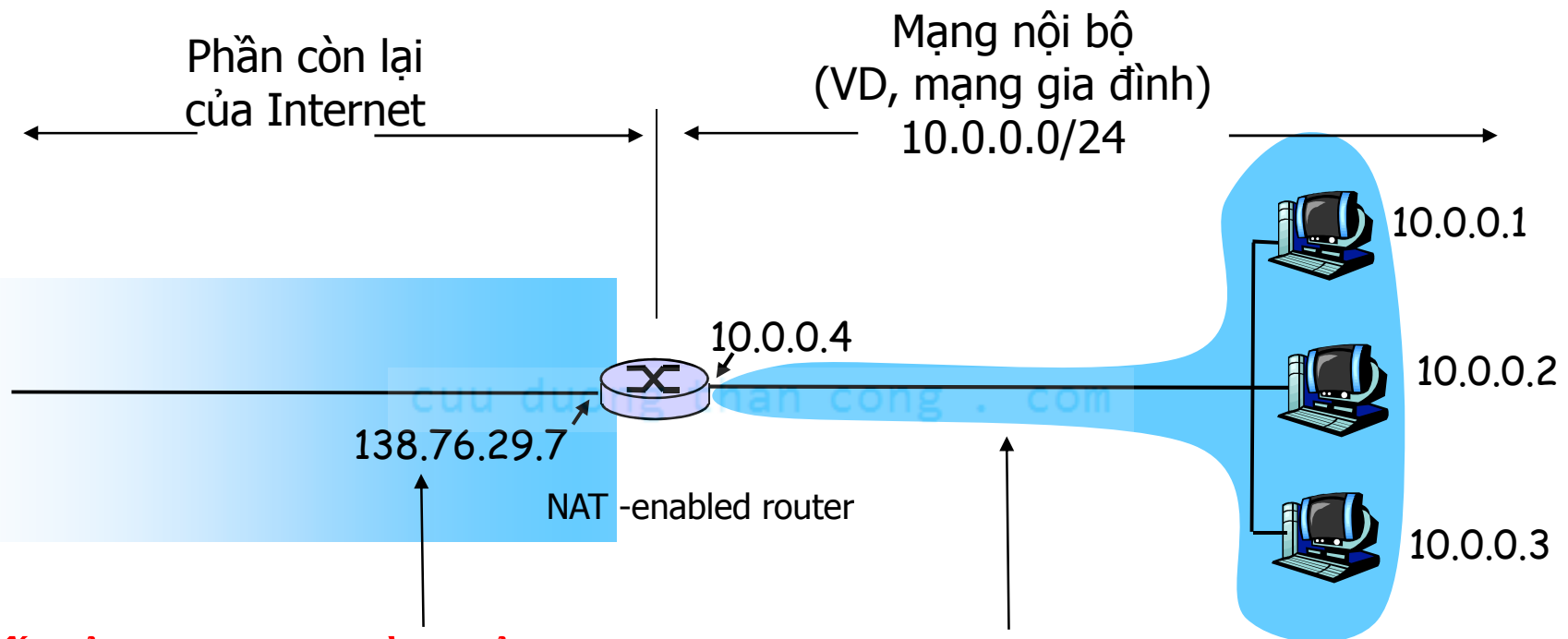
- Tình trạng thiếu địa chỉ IPv4
 - ❖ 16/11/2010: Chỉ còn khoảng 176 triệu địa chỉ (4%) (<http://www.ipv4exhaustioncounter.com/>)
 - ❖ Dự kiến 11/3/2011: Hết địa chỉ IPv4!
- Giải pháp:
 - ❖ NAT
 - ❖ IPv6

cuu duong than cong . com

IPv4: Tình trạng cạn kiệt và giải pháp

- RFC 1918: Dành riêng 3 vùng địa chỉ cho các thiết bị trong mạng nội bộ, gọi là các private address.
 - 10.0.0.0 - 10.255.255.255 (10/8 prefix)
 - 172.16.0.0 - 172.31.255.255 (172.16/12 prefix)
 - 192.168.0.0 - 192.168.255.255 (192.168/16 prefix)
- Q: Các địa chỉ private chỉ có giá trị trong mạng nội bộ. Làm thế nào để một máy tính sử dụng địa chỉ private, có thể gửi/nhận các datagram ra mạng Internet?
- A: có thể dùng NAT (Network Address Translation)

NAT: Network Address Translation



Tất cả datagrams *rời khỏi* mạng nội bộ, có *cùng* địa chỉ IP của NAT Server (138.76.29.7) ở field *source IP address*, nhưng có *source port #* khác nhau

địa chỉ 10.0.0.0/24 có thể được các thiết bị trong mạng nội bộ này sử dụng để liên lạc với nhau, nhưng không thể dùng để liên lạc với bên ngoài subnet

NAT: Network Address Translation

Implementation: Router đóng vai NAT Server phải:

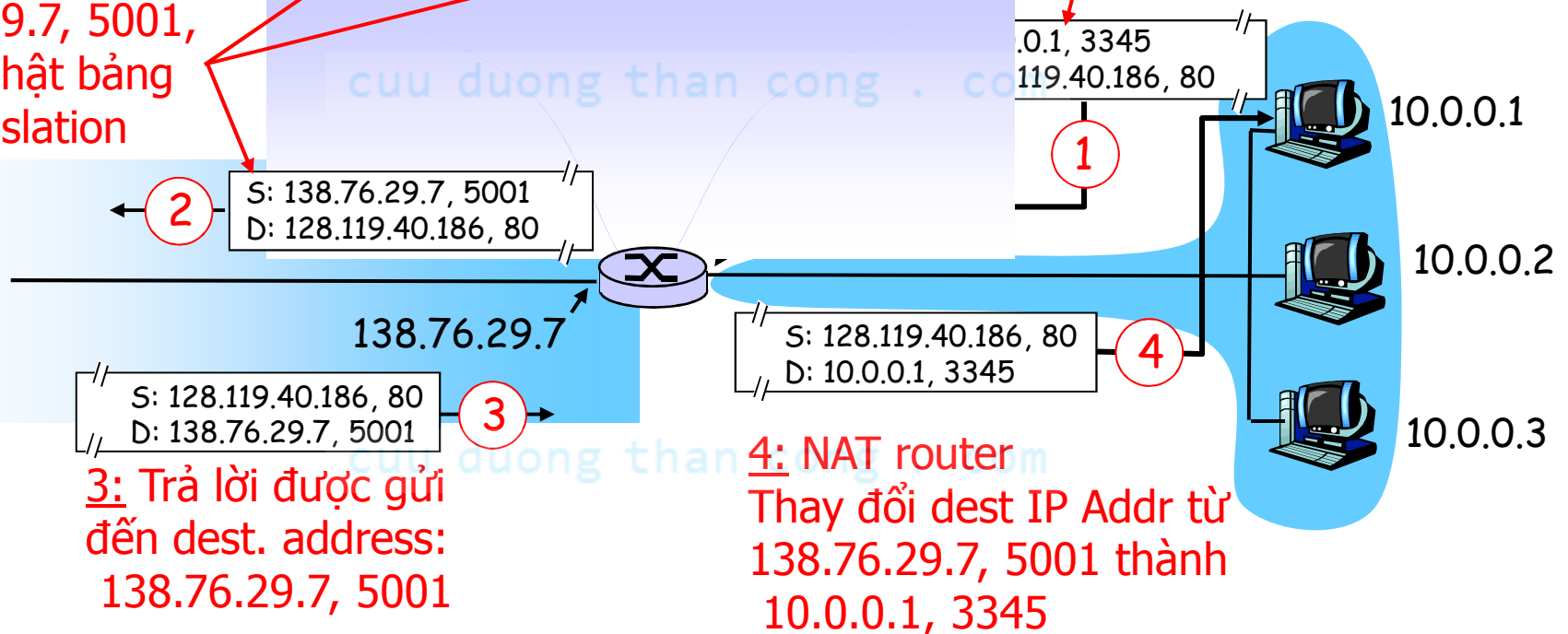
- ❖ *Với datagrams đi ra: thay thế* (source IP address, port #) thành (NAT IP address, new port #)
Các client/server ở xa sẽ dùng cặp (NAT IP address, new port #) để liên lạc.
- ❖ *Ghi nhận lại (trong bảng NAT Translation)* tất cả các cặp chuyển đổi (source IP address, port #) , (NAT IP address, new port #).
- ❖ *Với datagrams đi vào: thay thế* (NAT IP address, new port #) ở destination IP Address thành (source IP address, port #) tương ứng (tra từ bảng NAT translation)

NAT: Network Address Translation

2: NAT router
thay đổi source IP
addr từ 10.0.0.1,
3345 thành
138.76.29.7, 5001,
Và cập nhật bảng
NAT translation

Bảng NAT translation	
WAN side addr	LAN side addr
138.76.29.7, 5001	10.0.0.1, 3345
.....

1: host 10.0.0.1
gửi datagram đến
128.119.40.186, 80

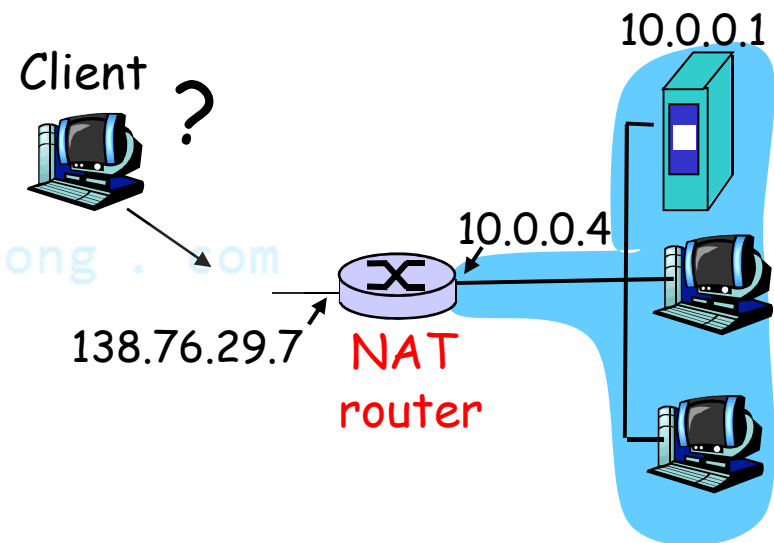


NAT: Network Address Translation

- Field port-number của datagram là 16 bit
 - ❖ Có trên 60,000 kết nối đồng thời từ 1 địa chỉ trong mạng nội bộ!
- NAT là giải pháp còn gây tranh cãi:
 - ❖ Port# là phương tiện để định vị tiến trình chứ không nên dùng để định vị host.
 - ❖ Routers chỉ nên xử lý đến tầng 3, ko nên chạm đến khái niệm port#!
 - ❖ Vi phạm nguyên tắc: các host nên truyền thông trực tiếp được với nhau, mà ko cần 1 thành phần trung gian nào làm thay đổi địa chỉ/port#
 - ❖ Sự cạn kiệt địa chỉ IPv4 nên được giải quyết bằng IPv6

NAT traversal problem

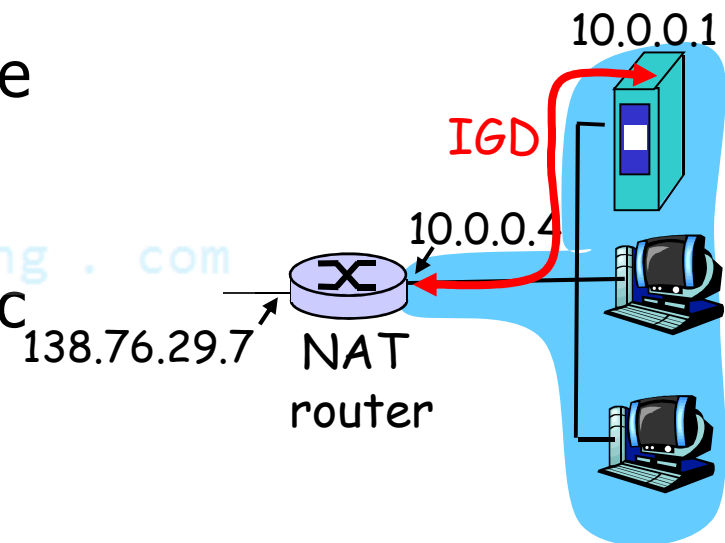
- Client muốn kết nối đến server có địa chỉ 10.0.0.1
 - ❖ Địa chỉ server 10.0.0.1 chỉ có thể dùng để định vị trong LAN (client từ ngoài LAN ko thể sử dụng địa chỉ này)
 - ❖ Chỉ có 1 địa chỉ mà client bên ngoài có thể thấy được, là địa chỉ NAT 138.76.29.7
- Giải pháp 1: **cấu hình tĩnh NAT** cho phép chuyển tiếp các request từ client đến NAT, đến 1 port nhất định trên server.
 - ❖ e.g., (138.76.29.7, port 2500) always forwarded to 10.0.0.1 port 25000



NAT traversal problem

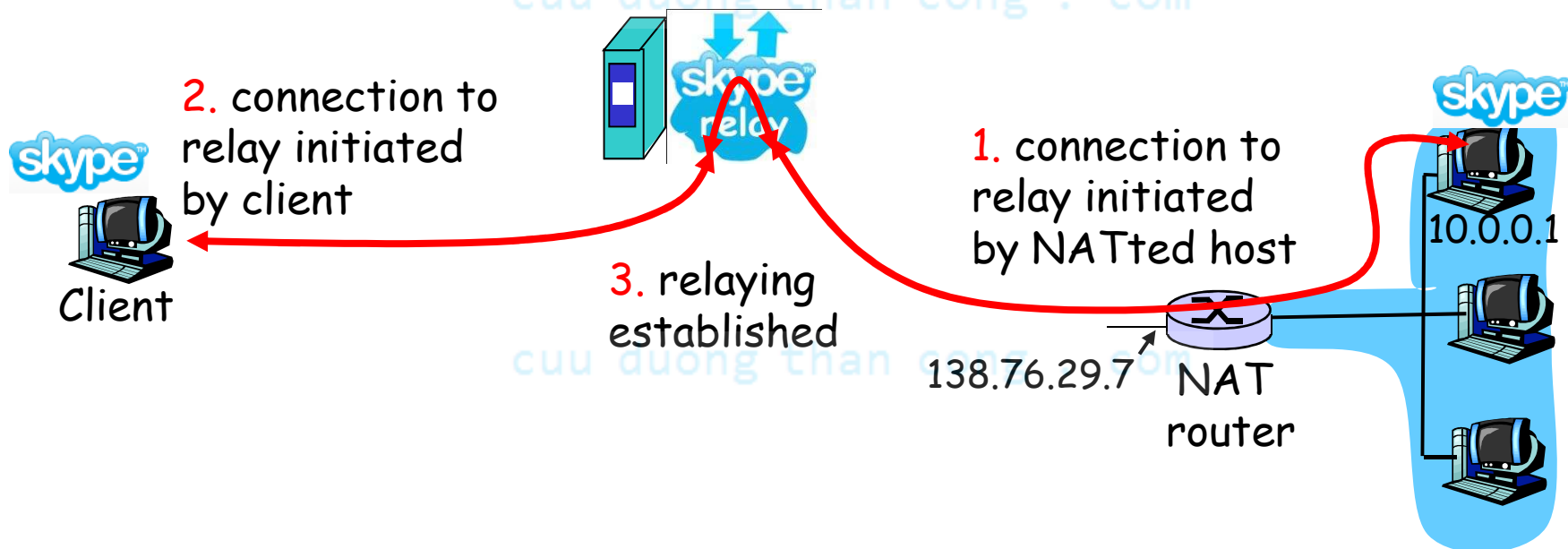
- Giải pháp 2: Sử dụng các giao thức Universal Plug and Play (UPnP), Internet Gateway Device (IGD) Protocol. Cho phép ứng dụng trên host có thể:

- ❖ Thêm/xóa ánh xạ giữa (Public IP Addr, Public Port#) của NAT và (Private IP Addr, Private Port#) của host
- ❖ Biết được (Public IP Addr, Public Port#) để quảng bá ra Internet.



NAT traversal problem

- Giải pháp3: relaying (used in Skype)
 - ❖ NATed client thiết lập kết nối đến máy trung chuyển
 - ❖ External client kết nối đến máy trung chuyển
 - ❖ Máy trung chuyển bắt cầu cho các gói tin giữa 2 bên



Chapter 4: Network Layer

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- **4.4 IP: Internet Protocol**
 - ❖ Datagram format
 - ❖ IPv4 addressing
 - ❖ **ICMP**
 - ❖ IPv6
- 4.5 Routing algorithms
 - ❖ Link state
 - ❖ Distance Vector
 - ❖ Hierarchical routing
- 4.6 Routing in the Internet
 - ❖ RIP
 - ❖ OSPF
 - ❖ BGP
- 4.7 Broadcast and multicast routing

ICMP: Internet Control Message Protocol

- Được các host và router sử dụng để truyền thông ở tầng Network

- ❖ Tường trình lỗi: unreachable host, network, port, protocol
- ❖ echo request/reply (được lệnh ping sử dụng)

- Thuộc tầng Network, nhưng “nằm trên” IP:

- ❖ Các thông điệp ICMP được gói trong các IP datagrams

- **Cấu trúc ICMP message:** Kiểu (type), Mã (code) cùng với 8 byte đầu tiên của IP datagram gây ra lỗi

<u>Type</u>	<u>Code</u>	<u>description</u>
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	6	dest network unknown
3	7	dest host unknown
4	0	source quench (congestion control - not used)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header

Traceroute and ICMP

- Bên truyền gửi một loạt segment UDP đến bên Nhận
 - ❖ First has TTL =1
 - ❖ Second has TTL=2, etc.
 - ❖ Unlikely port number
 - Khi datagram thứ n đến router thứ n:
 - ❖ Router vứt bỏ datagram
 - ❖ Và gửi 1 thông điệp ICMP (type 11, code 0) đến bên truyền
 - ❖ thông điệp chứa cả tên và địa chỉ của router
 - Khi thông điệp ICMP đến, bên gửi tính RTT
 - Traceroute thực hiện quá trình này 3 lần.
- Điều kiện dừng
- Thông thường, UDP segment đến được host nhận.
 - Host nhận trả về một thông điệp ICMP "host unreachable" (type 3, code 3)
 - Khi bên truyền nhận được thông điệp ICMP này, nó sẽ ngưng gửi.

Chapter 4: Network Layer

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- **4.4 IP: Internet Protocol**
 - ❖ Datagram format
 - ❖ IPv4 addressing
 - ❖ ICMP
 - ❖ **IPv6**
- 4.5 Routing algorithms
 - ❖ Link state
 - ❖ Distance Vector
 - ❖ Hierarchical routing
- 4.6 Routing in the Internet
 - ❖ RIP
 - ❖ OSPF
 - ❖ BGP
- 4.7 Broadcast and multicast routing

IPv6

- **Nguyên nhân ra đời:** giải pháp thay thế cho địa chỉ IPv4 (32-bit), sẽ sớm cạn kiệt.
- Các nguyên nhân khác: nhu cầu cải tiến header
 - ❖ Nhằm giúp cho việc xử lý/chuyển tiếp gói tin nhanh hơn.
 - ❖ Nhằm giúp thực hiện việc *Đảm bảo chất lượng dịch vụ* (Quality of Services – QoS) dễ dàng hơn.

Khuôn dạng IPv6 datagram:

- ❖ Phần header có kích thước cố định 40 byte
- ❖ Không cho phép phân mảnh.

IPv6 Header (Cont)



Traffic class: tương tự TOS trong IPv4.

Flow Label: xác định các datagram trong cùng “dòng” (flow).

Payload Length: Số byte dữ liệu theo sau header

Next header: xác định giao thức ở tầng trên.

Hop limit: giảm 1 khi đi qua 1 router, khi Hop limit = 0 thì datagram bị vứt bỏ

Other Changes from IPv4

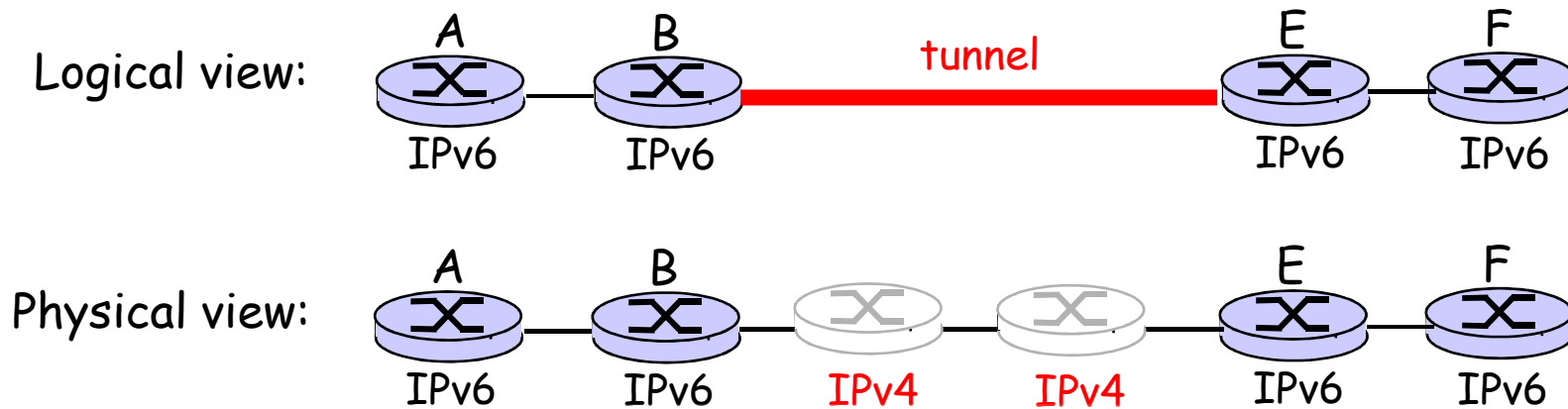
- *Checksum*: bị loại bỏ, nhằm giảm thời gian xử lý ở mỗi router.
- *Options*: vẫn còn, nhưng nằm ngoài header, được chỉ định bởi trường "Next Header"
- *ICMPv6*: phiên bản mới của ICMP
 - ❖ additional message types, e.g. "Packet Too Big"
 - ❖ multicast group management functions

cuu duong than cong . com

Transition From IPv4 To IPv6

- Không phải tất cả các router đều có thể được nâng cấp cùng lúc.
 - ❖ Không thể hẹn ngày rồi đồng loạt thay đổi.
 - ❖ Mạng sẽ hoạt động ra sao khi các router dành cho IPv4 và IPv6 cùng tồn tại?
- ***Tunneling***: Giải pháp kỹ thuật để sử dụng IPv6 trên nền IPv4: Khi đi qua các router IPv4, toàn bộ gói tin IPv6 được gói trong datagram IPv4

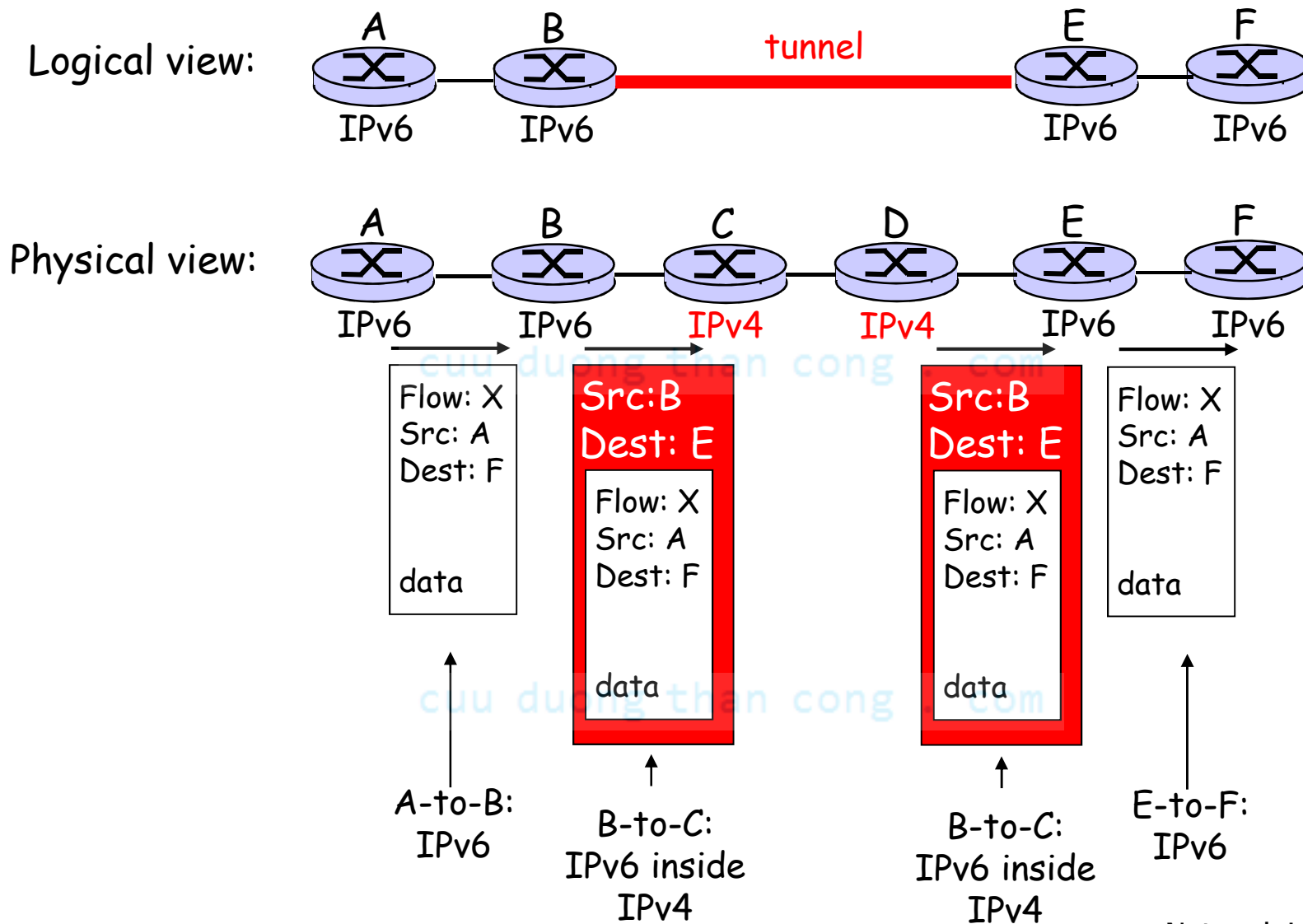
Tunneling



cuu duong than cong . com

cuu duong than cong . com

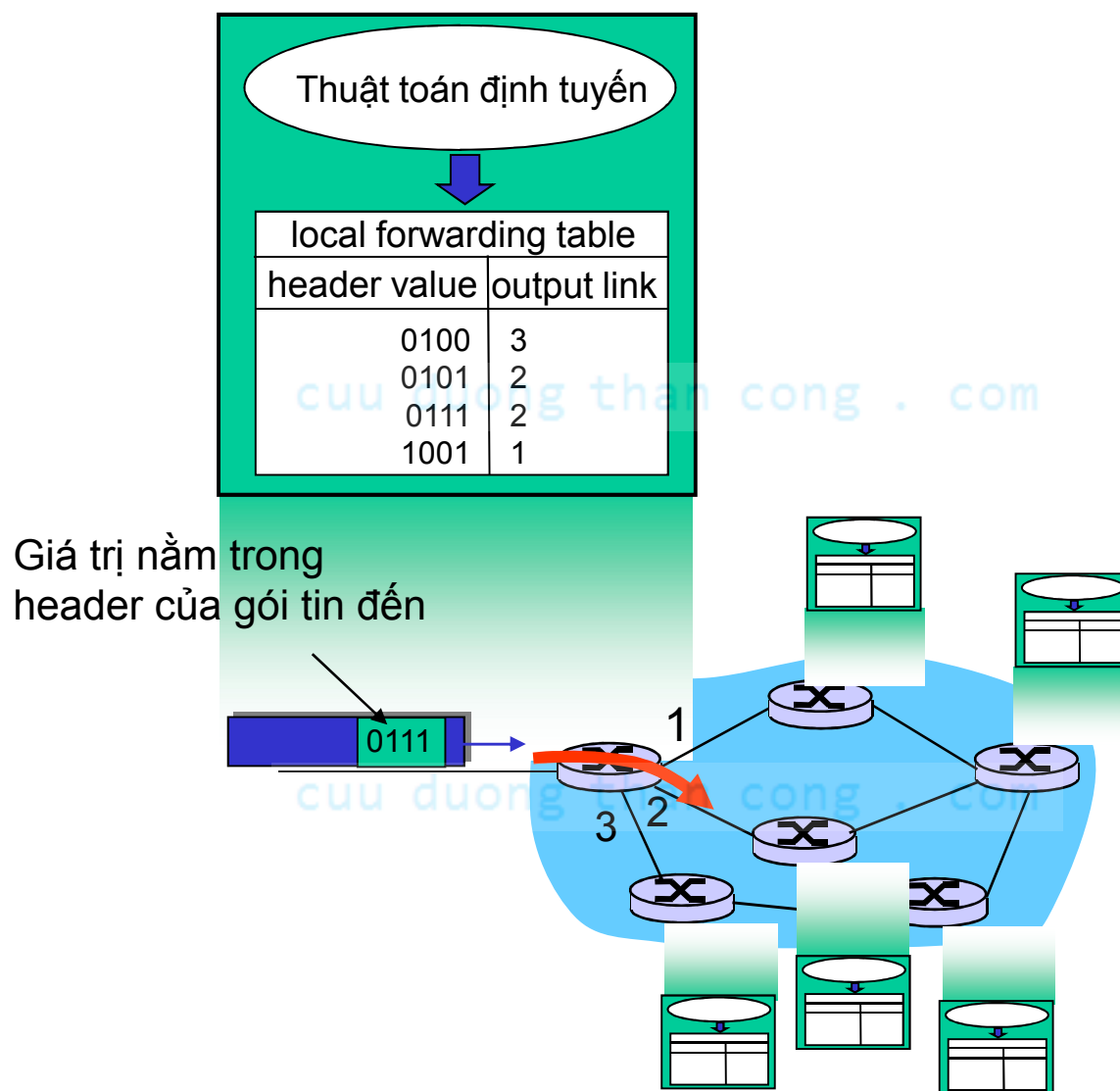
Tunneling



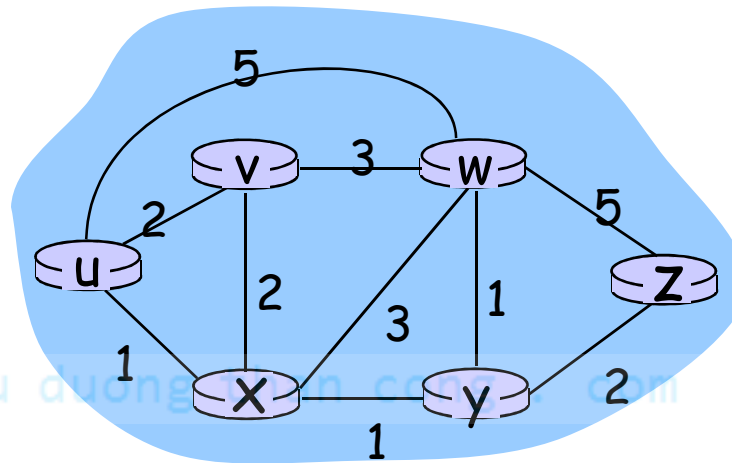
Chapter 4: Network Layer

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- 4.4 IP: Internet Protocol
 - ❖ Datagram format
 - ❖ IPv4 addressing
 - ❖ ICMP
 - ❖ IPv6
- 4.5 **Routing algorithms**
 - ❖ Link state
 - ❖ Distance Vector
 - ❖ Hierarchical routing
- 4.6 Routing in the Internet
 - ❖ RIP
 - ❖ OSPF
 - ❖ BGP
- 4.7 Broadcast and multicast routing

Interplay between routing, forwarding



Graph abstraction



Đồ thị: $G = (N, E)$

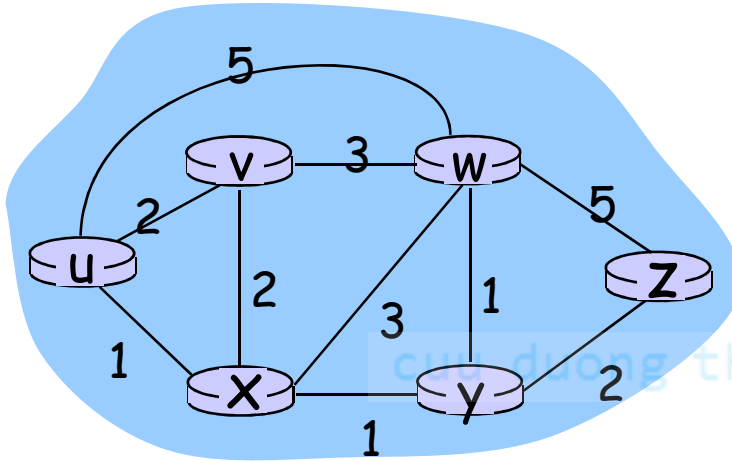
N = tập các routers = $\{ u, v, w, x, y, z \}$

E = tập các đường truyền = $\{ (u,v), (u,x), (v,x), (v,w), (x,w), (x,y), (w,y), (w,z), (y,z) \}$

Remark: Graph abstraction is useful in other network contexts

Example: P2P, where N is set of peers and E is set of TCP connections

Graph abstraction: costs



- $c(x,x')$ = chi phí đường truyền (x,x')

- e.g., $c(w,z) = 5$

chi phí đường truyền $(x_1, x_2, x_3, \dots, x_p) = c(x_1, x_2) + c(x_2, x_3) + \dots + c(x_{p-1}, x_p)$

Câu hỏi: Đâu là đường truyền có chi phí thấp nhất giữa u và z?

Thuật toán định tuyến: thuật toán tìm đường truyền có chi phí thấp nhất

Routing Algorithm classification

Global or decentralized information?

Global:

- Tất cả các router đều có đầy đủ thông tin về đầu nối và chi phí đường truyền
- "link state" algorithms

Decentralized:

- Các router chỉ biết các router láng giềng có nối trực tiếp, và chi phí đường truyền đến láng giềng.
- Quá trình lặp lại việc tính toán, trao đổi thông tin giữa các láng giềng.
- "distance vector" algorithms

Static or dynamic?

Static:

- Các tuyến thay đổi chậm.

Dynamic:

- Các tuyến thay đổi rất nhanh
 - ❖ Được cập nhật định kỳ.
 - ❖ Để phản ứng lại với các thay đổi về chi phí đường truyền.

Chapter 4: Network Layer

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- 4.4 IP: Internet Protocol
 - ❖ Datagram format
 - ❖ IPv4 addressing
 - ❖ ICMP
 - ❖ IPv6
- 4.5 Routing algorithms
 - ❖ Link state
 - ❖ Distance Vector
 - ❖ Hierarchical routing
- 4.6 Routing in the Internet
 - ❖ RIP
 - ❖ OSPF
 - ❖ BGP
- 4.7 Broadcast and multicast routing

A Link-State Routing Algorithm

Dijkstra's algorithm

- Tất cả các nút đều biết về thông tin đầu nối và chi phí đường truyền
 - ❖ Được thực hiện bằng cách quảng bá trên toàn đường truyền.
 - ❖ Tất cả các nút đều có cùng thông tin.
- Tính đường truyền có chi phí tối thiểu từ một nút ('nguồn') đến tất cả các nút khác.
 - ❖ Cung cấp các **forwarding table** cho nút đó.
- Lặp lại: sau k lần lặp, xác định được đường truyền chi phí nhỏ nhất đến k đích.

Ký hiệu:

- **$c(x,y)$** : chi phí từ nút x đến nút y; bằng ∞ nếu x không có đường nối trực tiếp.
- **$D(v)$** : chi phí đường truyền nhỏ nhất hiện hành, từ nút nguồn đến nút v.
- **$p(v)$** : nút đứng liền trước nút v trong đường truyền nhỏ nhất hiện hành, từ nguồn đến v.
- **N'** : tập các nút có đường truyền nhỏ nhất đã được xác định rõ ràng.

Dijkstra's Algorithm

1 **Khởi tạo:**

2 $N' = \{u\}$

3 Với mỗi nút $v \in N$

4 Nếu v nằm kề nút u

5 thì $D(v) = c(u, v)$

6 Ngược lại $D(v) = \infty$

7

8 **Lặp lại**

9 Tìm $w \notin N'$ sao cho $D(w)$ là nhỏ nhất

10 Thêm w vào N'

11 Cập nhật $D(v)$ cho mọi nút v nằm kề với w và chưa thuộc N' :

12 $D(v) = \min(D(v), D(w) + c(w, v))$

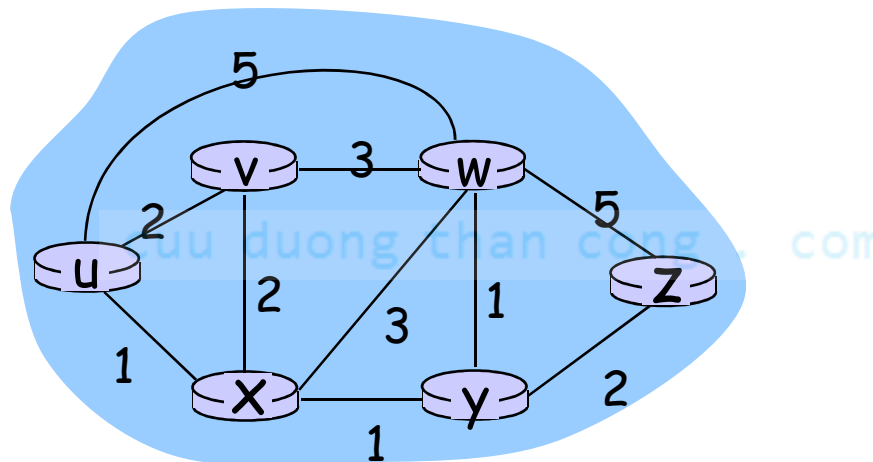
13 /* new cost to v is either old cost to v or known

14 shortest path cost to w plus cost from w to v */

15 **Cho đến khi tất cả các nút đều xuất hiện trong N'**

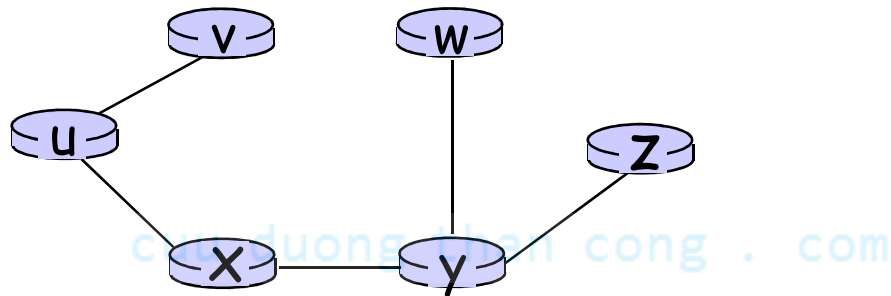
Dijkstra's algorithm: example

Step	N'	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					



Dijkstra's algorithm: example (2)

Cây biểu diễn các đường truyền có chi phí nhỏ nhất từ u:



Dẫn đến forwarding table ở nút u là:

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

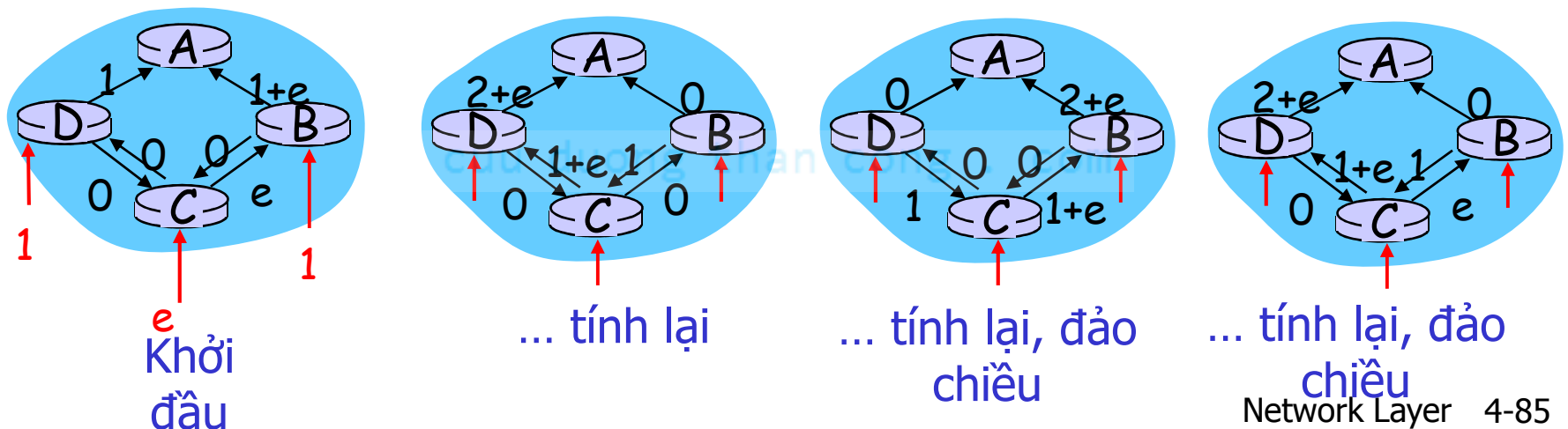
Dijkstra's algorithm, discussion

Độ phức tạp của thuật toán: với n nút

- Mỗi bước lặp: trường hợp xấu nhất là phải kiểm tra tất cả các nút w còn lại $\in N$, nhưng $\notin N'$, có $D(w)$ min
- \Rightarrow Chi phí $n(n+1)/2$
- \Rightarrow Độ phức tạp $O(n^2)$
- Nếu cài hiệu quả hơn, sử dụng cây trong bước 9: $O(n \log n)$

Sự dao động giữa các phương án đường đi chi phí nhỏ nhất:

- e.g., chi phí đường truyền = lượng dữ liệu truyền



Chapter 4: Network Layer

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- 4.4 IP: Internet Protocol
 - ❖ Datagram format
 - ❖ IPv4 addressing
 - ❖ ICMP
 - ❖ IPv6
- 4.5 Routing algorithms
 - ❖ Link state
 - ❖ Distance Vector
 - ❖ Hierarchical routing
- 4.6 Routing in the Internet
 - ❖ RIP
 - ❖ OSPF
 - ❖ BGP
- 4.7 Broadcast and multicast routing

Distance Vector Algorithm

Bellman-Ford Equation (dynamic programming)

Đặt

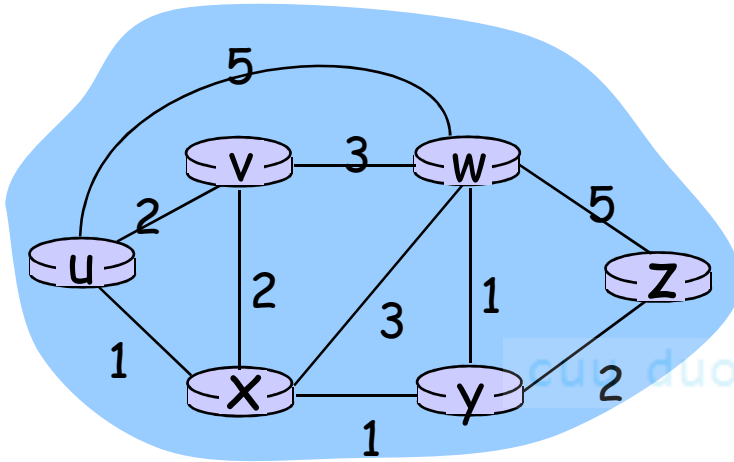
$d_x(y) :=$ chi phí nhỏ nhất của đường truyền từ x
đến y

Do đó

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

Trong đó min được lấy trên tất cả các láng giềng
v của x.

Bellman-Ford example



Rõ ràng, $d_v(z) = 5$, $d_x(z) = 3$, $d_w(z) = 3$

Theo pt Bellman-Ford:

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

Node that achieves minimum is next
hop in shortest path → forwarding table

Distance Vector Algorithm

- $D_x(y)$ = chi phí ước lượng là tối thiểu từ x đến y
- Nút x biết chi phí đến từng nút láng giềng v: $c(x,v)$
- Nút x duy trì distance vector $\mathbf{D}_x = [D_x(y): y \in N]$
- Nút x cũng duy trì những distance vector của các láng giềng.
 - ❖ Với mỗi láng giềng v, x duy trì

$$\mathbf{D}_v = [D_v(y): y \in N]$$

Distance vector algorithm (4)

Ý tưởng chính:

- Theo thời gian, mỗi nút gửi distance vector của chính nó cho các nút láng giềng.
- Việc một nút gửi DV cho các nút láng giềng là độc lập với các nút khác, ko bắt buộc các nút khác phải làm như vậy vào lúc đó.
- Khi 1 nút x nhận được DV từ 1 nút láng giềng, nó sẽ tự cập nhật lại DV của nó, theo công thức B-F:

$$D_x(y) \leftarrow \min_v \{c(x,v) + D_v(y)\} \quad \text{for each node } y \in N$$

- Bình thường, ước lượng $D_x(y)$ hội tụ về chi phí bé nhất $d_x(y)$

Distance Vector Algorithm (5)

Iterative, asynchronous:

việc tính toán lại ở mỗi nút được kích hoạt do:

- chi phí truyền ở cục bộ thay đổi.
- Nhận được DV từ láng giềng.

Distributed:

- Mỗi nút thông báo cho láng giềng chỉ khi DV của nó thay đổi.
 - ❖ Sau đó, các láng giềng cũng sẽ thông báo cho láng giềng của nó nếu cần thiết.

Ở mỗi nút:



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

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

node x table

		cost to		
		x	y	z
from	x	0	2	7
	y	∞	∞	∞
	z	∞	∞	∞

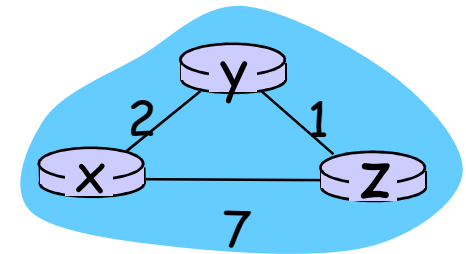
node y table

		cost to		
from		x	y	z
	x	∞	∞	∞
	y	2	0	1
	z	∞	∞	∞

node z table

		cost to		
		x	y	z
from	x	∞	∞	∞
	y	∞	∞	∞
	z	7	1	0

		cost to		
		x	y	z
from	x	0	2	3
	y	2	0	1
	z	7	1	0



time

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

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

node x table

		cost to		
		x	y	z
from	x	0	2	7
	y	∞	∞	∞
	z	∞	∞	∞

node y table

		cost to		
		x	y	z
from	x	∞	∞	∞
	y	2	0	1
	z	∞	∞	∞

node z table

		cost to		
		x	y	z
from	x	∞	∞	∞
	y	∞	∞	∞
	z	7	1	0

		cost to		
		x	y	z
from	x	0	2	3
	y	2	0	1
	z	7	1	0

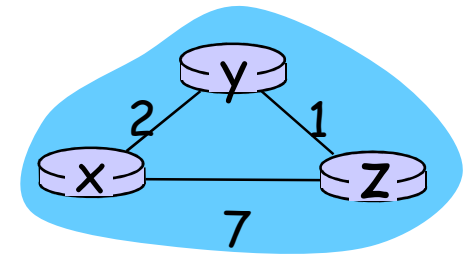
		cost to		
		x	y	z
from	x	0	2	7
	y	2	0	1
	z	7	1	0

		cost to		
		x	y	z
from	x	0	2	7
	y	2	0	1
	z	3	1	0

		cost to		
		x	y	z
from	x	0	2	3
	y	2	0	1
	z	3	1	0

		cost to		
		x	y	z
from	x	0	2	3
	y	2	0	1
	z	3	1	0

		cost to		
		x	y	z
from	x	0	2	3
	y	2	0	1
	z	3	1	0

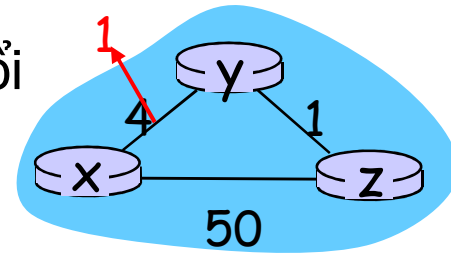


time

Distance Vector: link cost changes

Link cost changes:

- Nút phát hiện chi phí truyền cục bộ thay đổi
- Cập nhật thông tin định tuyến, tính lại DV
- Nếu DV thay đổi, thông báo cho các láng giềng



cuu duong than cong . com

“good
news
travels
fast”

• Ở thời điểm t_0 , y phát hiện chi phí truyền thay đổi, cập nhật DV của nó, và thông báo cho láng giềng của nó.

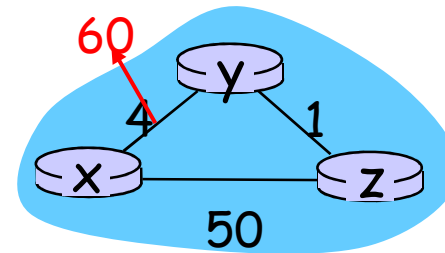
• Ở thời điểm t_1 , z phát hiện các cập nhật từ y và cập nhật các bảng của nó. Nó tính lại chi phí tối thiểu tới x và gửi DV của nó cho các láng giềng.

• Ở thời điểm t_2 , y nhận được các cập nhật của z và cập nhật lại các bảng của nó. Do chi phí tối thiểu của y không thay đổi và do đó y không gửi bất kỳ thông tin nào cho z .

Distance Vector: link cost changes

Link cost changes:

- good news travels fast
- bad news travels slow - "count to infinity" problem!
- 44 iterations before algorithm stabilizes: see text



Poisoned reverse:

- Nếu Z hướng gói tin đi qua Y để đến X:
 - Z báo cho Y biết khoảng cách từ Z đến X (mà Z quản) là vô hạn (do đó Y sẽ không hướng gói tin đến X ngang qua Z)
- will this completely solve count to infinity problem?

Comparison of LS and DV algorithms

Message complexity

- LS: với n nút, E đường truyền, $O(nE)$ thông điệp được gửi
- DV: chỉ có các trao đổi giữa các láng giềng
 - ❖ Thời gian hội tụ biến động

Speed of Convergence

- LS: thuật toán $O(n^2)$ đòi hỏi $O(nE)$ thông điệp
 - ❖ Có thể có sự dao động
- DV: Thời gian hội tụ biến động
 - ❖ Vấn đề routing loop
 - ❖ Vấn đề count-to-infinity

Robustness: điều gì sẽ xảy ra nếu router hỏng hóc?

LS:

- ❖ Nút có thể quảng cáo chi phí đường truyền không chính xác.
- ❖ Mỗi nút chỉ tính bảng của *chính* nó

DV:

- ❖ Nút có thể quảng cáo chi phí tuyến đường không chính xác
- ❖ Bảng của từng nút được các nút khác sử dụng
 - Lỗi lan truyền qua mạng

Chapter 4: Network Layer

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- 4.4 IP: Internet Protocol
 - ❖ Datagram format
 - ❖ IPv4 addressing
 - ❖ ICMP
 - ❖ IPv6
- 4.5 **Routing algorithms**
 - ❖ Link state
 - ❖ Distance Vector
 - ❖ **Hierarchical routing**
- 4.6 Routing in the Internet
 - ❖ RIP
 - ❖ OSPF
 - ❖ BGP
- 4.7 Broadcast and multicast routing

Hierarchical Routing

Các giả định nghiên cứu trước đây còn lý tưởng:

- Tất cả các routers đều giống nhau
- network “flat” (i.e không phân cấp)

... *không* khả thi, vì:

scale: với 200 triệu đích đến:

- Không thể lưu tất cả các đích đến trong routing table!
- Việc trao đổi các routing table có thể làm cho các đường truyền tràn ngập thông tin định tuyến, ko còn chỗ cho truyền dữ liệu!

administrative autonomy

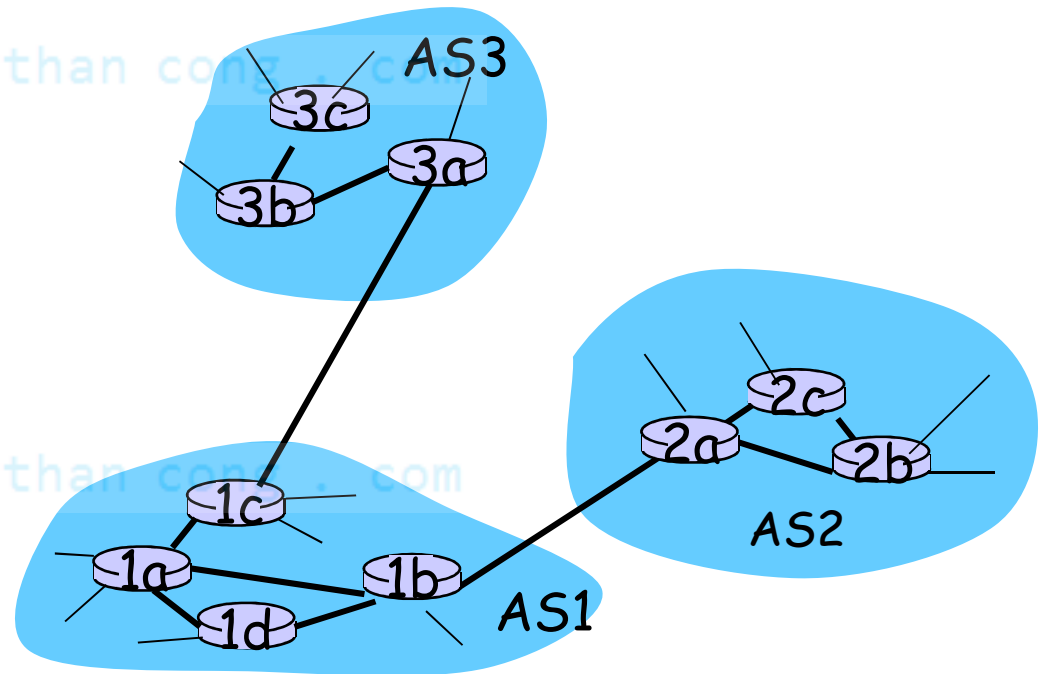
- internet = mạng của các mạng
- Mỗi quản trị viên mạng có thể muốn kiểm soát việc định tuyến trong mạng của họ.

Hierarchical Routing

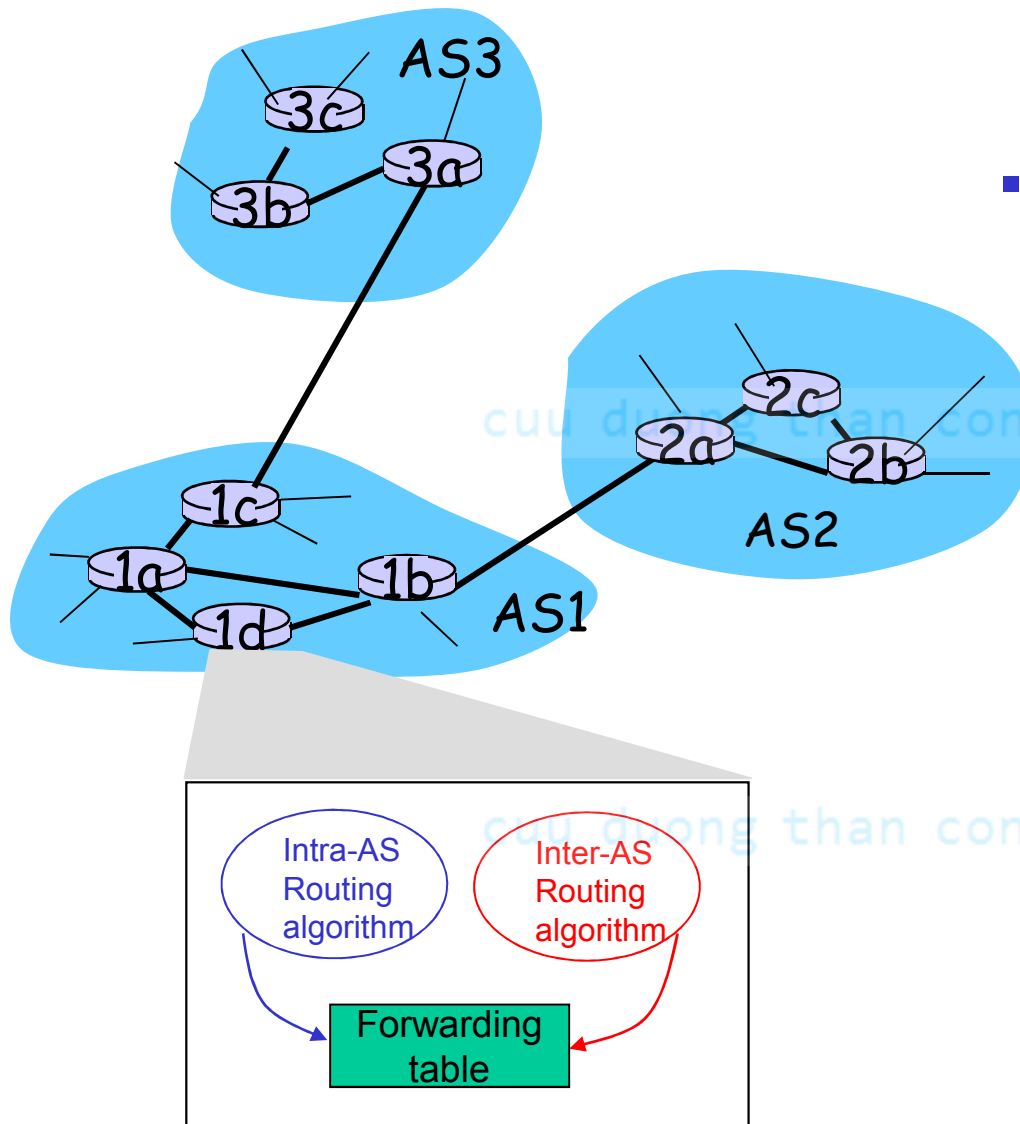
- Gom các router vào trong 1 vùng, gọi là “autonomous systems” (AS)
- Các router trong cùng AS sẽ chạy cùng 1 routing protocol
 - ❖ “intra-AS” routing protocol
 - ❖ Các router trong các AS khác nhau có thể chạy các intra-AS routing protocol khác nhau.

Gateway router

- Nối trực tiếp đến router trong AS khác.



Interconnected ASes



- Forwarding table được cấu hình bởi cả intra- và inter- AS routing protocol
 - ❖ Intra-AS thiết lập các dòng định tuyến cho các đích đến bên trong AS.
 - ❖ Inter-AS và intra-AS thiết lập các dòng định tuyến cho các đích đến bên ngoài.

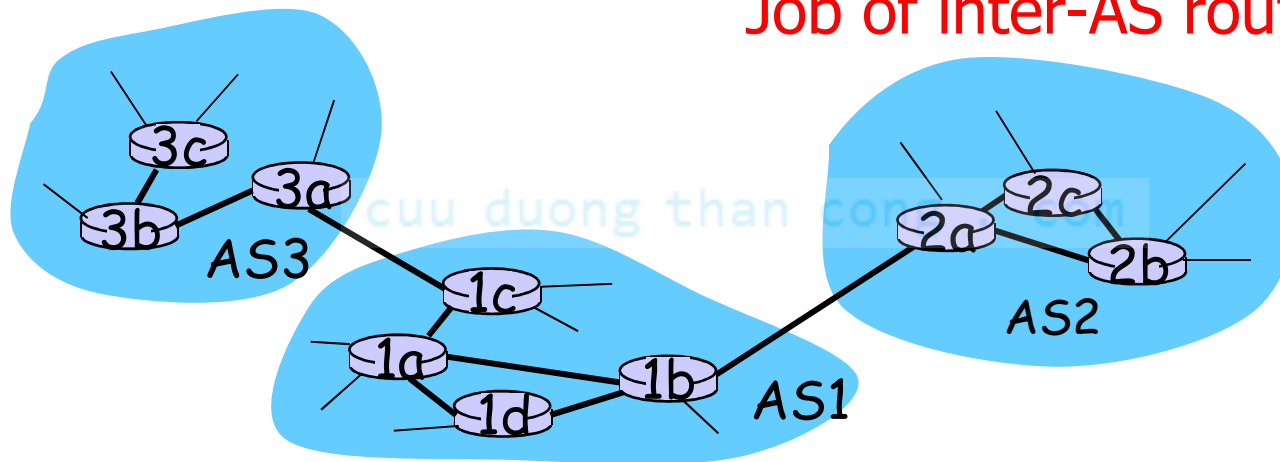
Inter-AS tasks

- Giả sử rằng router trong AS1 nhận được datagram hướng ra bên ngoài AS1:
 - ❖ Router cần chuyển gói tin đến gateway router nào?

AS1 phải:

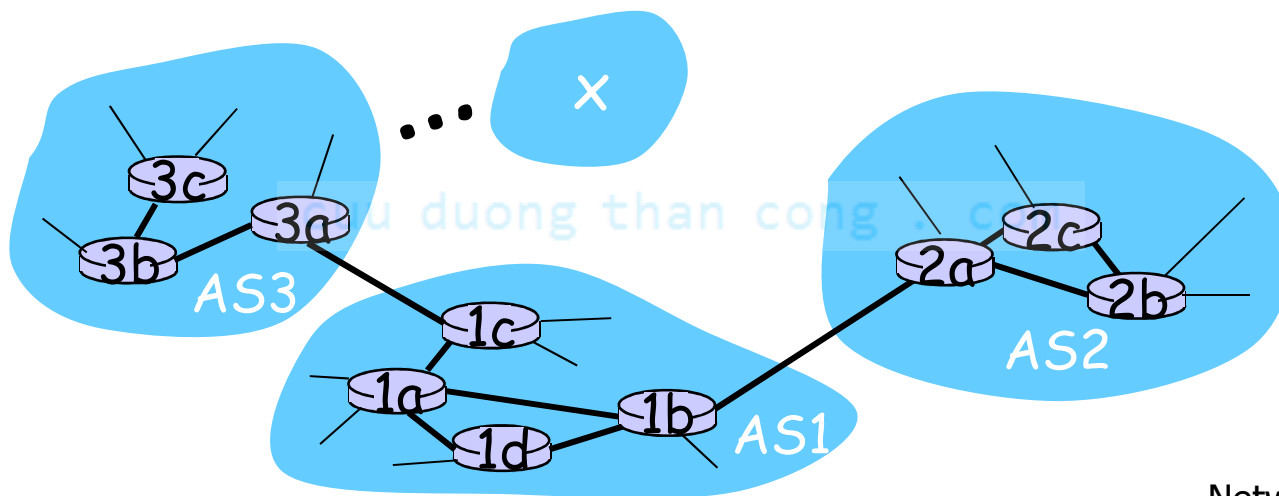
1. Tìm hiểu xem đích đến nào là có thể chạm tới được thông qua AS2, hoặc AS3.
2. Truyền bá khả năng này cho tất cả các router trong AS1.

Job of inter-AS routing!



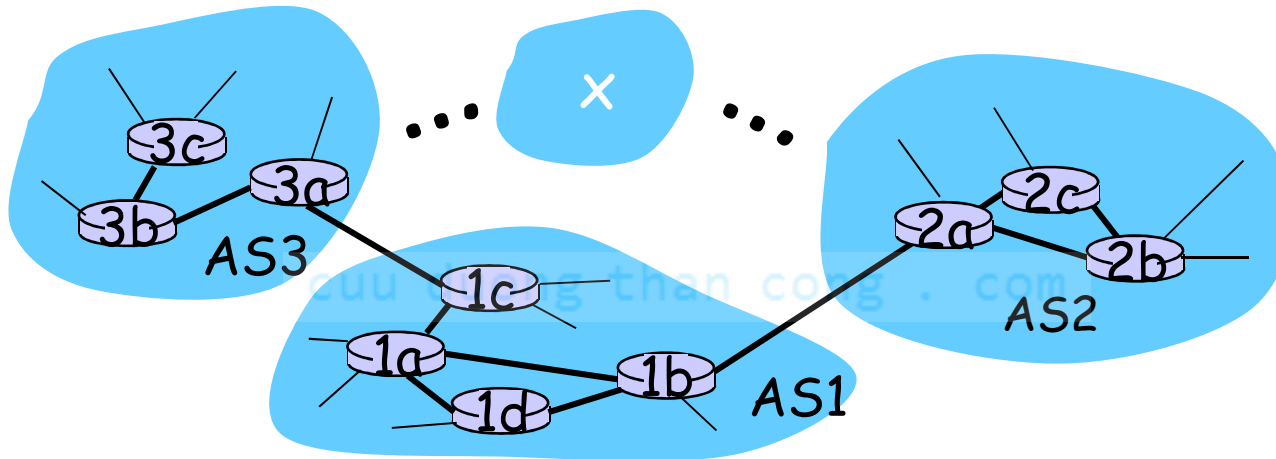
Example: Setting forwarding table in router 1d

- Giả sử AS1 học được thông tin (thông qua inter-AS protocol) là có thể đến được subnet **x** thông qua AS3 (gateway 1c) nhưng không qua AS2.
- inter-AS protocol truyền thông tin này đến tất cả các router bên trong.
- Router 1d xác định được rằng interface **I** của nó là nằm trên đường ngắn nhất đến 1c.
 - ❖ Thêm/cập nhật dòng **(x,I)** vào forwarding table



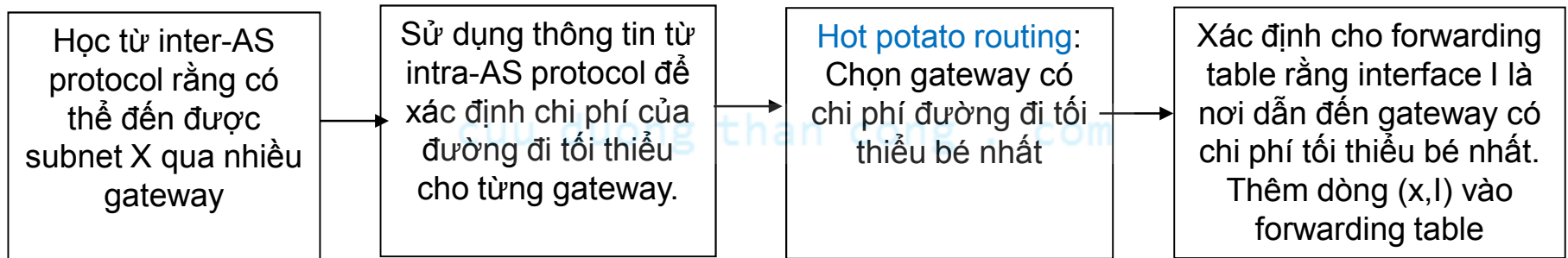
Example: Choosing among multiple ASes

- Giả sử AS1 học được từ inter-AS protocol rằng có thể đến được subnet **x** bằng cách đi qua AS3 *hoặc* qua AS2.
- Để cấu hình forwarding table, router 1d phải xác định gateway nào sẽ chịu trách nhiệm chuyển gói tin đến mạng **x**.
 - ❖ Đây cũng là công việc của inter-AS routing protocol!



Example: Choosing among multiple ASes

- Giả sử AS1 học được từ inter-AS protocol rằng có thể đến được subnet **X** bằng cách đi qua AS3 *hoặc qua* AS2.
- Để cấu hình forwarding table, router 1d phải xác định gateway nào sẽ chịu trách nhiệm chuyển gói tin đến mạng **X**.
 - ❖ Đây cũng là công việc của inter-AS routing protocol!
- **hot potato routing**: gửi gói tin về hướng router gần nhất.



Chapter 4: Network Layer

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- 4.4 IP: Internet Protocol
 - ❖ Datagram format
 - ❖ IPv4 addressing
 - ❖ ICMP
 - ❖ IPv6
- 4.5 Routing algorithms
 - ❖ Link state
 - ❖ Distance Vector
 - ❖ Hierarchical routing
- 4.6 Routing in the Internet
 - ❖ RIP
 - ❖ OSPF
 - ❖ BGP
- 4.7 Broadcast and multicast routing

Intra-AS Routing

- Còn gọi là **Interior Gateway Protocols (IGP)**
- Các giao thức thường gặp nhất thuộc loại này:
 - ❖ RIP: Routing Information Protocol
 - ❖ OSPF: Open Shortest Path First
 - ❖ IGRP: Interior Gateway Routing Protocol (Cisco proprietary)

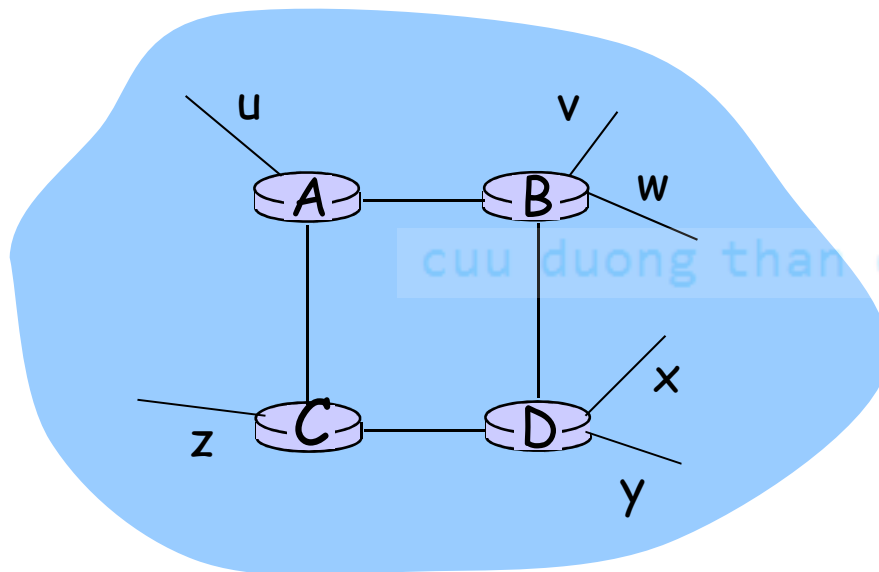
cuu duong than cong . com

Chapter 4: Network Layer

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- 4.4 IP: Internet Protocol
 - ❖ Datagram format
 - ❖ IPv4 addressing
 - ❖ ICMP
 - ❖ IPv6
- 4.5 Routing algorithms
 - ❖ Link state
 - ❖ Distance Vector
 - ❖ Hierarchical routing
- 4.6 Routing in the Internet
 - ❖ RIP
 - ❖ OSPF
 - ❖ BGP
- 4.7 Broadcast and multicast routing

RIP (Routing Information Protocol)

- Là thuật toán distance vector
- Xuất hiện trong bản BSD-UNIX Distribution vào năm 1982
- Độ đo khoảng cách: số lượng hop (tối đa trong 1 AS= 15 hops)
 - ❖ Hop: số subnet phải đi qua trên con đường ngắn nhất từ router nguồn đến subnet đích, kể cả subnet đích



Từ router A đến các subnets:

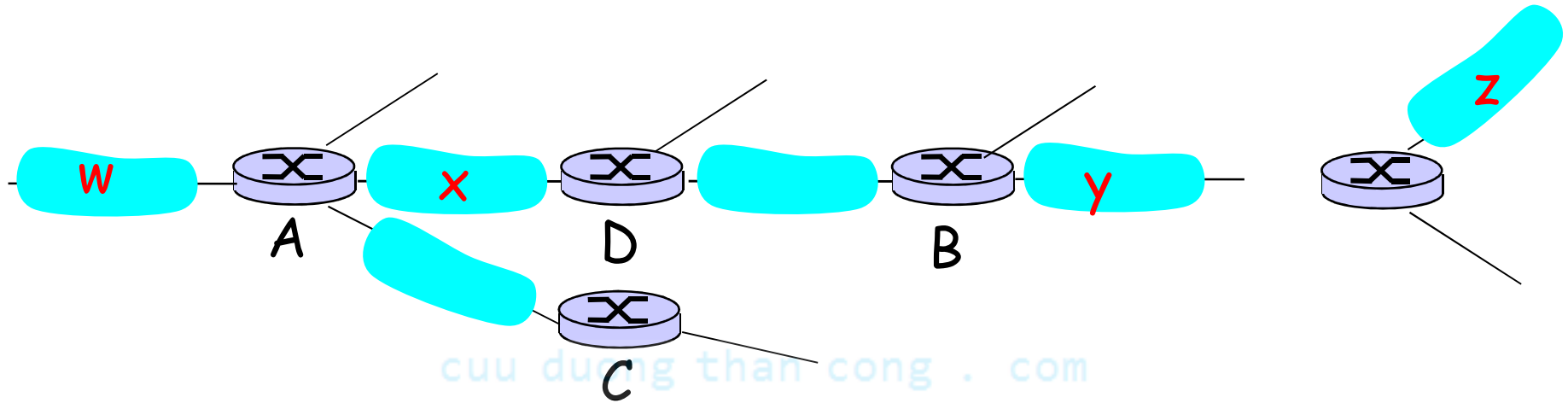
<u>Subnet đích</u>	<u>hops</u>
u	1
v	2
w	2
x	3
y	3
z	2

RIP advertisements

- distance vectors: được các router láng giềng trao đổi sau mỗi 30 giây, được chứa trong Response Message (còn được gọi là **advertisement**)
- Mỗi advertisement: danh sách lên tới 25 subnets đích bên trong AS

cuu duong than cong . com

RIP: Example



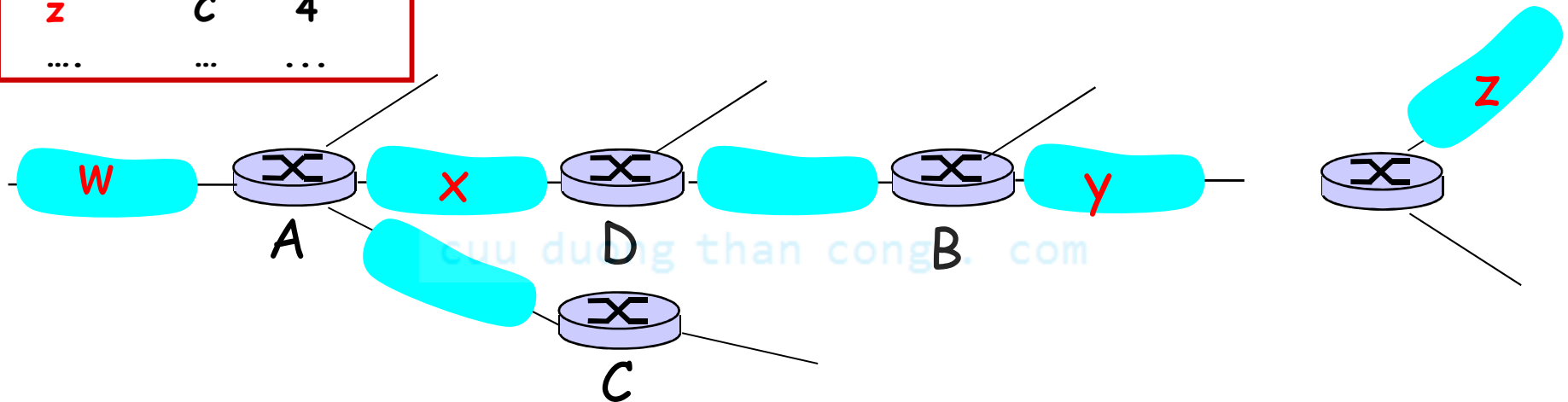
Destination Network	Next Router	Num. of hops to dest.
W	A	2
Y	B	2
Z	B	7
X	--	1
...

Routing/Forwarding table in D

RIP: Example

Dest	Next	hops
w	-	1
x	-	1
z	C	4
...

Advertisement
from A to D



Destination Network	Next Router	Num. of hops to dest.
w	A	2
y	B	2
z	B A	7 5
x	--	1
...

Routing/Forwarding table in D

Network Layer 4-111

RIP: Link Failure and Recovery

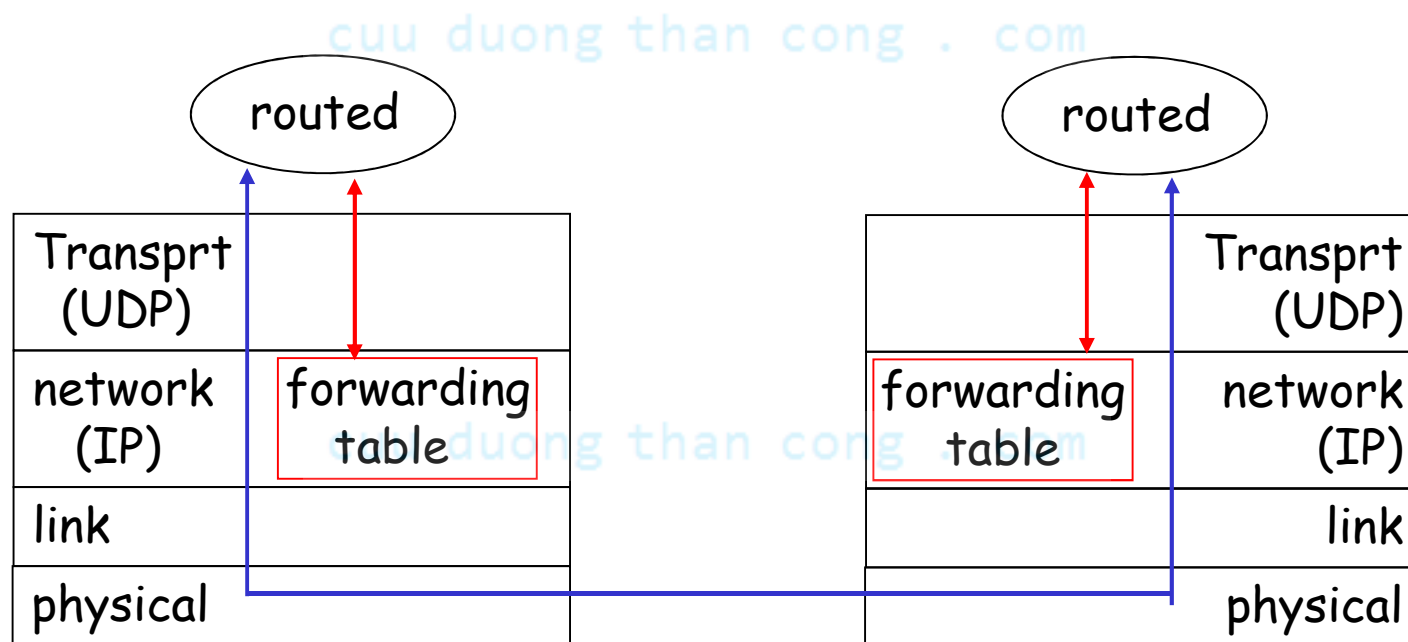
Nếu 1 router không thấy advertisement nào từ 1 router láng giềng của nó trong vòng 180 giây, → xem như router láng giềng đó đã không còn liên lạc được (có thể do router đã tắt/hư hoặc đường truyền bị gián đoạn)

- ❖ Các con đường đi qua router láng giềng đó bị xem là bất hợp lệ.
- ❖ Cần một advertisements mới để thông báo cho các router láng giềng (còn liên lạc được) biết điều này.
- ❖ Nếu routing table ở láng giềng có thay đổi, nó cũng sẽ gửi advertisement để thông báo.
- ❖ Các bảng học về đường truyền được lan truyền nhanh chóng trên toàn mạng (về nguyên tắc).
- ❖ Kỹ thuật *poison reverse* được dùng để ngăn tình trạng ping-pong loops (infinite distance = 16 hops)

Network Layer 4-112

RIP Table processing

- UNIX: RIP routing tables được một tiến trình của tầng **application** quản lý, được gọi là route-d (daemon)
- Các advertisements được gửi định kỳ bằng UDP



Chapter 4: Network Layer

- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- 4.4 IP: Internet Protocol
 - ❖ Datagram format
 - ❖ IPv4 addressing
 - ❖ ICMP
 - ❖ IPv6
- 4.5 Routing algorithms
 - ❖ Link state
 - ❖ Distance Vector
 - ❖ Hierarchical routing
- 4.6 Routing in the Internet
 - ❖ RIP
 - ❖ OSPF
 - ❖ BGP
- 4.7 Broadcast and multicast routing

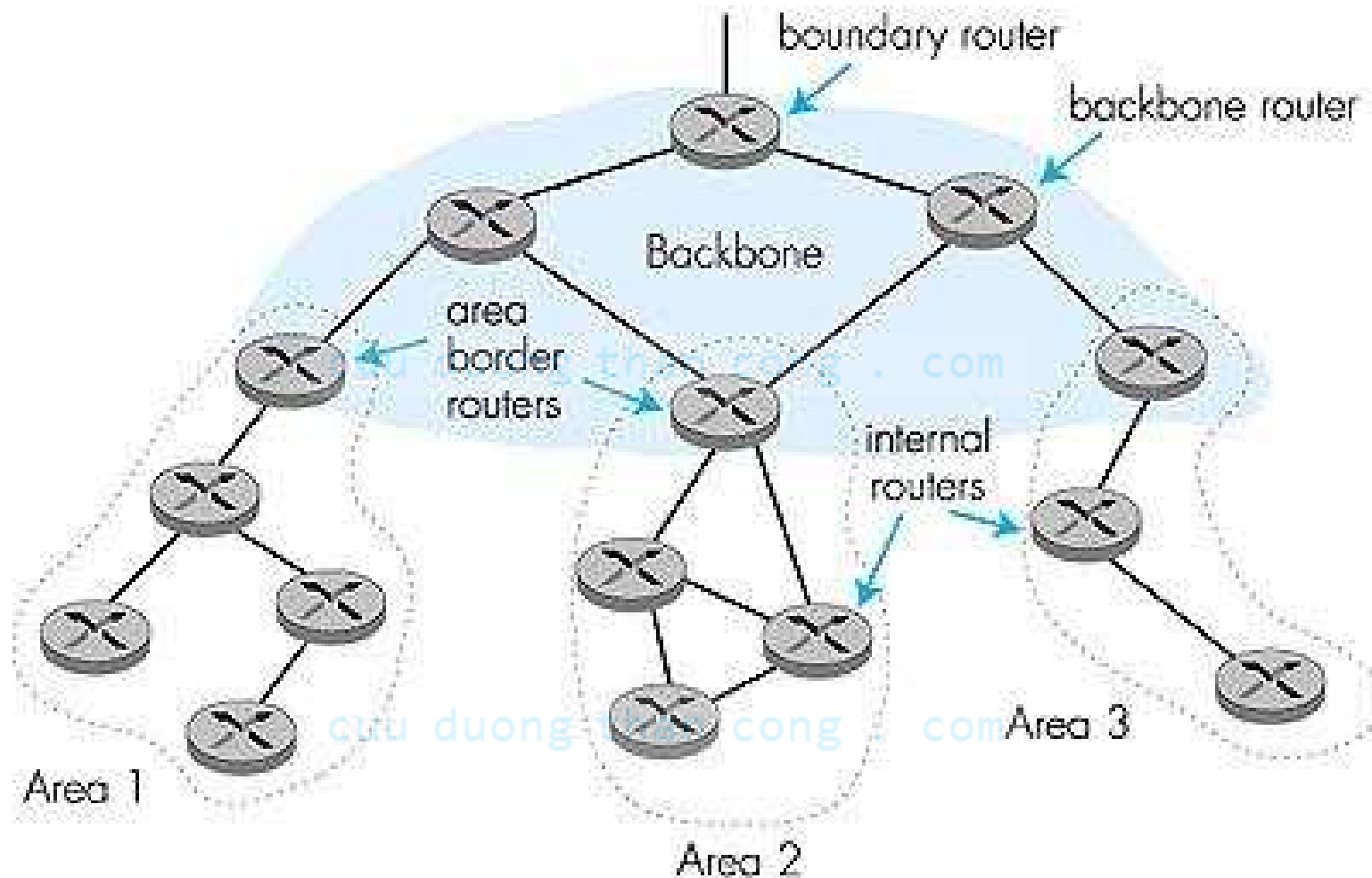
OSPF (Open Shortest Path First)

- “open”: publicly available
- Sử dụng thuật toán Link State
 - ❖ Gửi quảng bá các gói tin LS đến đến toàn bộ AS
 - Khi router phát hiện thấy có thay đổi (chi phí, tình trạng mạng).
 - Định kỳ sau mỗi 30 giây dù tình trạng mạng (Link state) có thay đổi hay không
 - ❖ Mỗi nút phải xây dựng bản đồ mạng cho riêng nó
 - ❖ Tính toán đường đi dựa vào giải thuật Dijkstra
- OSPF advertisement carries one entry per neighbor router
- Các advertisements được phát tán đến **toàn bộ** AS, chứ ko chỉ là các router láng giềng
 - ❖ Được thực hiện nhờ các OSPF message, trên nền IP (thay vì TCP hay UDP)

OSPF “advanced” features (not in RIP)

- **security**: tất cả các OSPF message đều được chứng thực (nhằm ngăn chặn những xâm nhập phá hoại)
- Cho phép **nhiều** đường đến đích có chi phí bằng nhau (nhưng trong RIP thì chỉ có 1 đường)
- Với mỗi đường truyền, có nhiều độ đo chi phí ứng với các kiểu dịch vụ (TOS – type of service) khác nhau (e.g., satellite link cost set “low” for best effort; high for real time)
- Cho phép tích hợp unicast và **multicast**:
 - ❖ Multicast OSPF (MOSPF) sử dụng cùng dữ liệu về tình trạng mạng như OSPF
- Cho phép AS cỡ lớn được cấu trúc **phân cấp (hierarchical)** và làm việc theo OSPF

Hierarchical OSPF



Hierarchical OSPF

- **Phân cấp 2 mức:** vùng cục bộ, backbone.
 - ❖ Các link-state advertisement chỉ được gửi trong vùng cục bộ
 - ❖ each nodes has detailed area topology; only know direction (shortest path) to nets in other areas.
- **area border routers:** “summarize” distances to nets in own area, advertise to other Area Border routers.
- **backbone routers:** run OSPF routing limited to backbone.
- **boundary routers:** connect to other AS's.

Chapter 4: Network Layer

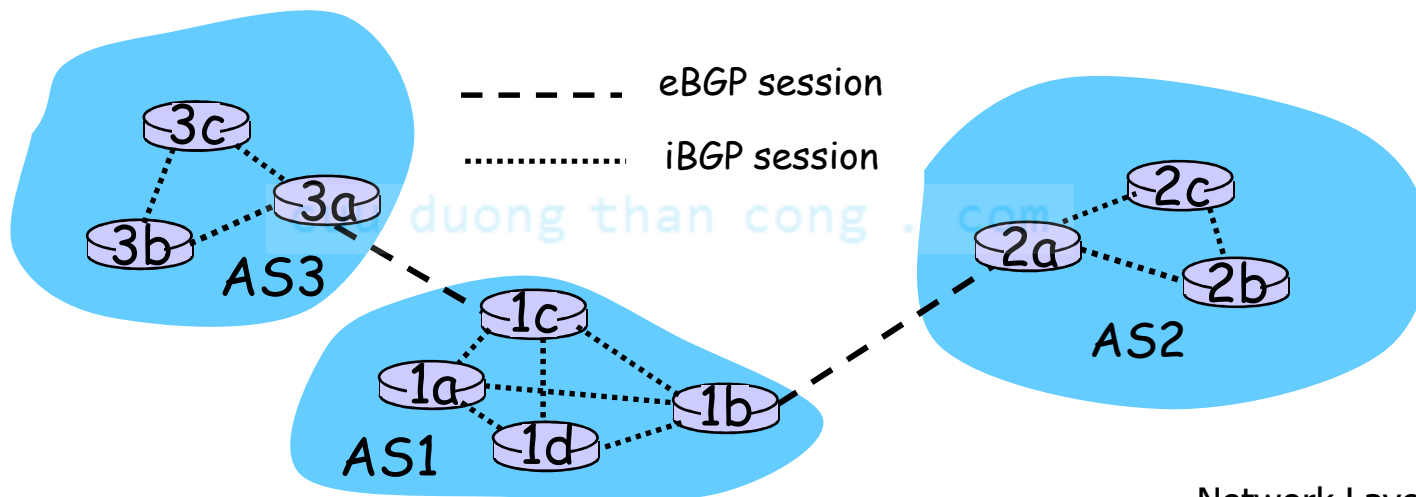
- 4.1 Introduction
- 4.2 Virtual circuit and datagram networks
- 4.3 What's inside a router
- 4.4 IP: Internet Protocol
 - ❖ Datagram format
 - ❖ IPv4 addressing
 - ❖ ICMP
 - ❖ IPv6
- 4.5 Routing algorithms
 - ❖ Link state
 - ❖ Distance Vector
 - ❖ Hierarchical routing
- 4.6 Routing in the Internet
 - ❖ RIP
 - ❖ OSPF
 - ❖ BGP
- 4.7 Broadcast and multicast routing

Internet inter-AS routing: BGP

- **BGP (Border Gateway Protocol):** *the de facto standard*
- BGP cung cấp cho mỗi AS phương tiện để:
 1. **Thu thập thông tin** về các subnet có thể liên lạc được từ các AS láng giềng.
 2. **Lan truyền các thông tin** về các subnet có thể liên lạc được cho tất cả các router bên trong AS
 3. **Xác định con đường "tốt"** đến các subnet dựa trên các thông tin về các subnet có thể liên lạc được và chính sách của nó.
- Cho phép 1 subnet quảng cáo sự tồn tại của nó đến phần còn lại của Internet: **"Có tôi đây"**

BGP basics

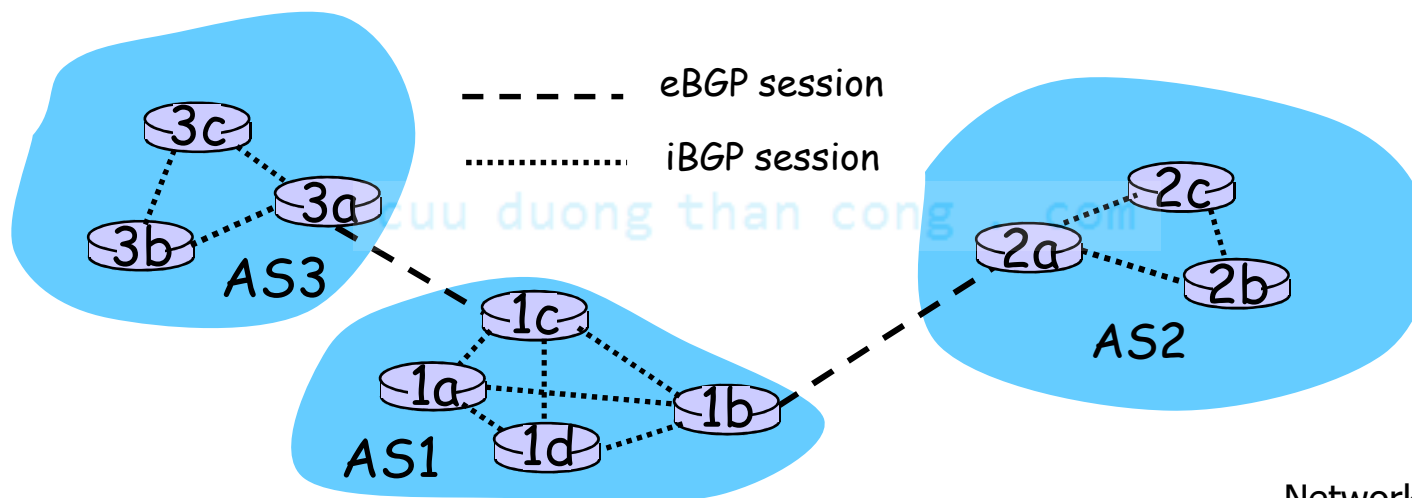
- Các cặp router (BGP peers) trao đổi thông tin định tuyến thông qua kết nối TCP “semi-permanent”: **BGP sessions**.
 - ❖ external BGP session vs. internal BGP session
 - ❖ Các BGP session không cần phải tương ứng với đường truyền vật lý
- Trong BGP: đích đến không phải là host, mà là subnet, biểu diễn bằng **prefix** (Vd 138.16.64/24)
- Khi AS2 quảng cáo một prefix đến AS1:
 - ❖ AS2 **cam kết** là nó sẽ chuyển tiếp các datagram về mạng con ứng với prefix.
 - ❖ AS2 có thể gộp các prefix trong gói tin advertisements của nó.



Network Layer 4-121

Distributing reachability info

- Sử dụng eBGP session giữa 3a và 1c, AS3 gửi các thông tin về mạng con mà có thể liên lạc được (reachable) cho AS1
 - ❖ Sau đó, 1c có thể sử dụng iBGP để phân phối thông tin này đến tất cả các router trong AS1.
 - ❖ Sau đó, 1b có thể quảng cáo tiếp thông tin này đến AS2, nhờ eBGP session từ 1b đến 2a.
- Khi router biết được thông tin mới, nó tạo ra một dòng cho một prefix trong forwarding table của nó.



Path attributes & BGP routes

- Các prefix được quảng cáo kèm theo các thuộc tính BGP.
 - ❖ prefix + attributes = "route"
- Hai thuộc tính quan trọng:
 - ❖ **AS-PATH**: chứa các AS mà các prefix advertisement đi qua, VD: AS 67, AS 17
 - ❖ **NEXT-HOP**: chứa địa chỉ IP của interface của gateway router (của AS đầu tiên đi ra subnet được xác định bởi prefix)
- Khi một gateway router nhận được một advertisement, có sử dụng **import policy** để chấp nhận hoặc từ chối.

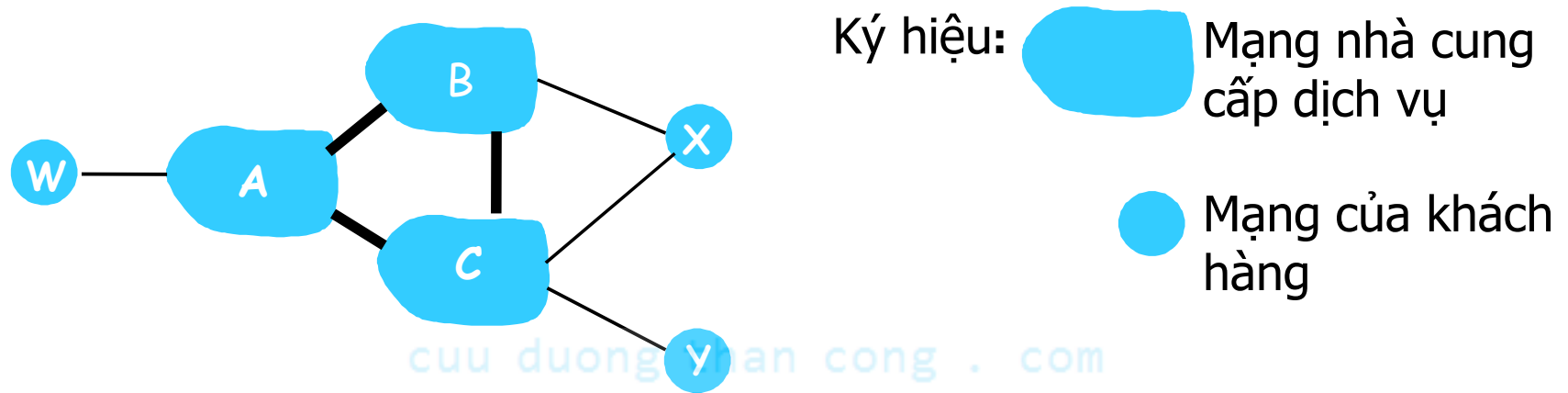
BGP route selection

- Router có thể học được nhiều hơn một tuyến đường đến 1 subnet ứng với 1 vài prefix. Router phải lựa chọn 1 tuyến đường.
- Quy tắc loại trừ:
 1. local preference value attribute: policy decision
 2. shortest AS-PATH
 3. closest NEXT-HOP router: hot potato routing
 4. additional criteria

BGP messages

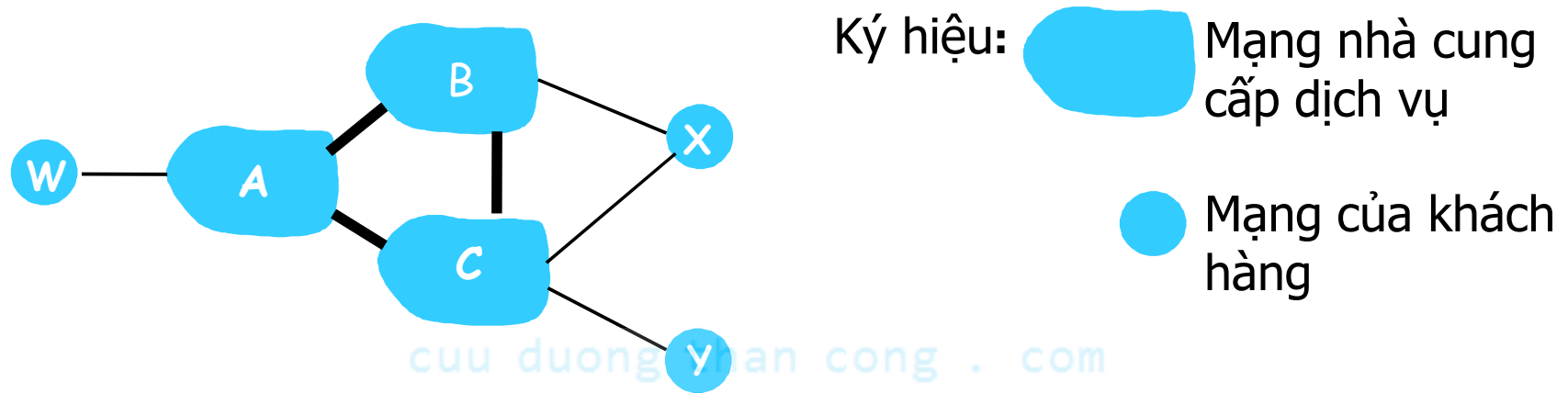
- Các BGP messages được trao đổi giữa các router, sử dụng TCP.
- BGP messages:
 - ❖ **OPEN:** mở kết nối TCP đến peer và ủy quyền cho bên gửi.
 - ❖ **UPDATE:** quảng cáo tuyến đường mới (hoặc rút lại tuyến cũ)
 - ❖ **KEEPALIVE** keeps connection alive in absence of UPDATES; also ACKs OPEN request
 - ❖ **NOTIFICATION:** báo lỗi cho các msg trước đó; cũng được dùng để đóng kết nối.

BGP routing policy



- A,B,C là các mạng của **nhà cung cấp dịch vụ**
- X,W,Y là các mạng khách hàng (của nhà cung cấp)
- X là **dual-homed**: được nối vào 2 mạng
 - X không muốn các tuyến đi từ B đến C qua nó
 - ... do đó X không quảng cáo đến B con đường tới C

BGP routing policy (2)



- A quảng cáo đến B con đường AW
- B quảng cáo đến X đường BAW
- B có nên quảng cáo đến C đường BAW?
 - No way! B không có “lợi nhuận” từ việc gánh luồng CBAW, do cả W lẫn C đều không phải là khách hàng của B
 - B muốn buộc C hướng luồng dữ liệu đến w thông qua A
 - B chỉ muốn phục vụ cho khách hàng của nó!

Network Layer 4-127