KHOA CÔNG NGHỆ THÔNG TIN-ĐHKHTN
CSC11004 - MẠNG MÁY TÍNH NÂNG CAO

INTERNET ROUTING

Lê Ngọc Sơn

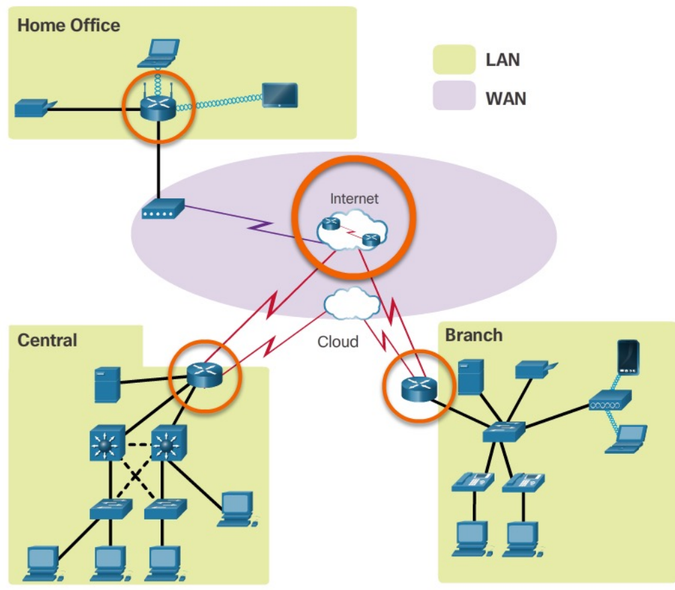
TPHCM, 9-2021

KHOA CÔNG NGHỆ THÔNG TIN
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

fit@hcmus

Router Functions

The router is responsible for the routing of traffic between networks



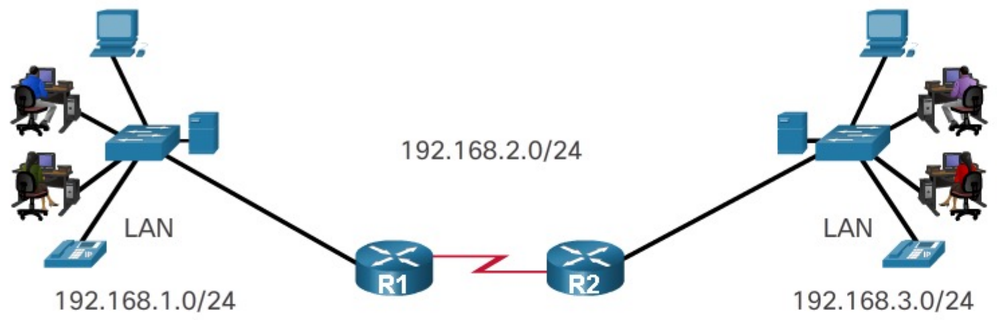
Routing Concepts

KHOA CÔNG NGHỆ THÔNG TIN
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

fit@hcmus

Routers Choose Best Paths

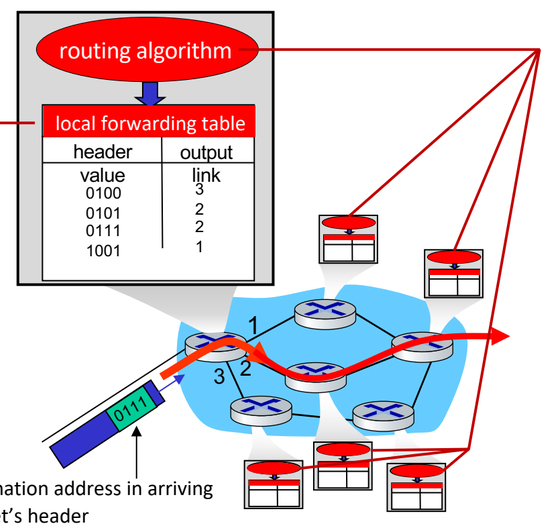
- Routers use static routes and dynamic routing protocols to learn about remote networks and build their routing tables.
- Routers use routing tables to determine the best path to send packets.
- Routers encapsulate the packet and forward it to the interface indicated in routing table.



Two key network-core functions

Forwarding:

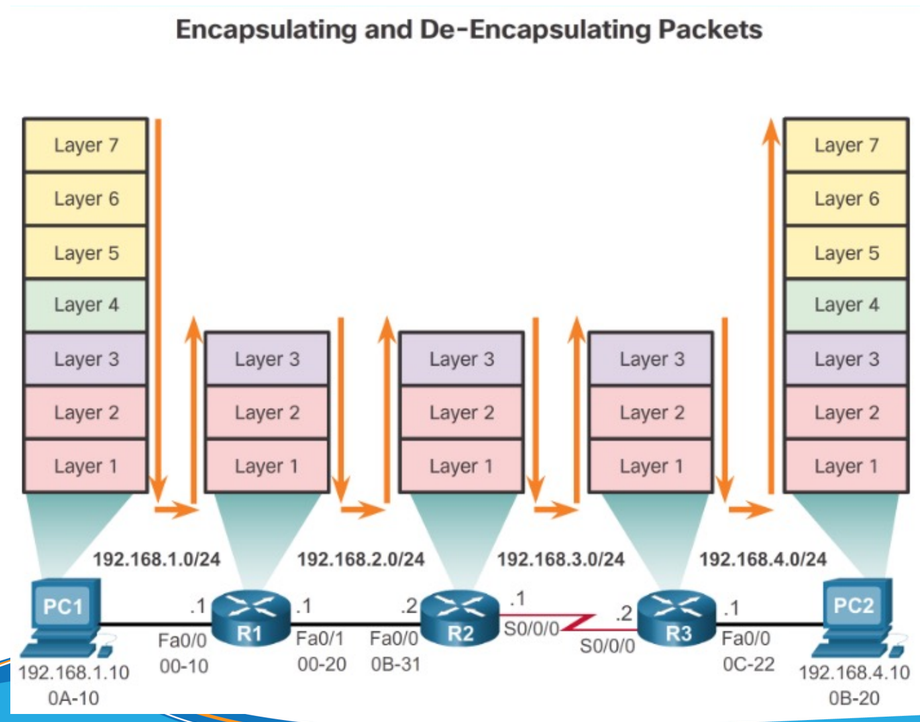
- aka "switching"
- *local* action: move arriving packets from router's input link to appropriate router output link



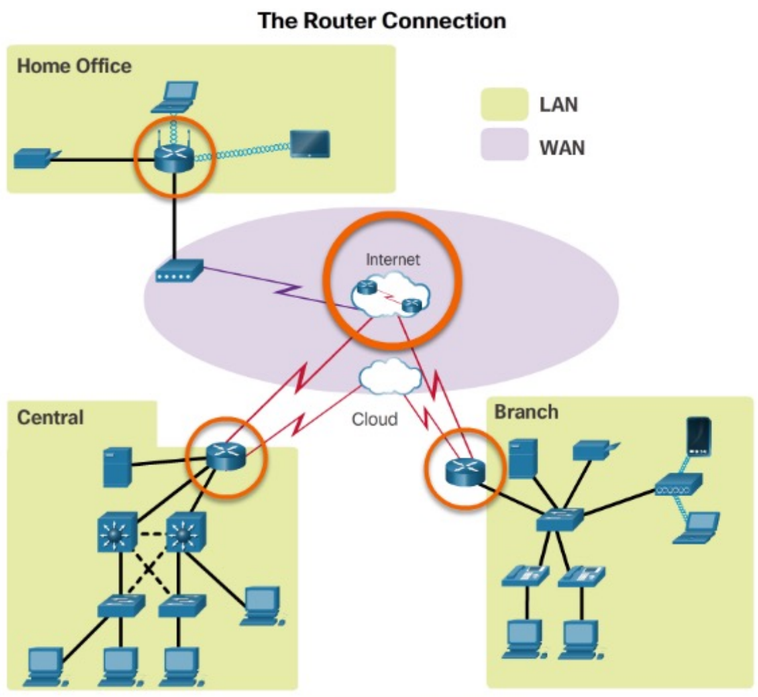
Routing:

- *global* action: determine source-destination paths taken by packets
- routing algorithms

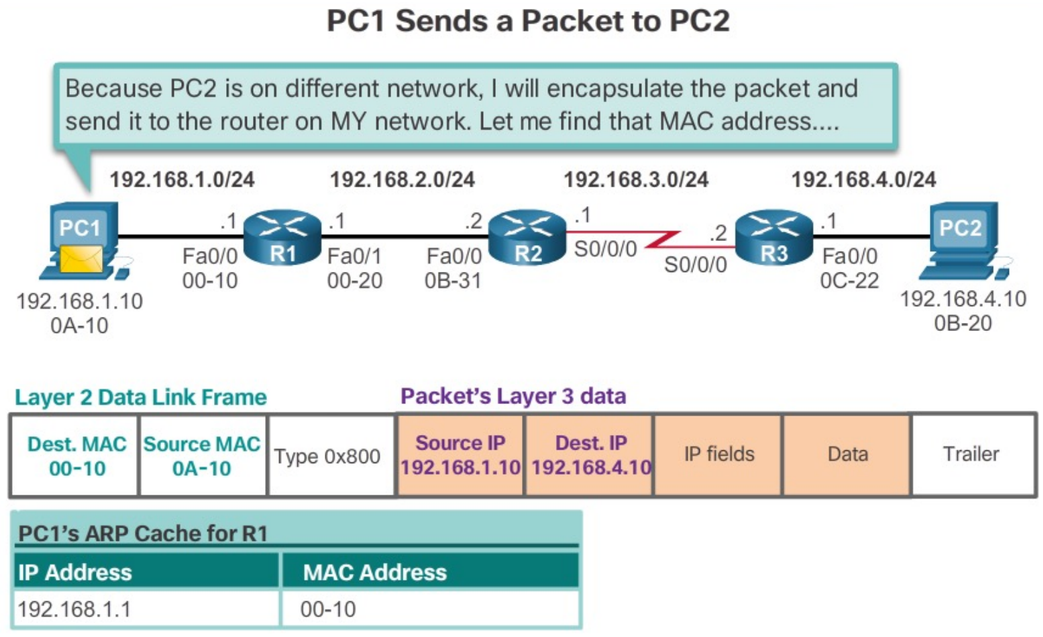
Router Switching Function



Routers Interconnect Networks

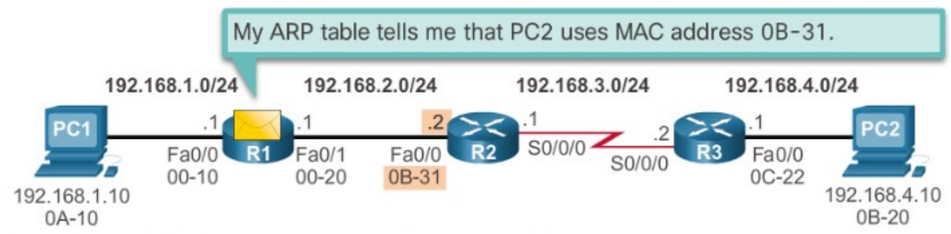


Send a Packet



Forward to Next Hop

R1 Forwards the Packet to PC2



Layer 2 Data Link Frame | Packet's Layer 3 data

Dest. MAC 0B-31		Type 0x800	Source IP 192.168.1.10	Dest. IP 192.168.4.10	IP fields	Data	Trailer

R1's ARP Cache

IP Address	MAC Address
192.168.2.2	0B-31

R1's Routing Table

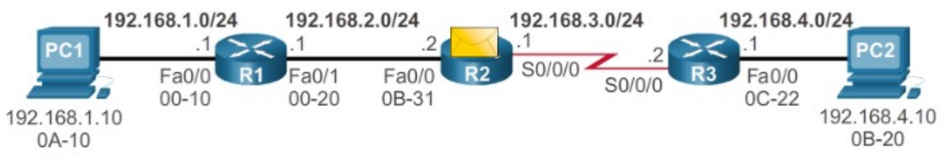
Network	Hops	Next-hop-IP	Exit Interface
192.168.1.0/24	0	Dir. Connect.	Fa0/0
192.168.2.0/24	0	Dir. Connect.	Fa0/1
192.168.3.0/24	1	192.168.2.2	Fa0/1
192.168.4.0/24	2	192.168.2.2	Fa0/1

Best Path

- **Best path is selected by a routing protocol based on the value or metric it uses to determine the distance to reach a network**:
  - A metric is the value used to measure the distance to a given network.
  - Best path to a network is the path with the lowest metric.
- **Dynamic routing protocols use their own rules and metrics to build and update routing tables**:
  - Routing Information Protocol (RIP) - Hop count
  - Open Shortest Path First (OSPF) - Cost based on cumulative bandwidth from source to destination
  - Enhanced Interior Gateway Routing Protocol (EIGRP) - Bandwidth, delay, load, reliability

Packet Routing

R2 Forwards the Packet to R3



Layer 2 Data Link Frame | Packet's Layer 3 data

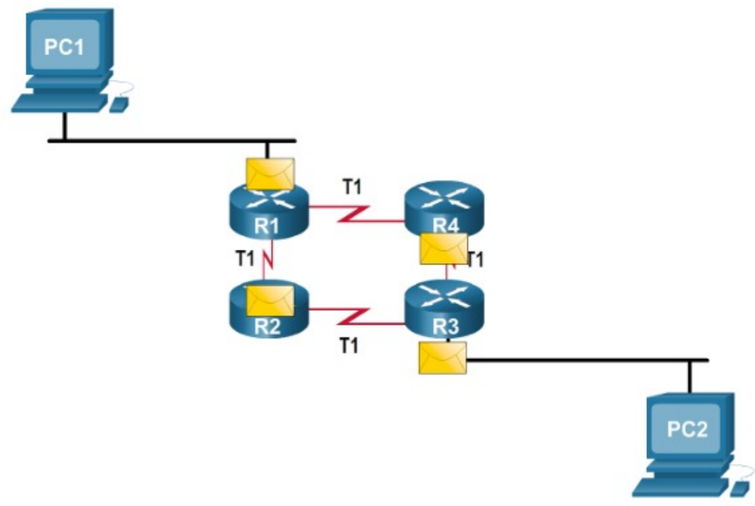
			Source IP 192.168.1.10	Dest. IP 192.168.4.10	IP fields	Data	Trailer

R2's Routing Table

Network	Hops	Next-hop-IP	Exit Interface
192.168.1.0/24	1	192.168.3.1	Fa/0/0
192.168.2.0/24	0	Dir. Connect.	Fa/0/0
192.168.3.0/24	0	Dir. Connect.	S0/0/0
192.168.4.0/24	1	192.162.3.2	S0/0/0

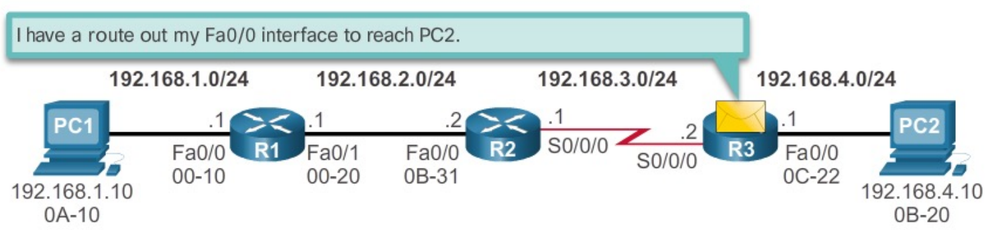
Load Balancing

- When a router has two or more paths to a destination with equal cost metrics, then the router forwards the packets using both paths equally:
  - Equal cost load balancing can improve network performance.
  - Equal cost load balancing can be configured to use both dynamic routing protocols and static routes.



Reach the Destination

R3 Forwards the Packet to PC2



Layer 2 Data Link Frame | Packet's Layer 3 data

		Type 0x800	Source IP 192.168.1.10	Dest. IP 192.168.4.10	IP fields	Data	Trailer

R3's Routing Table

Network	Hops	Next-hop-IP	Exit Interface
192.168.1.0/24	2	192.168.3.1	S0/0/0
192.168.2.0/24	1	192.168.3.1	S0/0/0
192.168.3.0/24	0	Dir. Connect.	S0/0/0
192.168.4.0/24	0	Dir. Connect.	Fa0/0

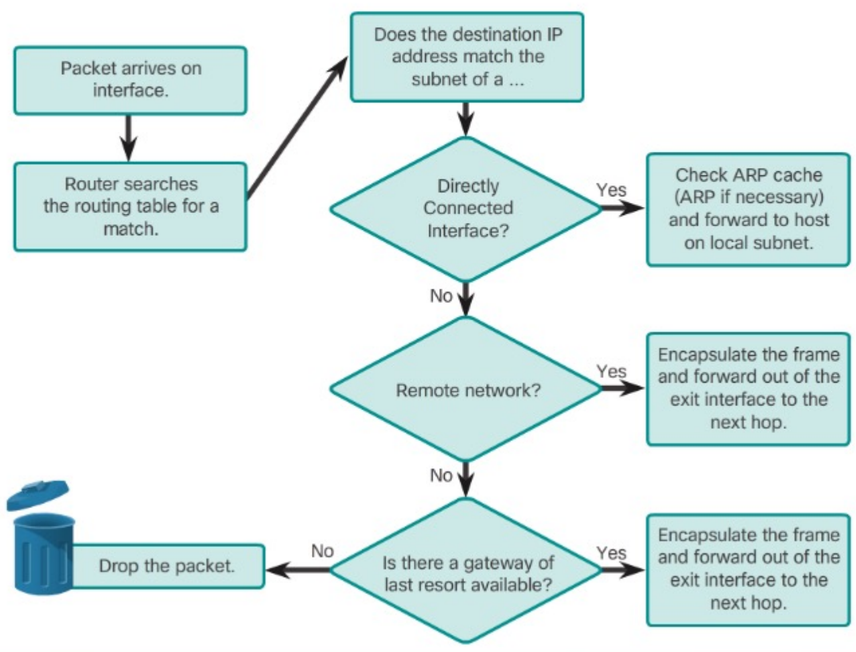
Administrative Distance

- If multiple paths to a destination are configured on a router, the path installed in the routing table is the one with the lowest Administrative Distance (AD):
  - A static route with an AD of 1 is more reliable than an EIGRP-discovered route with an AD of 90
  - A direc ute with an AD of 1.

Route Source	Administrative Distance
Connected	0
Static	1
EIGRP summary route	5
External BGP	20
Internal EIGRP	90
IGRP	100
OSPF	110
IS-IS	115
RIP	120
External EIGRP	170
Internal BGP	200

Routing Decisions

Packet Forwarding Decision Process



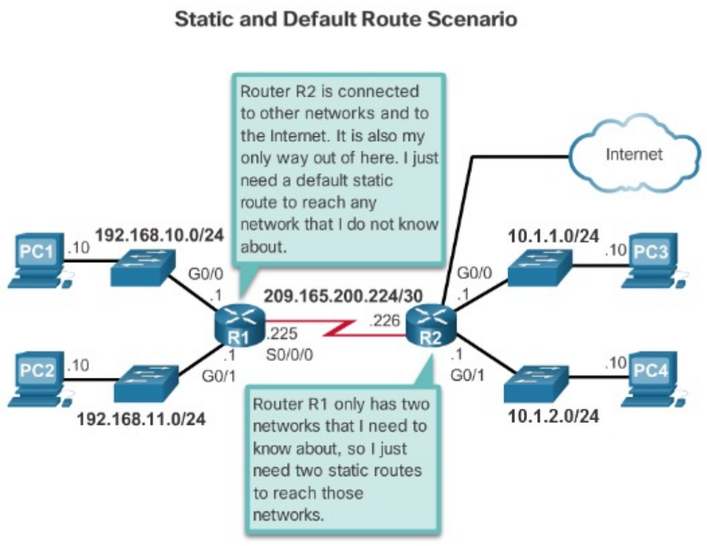
Static Routing

KHOA CÔNG NGHỆ THÔNG TIN
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

Reach Remote Networks

A router can learn about remote networks in one of two ways:

- **Manually** - Remote networks are manually entered into the route table using static routes.
- **Dynamically** - Remote routes are automatically learned using a dynamic routing protocol.



Why Use Static Routing?

Static routing provides some advantages over dynamic routing, including:

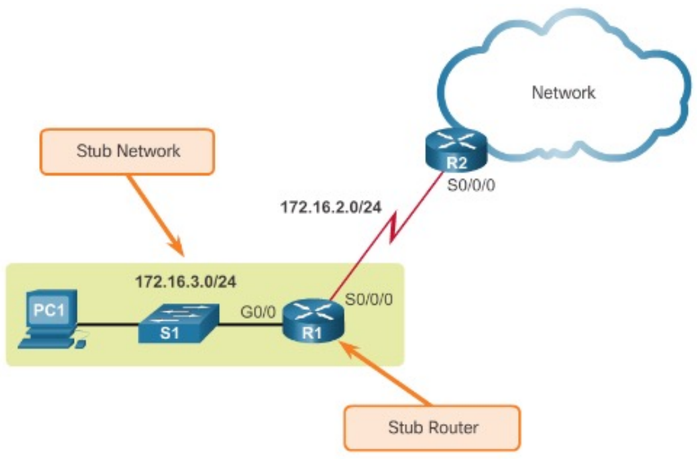
- Static routes are not advertised over the network, resulting in better security.
- Static routes use less bandwidth than dynamic routing protocols, no CPU cycles are used to calculate and communicate routes.
- The path a static route uses to send data is known.

	Dynamic Routing	Static Routing
Configuration Complexity	Generally independent of the network size	Increases with network size
Topology Changes	Automatically adapts to topology changes	Administrator intervention required
Scaling	Suitable for simple and complex topologies	Suitable for simple topologies
Security	Less secure	More secure
Resource Usage	Uses CPU, memory, link bandwith	No extra resources needed
Predictability	Route depends on the current topology	Route to destination is always the same

When to Use Static Routes

Static routing has three primary uses:

- Providing ease of routing table maintenance in smaller networks.
- Routing to and from stub networks. A stub network is a network accessed by a single route, and the router has no other neighbors.
- Using a single default route to represent a path to any network that does not have a more specific match with another route in the routing table.

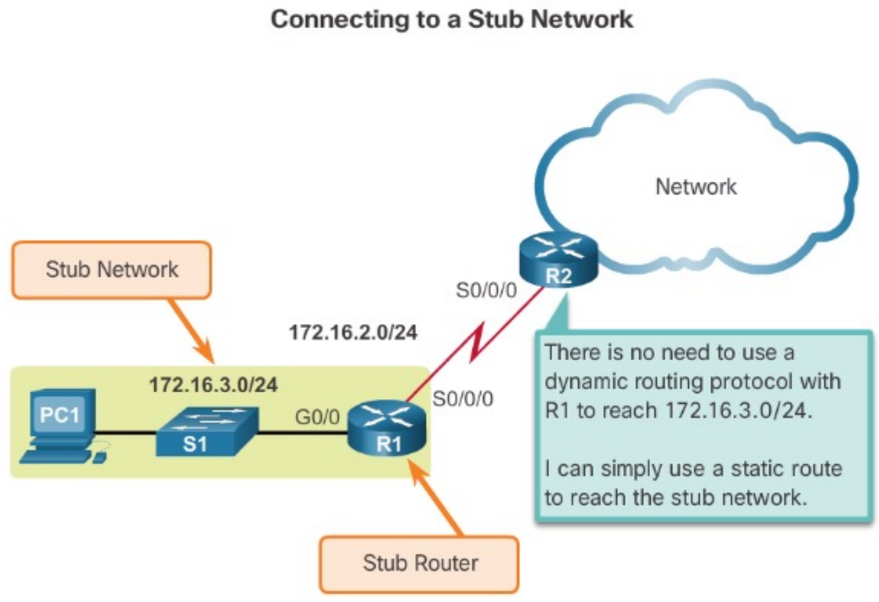


Static Route Applications

Static Routes are often used to:

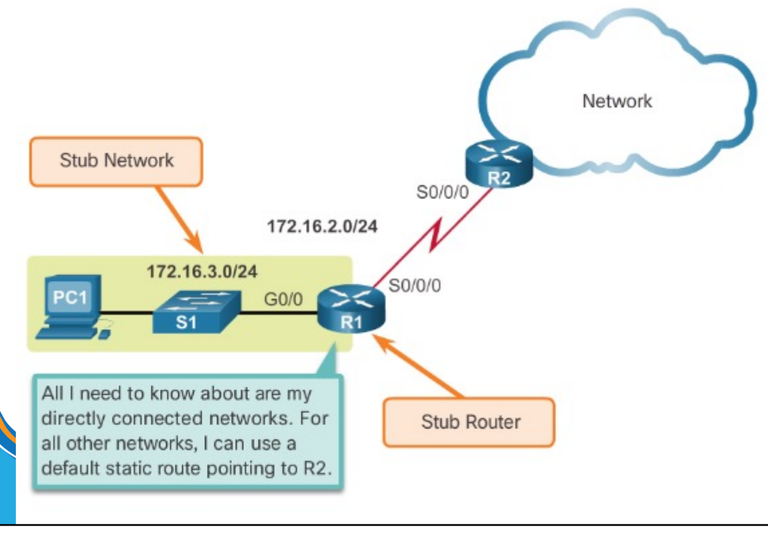
- Connect to a specific network.
- Provide a Gateway of Last Resort for a stub network.
- Reduce the number of routes advertised by summarizing several contiguous networks as one static route.
- Create a backup route in case a primary route link fails.

Standard Static Route

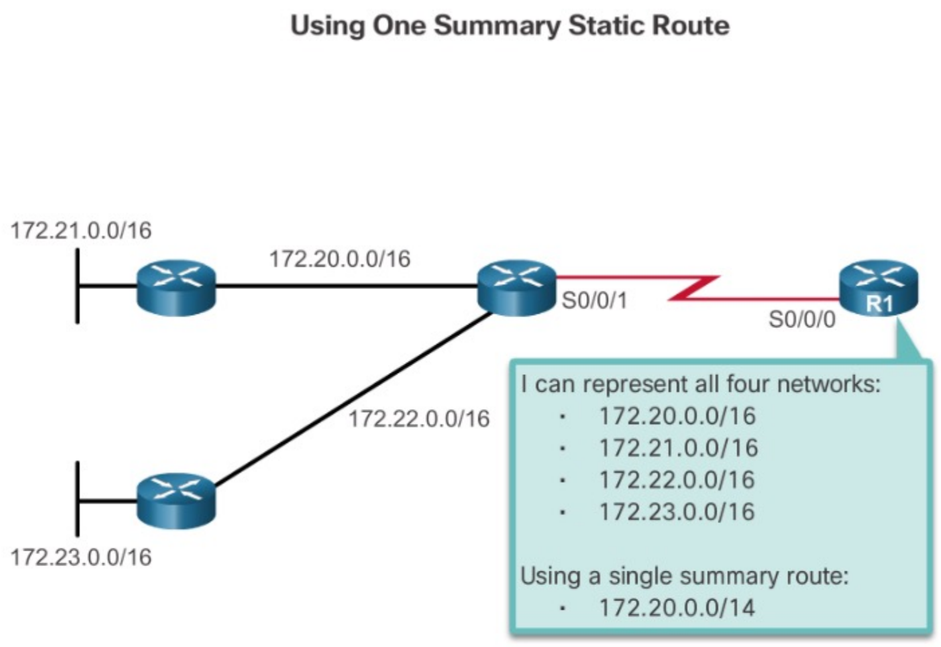


Default Static Route

- A default static route is a route that matches all packets.
- A default route identifies the gateway IP address to which the router sends all IP packets that it does not have a learned or static route.
- A default static route is simply a static route with 0.0.0.0/0 as the destination IPv4 address.



Summary Static Route



Static Route Applications

Static Routes are often used to:

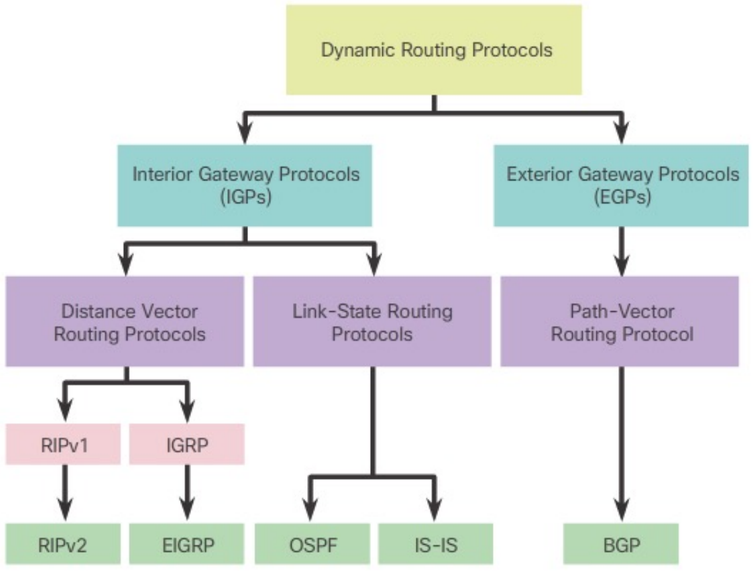
- Connect to a specific network.
- Provide a Gateway of Last Resort for a stub network.
- Reduce the number of routes advertised by summarizing several contiguous networks as one static route.
- Create a backup route in case a primary route link fails.

Dynamic Routing

KHOA CÔNG NGHỆ THÔNG TIN
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

Dynamic Routing Protocol Evolution

- Dynamic routing protocols have been used in networks since the late 1980s.
- Newer versions



Dynamic Routing Protocols Components

The purpose of dynamic routing protocols includes:

- Discovery of remote networks
- Maintaining up-to-date routing information
- Choosing the best path to destination networks
- Ability to find a new best path if the current path is no longer available

Dynamic Routing Protocols Components

Main components of dynamic routing protocols include:

- **Data structures -** Routing protocols typically use tables or databases for its operations. This information is kept in RAM.
- **Routing protocol messages -** Routing protocols use various types of messages to discover neighboring routers, exchange routing information, and other tasks to learn and maintain accurate information about the network.
- **Algorithm -** Routing protocols use algorithms for facilitating routing information for best path determination.

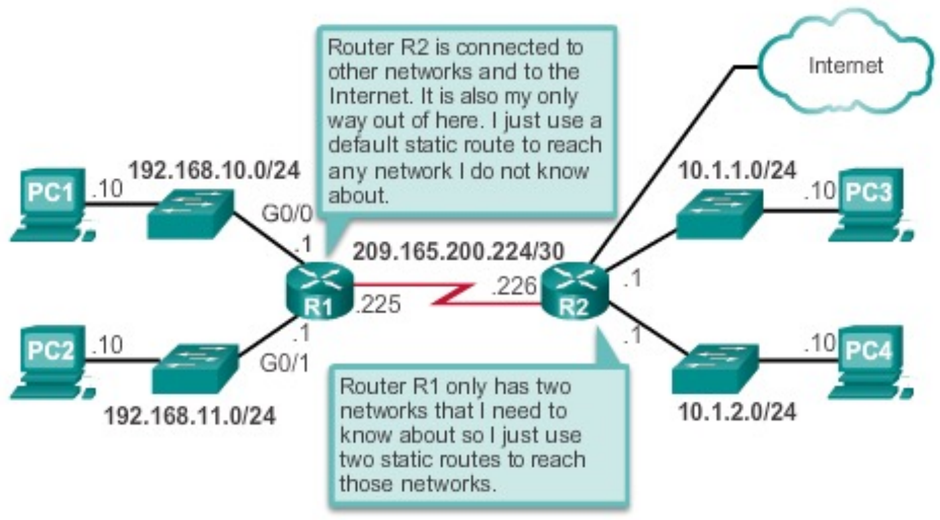
Static Routing Uses

Networks typically use a combination of both static and dynamic routing.

Static routing has several primary uses:

- Providing ease of routing table maintenance in smaller networks that are not expected to grow significantly.
- Routing to and from a stub network. A network with only one default route out and no knowledge of any remote networks.
- Accessing a single default router. This is used to represent a path to any network that does not have a match in the routing table.

Static Routing Uses (cont.)



Static Routing Advantages and Disadvantages

Advantages	Disadvantages
Easy to implement in a small network.	Suitable only for simple topologies or for special purposes such as a default static route.
Very secure. No advertisements are sent as compared to dynamic routing protocols.	Configuration complexity increases dramatically as network grows.
Route to destination is always the same.	Manual intervention required to re-route traffic.
No routing algorithm or update mechanism required; therefore, extra resources (CPU or RAM) are not required.	

Dynamic Routing Advantages & Disadvantages

Advantages	Disadvantages
Suitable in all topologies where multiple routers are required.	Can be more complex to implement.
Generally independent of the network size.	Less secure. Additional configuration settings are required to secure.
Automatically adapts topology to reroute traffic if possible.	Route depends on the current topology.
	Requires additional CPU, RAM, and link bandwidth.

Interior Gateway Protocols (IGP)

KHOA CÔNG NGHỆ THÔNG TIN
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

Distance Vector Routing Protocols

TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN TP. HỒ CHÍ MINH

KHOA CÔNG NGHỆ THÔNG TIN
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

fit@hcmus

Distance Vector Routing Protocols

- Routing Protocol Characteristics
  - Criteria used to compare routing protocols includes
    - Time to convergence
    - Scalability
    - Resource usage
    - Implementation & maintenance

Distance Vector Routing Protocols

- **Distance Vector Technology** - the Meaning of Distance Vector
  - A router using distance vector routing protocols knows 2 things:
    - Distance to final destination
    - Vector, or direction, traffic should be directed

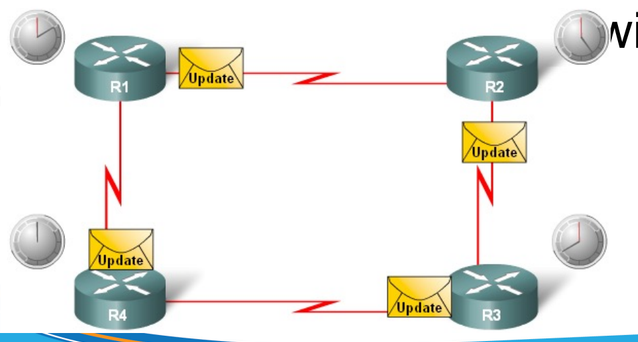
Distance Vector Routing Protocols

Advantages & Disadvantages of Distance Vector Routing Protocols

Advantages:	Disadvantages:
Simple implementation and maintenance. The level of knowledge required to deploy and later maintain a network with distance vector protocol is not high.	**Slow convergence.** The use of periodic updates can cause slower convergence. Even if some advanced techniques are used, like triggered updates which are discussed later, the overall convergence is still slower compared to link state routing protocols.
Low resource requirements. Distance vector protocols typically do not need large amounts of memory to store the information. Nor do they require a powerful CPU. Depending of the network size and the IP addressing implemented they also typically do not require a high level of link bandwidth to send routing updates. However, this can become an issue if you deploy a distance vector protocol in a large network.	**Limited scalability.** Slow convergence may limit the size of the network because larger networks require more time to propagate routing information.
	Routing loops. Routing loops can occur when inconsistent routing tables are not updated due to slow convergence in a changing network.

Distance Vector Routing Protocols

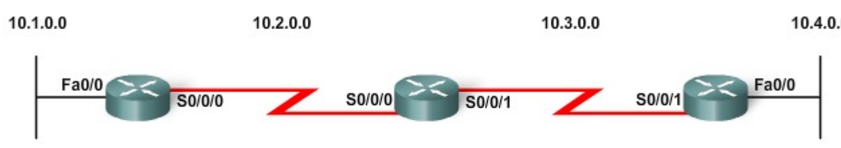
- **Characteristics of Distance Vector routing protocols**:
  - Periodic updates
  - Neighbors
  - Broadcast updates
  - Enti with routing upd



Network Discovery

- **Router initial start up** (Cold Starts)
  - **Initial network discovery**
    - Directly connected networks are initially placed in routing table

Network Discovery - Cold Start



Network	Interface	Hop
10.1.0.0	Fa0/0	0
10.2.0.0	S0/0/0	0

Network	Interface	Hop
10.2.0.0	S0/0/0	0
10.3.0.0	S0/0/1	0

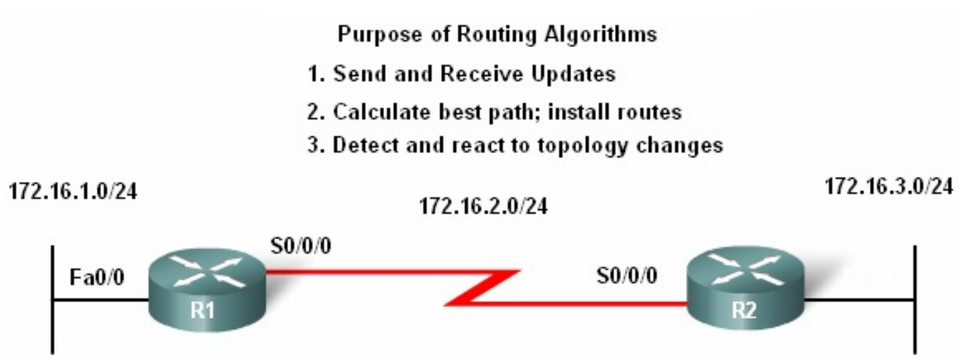
Network	Interface	Hop
10.3.0.0	S0/0/1	0
10.4.0.0	Fa0/0	0

Distance Vector Routing Protocols

- **Routing Protocol Algorithm**:
  - Defined as a procedure for accomplishing a certain task

Purpose of Routing Algorithms

1. Send and Receive Updates
2. Calculate best path; install routes
3. Detect and react to topology changes



Network	Interface	Hope
172.16.1.0/24	Fa0/0	0
172.16.2.0/24	S0/0/0	0

Network	Interface	Hope
172.16.2.0/24	S0/0/0	0
172.16.1.0/24	S0/0/0	1

Network Discovery

- **Initial Exchange** of Routing Information
  - If a routing protocol is configured then:
    - Routers will exchange routing information
    - Routing updates received from other routers
- Router checks update for new information
  - If the
    - M
    - Ne able

Network Discovery - Initial Exchange

Network	Interface	Hop
10.1.0.0	Fa0/0	0
10.2.0.0	S0/0/0	0
10.3.0.0	S0/0/0	1

Network	Interface	Hop
10.2.0.0	S0/0/0	0
10.3.0.0	S0/0/1	0
10.1.0.0	S0/0/0	1
10.4.0.0	S0/0/1	1

Network	Interface	Hop
10.3.0.0	S0/0/1	0
10.4.0.0	Fa0/0	0
10.2.0.0	S0/0/1	1

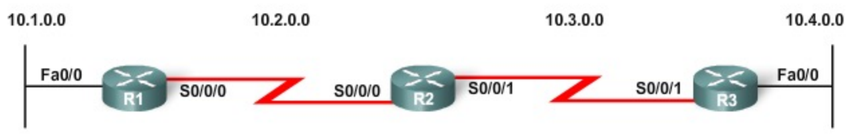
Network Discovery

Exchange of Routing Information

- **Router convergence** is reached when
  - All routing tables in the network contain the same network information
- Routers continue to exchange routing information

gence is

Network Discover - Next Update



Network	Interface	Hop
10.1.0.0	Fa0/0	0
10.2.0.0	S0/0/0	0
10.3.0.0	S0/0/0	1
10.4.0.0	S0/0/0	2

Network	Interface	Hop
10.2.0.0	S0/0/0	0
10.3.0.0	S0/0/1	0
10.1.0.0	S0/0/0	1
10.4.0.0	S0/0/1	1

Network	Interface	Hop
10.3.0.0	S0/0/1	0
10.4.0.0	Fa0/0	0
10.2.0.0	S0/0/1	1
10.1.0.0	S0/0/1	2

Link State Routing Protocols

fit@hcmus

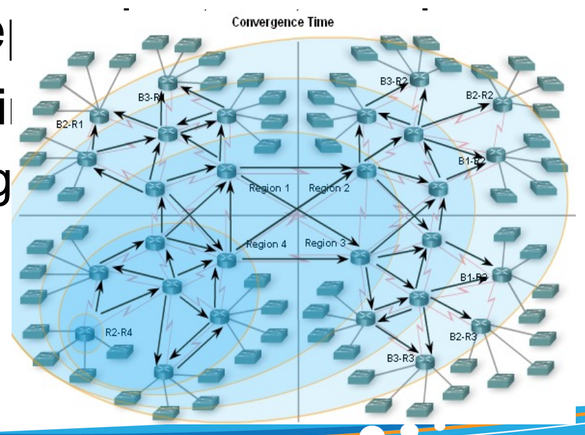
KHOA CÔNG NGHỆ THÔNG TIN
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

Network Discovery

Convergence must be reached before a network is considered completely operable

Speed of achieving convergence consists of 2 interde

- Speed of broadcasti
- Speed of calculating



Link-State Routing

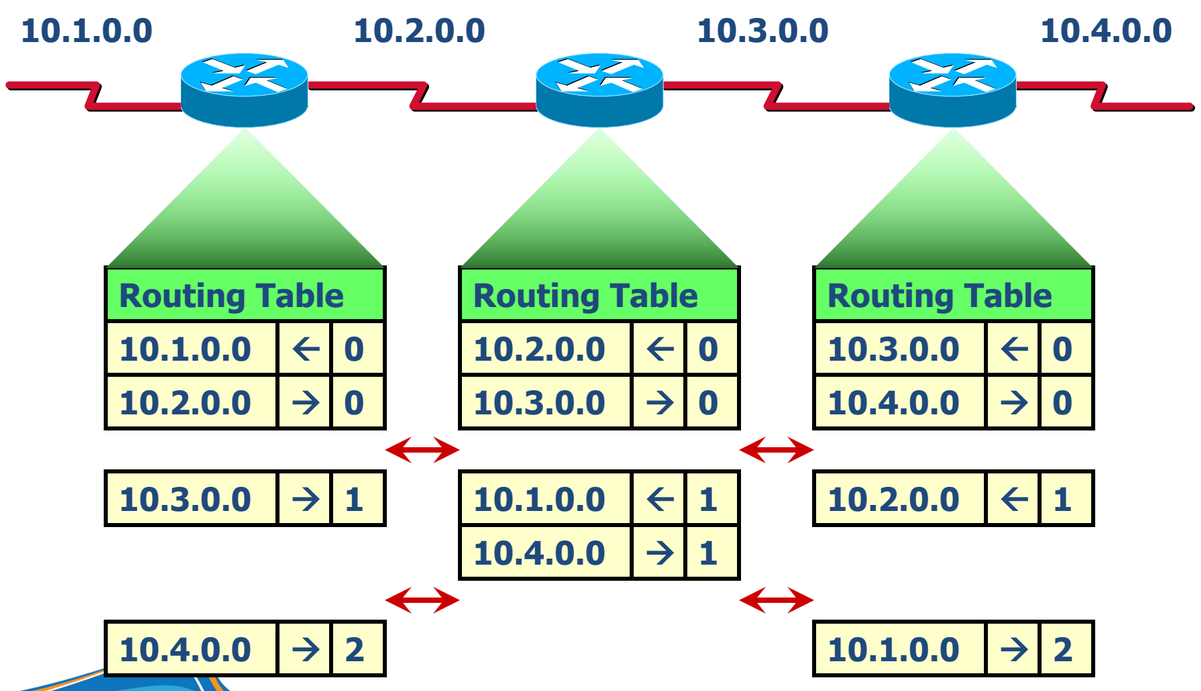
Link state routing protocols

- Also known as shortest path first algorithms
- These protocols built around Dijkstra's SPF

Classification of Routing Protocols

	Interior Gateway Protocols: Distance Vector Routing Protocols		Interior Gateway Protocols: Link State Routing Protocols		Exterior Gateway Protocols: Path Vector
Classful	RIP	IGRP			EGP
Classless	RIPv2	EIGRP	OSPFv2	IS-IS	BGPv4
IPv6	RIPng	EIGRP for IPv6	OSPFv3	IS-IS for IPv6	BGPv4 for IPv6

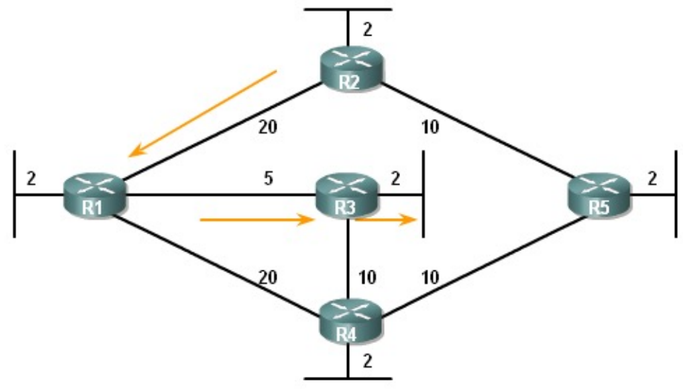
Network Discovery



Link-State Routing

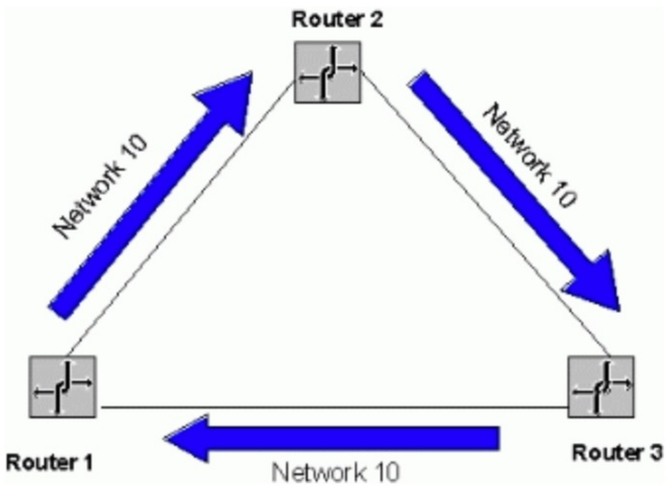
Dikjstra's algorithm also known as the shortest path first (SPF) algorithm

Dijkstra's Shortest Path First Algorithm



Shortest Path for host on R2 LAN to reach host on R3 LAN:
R2 to R1 (20) + R1 to R3 (5) + R3 to LAN (2) = 27

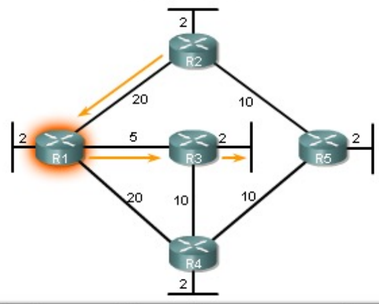
Routing Loop



Link-State Routing

The shortest path to a destination is not necessarily the path with the least number c

Introduction to the SPF Algorithm
SPF Tree for R1



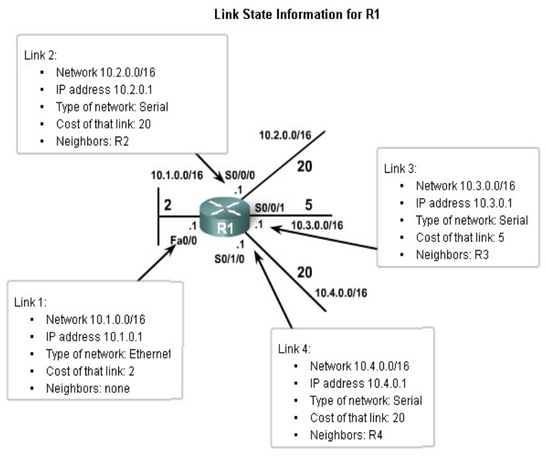
Destination	Shortest Path	Cost
R2 LAN	R1 to R2	22
R3 LAN	R1 to R3	7
R4 LAN	R1 to R3 to R4	17
R5 LAN	R1 to R3 to R4 to R5	27

Link-State Routing

- Link-State Routing Process
  - How routers using Link State Routing Protocols reach **convergence**
    - Each routers learns about its own directly connected networks
    - Link state routers exchange hello packet to “meet” other directly
    - Connected link state routers
    - Each router builds its own Link State Packet (LSP) which includes information about neighbors such as neighbor ID, link type, & bandwidth
    - After the LSP is created the router floods it to all neighbors who then store the information and then forward it until all routers have the same information
    - Once all the routers have received all the LSPs, the routers then construct a topological map of the network which is used to determine the best routes to a destination

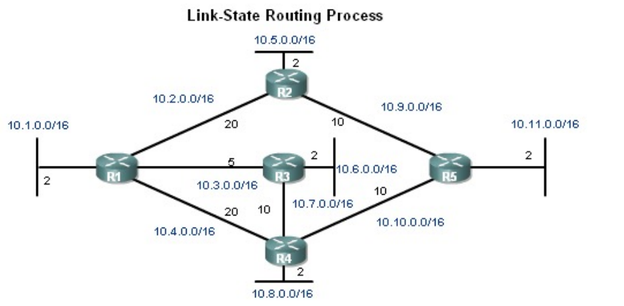
Link-State Routing

- Directly Connected Networks
- Link
  - This is an interface on a router
- Link state
  - This is the information about the state of the links



Link-State Routing

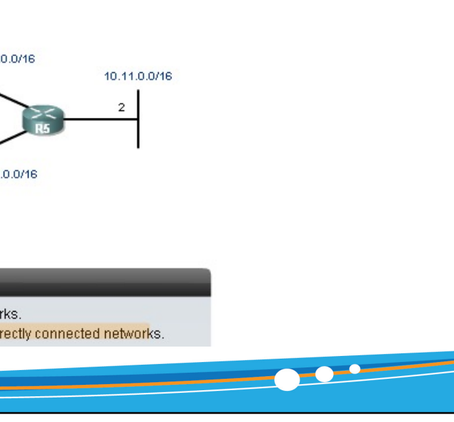
- Sending Hello Packets to Neighbors
  - Link state routing protocols use a hello protocol
  - Purpose of a hello protocol:
    - To discover neighbors (that use the same link state routing protocol) on its link



1. Each router learns about each of its own directly connected networks.
2. Each router is responsible for "saying hello" to its neighbors on directly connected networks.

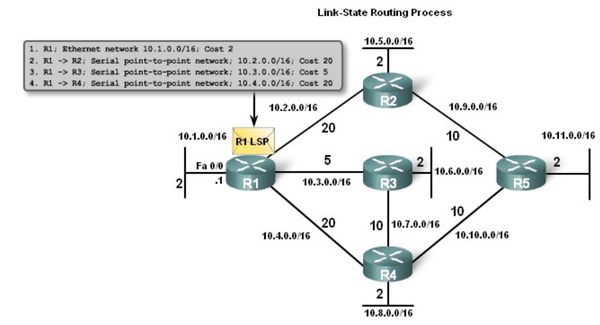
Link-State Routing

- Sending Hello Packets to Neighbors
  - Connected interfaces that are using the same link state routing protocols will exchange hello packets
  - Once routers learn it has neighbors they form an adjacency
    - 2 adjacent neighbors will exchange hello packets
    - These packets will serve as a keep alive



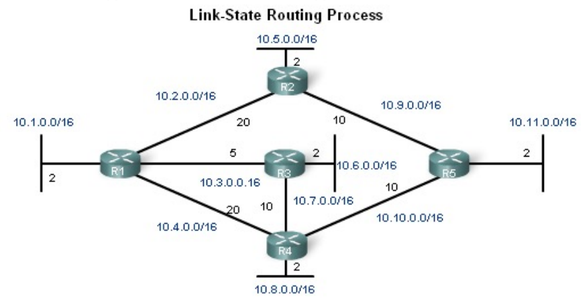
Link-State Routing

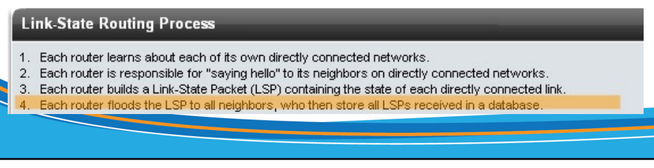
- Building the Link State Packet
  - Each router builds its own Link State Packet (LSP)
  - Contents of LSP:
    - State of each directly connected link
    - Includes information about neighbors such as neighbor ID, link type, & bandwidth



Link-State Routing

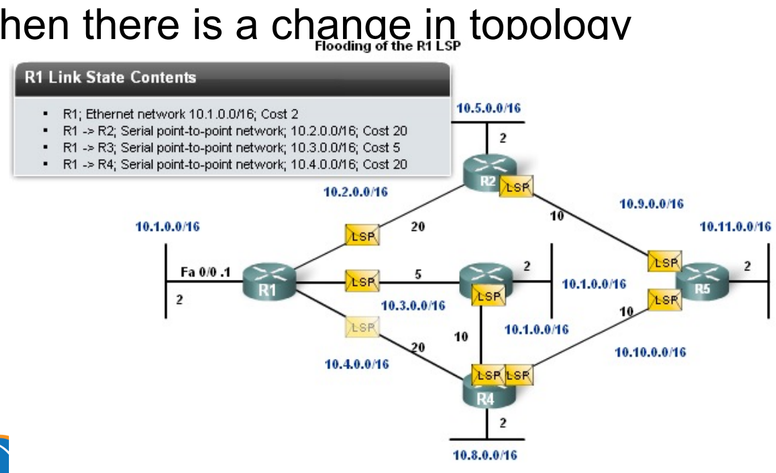
- Flooding LSPs to Neighbors
  - Once LSP are created they are forwarded out to neighbors
  - After receiving the LSP the neighbor continues to forward it throughout routing area





Link-State Routing

- LSPs are sent out under the following conditions:
  - Initial router start up or routing process
  - When there is a change in topology

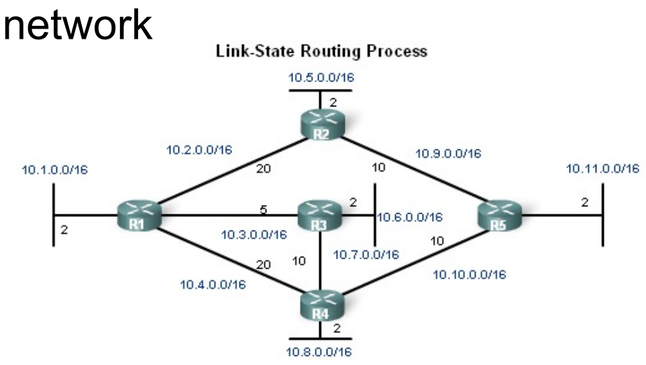


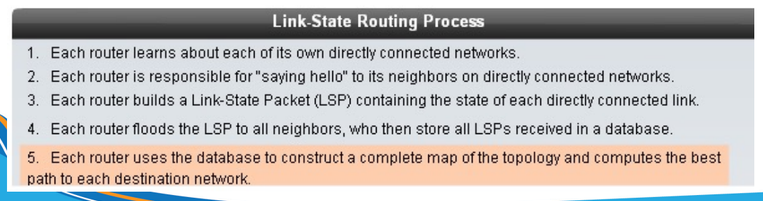
Link-State Routing

- Sending Hello Packets to Neighbors
  - Connected interfaces that are using the same link state routing protocols will exchange hello packets
  - Once routers learn it has neighbors they form an adjacency
    - 2 adjacent neighbors will exchange hello packets
    - These packets will serve as a keep alive

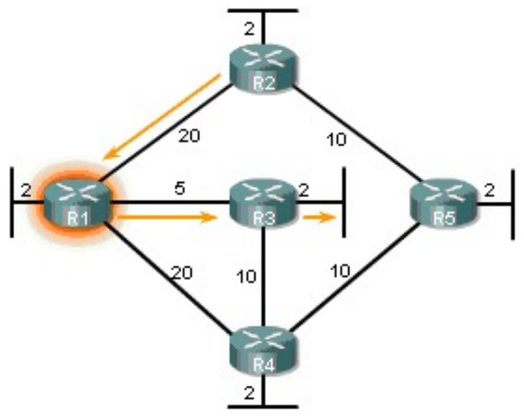
Link-State Routing

- Constructing a link state data base
  - Routers use a database to construct a topology map of the network





Link-State Routing



Destination	Shortest Path	Cost
R2 LAN	R1 to R2	22
R3 LAN	R1 to R3	7
R4 LAN	R1 to R3 to R4	17
R5 LAN	R1 to R3 to R4 to R5	27

R1 Link-State Database

R1s Link-State DatabaseLSPs from R2:
- Connected to neighbor R1 on network 10.2.0.0/16, cost of 20
- Connected to neighbor R5 on network 10.9.0.0/16, cost of 10
- Has a network 10.5.0.0/16, cost of 2

LSPs from R3:
- Connected to neighbor R1 on network 10.3.0.0/16, cost of 5
- Connected to neighbor R4 on network 10.7.0.0/16, cost of 10
- Has a network 10.6.0.0/16, cost of 2

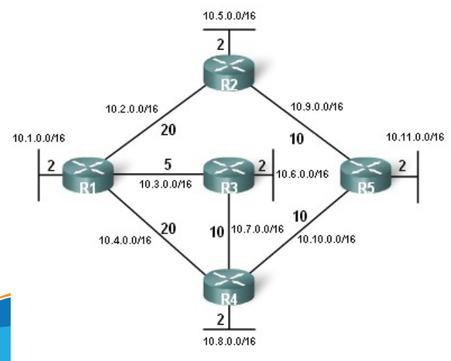
LSPs from R4:
- Connected to neighbor R1 on network 10.4.0.0/16, cost of 20
- Connected to neighbor R3 on network 10.7.0.0/16, cost of 10
- Connected to neighbor R5 on network 10.10.0.0/16, cost of 10
- Has a network 10.8.0.0/16, cost of 2

LSPs from R5:
- Connected to neighbor R2 on network 10.9.0.0/16, cost of 10
- Connected to neighbor R4 on network 10.10.0.0/16, cost of 10
- Has a network 10.11.0.0/16, cost of 2

R1 Link-states:
- Connected to neighbor R2 on network 10.2.0.0/16, cost of 20
- Connected to neighbor R3 on network 10.3.0.0/16, cost of 5
- Connected to neighbor R4 on network 10.4.0.0/16, cost of 20
- Has a network 10.1.0.0/16, cost of 2

Link-State Routing

- Shortest Path First (SPF) Tree
  - Building a **portion** of the SPF tree
  - Process begins by examining R2's LSP information
    - R1 *ignores* 1st LSP
    - Reason: R1 already knows it's connected to R2



R1s Link State Database

R1 Links-states:
- Connected to neighbor R2 on network 10.2.0.0/16, cost of 20
- Connected to neighbor R3 on network 10.3.0.0/16, cost of 5
- Connected to neighbor R4 on network 10.4.0.0/16, cost of 20
- Has a network 10.1.0.0/16, cost of 2

LSPs from R2:
- Connected to neighbor R1 on network 10.2.0.0/16, cost of 20
- Connected to neighbor R5 on network 10.9.0.0/16, cost of 10
- Has a network 10.5.0.0/16, cost of 2

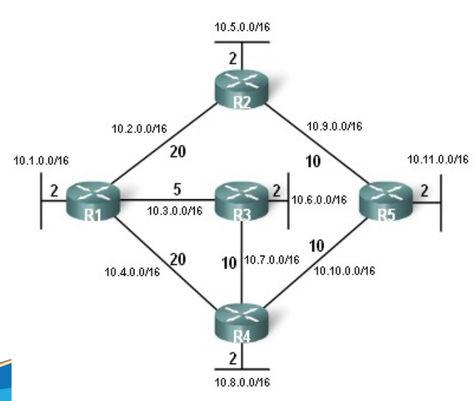
LSPs from R3:
- Connected to neighbor R1 on network 10.3.0.0/16, cost of 5
- Connected to neighbor R4 on network 10.7.0.0/16, cost of 10
- Has a network 10.6.0.0/16, cost of 2

LSPs from R4:
- Connected to neighbor R1 on network 10.4.0.0/16, cost of 20
- Connected to neighbor R3 on network 10.7.0.0/16, cost of 10
- Connected to neighbor R5 on network 10.10.0.0/16, cost of 10
- Has a network 10.8.0.0/16, cost of 2

LSPs from R5:
- Connected to neighbor R2 on network 10.9.0.0/16, cost of 10
- Connected to neighbor R4 on network 10.10.0.0/16, cost of 10
- Has a network 10.11.0.0/16, cost of 2

Link-State Routing

- Building a **portion** of the SPF tree
  - R1 uses 2nd LSP
    - Reason: R1 can create a link from R2 to R5 - this information is added to R1's SPF tree



R1s Link State Database

R1 Links-states:
- Connected to neighbor R2 on network 10.2.0.0/16, cost of 20
- Connected to neighbor R3 on network 10.3.0.0/16, cost of 5
- Connected to neighbor R4 on network 10.4.0.0/16, cost of 20
- Has a network 10.1.0.0/16, cost of 2

LSPs from R2:
- Connected to neighbor R1 on network 10.2.0.0/16, cost of 20
- Connected to neighbor R5 on network 10.9.0.0/16, cost of 10
- Has a network 10.5.0.0/16, cost of 2

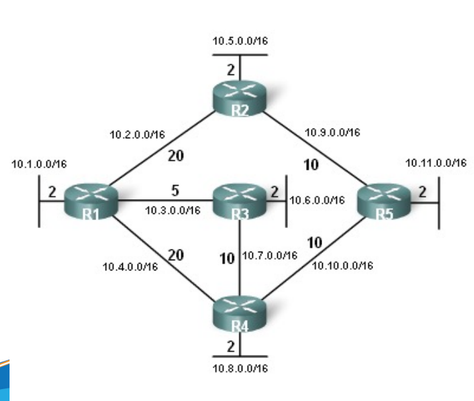
LSPs from R3:
- Connected to neighbor R1 on network 10.3.0.0/16, cost of 5
- Connected to neighbor R4 on network 10.7.0.0/16, cost of 10
- Has a network 10.6.0.0/16, cost of 2

LSPs from R4:
- Connected to neighbor R1 on network 10.4.0.0/16, cost of 20
- Connected to neighbor R3 on network 10.7.0.0/16, cost of 10
- Connected to neighbor R5 on network 10.10.0.0/16, cost of 10
- Has a network 10.8.0.0/16, cost of 2

LSPs from R5:
- Connected to neighbor R2 on network 10.9.0.0/16, cost of 10
- Connected to neighbor R4 on network 10.10.0.0/16, cost of 10
- Has a network 10.11.0.0/16, cost of 2

Link-State Routing

- Building a **portion** of the SPF tree
  - R1 uses 3rd LSP
    - Reason: R1 learns that R2 is connected to 10.5.0.0/16
    - This link is added to R1's SPF tree



R1s Link State Database

R1 Links-states:
- Connected to neighbor R2 on network 10.2.0.0/16, cost of 20
- Connected to neighbor R3 on network 10.3.0.0/16, cost of 5
- Connected to neighbor R4 on network 10.4.0.0/16, cost of 20
- Has a network 10.1.0.0/16, cost of 2

LSPs from R2:
- Connected to neighbor R1 on network 10.2.0.0/16, cost of 20
- Connected to neighbor R5 on network 10.9.0.0/16, cost of 10
- Has a network 10.5.0.0/16, cost of 2

LSPs from R3:
- Connected to neighbor R1 on network 10.3.0.0/16, cost of 5
- Connected to neighbor R4 on network 10.7.0.0/16, cost of 10
- Has a network 10.6.0.0/16, cost of 2

LSPs from R4:
- Connected to neighbor R1 on network 10.4.0.0/16, cost of 20
- Connected to neighbor R3 on network 10.7.0.0/16, cost of 10
- Connected to neighbor R5 on network 10.10.0.0/16, cost of 10
- Has a network 10.8.0.0/16, cost of 2

LSPs from R5:
- Connected to neighbor R2 on network 10.9.0.0/16, cost of 10
- Connected to neighbor R4 on network 10.10.0.0/16, cost of 10
- Has a network 10.11.0.0/16, cost of 2

Link-State Routing

- Building a **portion** of the SPF tree
  - R1 uses 3rd LSP
    - Reason: R1 learns that R2 is connected to 10.5.0.0/16
    - This link is added to R1's SPF tree

R1s Link State Database

R1 Links-states:
- Connected to neighbor R2 on network 10.2.0.0/16, cost of 20
- Connected to neighbor R3 on network 10.3.0.0/16, cost of 5
- Connected to neighbor R4 on network 10.4.0.0/16, cost of 20
- Has a network 10.1.0.0/16, cost of 2

LSPs from R2:
- Connected to neighbor R1 on network 10.2.0.0/16, cost of 20
- Connected to neighbor R5 on network 10.9.0.0/16, cost of 10
- Has a network 10.5.0.0/16, cost of 2

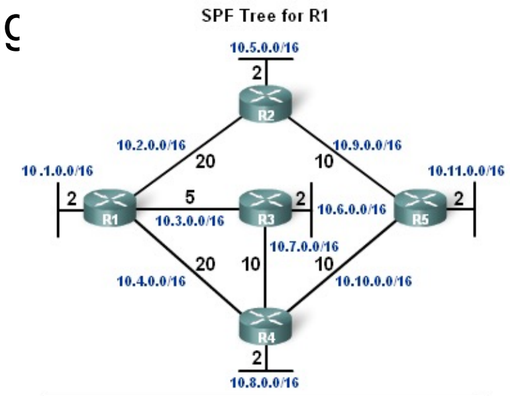
LSPs from R3:
- Connected to neighbor R1 on network 10.3.0.0/16, cost of 5
- Connected to neighbor R4 on network 10.7.0.0/16, cost of 10
- Has a network 10.6.0.0/16, cost of 2

LSPs from R4:
- Connected to neighbor R1 on network 10.4.0.0/16, cost of 20
- Connected to neighbor R3 on network 10.7.0.0/16, cost of 10
- Connected to neighbor R5 on network 10.10.0.0/16, cost of 10
- Has a network 10.8.0.0/16, cost of 2

LSPs from R5:
- Connected to neighbor R2 on network 10.9.0.0/16, cost of 10
- Connected to neighbor R4 on network 10.10.0.0/16, cost of 10
- Has a network 10.11.0.0/16, cost of 2

Link-State Routing

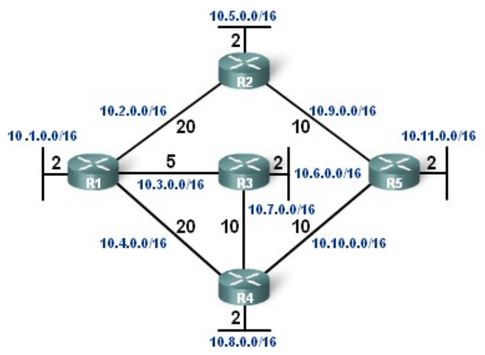
- Determining the shortest path
  - The shortest path to a destination determined by adding [illegible] e lowest cost



Destination	Shortest Path	Cost
R2 LAN	R1 to R2	22
R3 LAN	R1 to R3	7
R4 LAN	R1 to R3 to R4	17
R5 LAN	R1 to R3 to R4 to R5	27

Link-State Routing

- Once the SPF algorithm has determined the shortest path routes, these routes [illegible]



R1 Routing Table

SPF Information
- Network 10.5.0.0/16 via R2 serial 0/0/0 at a cost of 22
- Network 10.6.0.0/16 via R3 serial 0/0/1 at a cost of 7
- Network 10.7.0.0/16 via R3 serial 0/0/1 at a cost of 15
- Network 10.8.0.0/16 via R3 serial 0/0/1 at a cost of 17
- Network 10.9.0.0/16 via R2 serial 0/0/0 at a cost of 30
- Network 10.10.0.0/16 via R3 serial 0/0/1 at a cost of 25
- Network 10.11.0.0/16 via R3 serial 0/0/1 at a cost of 27

R1 Routing Table

Directly Connected Networks
- 10.1.0.0/16 Directly Connected Network
- 10.2.0.0/16 Directly Connected Network
- 10.3.0.0/16 Directly Connected Network
- 10.4.0.0/16 Directly Connected Network

Remote Networks
- 10.5.0.0/16 via R2 serial 0/0/0, cost = 22
- 10.6.0.0/16 via R3 serial 0/0/1, cost = 7
- 10.7.0.0/16 via R3 serial 0/0/1, cost = 15
- 10.8.0.0/16 via R3 serial 0/0/1, cost = 17
- 10.9.0.0/16 via R2 serial 0/0/0, cost = 30
- 10.10.0.0/16 via R3 serial 0/0/1, cost = 25
- 10.11.0.0/16 via R3 serial 0/0/1, cost = 27

Link-State Routing Protocols

- Advantages of a Link-State Routing Protocol

Routing protocol	Builds Topological map	Router can independently determine the shortest path to every network.	Convergence	A periodic/ event driven routing updates	Use of LSP
Distance vector	No	No	Slow	Generally No	No
Link State	Yes	Yes	Fast	Generally Yes	Yes

Link-State Routing Protocols

- Requirements for using a link state routing protocol
  - Memory requirements
    - Typically link state routing protocols use more memory
  - Processing Requirements
    - More CPU processing is required of link state routing protocols
  - Bandwidth Requirements
    - Initial startup of link state routing protocols can consume lots of bandwidth

Link-State Routing Protocols

- 2 link state routing protocols used for routing IP
  - Open Shortest Path First (OSPF)
  - Intermediate System-Intermediate System (IS

OSPF and IS-IS

OSPF

- OSPFv2: OSPF for IPv4 networks (RFC 1247 and RFC 2328)
- OSPFv3: OSPF for IPv6 networks (RFC 2740)
- OSPFv2 discussed in chapter 11

IS-IS

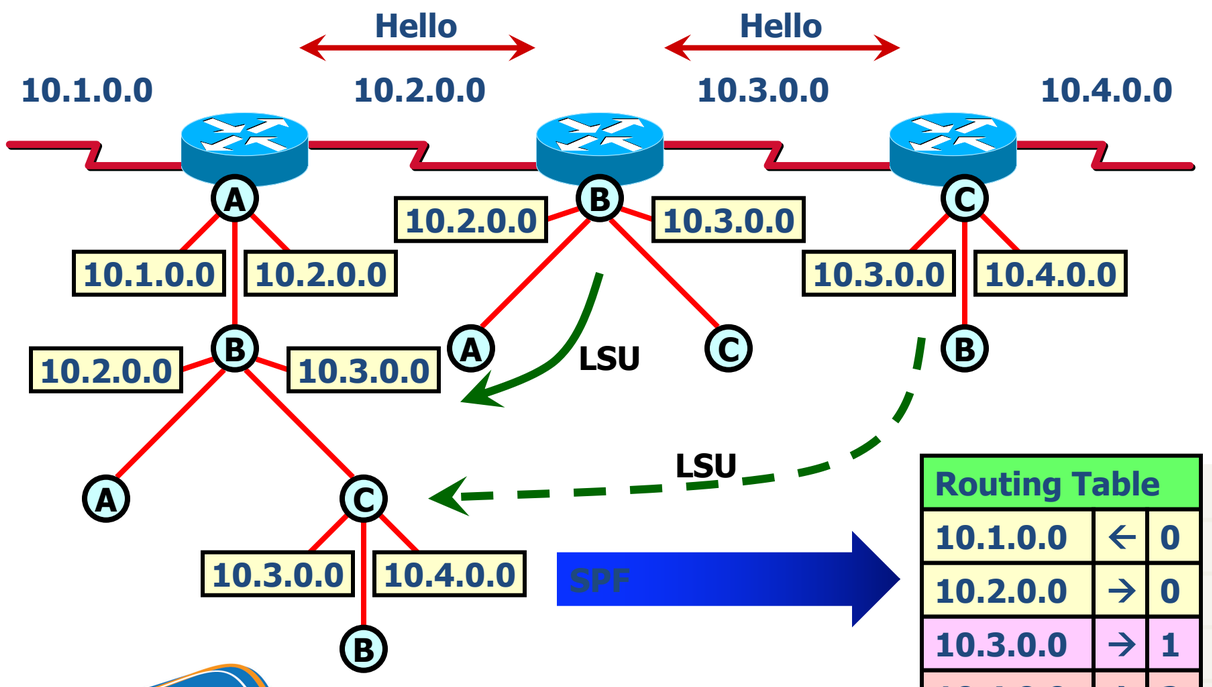
- ISO 10589
- Integrated IS-IS, Dual IS-IS supports IP networks
- Used mainly by ISPs and carriers
- Discussed in CCNP

Exterior Gateway Protocols (IGP)

fit@hcmus

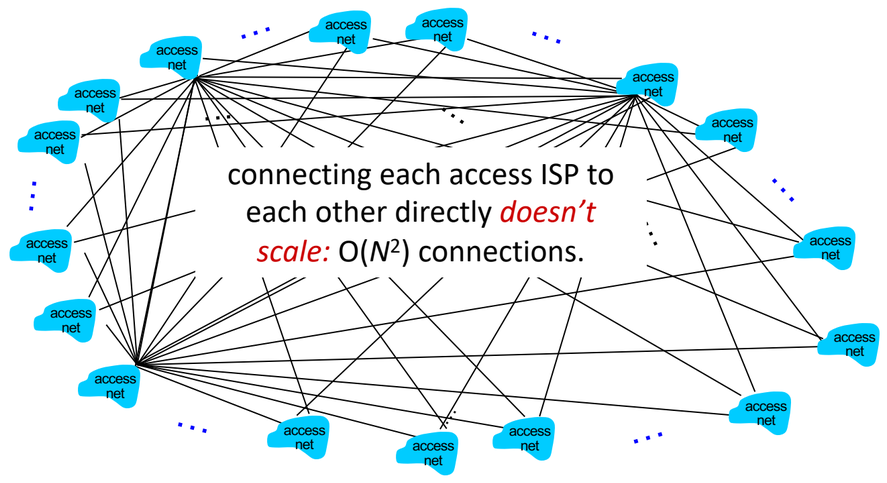
KHOA CÔNG NGHỆ THÔNG TIN
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

Link state



Internet structure: a “network of networks”

Question: given *millions* of access ISPs, how to connect them together?



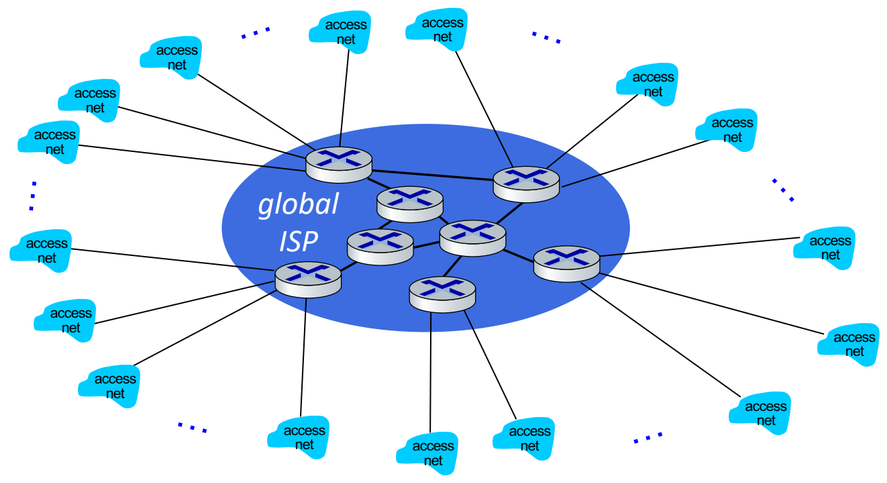
Routing Information Protocol (RIP)

fit@hcmus

KHOA CÔNG NGHỆ THÔNG TIN
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

Internet structure: a “network of networks”

Option: connect each access ISP to one global transit ISP?
Customer and provider ISPs have economic agreement.



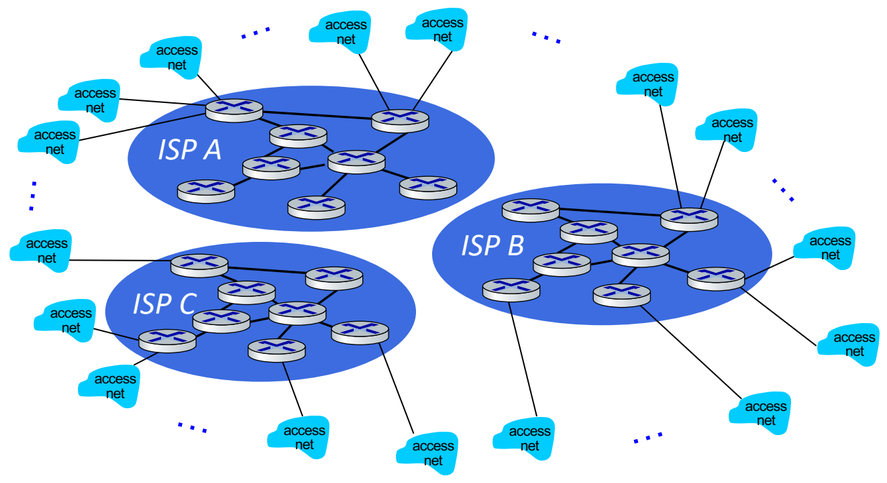
Open Shortest Path First (OSPF)

fit@hcmus

KHOA CÔNG NGHỆ THÔNG TIN
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

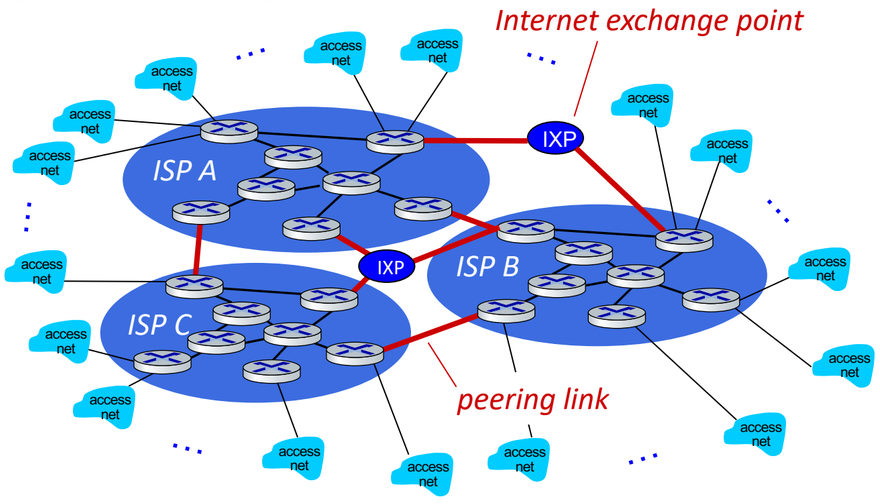
Internet structure: a “network of networks”

But if one global ISP is viable business, there will be competitors ….



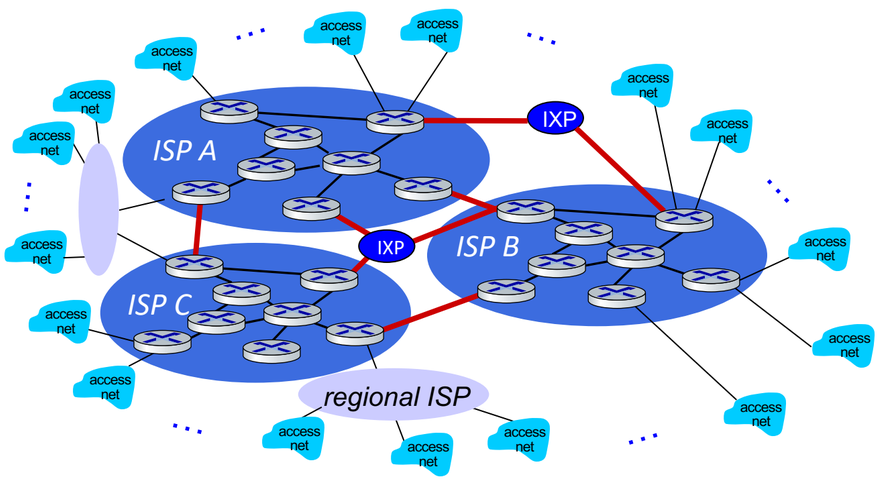
Internet structure: a “network of networks”

But if one global ISP is viable business, there will be competitors …. who will want to be connected



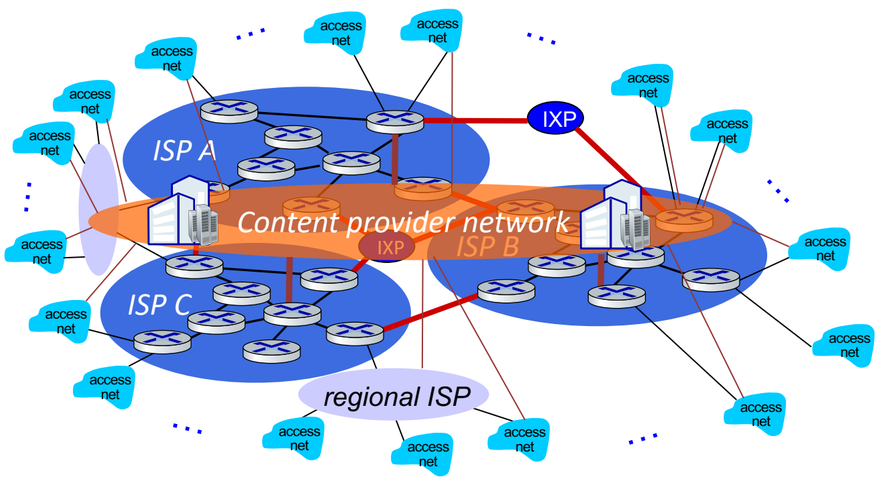
Internet structure: a “network of networks”

… and regional networks may arise to connect access nets to ISPs

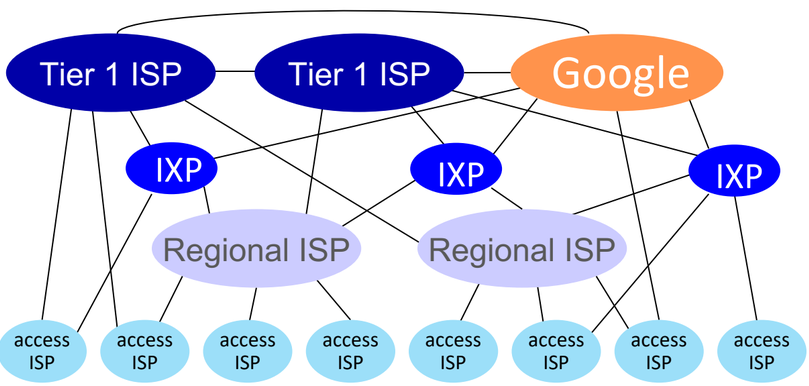


Internet structure: a “network of networks”

… and content provider networks (e.g., Google, Microsoft, Akamai) may run their own network, to bring services, content close to end users



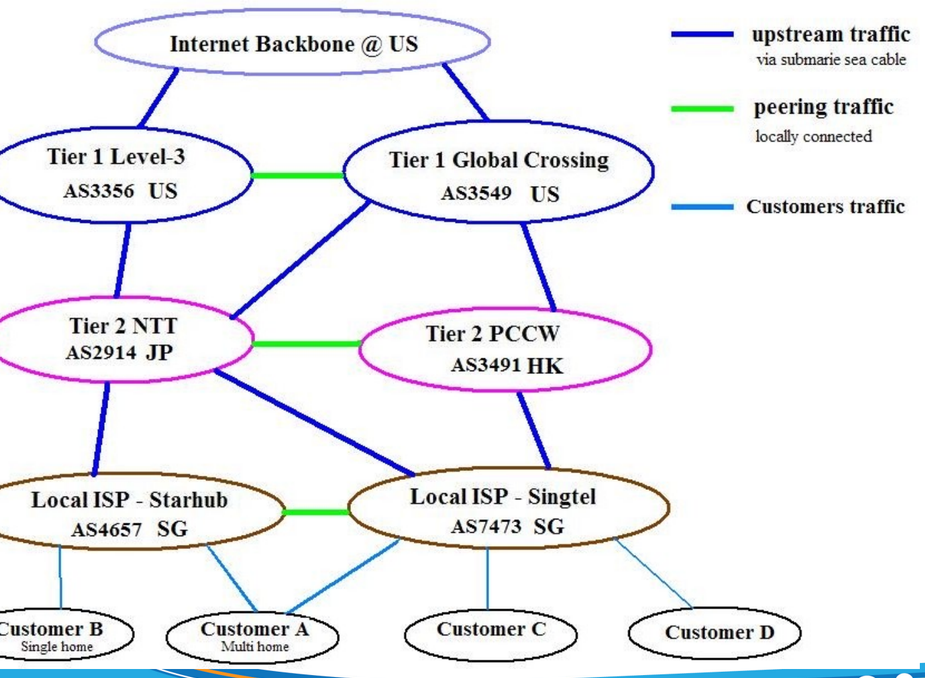
Internet structure: a “network of networks”



At “center”: small # of well-connected large networks

- “tier-1” commercial ISPs (e.g., Level 3, Sprint, AT&T, NTT), national & international coverage
- content provider networks (e.g., Google, Facebook): private network that connects its data centers to Internet, often bypassing tier-1, regional ISPs

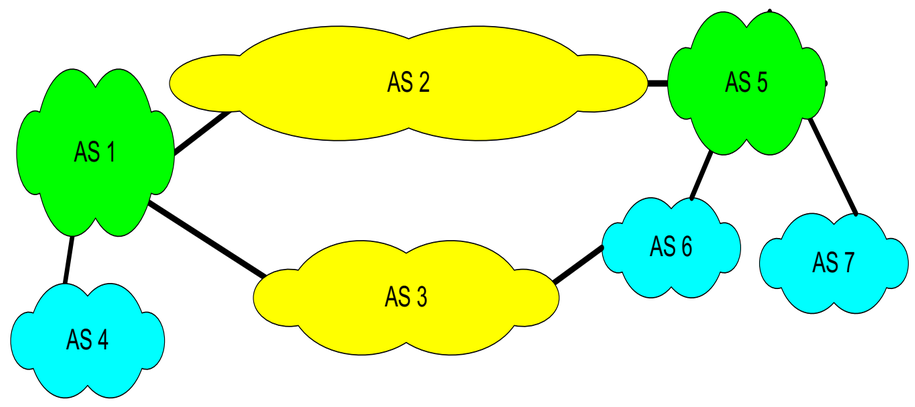
Internet structure: a “network of networks”



Autonomous Systems

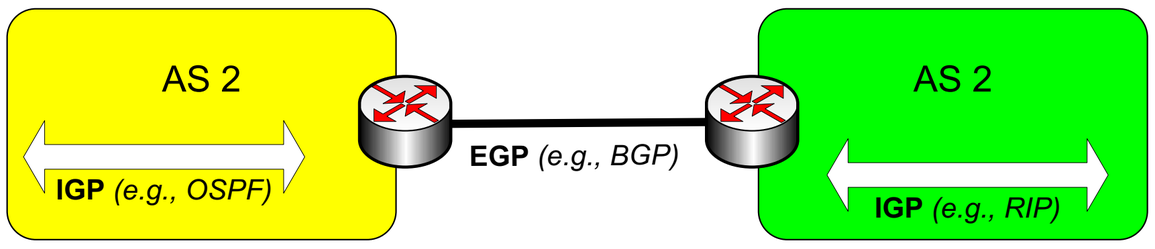
- An **autonomous system (AS)** is a region of the Internet that is administered by a single entity and that has a unified routing policy
- Each autonomous system is assigned an Autonomous System Number (**ASN**).
  - UofT’s campus network (AS239)
  - Rogers Cable Inc. (AS812)
  - Sprint (AS1239, AS1240, AS 6211, …)
- Interdomain routing is concerned with determining paths between autonomous systems **(interdomain routing**)
- Routing protocols for interdomain routing are called **exterior gateway protocols** (EGP)

Interdomain vs Intradomain Routing



- Routing protocols for intradomain routing are called interior gateway protocols (IGP)
  - Objective: shortest path
- Routing protocols for interdomain routing are called exterior gateway protocols (EGP)
  - Objective: satisfy policy of the AS

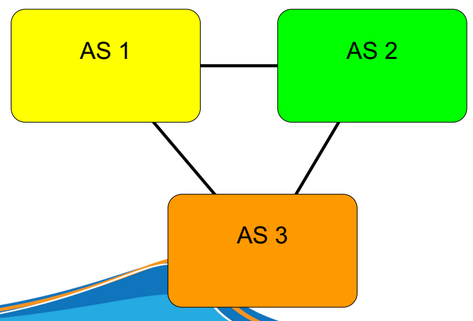
Interdomain vs Intradomain

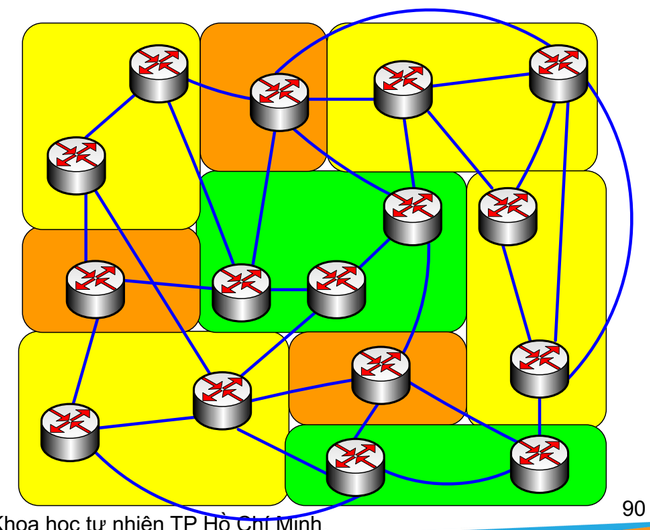


- Intradomain routing
  - Routing is done based on metrics
  - Routing domain is one autonomous system
- Interdomain routing
  - Routing is done based on policies
  - Routing domain is the entire Internet

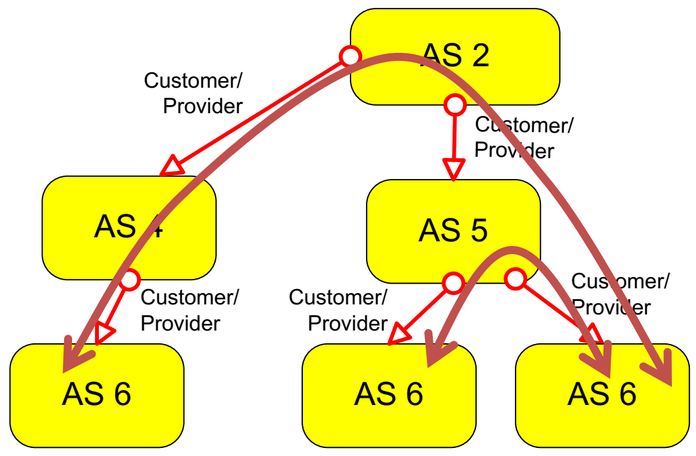
Interdomain Routing

- Interdomain routing is based on connectivity between autonomous systems
- Interdomain routing can ignore many details of router interconnection





Customer/Provider



- A stub network typically obtains access to the Internet through a transit network.
- Transit network that is a provider may be a customer for another network
- Customer pays provider for service

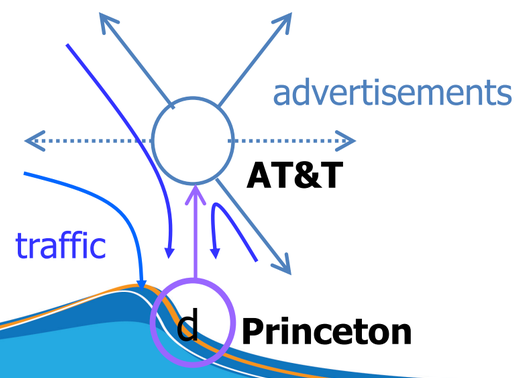
Autonomous Systems Terminology

- **local traffic** = traffic with source or destination in AS
- **transit traffic** = traffic that passes through the AS
- **Stub AS** = has connection to only one AS, only carry local traffic
- **Multihomed AS** = has connection to >1 AS, but does not carry transit traffic
- **Transit AS** = has connection to >1 AS and carries transit traffic

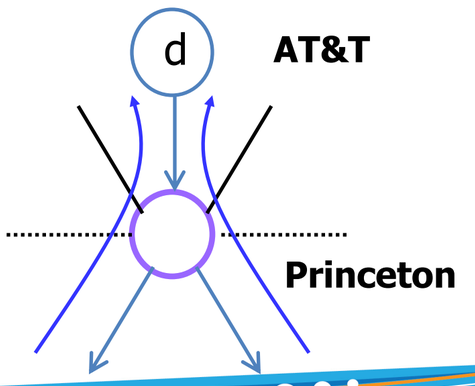
Customer-Provider Relationship

- Customer pays provider for access to Internet
  - Provider exports customer's routes to everybody
  - Customer exports provider's routes to customers

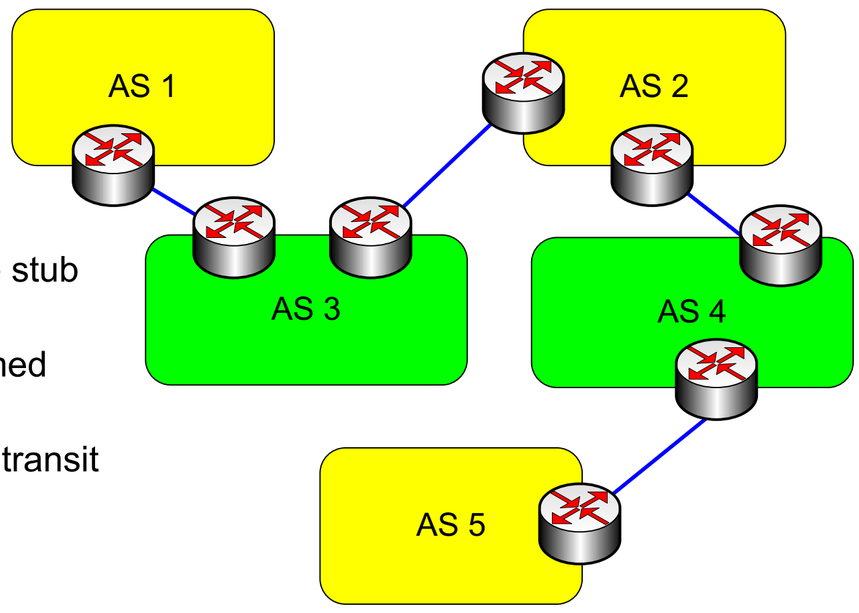
Traffic **to** the customer



Traffic **from** the customer

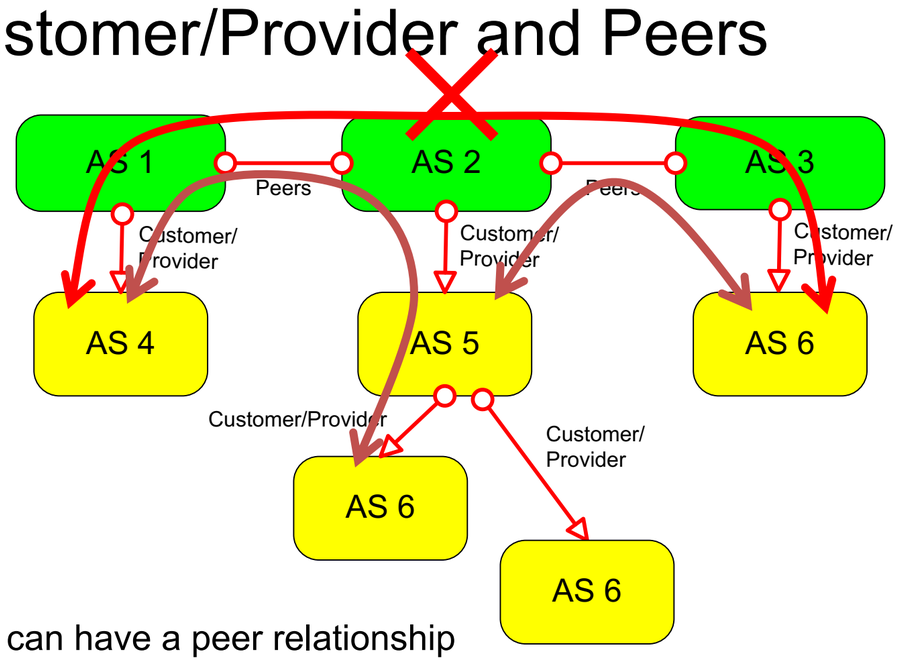


Stub and Transit Networks



- AS 1, and AS 5 are stub networks
- AS 2 is a multi-homed stub network
- AS 3 and AS 4 are transit networks

Customer/Provider and Peers



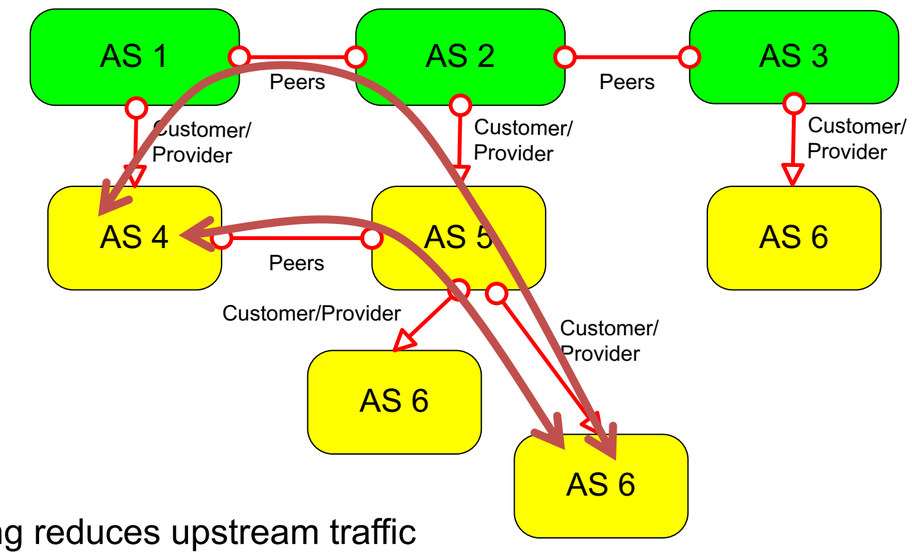
- Transit networks can have a peer relationship
- Peers provide transit between their respective customers
- Peers do not provide transit between peers
- Peers normally do not pay each other for service

Selective Transit

Example:

- Transit AS 3 carries traffic between AS 1 and AS 4 and between AS 2 and AS 4
- But AS 3 does **not** carry traffic between AS 1 and AS 2
- The example shows a routing policy.

Shortcuts through peering

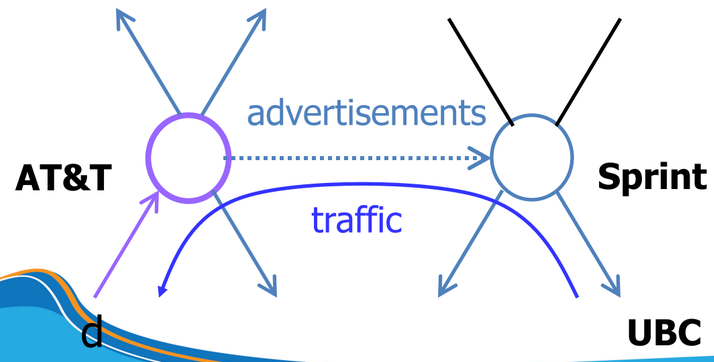


- Note that peering reduces upstream traffic
- Delays can be reduced through peering
- But: Peering may not generate revenue

Peer-Peer Relationship

- Peers exchange traffic between customers
  - AS exports *only* customer routes to a peer
  - AS exports a peer's routes *only* to its customers

Traffic to/from the peer and its customers



How Peering Decisions are Made?

Peer

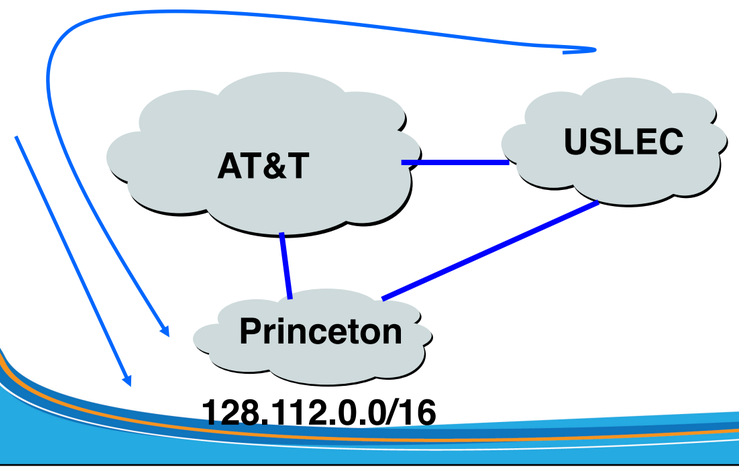
- Reduces upstream transit costs
- Can increase end-to-end performance
- May be the only way to connect your customers to some part of the Internet ("Tier 1")

Don't Peer

- You would rather have customers
- Peers are usually your competition
- Peering relationships may require periodic renegotiation

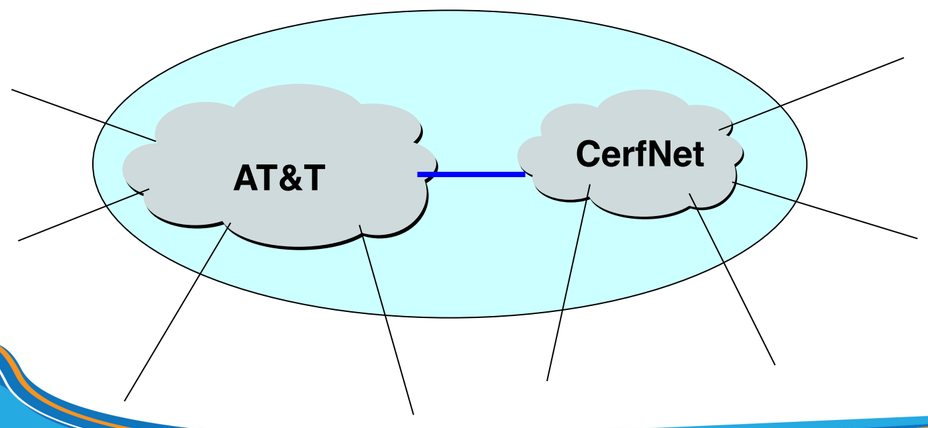
Backup Relationship

- Backup provider
  - Only used if the primary link fails
  - Routes through other paths



Sibling Relationship

- Two ASes owned by the same institution
  - E.g., two ASes that have merged
  - E.g., two ASes simply for scaling reasons
  - Essentially act as a single AS



Border Gateway Protocol

- A Routing Protocol used to exchange routing information between different networks
  - Exterior gateway protocol
- Described in RFC4271
  - RFC4276 gives an implementation report on BGP
  - RFC4277 describes operational experiences using BGP
- The Autonomous System is the cornerstone of BGP
  - It is used to uniquely identify networks with a common routing policy

Border Gateway Protocol

- Path Vector Protocol
- Incremental Updates
- Many options for policy enforcement
- Classless Inter Domain Routing (CIDR)
- Widely used for Internet backbone
- Autonomous systems

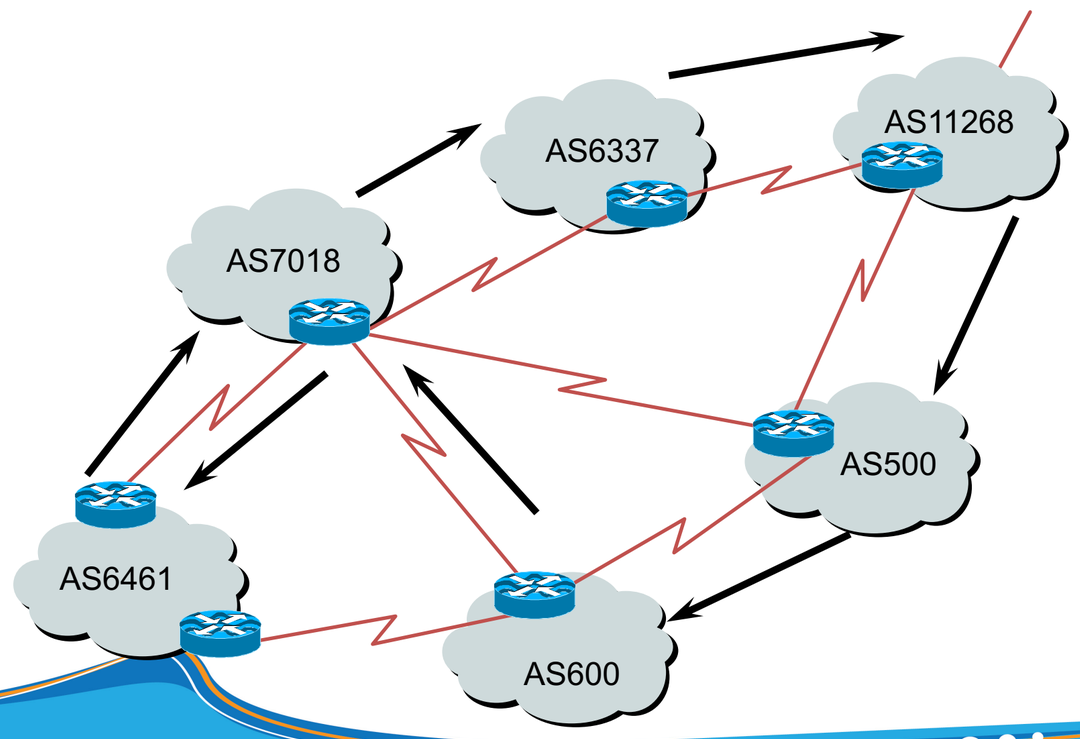
Path Vector Protocol

- BGP is classified as a *path vector* routing protocol (see RFC 1322)
  - A path vector protocol defines a route as a pairing between a destination and the attributes of the path to that destination.

```
12.6.126.0/24 207.126.96.43  1021  0 6461 7018 6337 11268 i
```

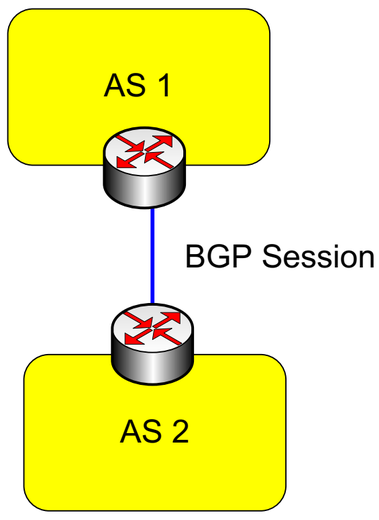
AS Path

Path Vector Protocol



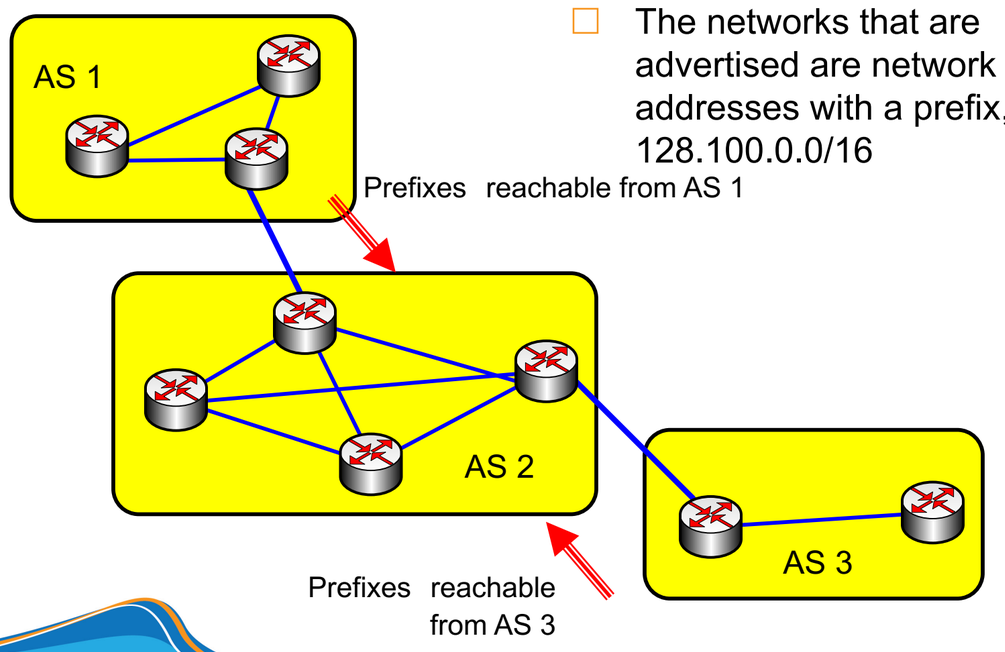
BGP interactions

- Router establishes a TCP connection (TCP port 175)
- Routers exchange BGP routes
- Periodically send updates
- BGP is executed between two routers
  - BGP session
  - BGP peers or BGP speakers
- **Note:** Not all autonomous systems need to run BGP. On many stub networks, the route to the provider can be statically configured



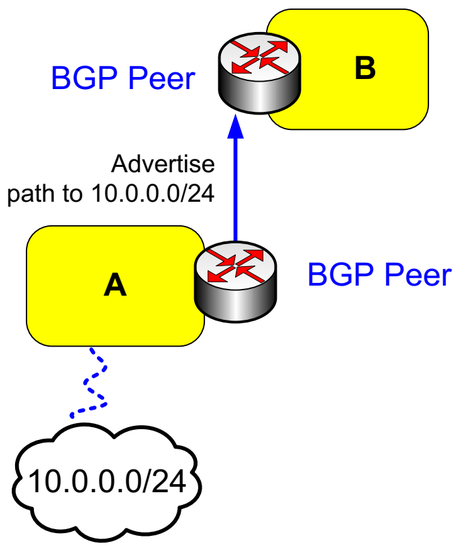
BGP interactions

- The networks that are advertised are network IP addresses with a prefix, E.g., 128.100.0.0/16



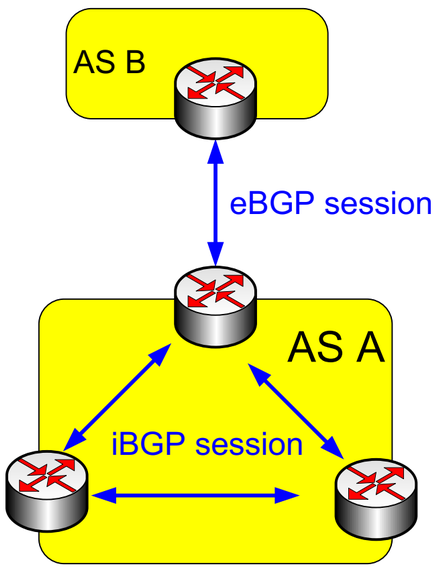
BGP interactions

- BGP peers advertise reachability of IP networks
- A advertises a path to a network (e.g., 10.0.0.0/8) to B only if it is willing to forward traffic going to that network
- **Path-Vector:**
  - A advertises the complete path to the advertised network
  - Path is sent as a list of AS's



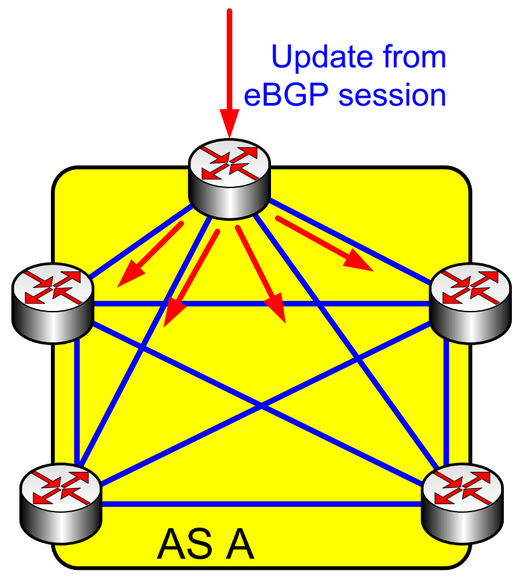
BGP Sessions

- **External BGP session (eBGP):**
  Peers are in different AS'es
- **Internal BGP session (iBGP)**
  Peers are in same AS
- Note that iBGP sessions are going over routes that are set up by an intradomain routing protocol!



iBGP sessions

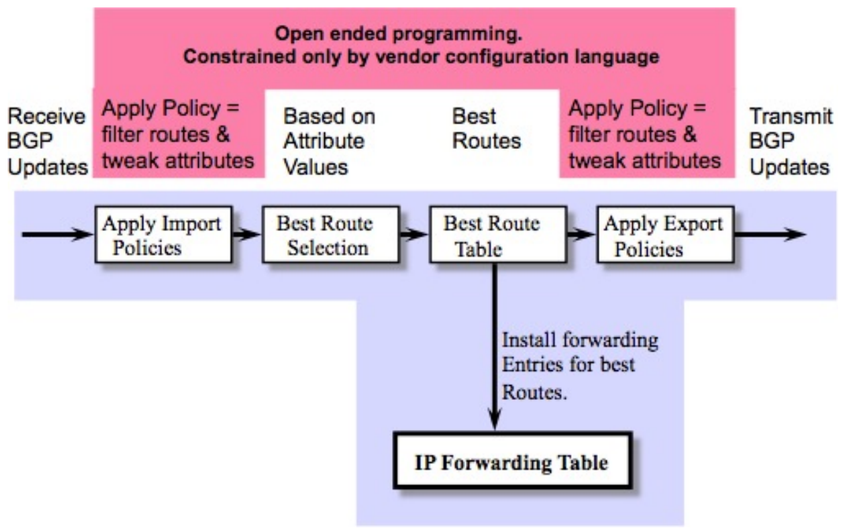
- All iBGP peers in the same autonomous system are fully meshed
- Peer announces routes received via eBGP to iBGP peers
- **But:** iBGP peers do not announce routes received via iBGP to other iBGP peers



BGP Message Types

- **Open**: Establishes a peering session
- **Keep Alive**: Handshake at regular intervals to maintain peering session
- **Notification**: Closes a peering session
- **Update**: Advertises new routes or withdraws previously announced routes. Each announced route is specified as a network prefix with attribute values

BGP General Operation

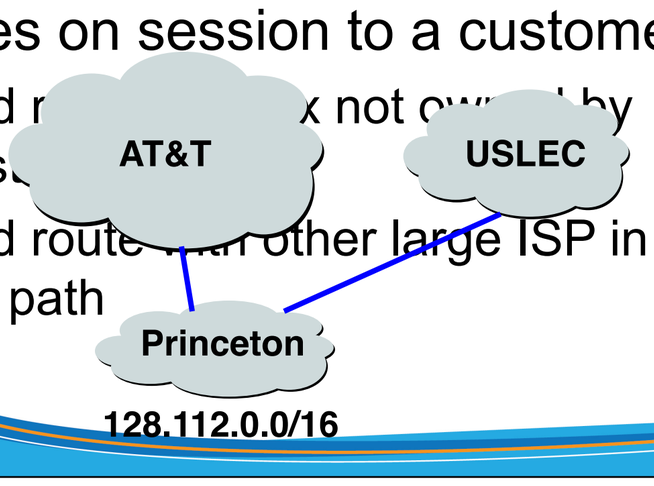


BGP Updates

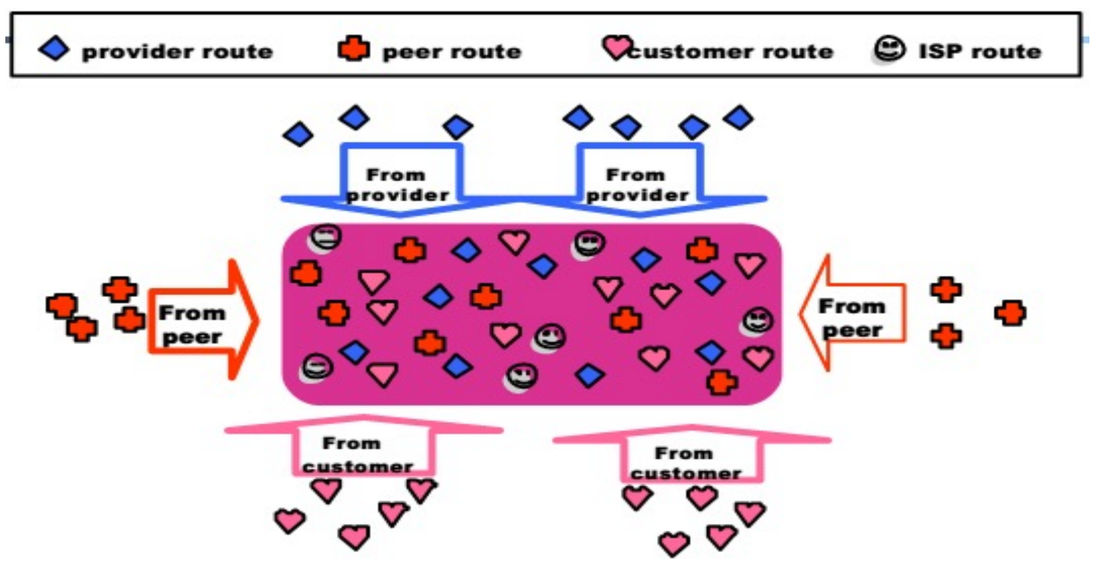
- BGP routers advertise **routes**
- Each route consists of a network prefix and a list of **attributes** that specify information about a route
- Mandatory attributes:
  - **ORIGIN**
  - **AS_PATH**
  - **NEXT_HOP**
- Many other attributes
  - **LOCAL_PREF**
  - **MULTI_EXIT_DISC**
  - **.....**

Import Policy: Filtering

- Discard some route announcements
  - Detect configuration mistakes and attacks
- Examples on session to a customer
  - Discard r x not ow d by the cus
  - Discard route with other large ISP in the AS path

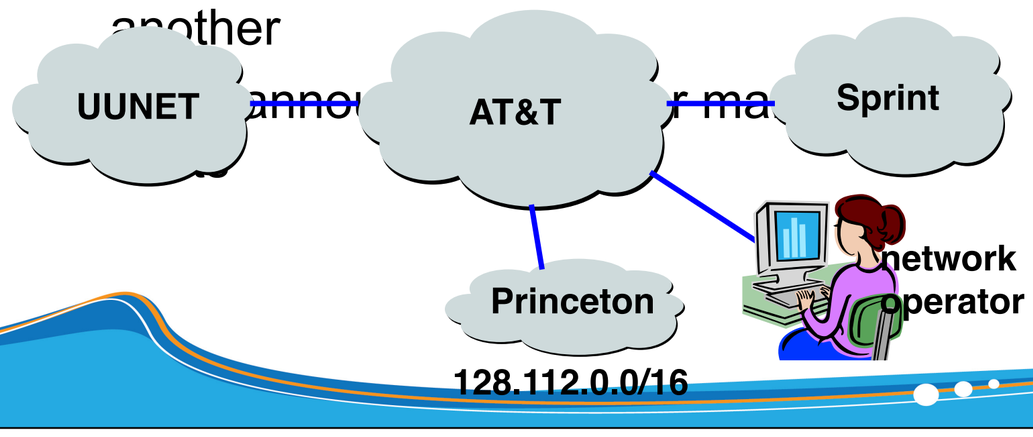


Import Rules

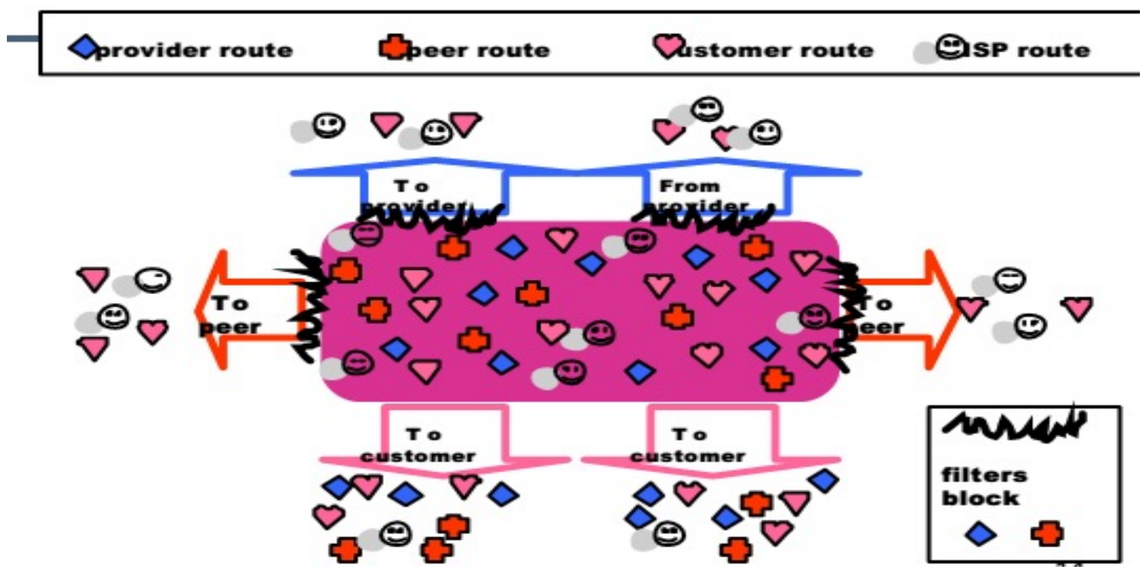


Export Policy: Filtering

- Discard some route announcements
  - Limit propagation of routing information
- Examples
  - Don't announce routes from one peer to another



Export Rules



BGP Attributes

Group Discussion

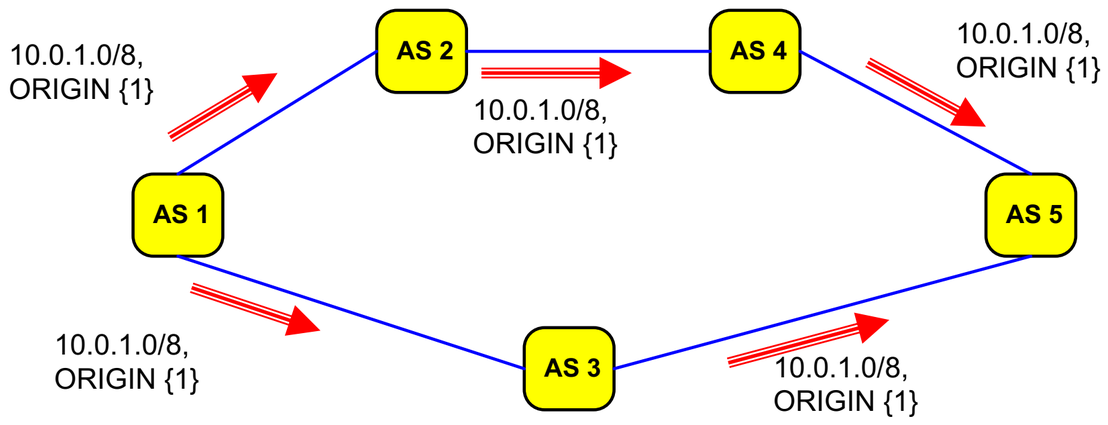
BGP Attributes

There are four basic types of attributes:

- **Well known mandatory attributes**; these attributes must be recognized by all BGP speakers, and must be included in all update messages. Almost all of the attributes impacting the path decision process, described in the next section, are well known mandatory attributes
  - ORIGIN, AS – PATH, NEXT HOP
- **Well known discretionary attributes**; these attributes must be recognized by all BGP speakers, and may be carried in updates, but are not required in every update.
  - LOCAL_PREF
- **Optional transitive attributes**; these attributes may be recognized by some BGP speakers, but not all. They should be preserved and advertised to all peers whether or not they are recognized.
  - COMMUNITIES
- **Optional non-transitive attributes**; these attributes may be recognized by some BGP speakers, but not all. If an update containing an optional transitive attribute is received, the update should be advertised to peers without the unrecognized attributes
  - Multiple Exit Discriminator (MED)

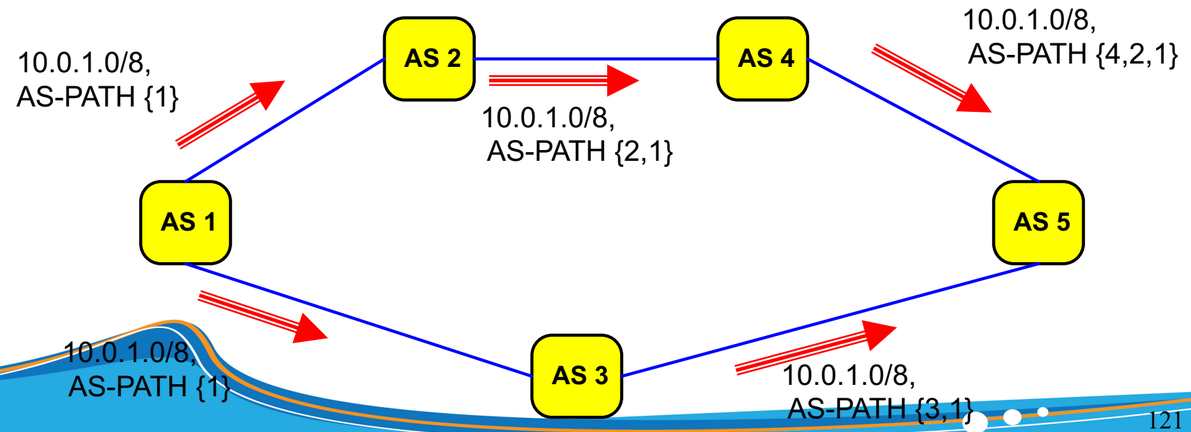
ORIGIN attribute

- Originating domain sends a route with ORIGIN attribute
- Three values: igp, egp, incomplete



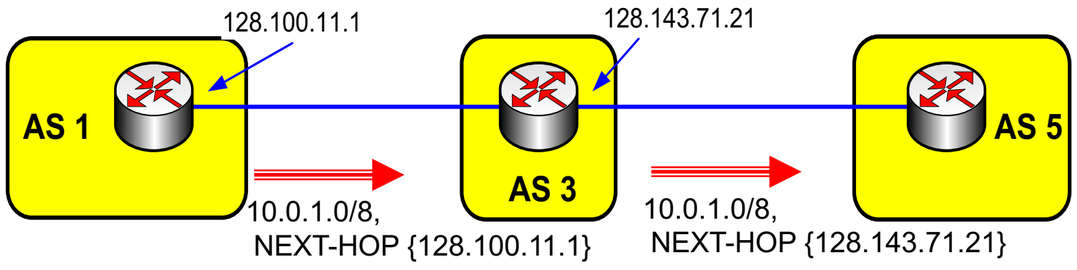
AS-PATH attributes

- Each AS that propagates a route prepends its own AS number
  - AS-PATH collects a path to reach the network prefix
- Path information prevents routing loops from occurring
- Path information also provides information on the length of a path (By default, a shorter route is preferred)



NEXT-HOP attributes

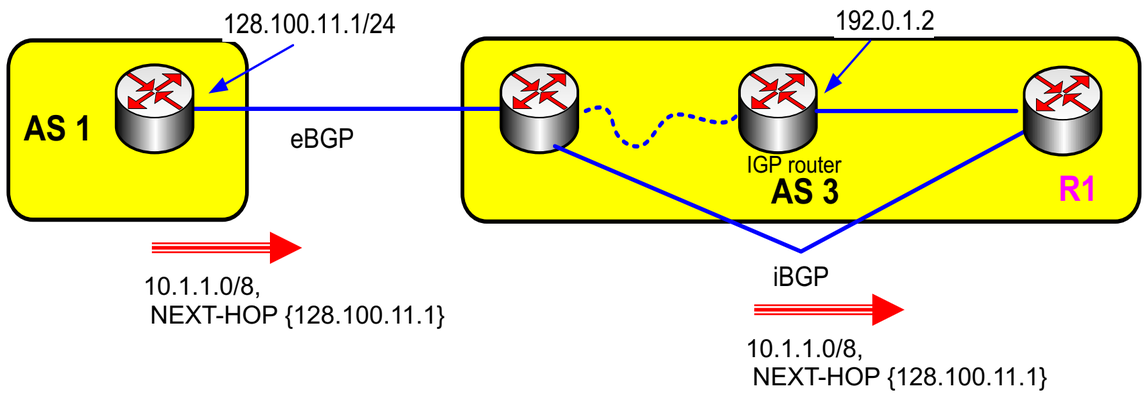
- Each router that sends a route advertisement it includes its own IP address in a NEXT-HOP attribute
- The attribute provides information for the routing table of the receiving router.



BGP Decision Process: Multiple Steps

- Highest local preference
  - *Set by import policies* upon receiving advertisement
- Shortest AS path
  - Included in the route advertisement
- Lowest origin type
  - IGP < EGP < incomplete
- Smallest multiple exit discriminator (MED)
  - Included in the advertisement or reset by import policy
- Smallest internal path cost to the next hop
  - Based on intradomain routing protocol (e.g., OSPF)
- Smallest next-hop router id
  - Final tie-break

Connecting NEXT-HOP with IGP information



At R1:

Routing table

Dest.	Next hop
128.100.11.0/24	192.0.1.2

BGP info

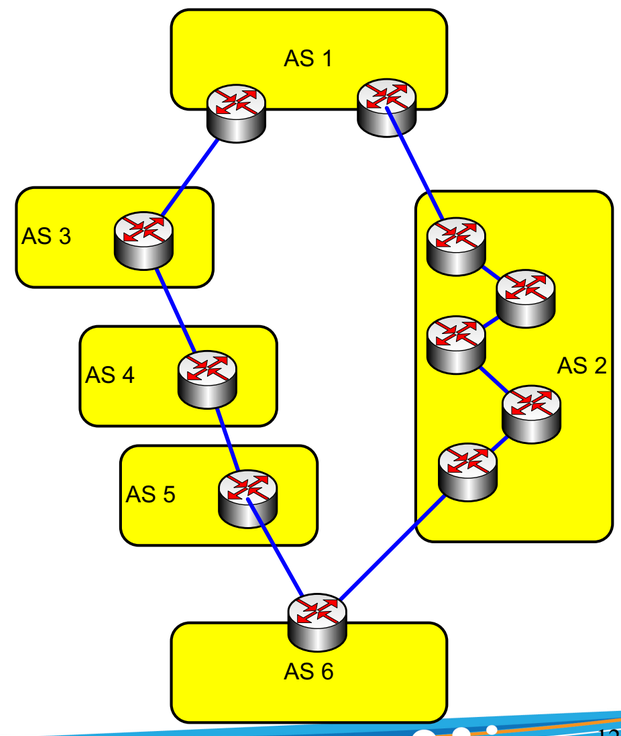
Dest.	Next hop
10.1.1.0/8	128.100.11.1

Routing table

Dest.	Next hop
128.100.11.0/24	192.0.1.2
10.1.1.0/8	192.0.1.2

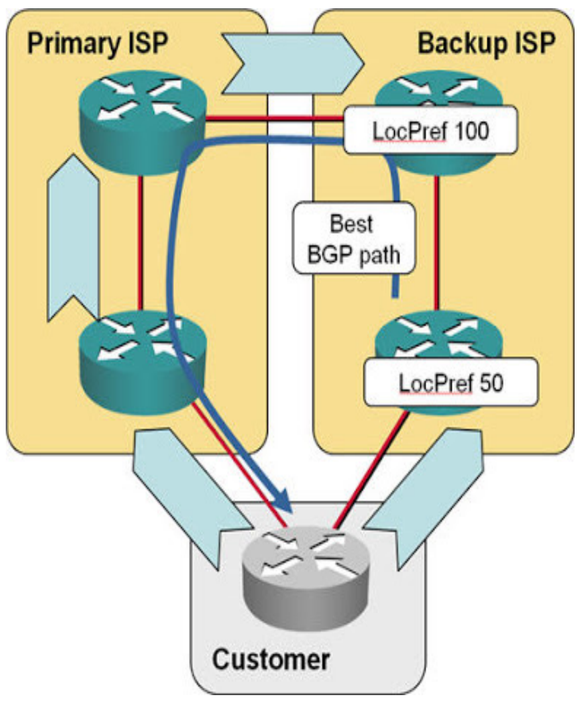
Short AS-PATH does not mean that route is short

- From AS 6's perspective
  - Path {AS2, AS1} is short
  - Path {AS5, AS4, AS3, AS1} is long
- But the number of traversed routers is larger when using the shorter AS-PATH



LOCAL_PREF attribute

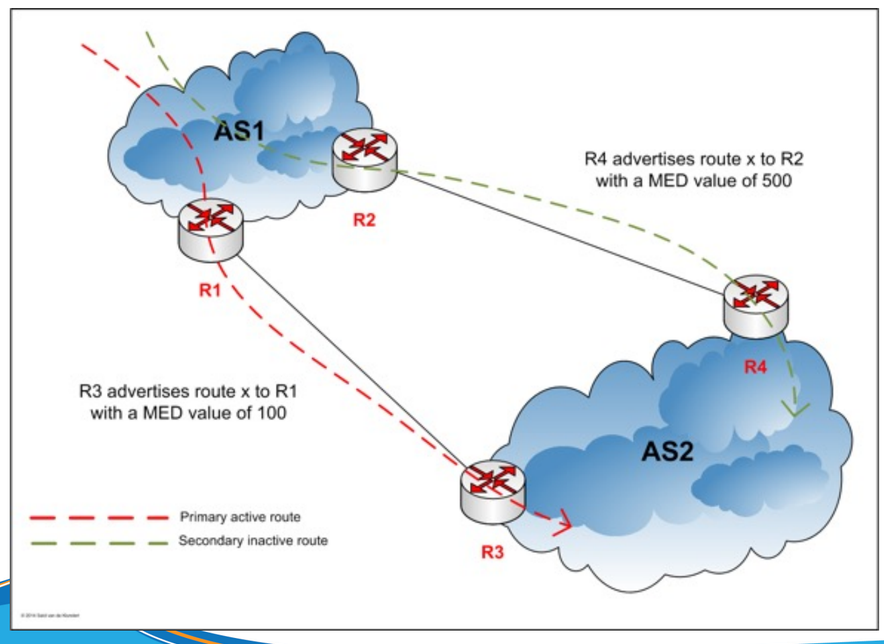
- Local Preference



Thank you

MED Attribute

- Multiple exit discriminator (MED)



Optional Slides

BGP Issues

- BGP is a simple protocol but it is very difficult to configure
- BGP has severe stability issue due to policies → BGP is known to not converge
- As of July 2005, 39,000 AS numbers (of available 64,510) are consumed

How Secure is Today's Internet Routing?

February 2008: Pakistan Telecom hijacks YouTube!

Corrigendum- Most Urgent

GOVERNMENT OF PAKISTAN
PAKISTAN TELECOMMUNICATION AUTHORITY
ZONAL OFFICE PESHAWAR
Plot-11, Sector A-3, Phase-V, Hayatabad, Peshawar.
Ph: 091-9217279- 5829177 Fax: 091-9217254
www.pta.gov.pk

NWFP-33-16 (BW)/06/PTA February ,2008

Subject: **Blocking of Offensive Website**

Reference: This office letter of even number dated 22.02.2008.

I am directed to request all ISPs to immediately block access to the following website

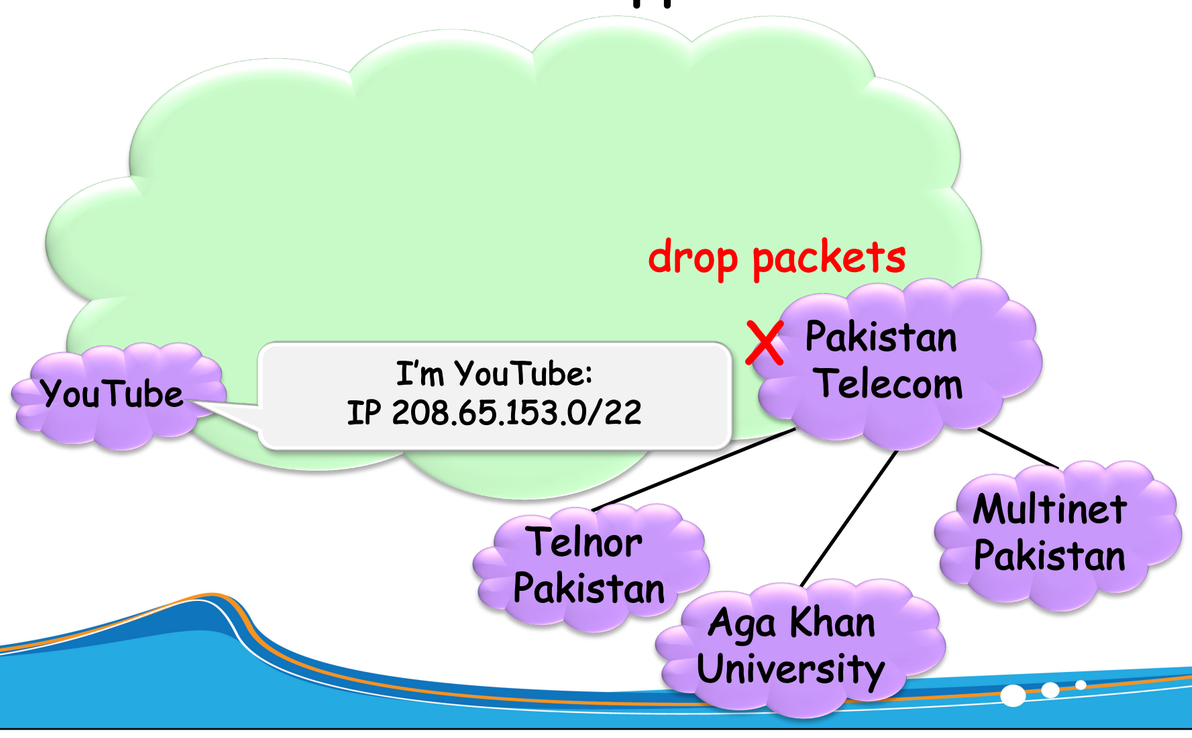
URL: http://www.youtube.com/watch?v=o3s8jtvvg00

IPs: 208.65.153.238, 208.65.153.253, 208.65.153.251

Compliance report should reach this office through return fax or at email peshawar@pta.gov.pk today please.

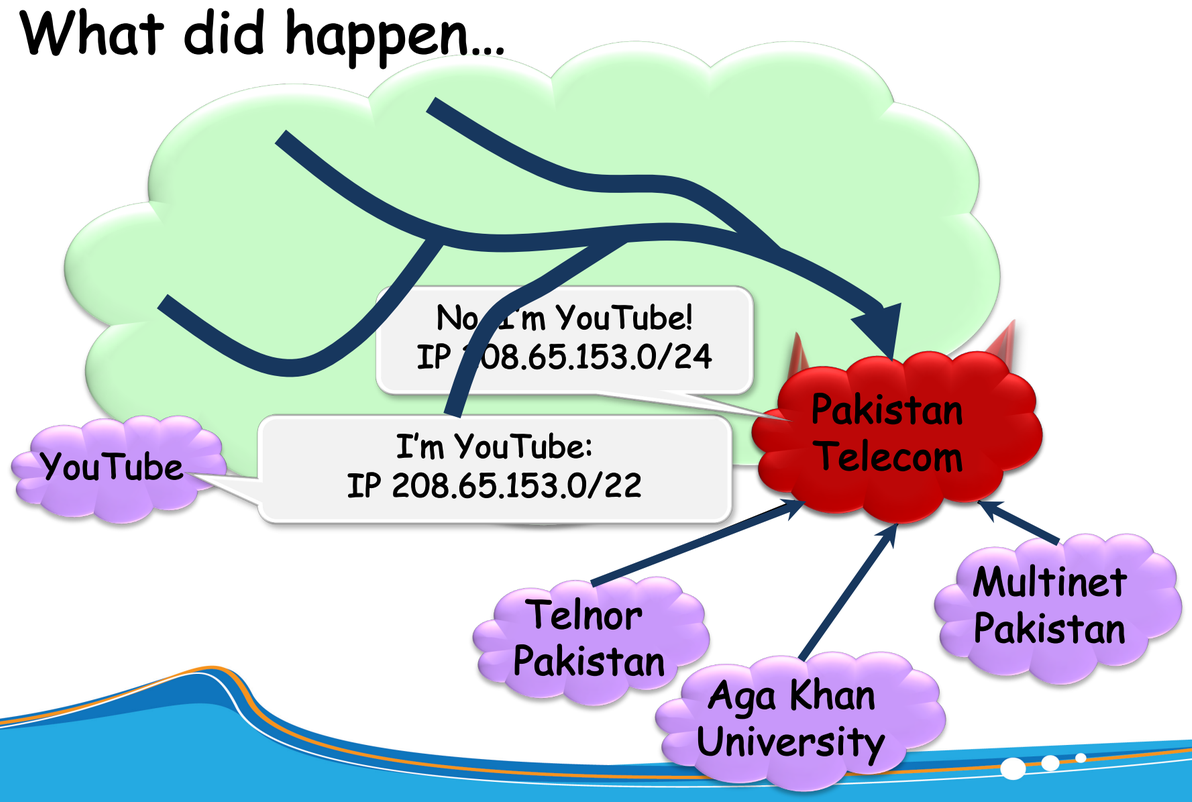
How Secure is Today's Internet Routing?

What should have happened…

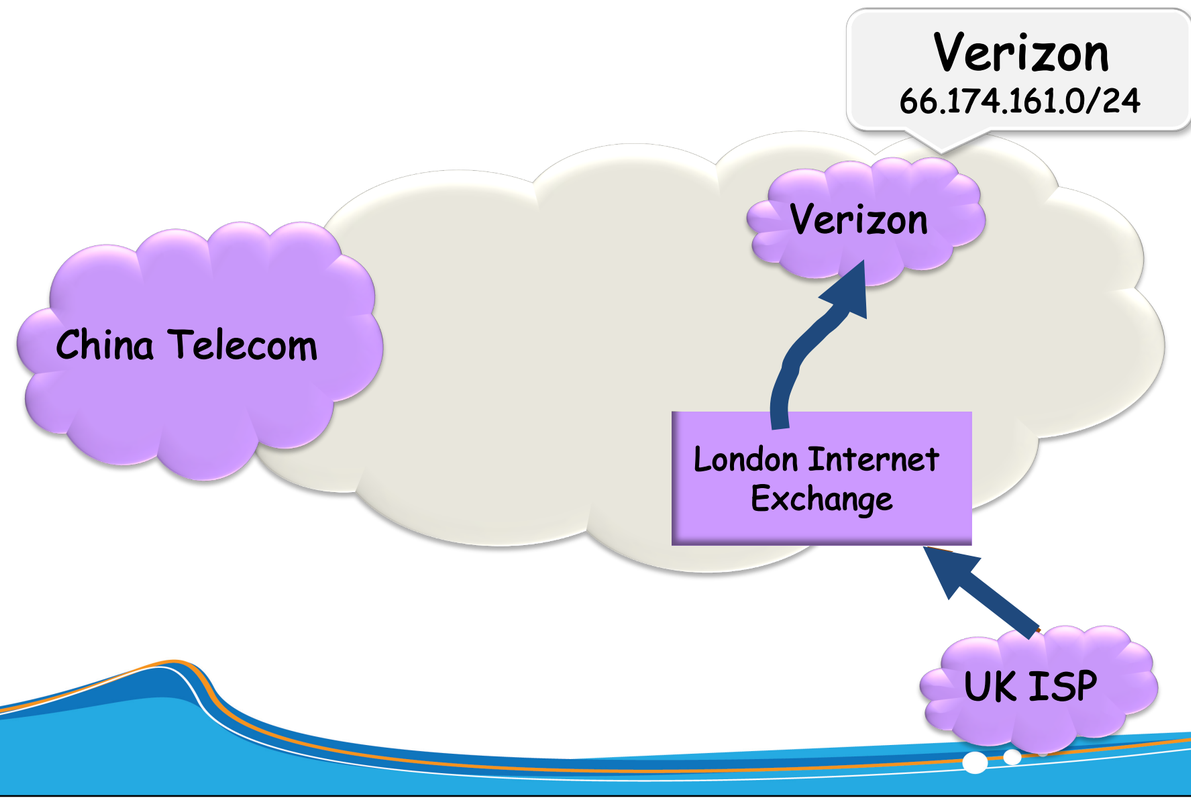


How Secure is Today's Internet Routing?

What did happen…

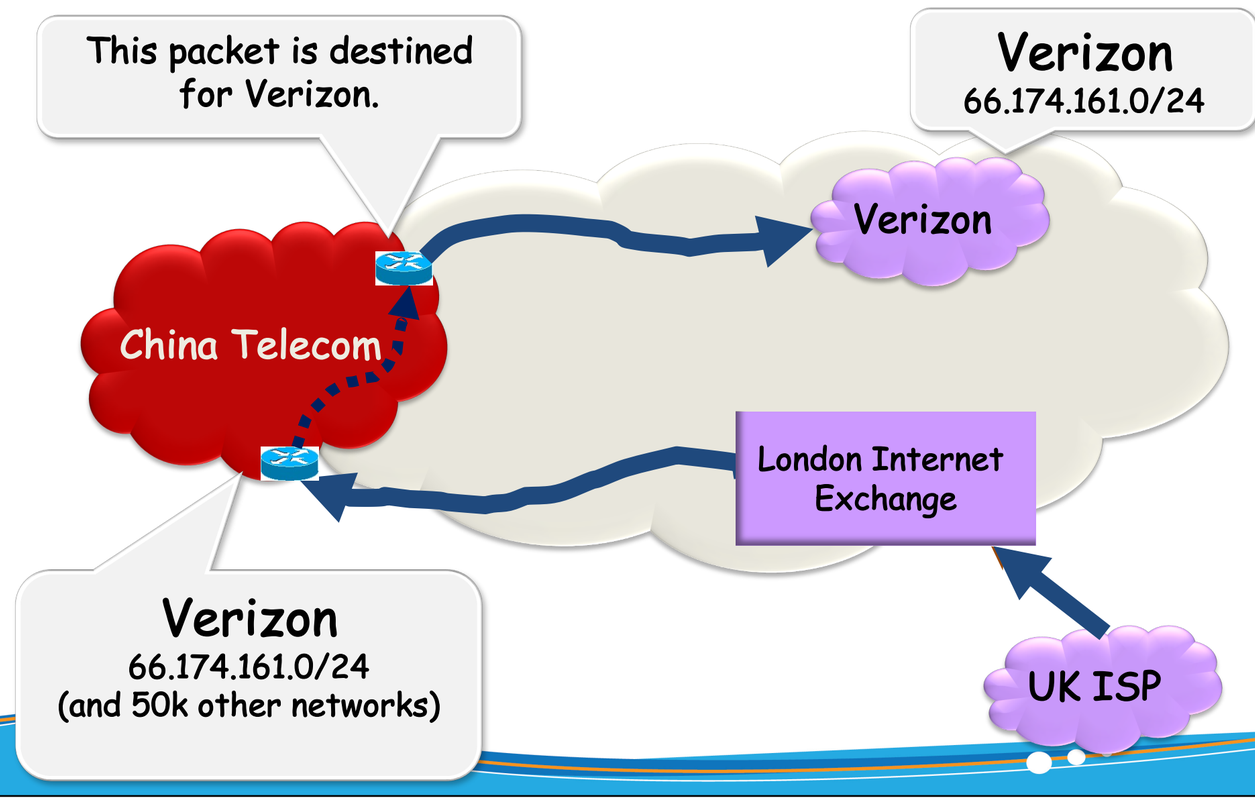


How Secure is Routing on the Internet Today? (2)



How Secure is Routing on the Internet Today? (2)

April 2010 : China Telecom intercepts traffic



BGP Security Today

- **Applying best common practices (BCPs)**
  - Filtering routes by prefix and AS path, *etc.*
- **This is not good enough!**
  - Depends on vigilant application of BCPs … and not making configuration mistakes!
  - Doesn't address fundamental problems, *e.g.*, prefix hijacking!

Securing Internet Routing

- How to secure Internet routing?
  - Long standing agenda in the standards and research communities.
- Over the past 15 years, several secure Internet routing protocols have been proposed.

Securing Internet Routing

- The U.S. federal government is accelerating its efforts to secure the Internet's routing system … The effort … will secure the Internet's core routing protocol known as the **Border Gateway Protocol (BGP)**.

- "**BGP is one of the largest threats on the Internet**. It's incredible, the insecurity of the routing system."

  (Danny McPherson, CSO at Arbor Networks, Jan 2009)

Secure Routing Protocols

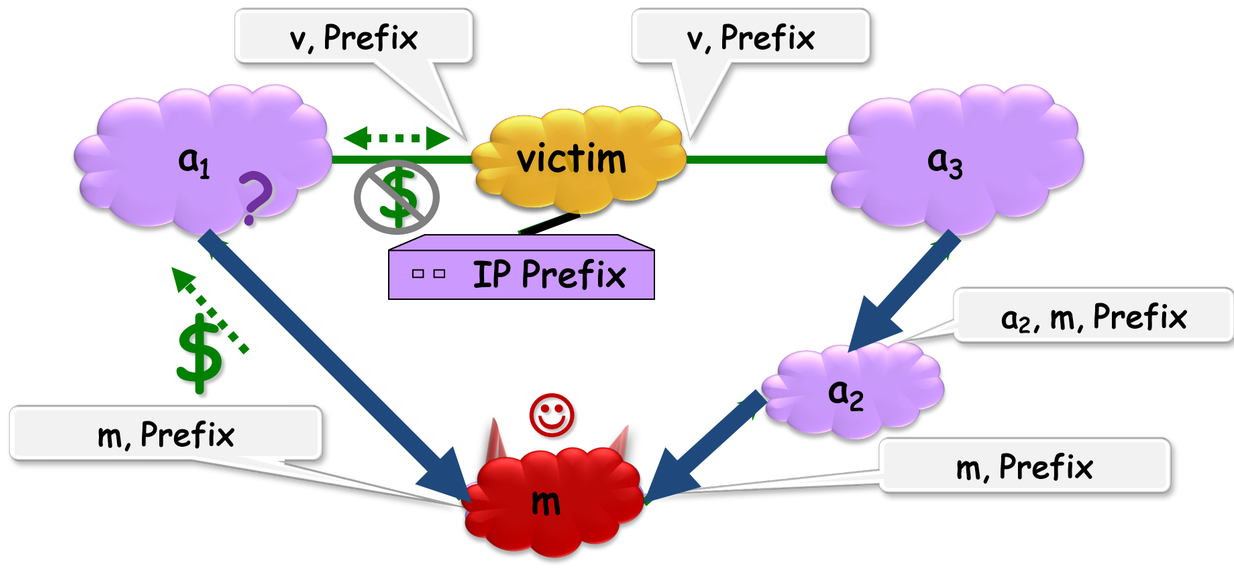
Prefix Hijacking and Origin Authentication

KHOA CÔNG NGHỆ THÔNG TIN
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

IP Address Ownership and Hijacking

- **IP address block assignment**
  - Regional Internet Registries (ARIN, RIPE, APNIC)
  - Internet Service Providers
- **Proper origination of a prefix into BGP**
  - By the AS who owns the prefix
  - … or, by its upstream provider(s) in its behalf
- **However, what's to stop someone else?**
  - Prefix hijacking: another AS originates the prefix
  - BGP does not verify that the AS is authorized
  - Registries of prefix ownership are inaccurate

Prefix Hijacking



Hijacking is Hard to Debug

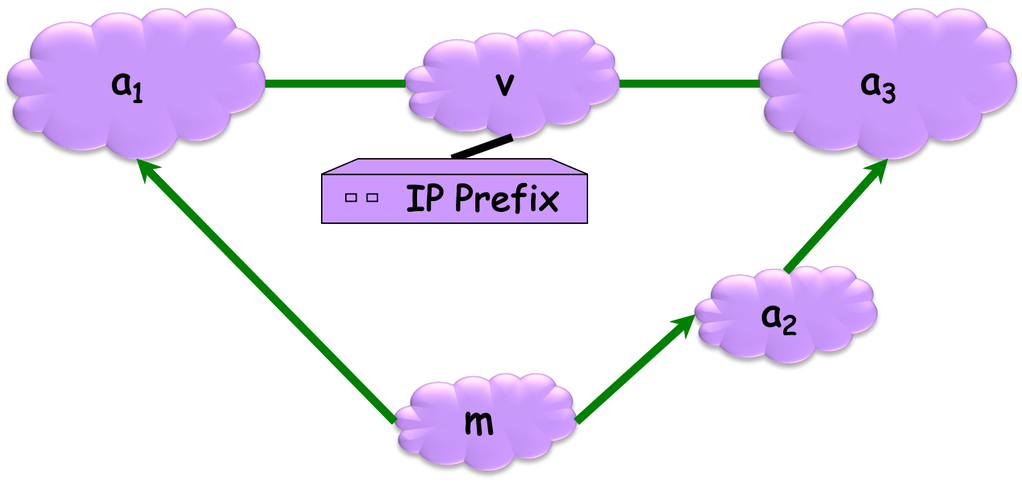
- **The victim AS doesn't see the problem**
  - Picks its own route
  - Might not even learn the bogus route
- **May not cause loss of connectivity**
  - *E.g.*, if the bogus AS snoops and redirects
  - … may only cause performance degradation
- **Or, loss of connectivity is isolated**
  - E.g., only for sources in parts of the Internet
- **Diagnosing prefix hijacking**
  - Analyzing updates from many vantage points
  - Launching traceroute from many vantage points

How to Hijack a Prefix

- **The hijacking AS has**
  - Router with BGP session(s)
  - Configured to originate the prefix
- **Getting access to the router**
  - Network operator makes configuration mistake
  - Disgruntled operator launches an attack
  - Outsider breaks in to the router and reconfigures
- **Getting other ASes to believe bogus route**
  - Neighbor ASes do not discard the bogus route
  - E.g., not doing protective filtering

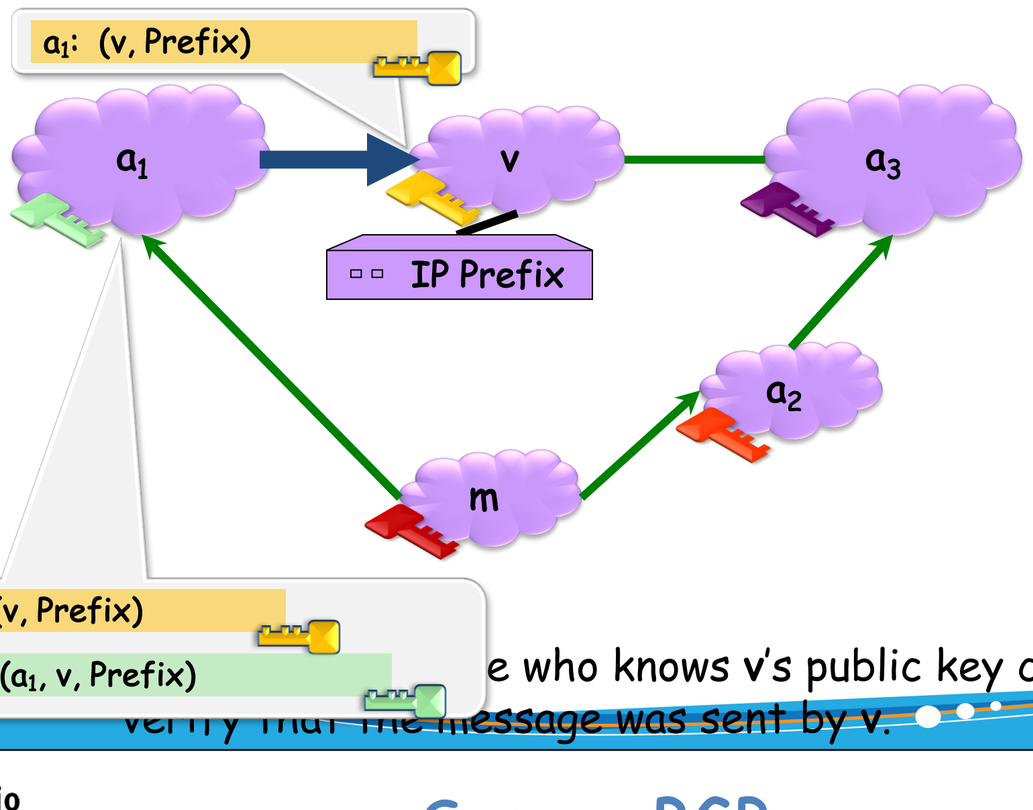
Origin Authentication

A secure database maps IP prefixes to owner ASes.



Secure BGP

Origin Authentication + cryptographic signatures



e who knows v's public key can verify that the message was sent by v.

Secure BGP

- S-BGP can validate the order in which ASes were traversed.
- S-BGP can validate that no intermediate ASes were added or removed.
- S-BGP can validate that the route is recent.

Incremental Deployment?

- There is a necessary transition period.
- S-BGP must be backwards compatible with BGP
- Who upgrades first? Why?

Pessimistic View

*ISPs would be the ones forced to **upgrade all of their equipment** to support this initiative, but **how would it benefit them**? As commercial companies, if there is little to no benefit (potential to increase profit), why would they implement a potentially costly solution? The answer is **they won't**.*

[http://www.omninerd.com/articles/Did_China_Hijack_15_of_the_Internet_Routers_BGP_and_Ignorance]

unless everyone else does?

S-BGP = IPV6?

Are We There Yet?

KHOA CÔNG NGHỆ THÔNG TIN
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

Conclusions

- **Internet protocols designed based on trust**
  - The insiders are good guys
  - All bad guys are outside the network
- **Border Gateway Protocol is very vulnerable**
  - Glue that holds the Internet together
  - Hard for an AS to locally identify bogus routes
  - Attacks can have very serious global consequences
- **Proposed solutions/approaches**
  - Secure variants of the Border Gateway Protocol

S-BGP Deployment Challenges

- **Complete, accurate registries**
  - E.g., of prefix ownership
- **Public Key Infrastructure**
  - To know the public key for any given AS
- **Cryptographic operations**
  - *E.g.*, digital signatures on BGP messages
- **Need to perform operations quickly**
  - To avoid delaying response to routing changes
- **Difficulty of incremental deployment**
  - Hard to have a "flag day" to deploy S-BGP

One last thing…

Harming Internet Routing Without Attacking BGP

KHOA CÔNG NGHỆ THÔNG TIN
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

Attacks on TCP and Data-Plane Attacks

- **Attack TCP!**
  - A BGP session runs over TCP.
- **Do not forward traffic as advertised!**
  - Drop packets!
  - Route packets along unannounced routes!