Routing Fundamental

PacNOG20 July 3, 2017. Suva, Fiji.

Jessica Wei

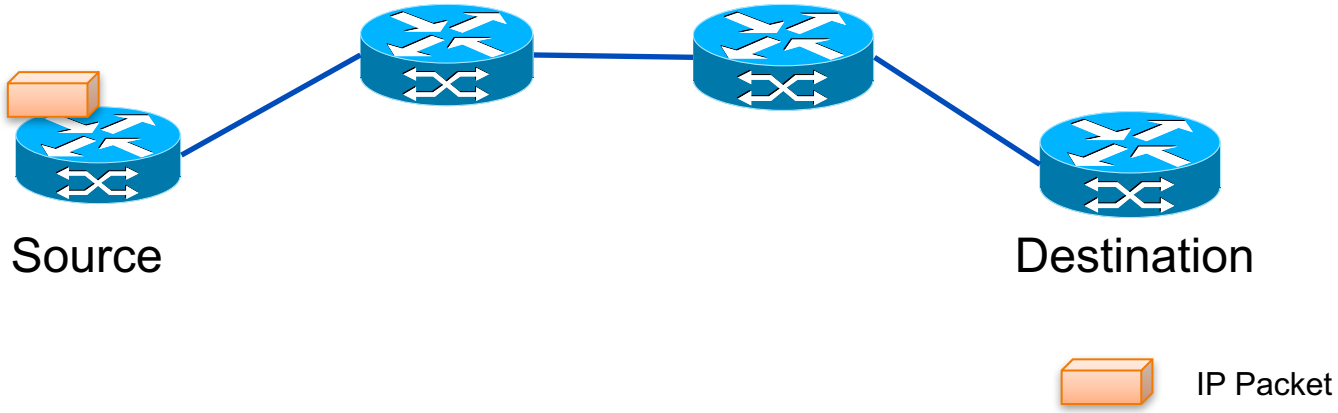
Issue Date: [Date[

Revision: [xx]

APNIC

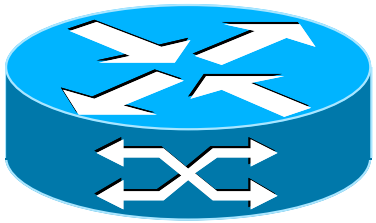
What is IP Routing

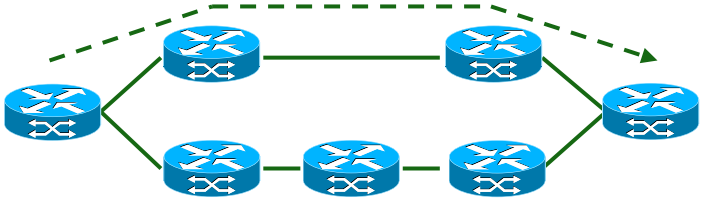
- IP Routing is to guide IP packets from source to destination.
- A route is a path along which packets are sent from the source to the destination.



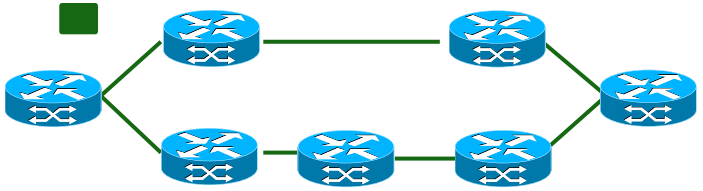
Routers

- Two key roles:



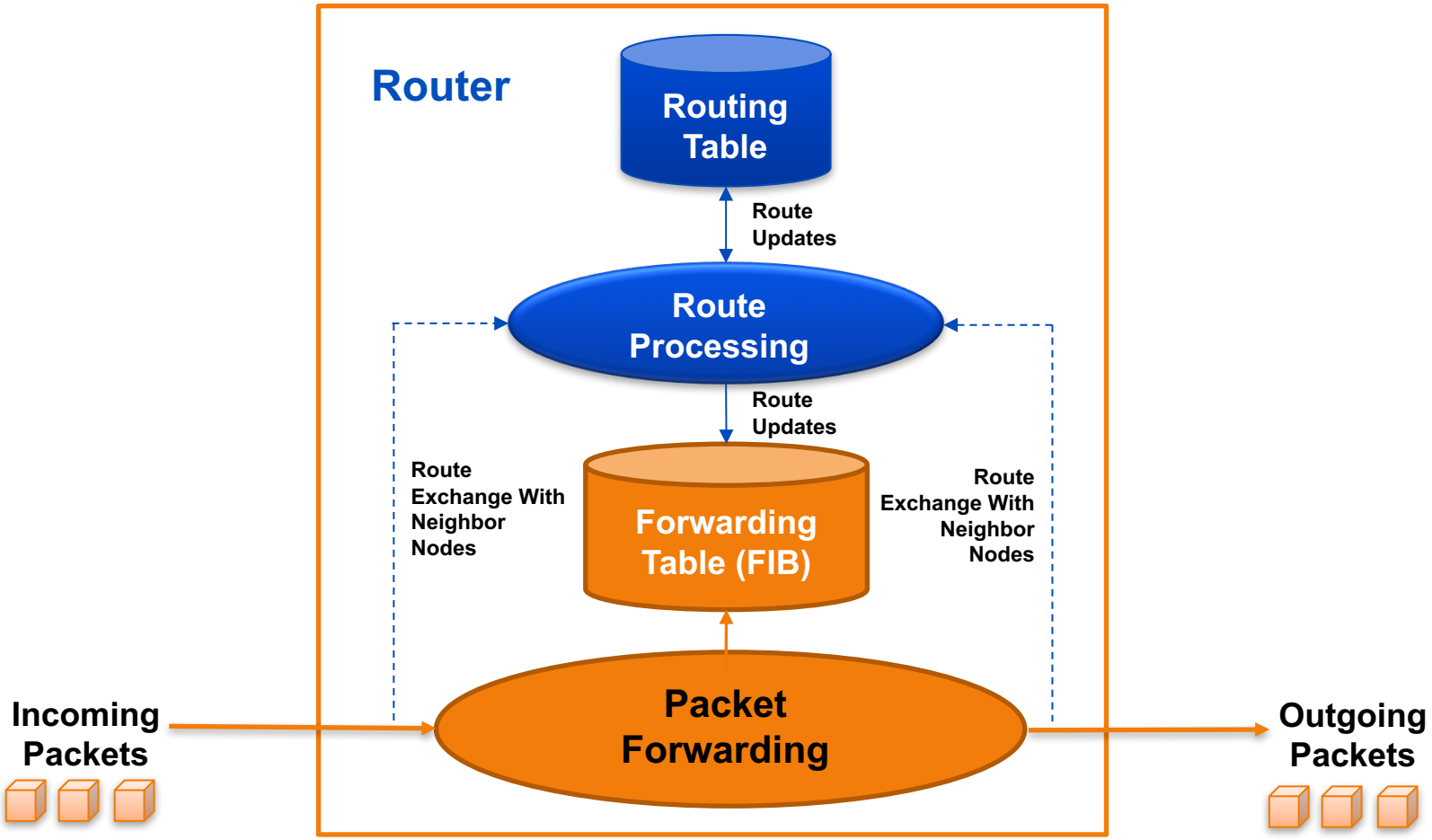


Determining network paths



Packet forwarding

Router Architecture

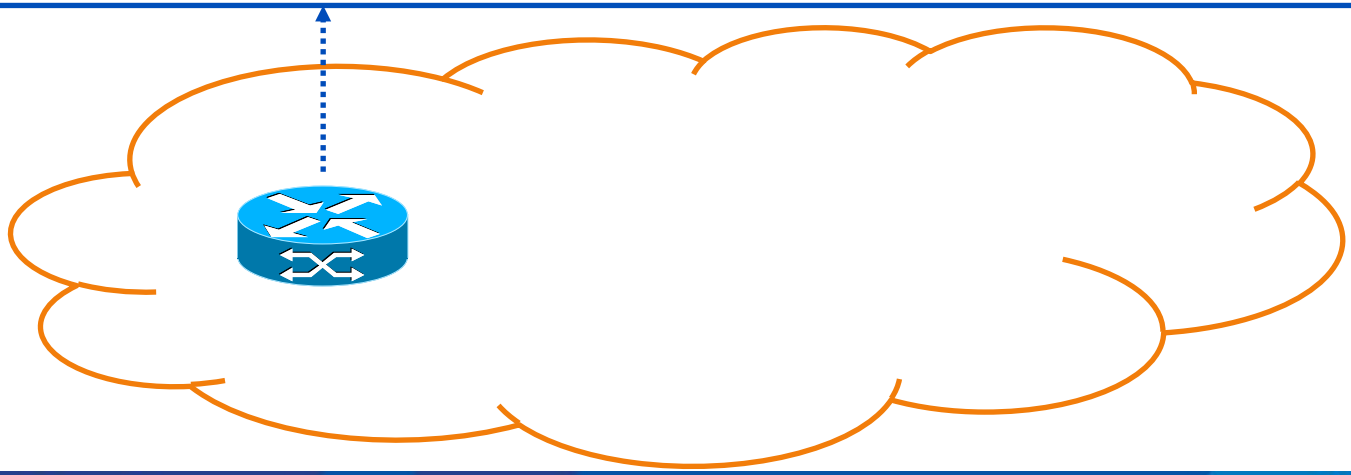


Source: Deepankar Medhi & Karthikeyan Ramasamy "Network Routing: Algorithms, Protocols, and Architectures"

What is Routing Table

Cisco IOS

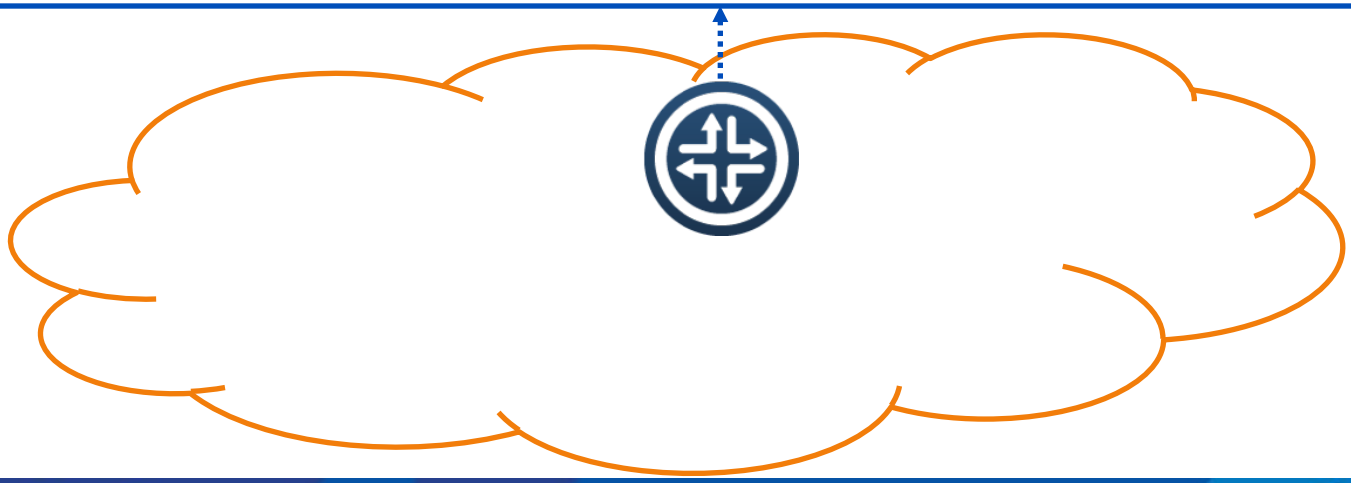
```
Router#show ip route
Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area ……
……
 172.16.0.0/16 is variably subnetted, 18 subnets, 3 masks
C        172.16.10.0/30 is directly connected, GigabitEthernet0/0/0
L        172.16.10.2/32 is directly connected, GigabitEthernet0/0/0
O        172.16.10.24/30 [110/5] via 172.16.10.1, 7w0d, GigabitEthernet0/0/0
O        172.16.10.28/30 [110/7] via 172.16.10.1, 7w0d, GigabitEthernet0/0/0
O        172.16.10.32/30 [110/6] via 172.16.10.1, 7w0d, GigabitEthernet0/0/0
S        172.16.11.0/27 is directly connected, Null0
O        172.16.11.64/30 [110/6] via 172.16.10.1, 7w0d, GigabitEthernet0/0/0
……
```



What is Routing Table

Juniper
JUNOS

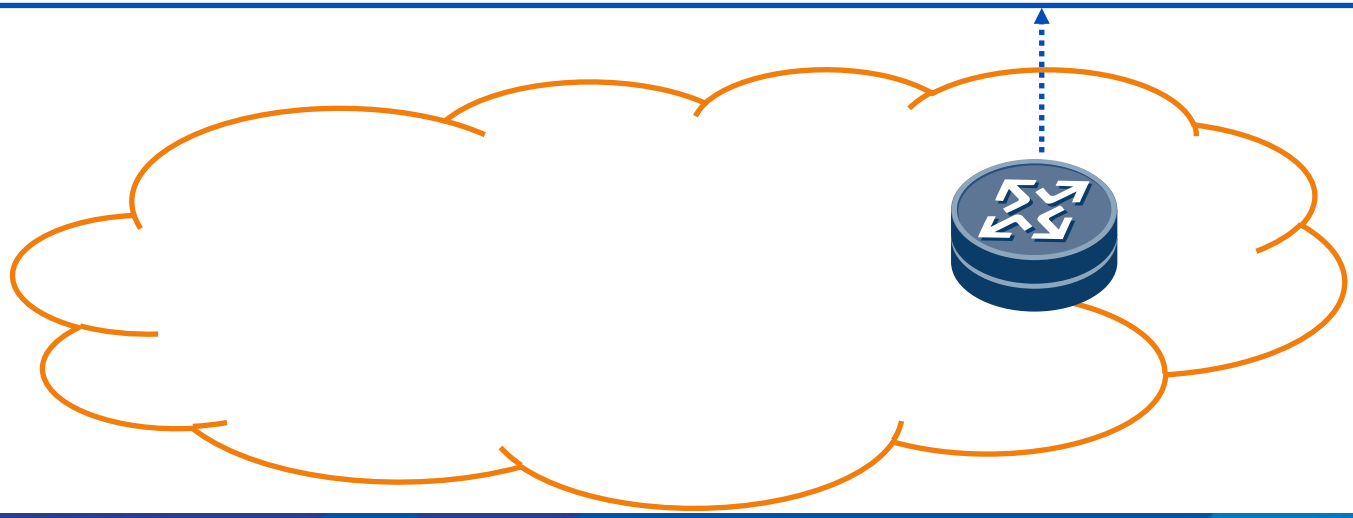
```
root@RT> show route
inet.0: 5 destinations, 5 routes (5 active, 0 holddown, 0 hidden)
+ = Active Route, - = Last Active, * = Both
10.0.1.0/30        *[Direct/0] 00:44:34
                    > via em1.0
10.0.1.1/32        *[Local/0] 00:44:34
                       Local via em1.0
10.0.2.0/30        *[BGP/170] 00:04:23, localpref 100
                       AS path: 100 I
                    > to 10.0.1.2 via em1.0
200.1.1.0/24       *[BGP/170] 00:04:24, localpref 100
                       AS path: 100 65002 I
                    > to 10.0.1.2 via em1.0
......
```



What is Routing Table

Huawei VRP

```
<Router>display ip routing-table
Route Flags: R - relay, D - download to fib
------------------------------------------------------------------------------
Routing Tables: Public
         Destinations : 43        Routes : 55
Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface
      127.0.0.0/8   Direct  0    0           D   127.0.0.1       InLoopBack0
      127.0.0.1/32  Direct  0    0           D   127.0.0.1       InLoopBack0
127.255.255.255/32  Direct  0    0           D   127.0.0.1       InLoopBack0
     172.16.0.0/24  IBGP    255  0          RD   172.16.15.2     GigabitEthernet4/0/1
                    IBGP    255  0          RD   172.16.15.2     GigabitEthernet4/0/0
     172.16.4.0/24  Static  60   0           D   0.0.0.0         NULL0
……
```



Destination

Cisco IOS

```
Router#show ip route

Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area ……
……
 172.16.0.0/16 is variably subnetted, 18 subnets, 3 masks
C        172.16.10.0/30 is directly connected, GigabitEthernet0/0/0
L        172.16.10.2/32 is directly connected, GigabitEthernet0/0/0
O        172.16.10.24/30 [110/5] via 172.16.10.1, 7w0d, GigabitEthernet0/0/0
O        172.16.10.28/30 [110/7] via 172.16.10.1, 7w0d, GigabitEthernet0/0/0
O        172.16.10.32/30 [110/6] via 172.16.10.1, 7w0d, GigabitEthernet0/0/0
S        172.16.11.0/27 is directly connected, Null0
O        172.16.11.64/30 [110/6] via 172.16.10.1, 7w0d, GigabitEthernet0/0/0
……
```

- Destination identifies the destination IP address or the destination network address of an IP packet.

Next-Hop

Cisco IOS

```
Router#show ip route

Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area ……
……
 172.16.0.0/16 is variably subnetted, 18 subnets, 3 masks
C        172.16.10.0/30 is directly connected, GigabitEthernet0/0/0
L        172.16.10.2/32 is directly connected, GigabitEthernet0/0/0
O        172.16.10.24/30 [110/5] via 172.16.10.1, 7w0d, GigabitEthernet0/0/0
O        172.16.10.28/30 [110/7] via 172.16.10.1, 7w0d, GigabitEthernet0/0/0
O        172.16.10.32/30 [110/6] via 172.16.10.1, 7w0d, GigabitEthernet0/0/0
S        172.16.11.0/27 is directly connected, Null0
O        172.16.11.64/30 [110/6] via 172.16.10.1, 7w0d, GigabitEthernet0/0/0
……
```

- Next-hop indicates the IP address of the next router that an IP packet passes through.

Outgoing Interface

Cisco IOS

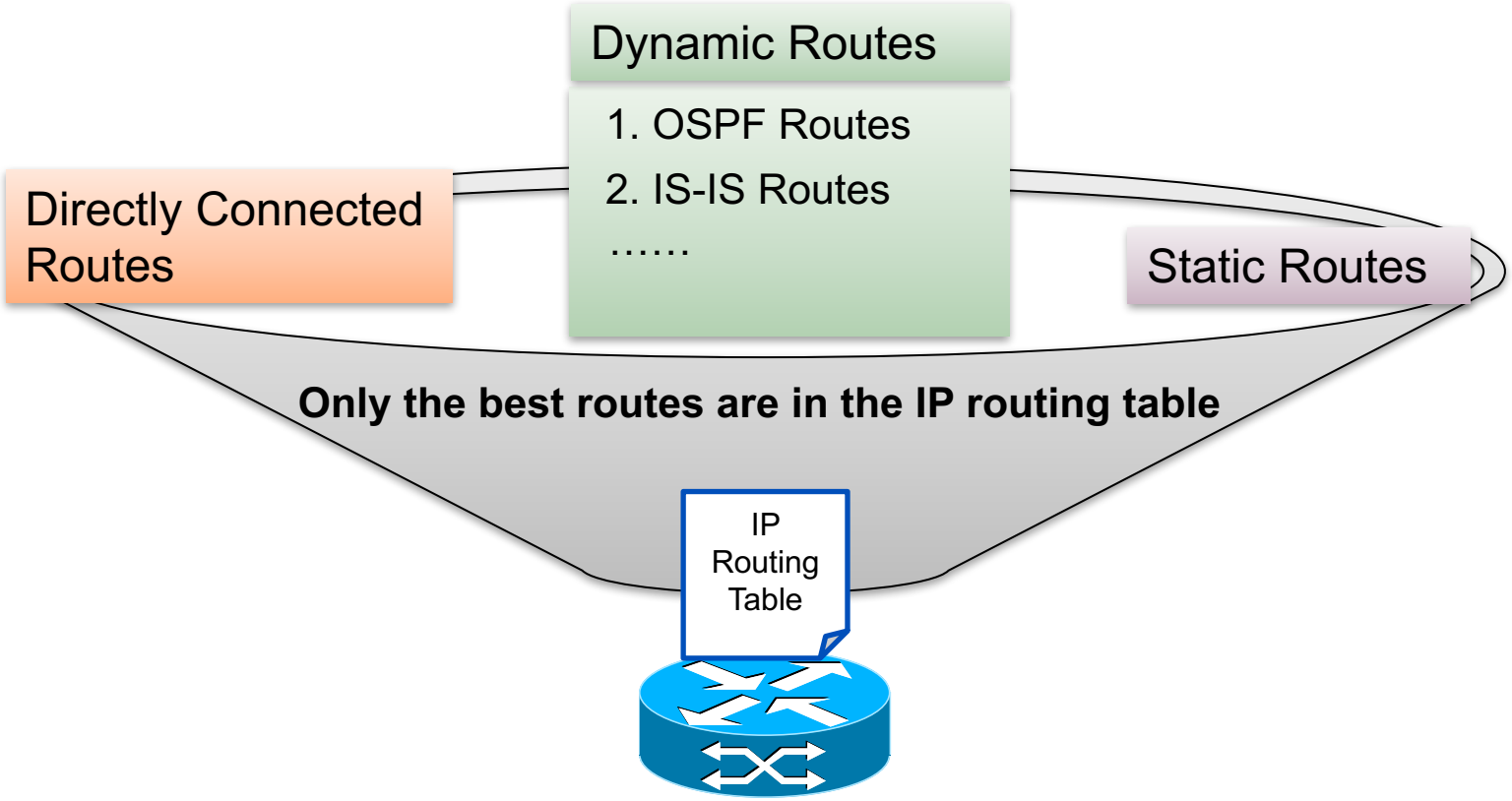
```
Router#show ip route

Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area ……
……
 172.16.0.0/16 is variably subnetted, 18 subnets, 3 masks
C        172.16.10.0/30 is directly connected, GigabitEthernet0/0/0
L        172.16.10.2/32 is directly connected, GigabitEthernet0/0/0
O        172.16.10.24/30 [110/5] via 172.16.10.1, 7w0d, GigabitEthernet0/0/0
O        172.16.10.28/30 [110/7] via 172.16.10.1, 7w0d, GigabitEthernet0/0/0
O        172.16.10.32/30 [110/6] via 172.16.10.1, 7w0d, GigabitEthernet0/0/0
S        172.16.11.0/27 is directly connected, Null0
O        172.16.11.64/30 [110/6] via 172.16.10.1, 7w0d, GigabitEthernet0/0/0
……
```

- Outgoing interface is the interface from which the packet will be sent out.

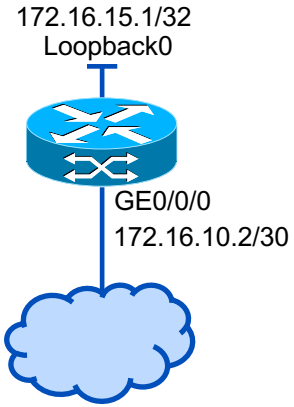
Populate a Routing Table

- Routes in the routing table are from ?



Directly Connected Route

- Direct route is discovered by link layer protocols

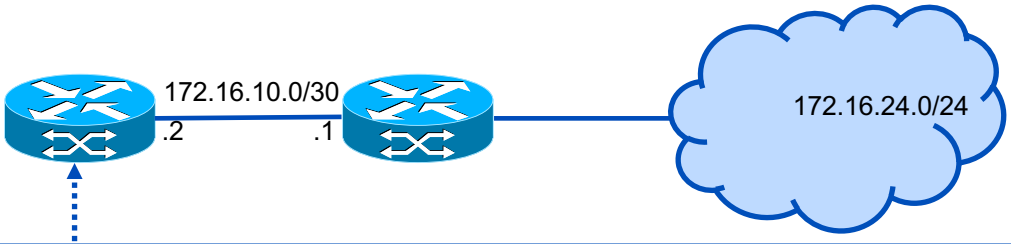



```
Router# show ip route
Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area ……
……
C        172.16.10.0/30 is directly connected, GigabitEthernet0/0/0
L        172.16.10.2/32 is directly connected, GigabitEthernet0/0/0
C        172.16.15.1/32 is directly connected, Loopback0
……
```

Cisco IOS

Static Route

- Static routes are configured manually by network engineers.
- Easy to configure, but cannot automatically adapt to network topology changes, require subsequent maintenance.




```
Router (config)#ip route 172.16.24.0 255.255.255.0 172.16.10.1
```

- There is a special static route --- default route

```
Router (config)#ip route 0.0.0.0 0.0.0.0 172.16.10.1
```

RIB and FIB

RIB

```
Router# show ip route                                                    Cisco IOS
Codes: L - local, C - connected, S - static, R - RIP, M - mobile, B - BGP
D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area ……
……
L        172.16.10.2/32 is directly connected, GigabitEthernet0/0/0
O        172.16.10.24/30 [110/5] via 172.16.10.1, 7w0d, GigabitEthernet0/0/0
S        172.16.11.0/27 is directly connected, Null0
O        172.16.11.64/30 [110/6] via 172.16.10.1, 7w0d, GigabitEthernet0/0/0
……
```

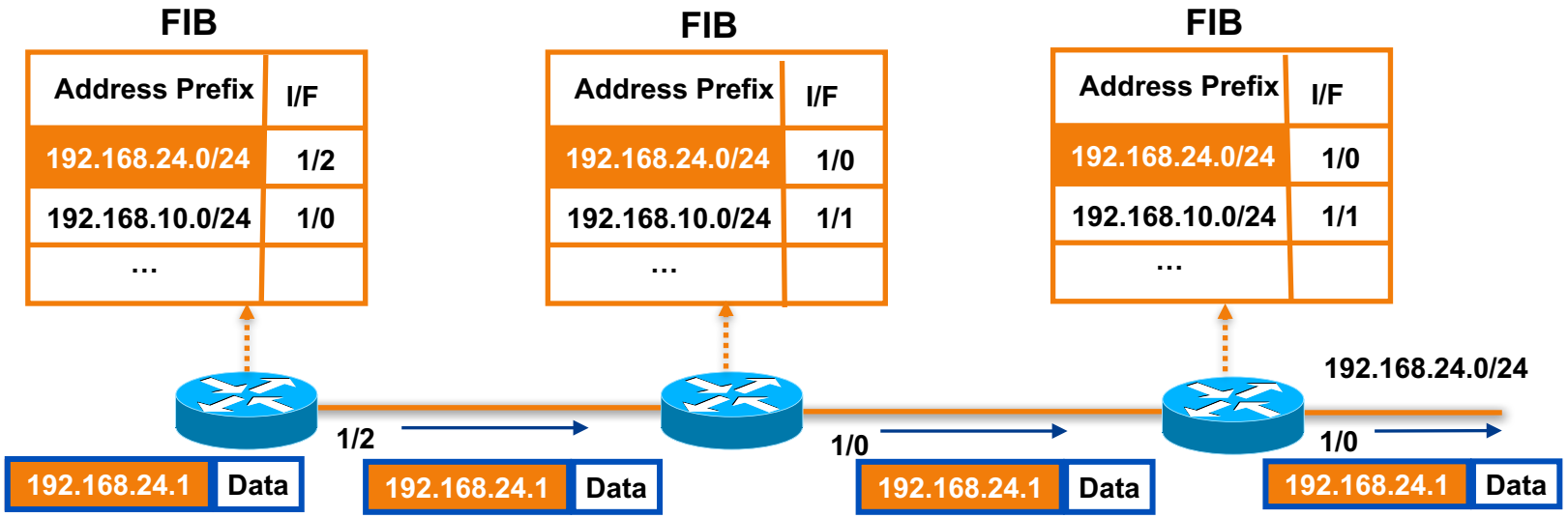
FIB

```
Router# show ip cef                                                      Cisco IOS
Prefix               Next Hop            Interface
……
172.16.10.2/32       receive             GigabitEthernet0/0/0
172.16.10.24/30      172.16.10.1         GigabitEthernet0/0/0
172.16.11.0/27       attached            Null0
172.16.11.64/30      172.16.10.1         GigabitEthernet0/0/0
……
```

- Routers send active routes in the routing table to the FIB table. A router searches the FIB table for the optimal route to forward the packet.

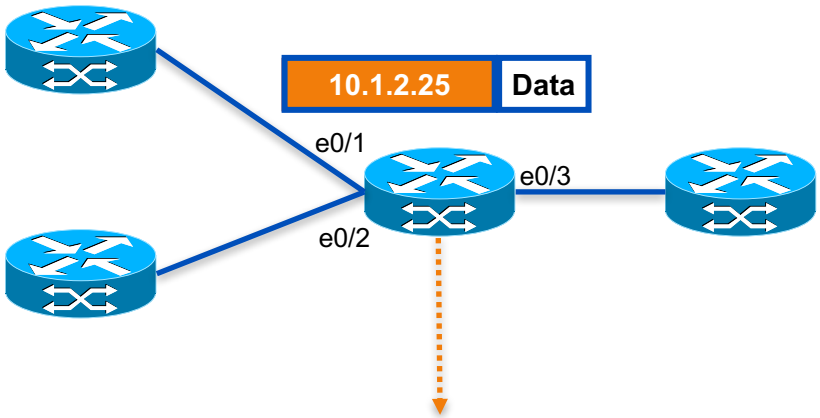
Routing and Forwarding

- **Routing** process constructs a view of the network topology and computes the best paths.
- **Packet forwarding** includes to transfer packets between interfaces and also some additional processes, i.e. classification, translation and so on.



Route Lookup

1. Destination IP Address in the packet Logical AND Subnet mask.
2. Compare the result with the destination IP address in FIB.



AND

00001010 00000001 00000010 00011001 ← 10.1.2.25

00000000 00000000 00000000 00000000 ← Subnet Mask(/0)

=

00000000 00000000 00000000 00000000

0.0.0.0

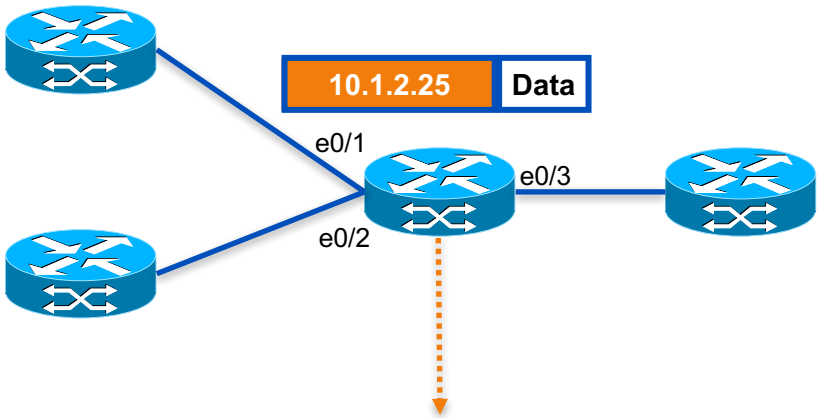
0.0.0.0

Match

Destination	I/F	
0.0.0.0/0	**e0/1**	Match
10.1.0.0/16	e0/2	
10.1.1.0/24	e0/3	
10.1.2.0/24	e0/2	
10.2.0.0/16	e0/1	

Route Lookup

1. Destination IP Address in the packet Logical AND Subnet mask.
2. Compare the result with the destination IP address in FIB.



AND

00001010 00000001 00000010 00011001 ← 10.1.2.25

11111111 11111111 00000000 00000000 ← Subnet Mask(/16)

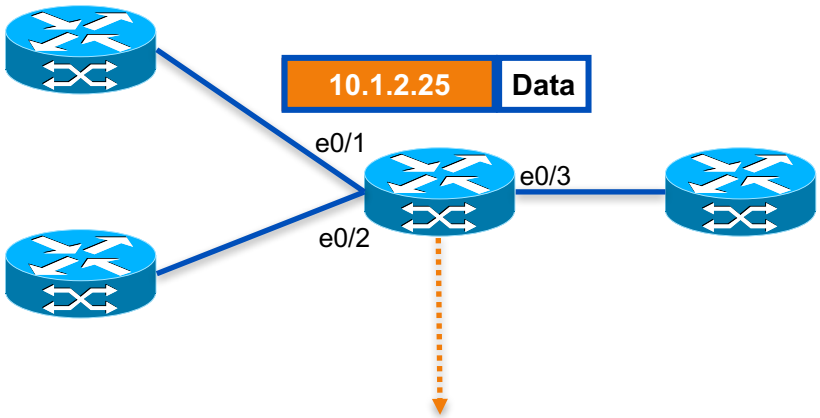
= 00001010 00000001 00000000 00000000 → 10.1.0.0

10.1.0.0 → Match

Destination	I/F	
0.0.0.0/0	e0/1	Match
10.1.0.0/16	**e0/2**	Match
10.1.1.0/24	e0/3	
10.1.2.0/24	e0/2	
10.2.0.0/16	e0/1	

Route Lookup

1. Destination IP Address in the packet Logical AND Subnet mask.
2. Compare the result with the destination IP address in FIB.



AND

00001010 00000001 00000010 00011001 ← 10.1.2.25

11111111 11111111 11111111 00000000 ← Subnet Mask(/24)

=

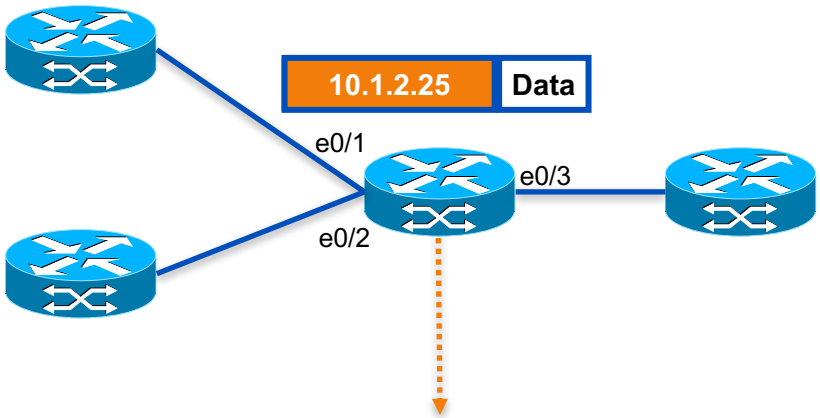
00001010 00000001 00000010 00000000 → 10.1.2.0

10.1.1.0 ← Not Match

Destination	I/F	
0.0.0.0/0	e0/1	Match
10.1.0.0/16	e0/2	Match
10.1.1.0/24	**e0/3**	Not Match
10.1.2.0/24	e0/2	
10.2.0.0/16	e0/1	

Route Lookup

1. Destination IP Address in the packet Logical AND Subnet mask.
2. Compare the result with the destination IP address in FIB.



AND

00001010 00000001 00000010 00011001	←	10.1.2.25
11111111 11111111 11111111 00000000	←	Subnet Mask(/24)
00001010 00000001 00000010 00000000		10.1.2.0

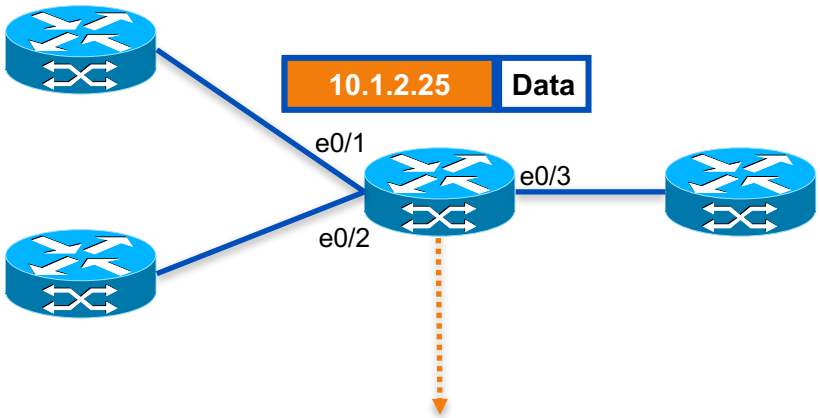
Match

10.1.2.0 ←

Destination	I/F	
0.0.0.0/0	e0/1	Match
10.1.0.0/16	e0/2	Match
10.1.1.0/24	e0/3	Not Match
10.1.2.0/24	**e0/2**	Match
10.2.0.0/16	e0/1	

Route Lookup

1. Destination IP Address in the packet Logical AND Subnet mask.
2. Compare the result with the destination IP address in FIB.



AND

00001010 00000001 00000010 00011001 ← 10.1.2.25

11111111 11111111 00000000 00000000 ← Subnet Mask(/16)

=

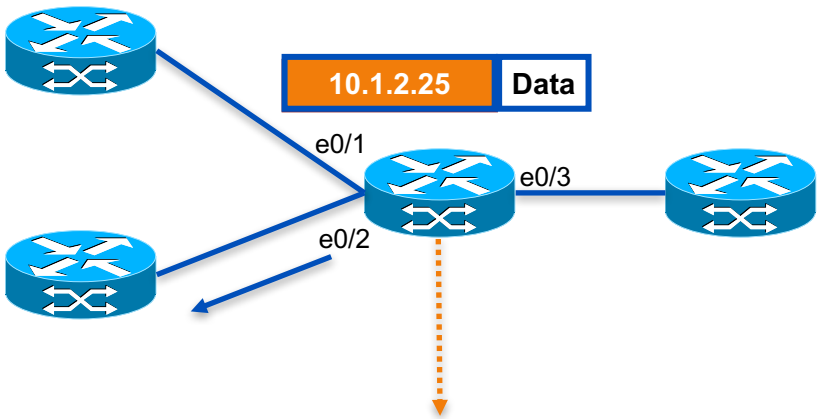
00001010 00000001 00000000 00000000 → 10.1.0.0

10.2.0.0 → Not Match

Destination	I/F	
0.0.0.0/0	e0/1	Match
10.1.0.0/16	e0/2	Match
10.1.1.0/24	e0/3	Not Match
10.1.2.0/24	e0/2	Match
10.2.0.0/16	**e0/1**	Not Match

Route Lookup – Longest Matching

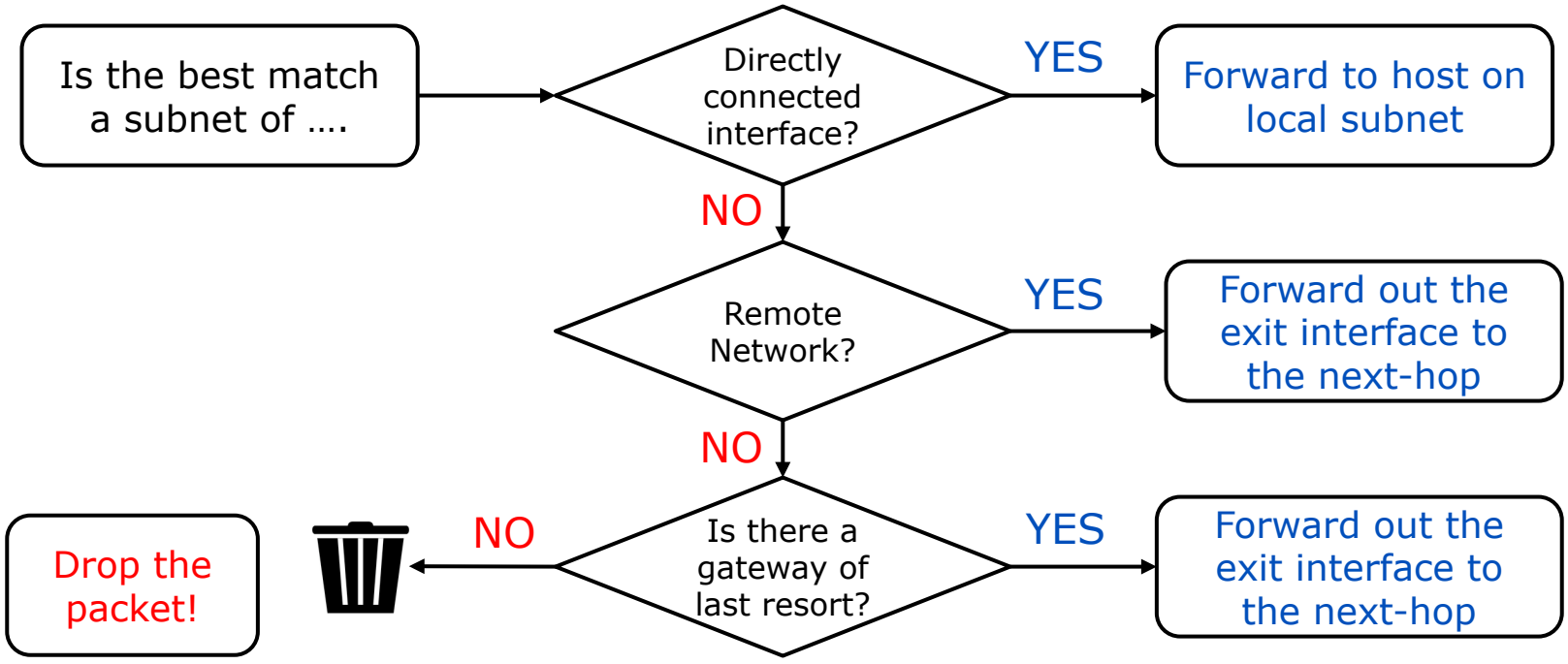
1. Destination IP Address in the packet Logical AND Subnet mask
2. Compare the result with the destination IP address in FIB.
3. Router selects the route with **longest** prefix length of subnet mask.



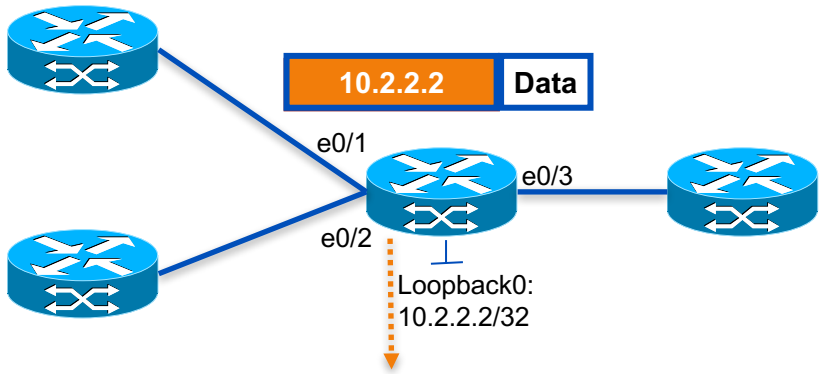
Destination	I/F	
0.0.0.0/0	e0/1	Match
10.1.0.0/16	e0/2	Match
10.1.1.0/24	e0/3	Not Match
10.1.2.0/**24**	e0/2	Match
10.2.0.0/16	e0/1	Not Match

Forwarding Decision

- If a best match is found, the router determines
  - the correct exit interface to reach the next-hop/destination



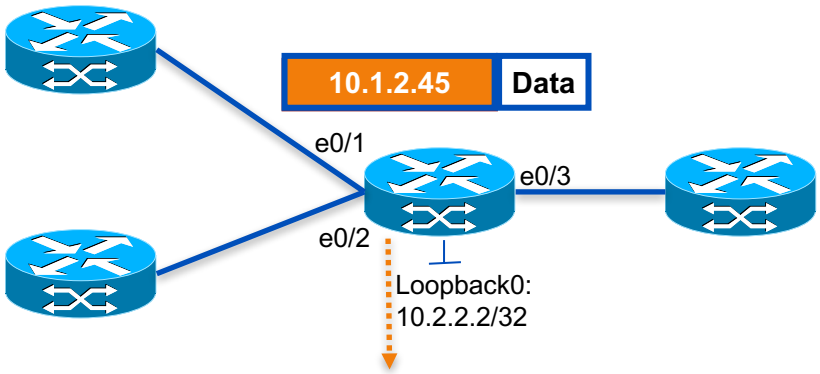
Forwarding Decision (1)



Destination	I/F
0.0.0.0/0	ethernet0/1
10.1.0.0/16	ethernet0/2
10.1.1.0/24	ethernet0/3
10.1.2.0/24	ethernet0/2
10.2.2.2/32	Loopback0

Local interface

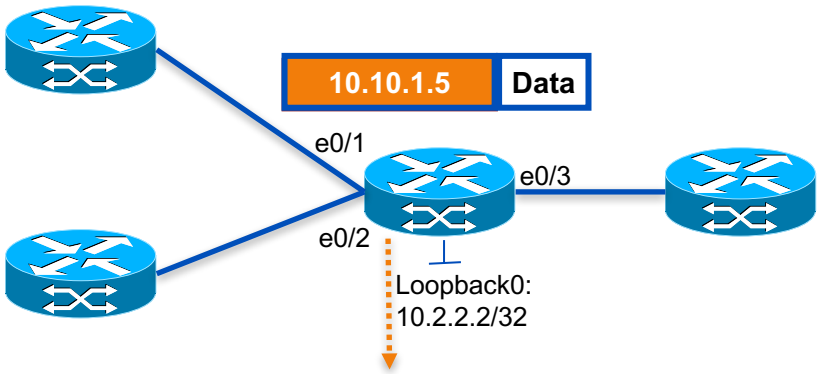
Forwarding Decision (2)



Destination	I/F
0.0.0.0/0	ethernet0/1
10.1.0.0/16	ethernet0/2
10.1.1.0/24	ethernet0/3
10.1.2.0/24	ethernet0/2
10.2.2.2/32	Loopback0

Send to nexthop

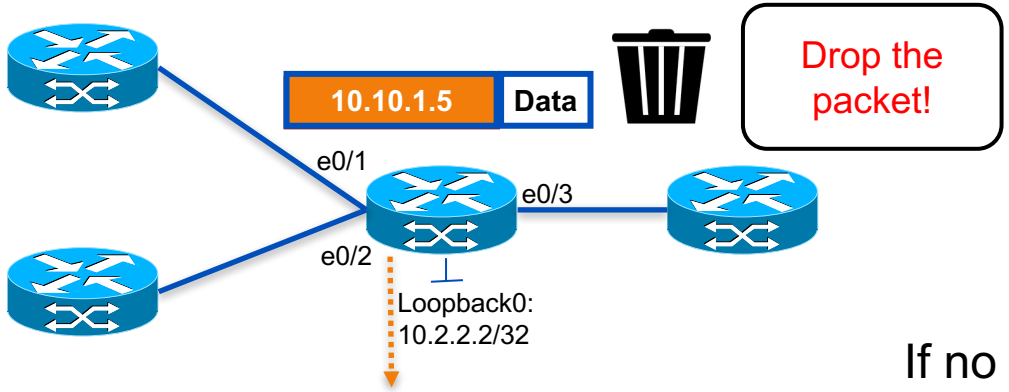
Forwarding Decision (3)



Destination	I/F
0.0.0.0/0	ethernet0/1
10.1.0.0/16	ethernet0/2
10.1.1.0/24	ethernet0/3
10.1.2.0/24	ethernet0/2
10.2.2.2/32	Loopback0

Use the default route

Forwarding Decision (4)



Destination	I/F
10.1.0.0/16	ethernet0/2
10.1.1.0/24	ethernet0/3
10.1.2.0/24	ethernet0/2
10.2.2.2/32	Loopback0

If no default route, the packet will be dropped.

Internet Routing

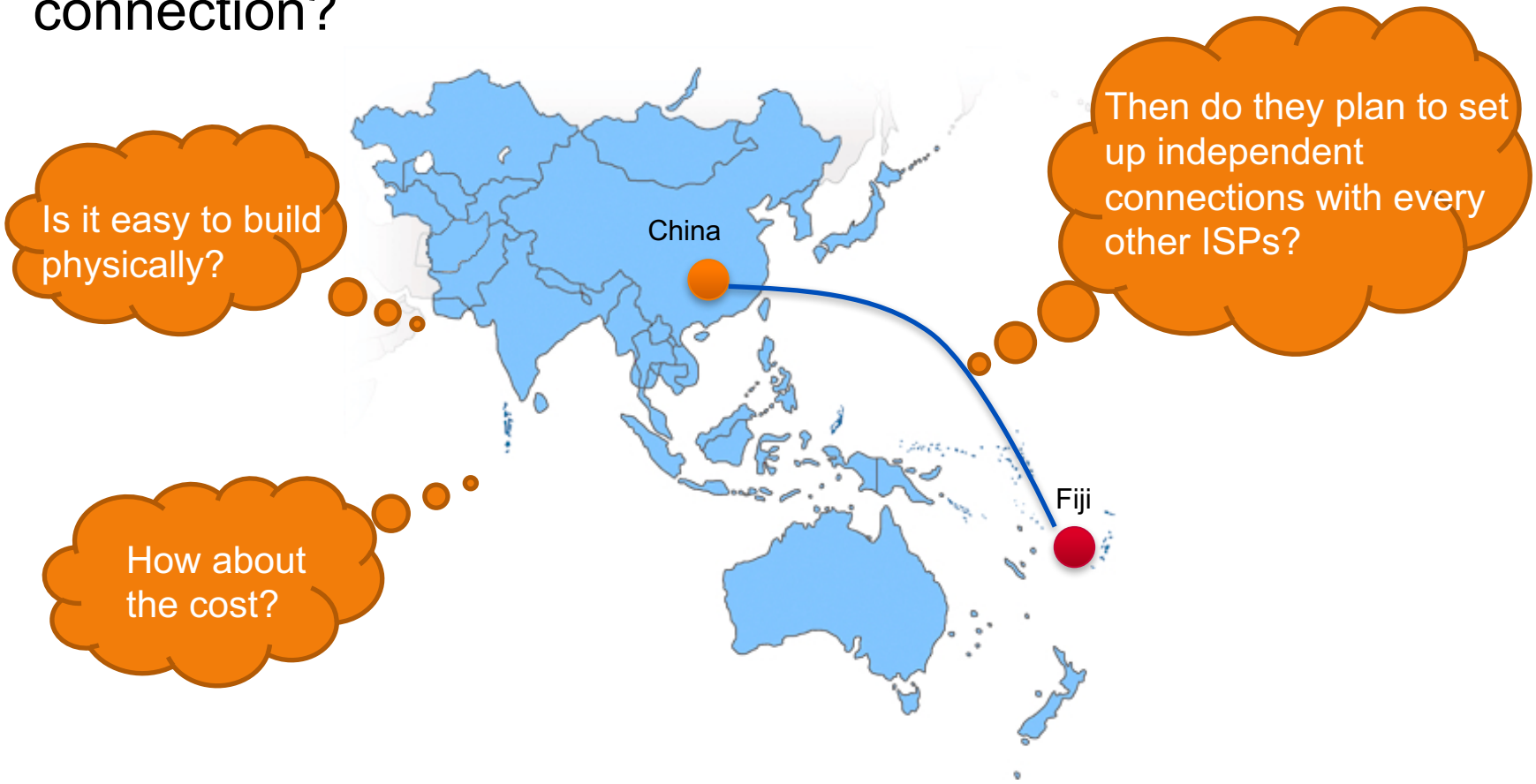
APNIC

How does Routing Work in Internet?

- If ISP in Fiji wants to tell ISP in China about their customer, how can they communicate?

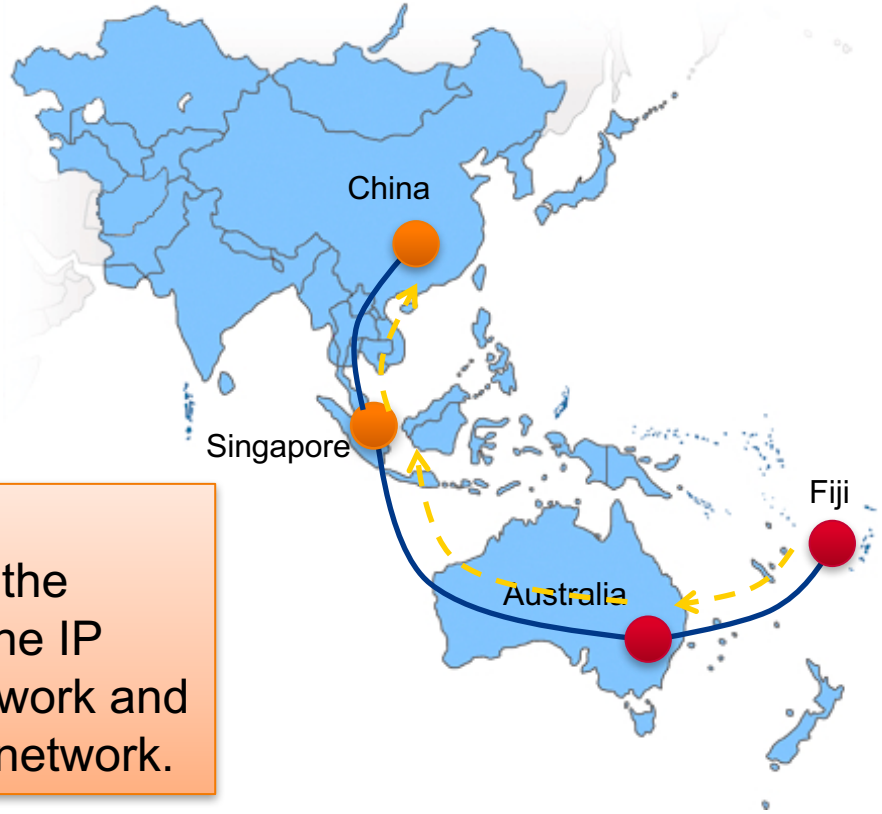
How does Routing Work in Internet?

- Should the ISPs in Fiji and China build the direct connection?



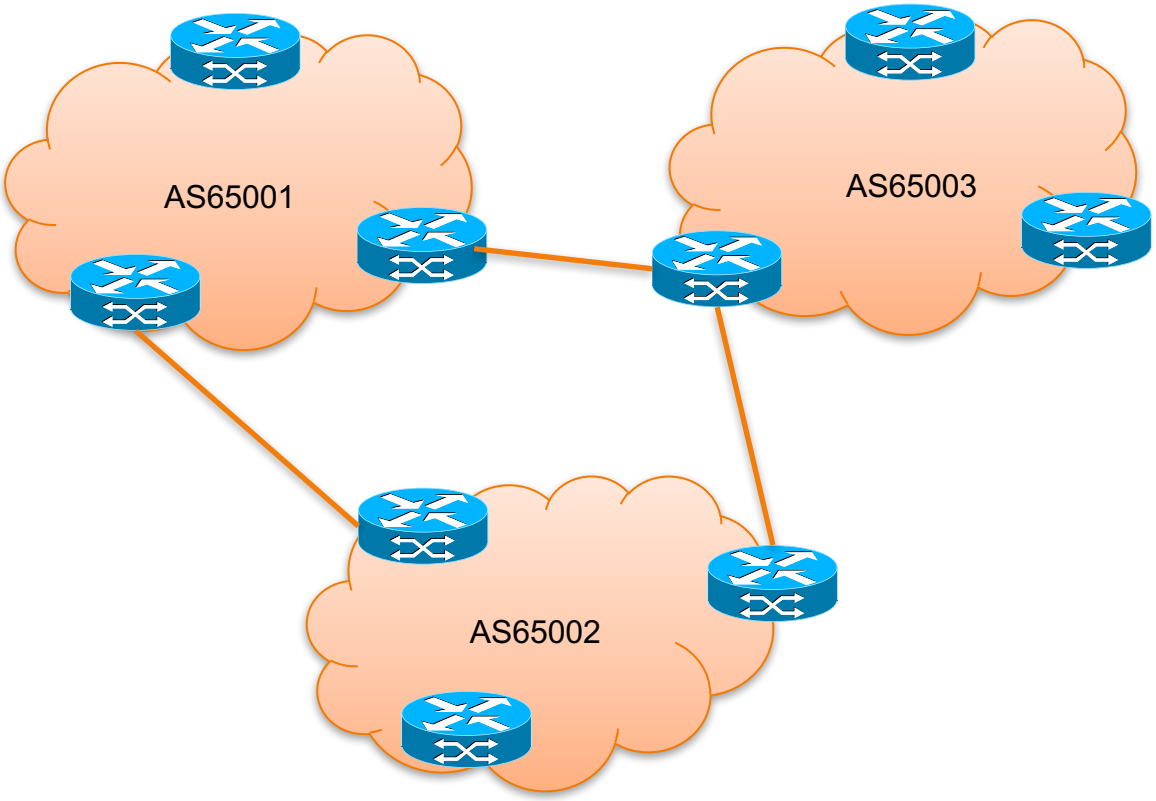
How does Routing Work in Internet?

- In fact, ISP in Fiji talks to its neighboring ISP about its customer, and the neighboring ISPs pass this information on to their neighbors, and so on. Vice versa for ISP in China.

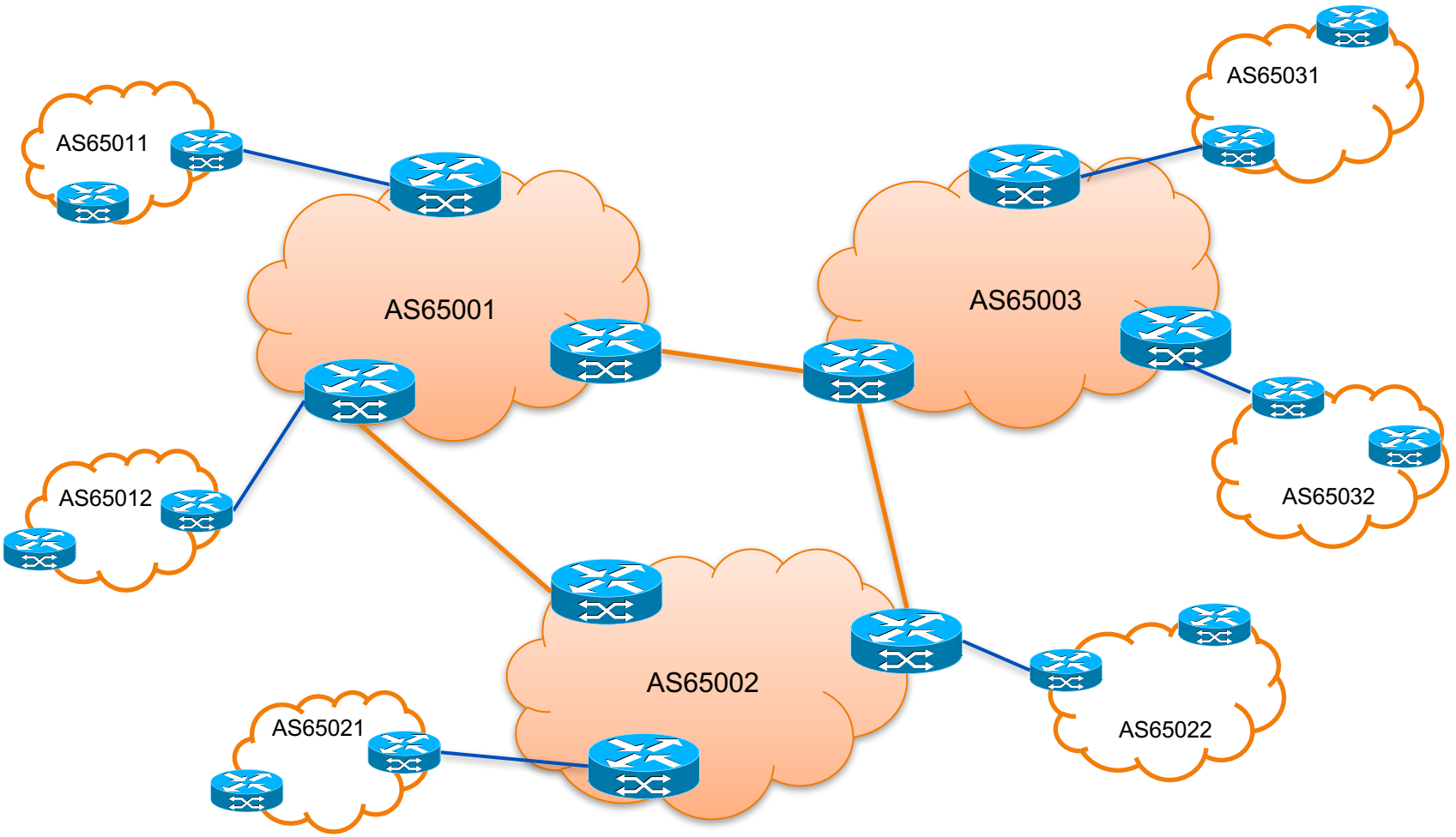


Routing:
ISP in Fiji tells the neighbor ISP the IP block of its network and its customers' network.

An Example of Internet

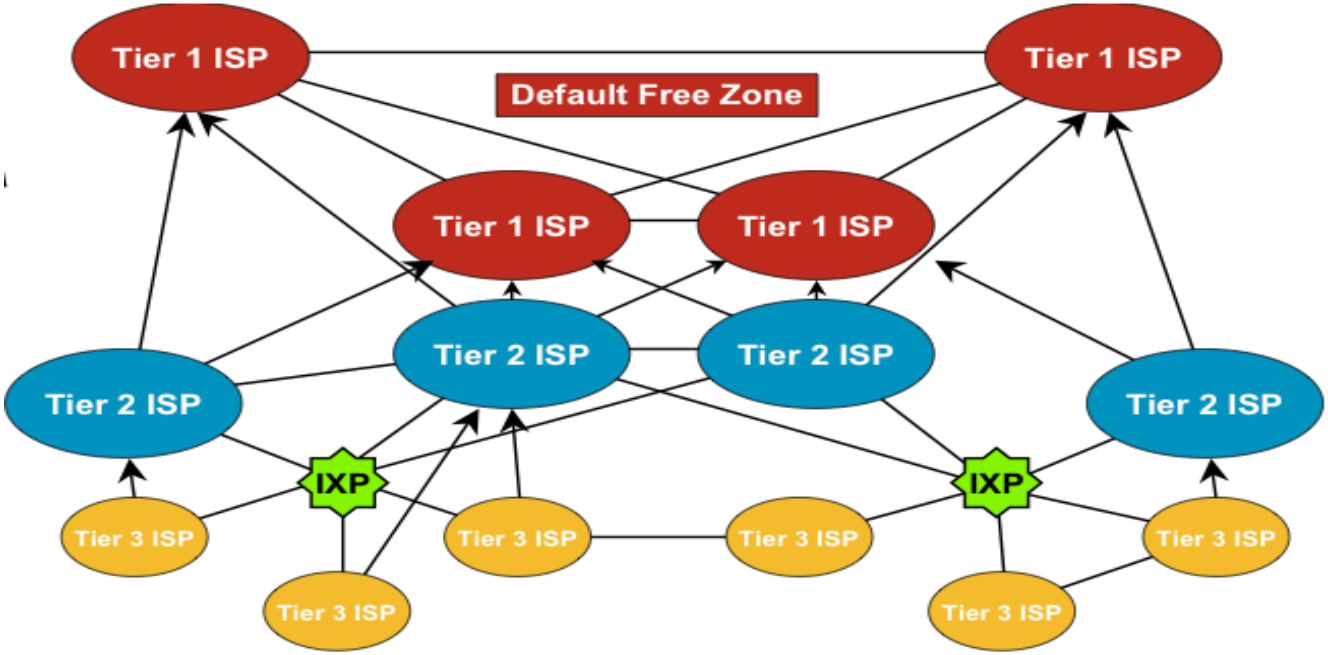


An Example of Internet



Internet Hierarchy

- Default free zone
  - Made of Tier-1 ISPs who have explicit routes to every network on the Internet
    - No need for default routes!

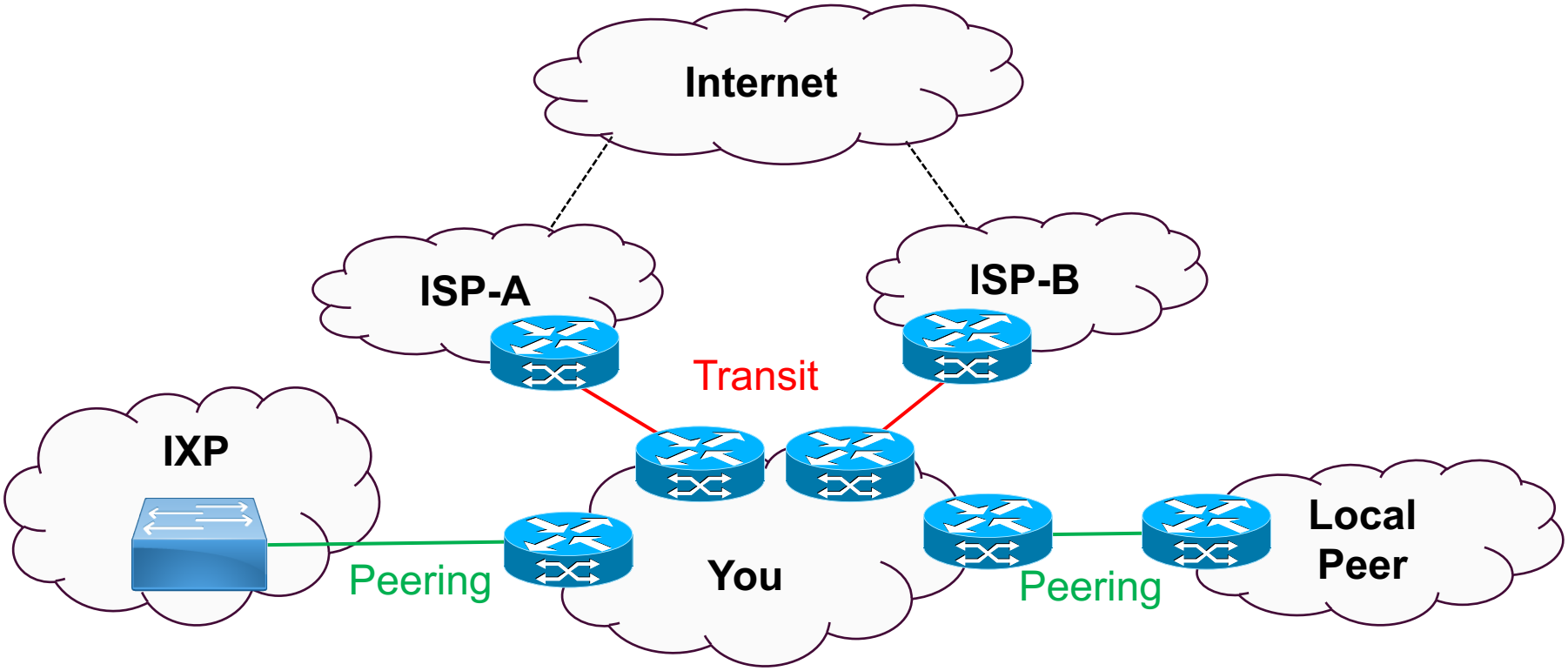


Exchanging Routes

- Pay someone to advertise your networks
  - TRANSIT
  - Make sure they have good onward peering/transit!
- Interconnect with as other ASes to exchange locally originated routes and traffic
  - PEERING
  - Private Peering
    - Between two ASes
  - Public Peering
    - at an IXP (domestic/global)

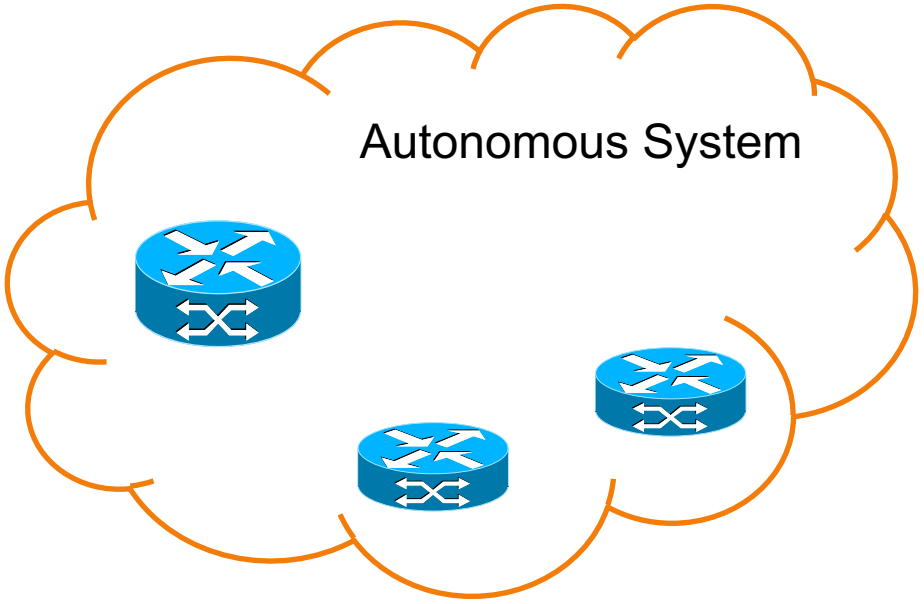
Example of ISP Connection

- More than one upstream ISP with local and public peering



Autonomous System

- An AS is a group of IP networks operated by one or more network operator(s) that has a single and clearly defined external routing policy.



Autonomous System Number

- An AS has a globally unique number associated with it; this number is used in both the exchange of exterior routing information and as an identifier of the AS itself.
  - 16-bit ASN: range from 0 to 65535
  - 32-bit ASN: range from 65536 to 4294967295
- One organization can get a public AS number from its upstream provider or from RIR.

Reserved ASN

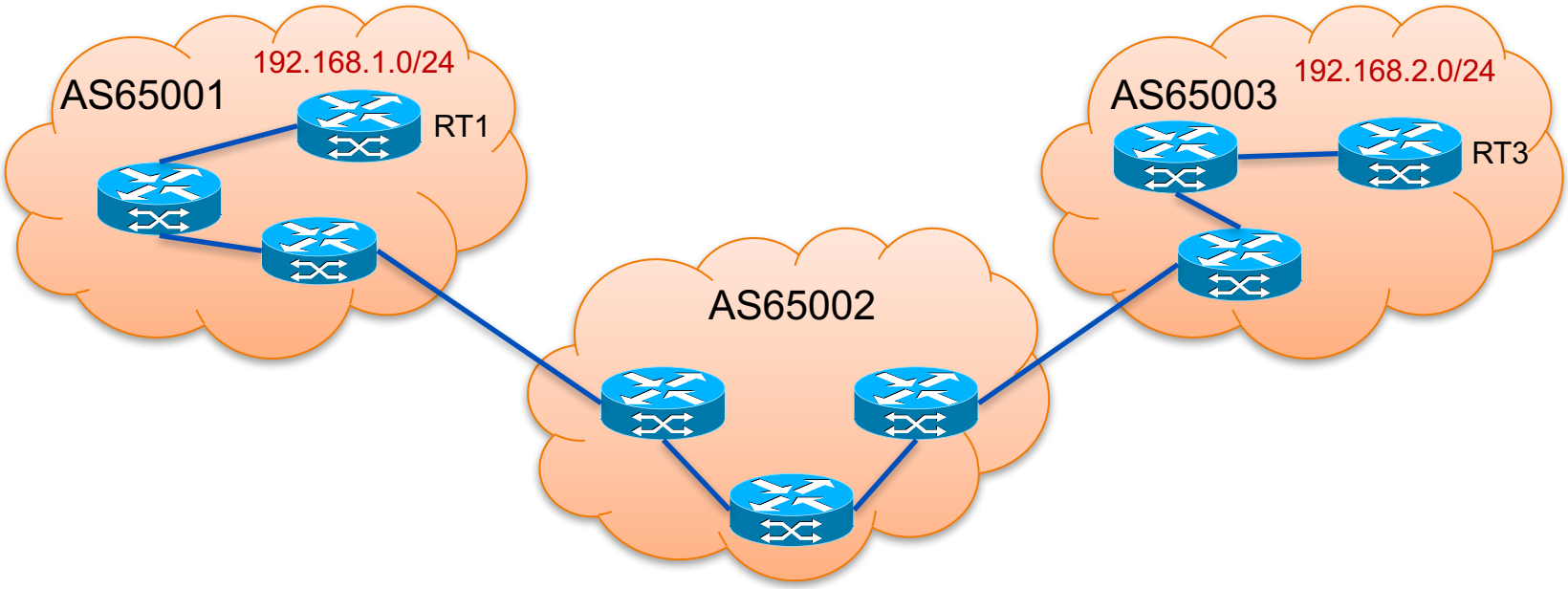
- ASN for special use are as follows, the others are for public allocation.

AS Number Range	Description	Reference
0	Reserved	RFC7607
23456	AS_TRANS	RFC6793
64496-64511	Reserved for use in documentation and sample code	RFC5398
64512-65534	Reserved for Private Use	RFC6996
65535	Reserved	RFC7300
65536-65551	Reserved for use in documentation and sample code	RFC5398
65552-131071	Reserved	
4200000000-4294967294	Reserved for Private Use	RFC6996
4294967295	Reserved	RFC7300

https://www.iana.org/assignments/as-numbers/as-numbers.xhtml

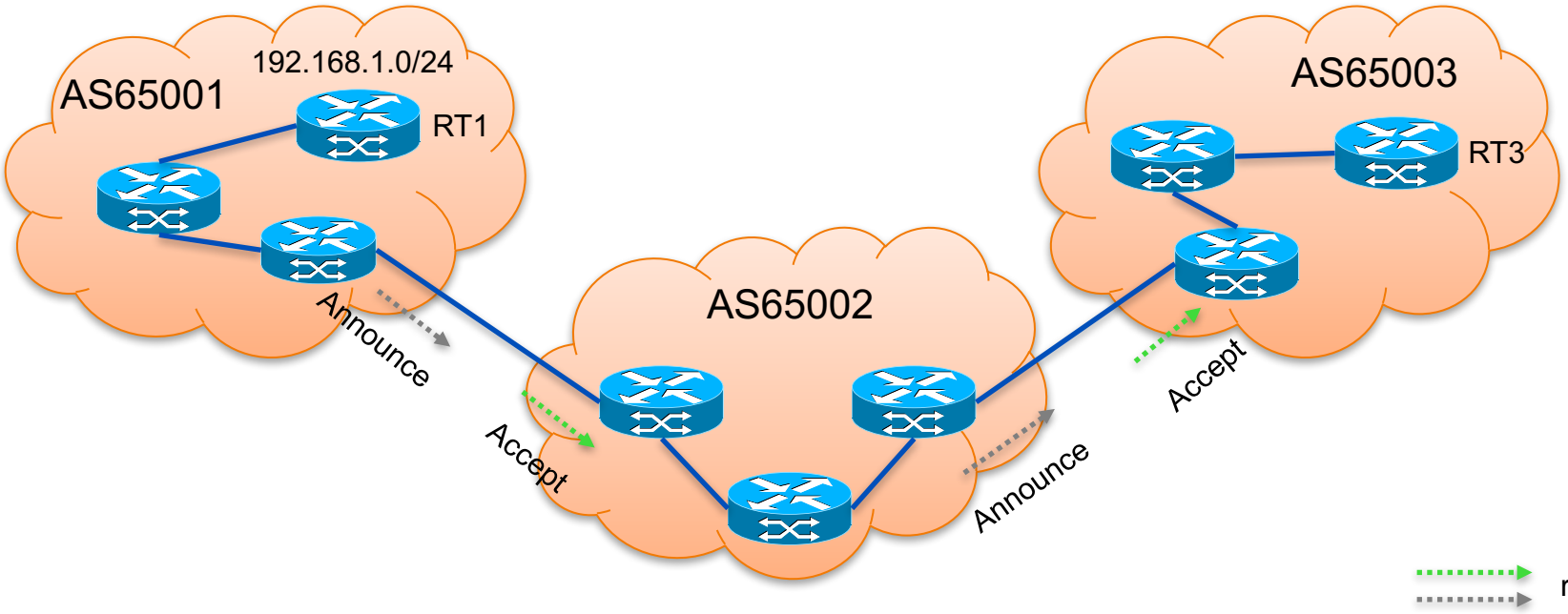
Origination

- Route 192.168.1.0/24 is originated from AS65001
- Route 192.168.2.0/24 is originated from AS65003



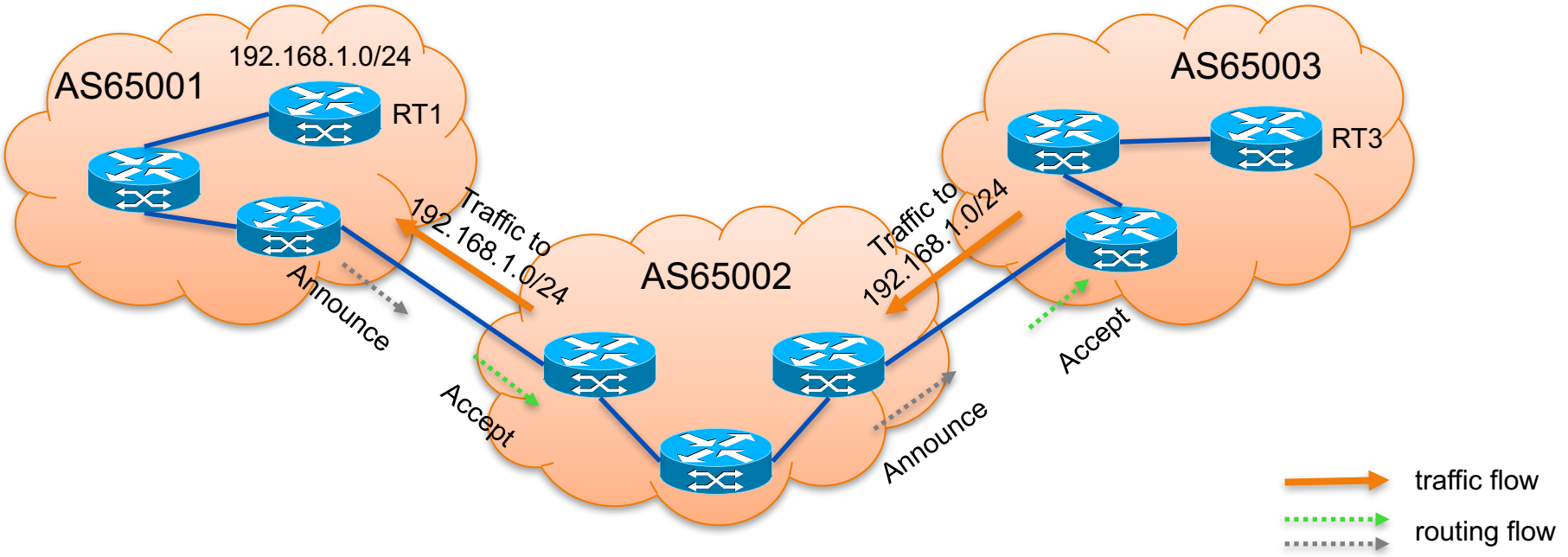
Announcement

- AS65001 announces 192.168.1.0/24 to AS65002, AS65002 accepts this route.
- AS65002 announces it to AS65003, AS65003 accepts it.



Routing Flow & Traffic Flow

- Traffic and routing information flow in opposite directions.
- After AS65003 has learnt this route, traffic will be sent from AS65003 to AS65002, from AS65002 to AS65001.



Routing Policy

- Routing policy is used to solve following problems:

- ? What are the routes I should receive or should not receive from other AS
- ? What are the routes I should send or should not send to other AS
- ? How to classify different types of routes in my AS

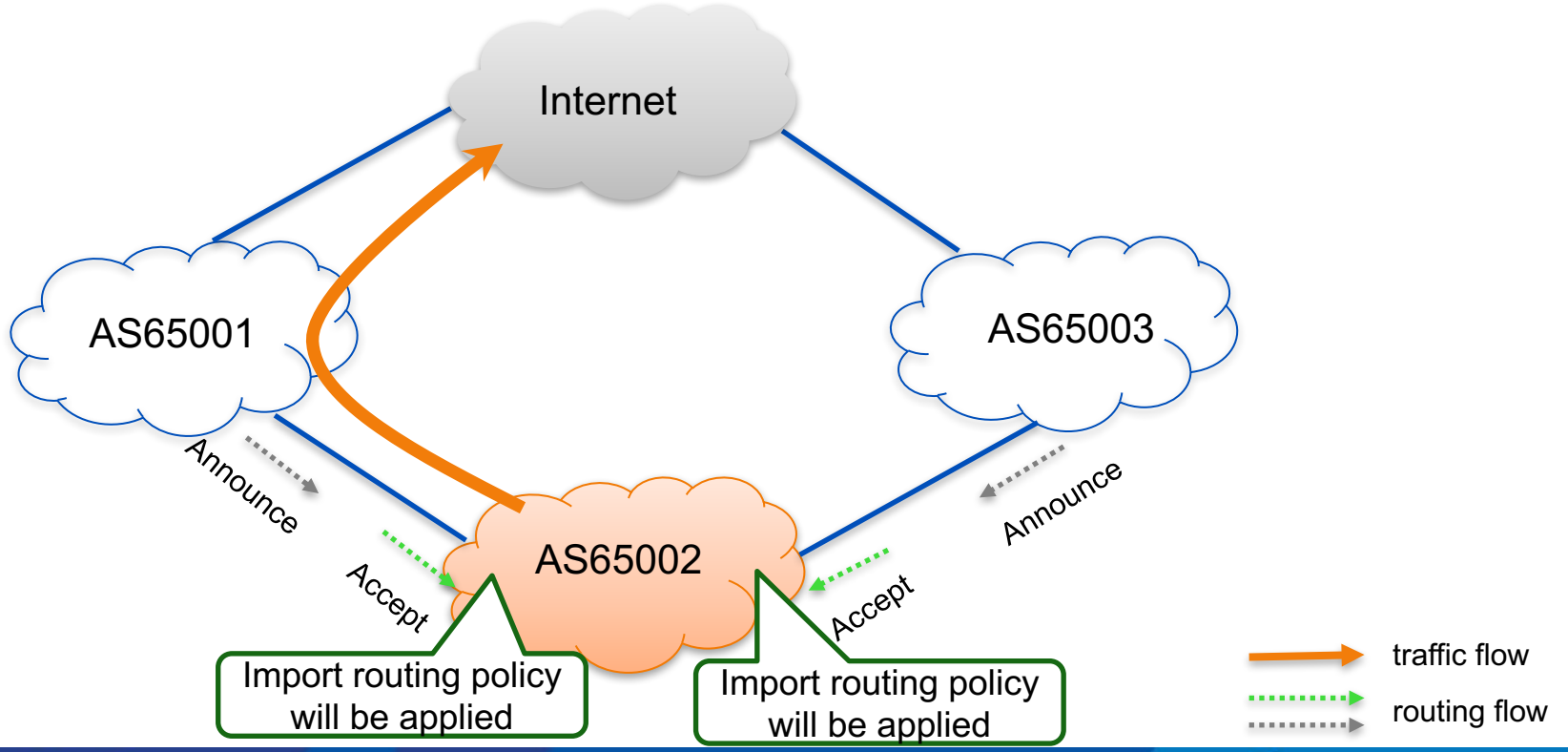
Different Types of Routing Policy

- **Business relationship policy**
  - From economic or political relationships an ISP has with its neighbor,
- **Traffic engineering policy**
  - Control traffic flow within an ISP and across peering links to avoid congestion
  - Provide good service quality,
- **Policies for scalability**
  - Reduce routing information and avoid overloading routers,
- **Security-related policies**
  - Protect an ISP against malicious or accidental attacks.

https://www.cs.princeton.edu/~jrex/papers/policies.pdf

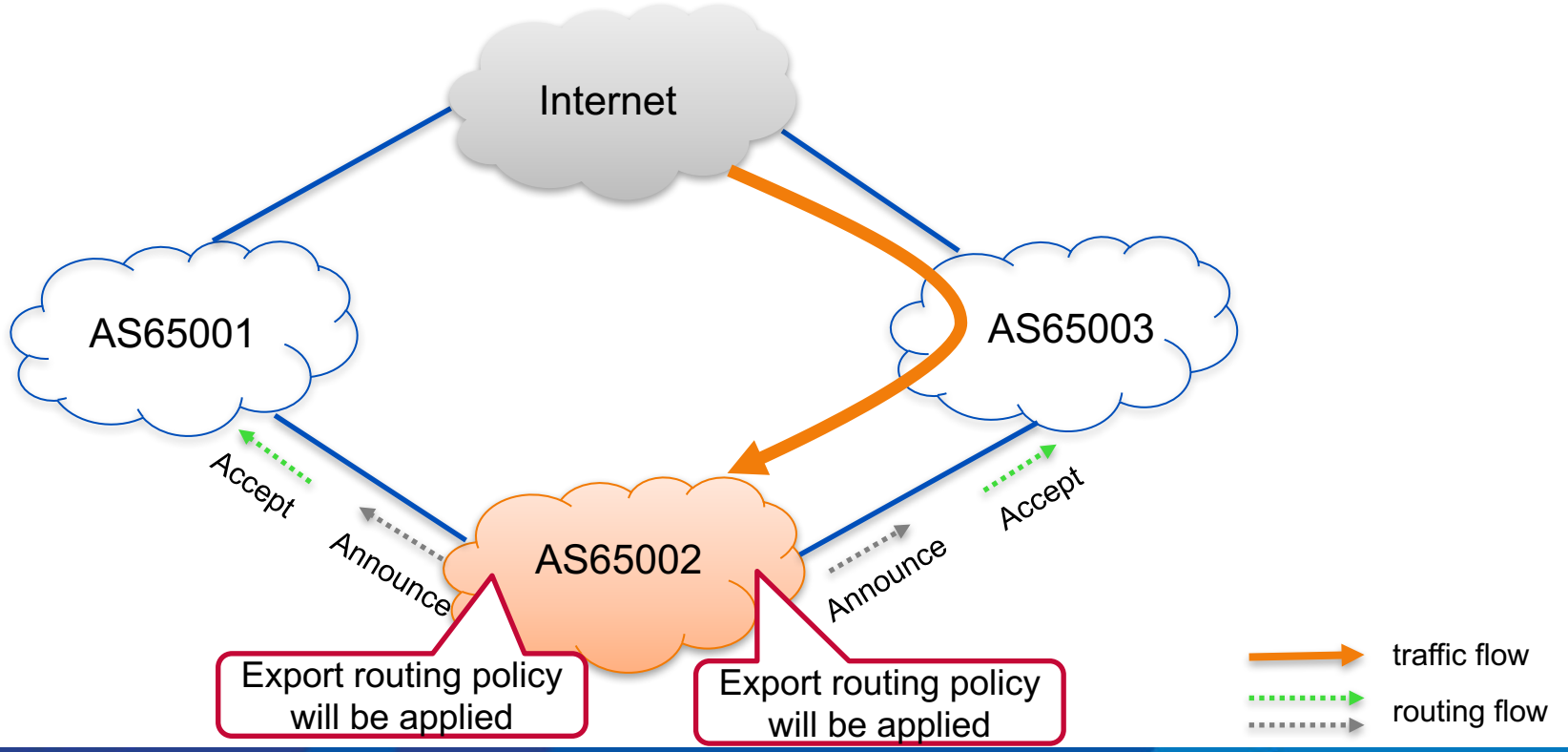
Import Policy & Outbound Traffic

- **Import policy** is applied to determine which routes should be filtered and hence eliminated from consideration, and may append or modify attributes.
- Operators can influence outbound traffic flow by configuring import policies.



Export Policy & Inbound Traffic & Cooperation

- Operators can influence inbound traffic flow by configuring export policies. But any intermediate AS need to cooperate.



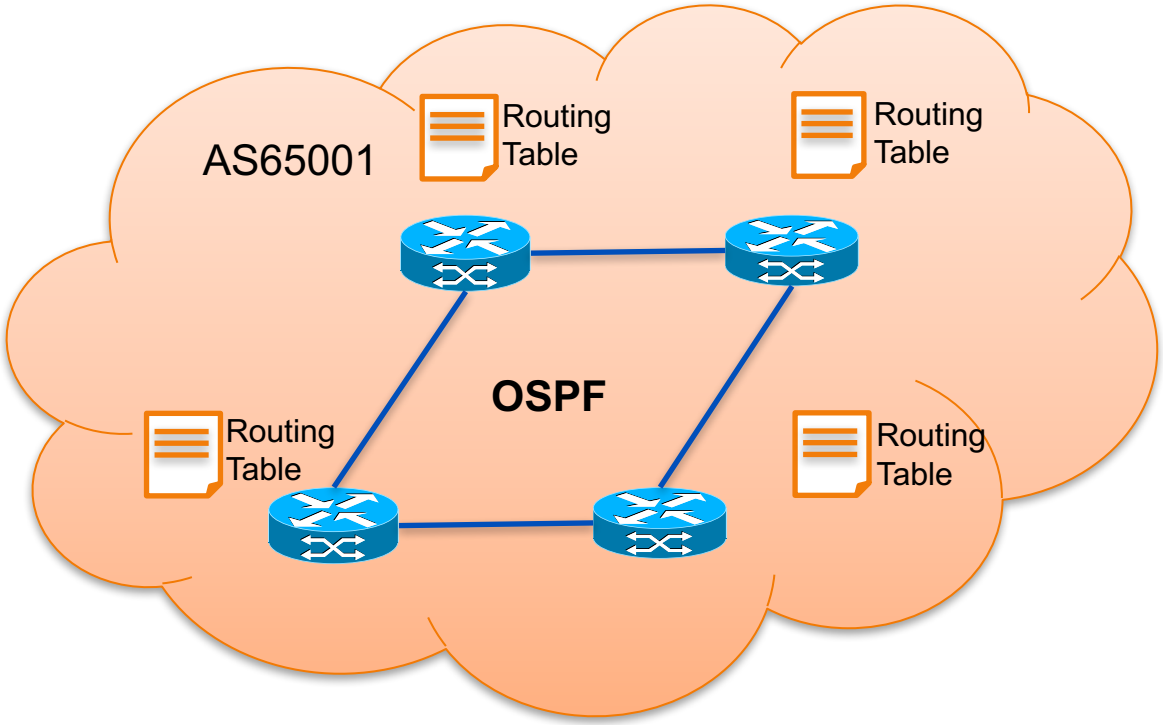
Dynamic Routing Protocols

- How do routers exchange network information with each other?

- IGP
  - Interior Gateway Protocol
  - exchange routing information within a single AS
  - OSPF, IS-IS
- EGP
  - Exterior Gateway Protocol
  - exchange routing information between ASes
  - BGP

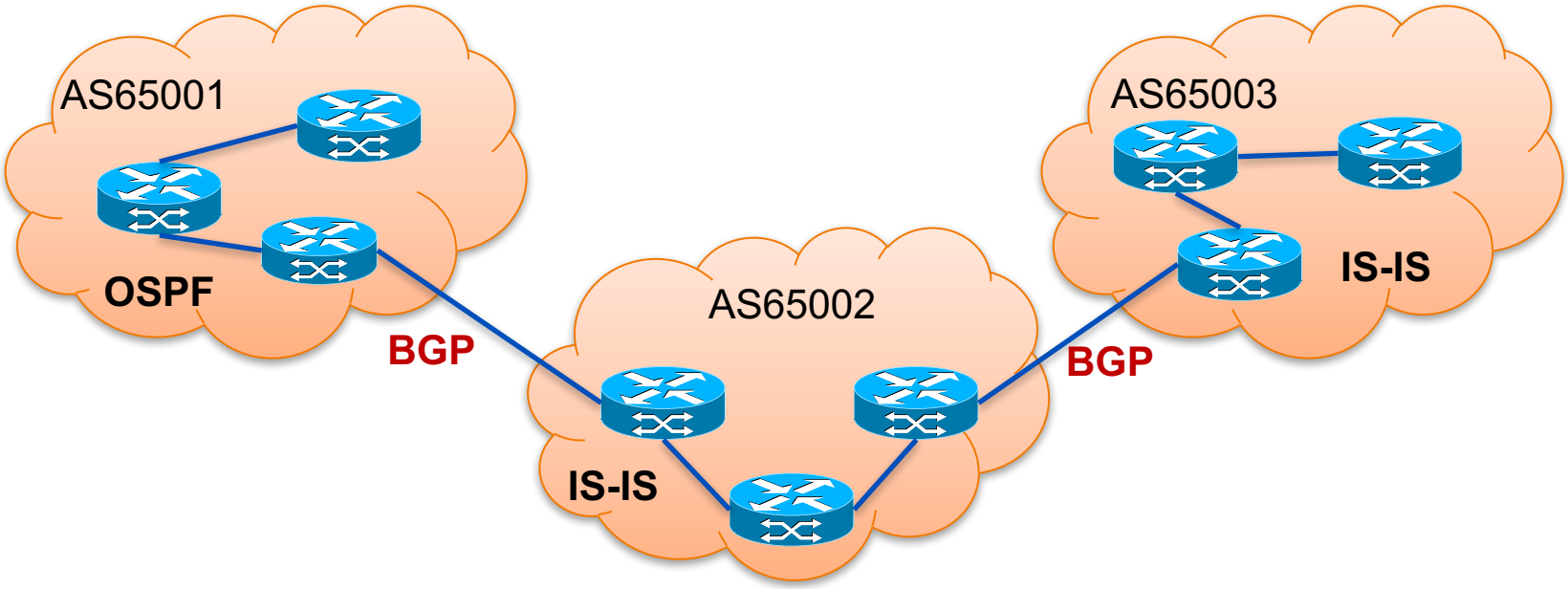
Interior Gateway Protocol

- In the same AS, routers will exchange route information with each other by using IGP to build the routing table.
- IGP carries ISP infrastructure routes.



Exterior Gateway Protocols

- BGP is the only EGP, used to connect different ASes in Internet. BGP carries the routes between different ASes.



Cisco IOS Default Administrative Distances

Route Source	Default Distance
Connected interface	0
Static route	1
Enhanced Interior Gateway Routing Protocol (EIGRP) summary route	5
External Border Gateway Protocol (BGP)	20
Internal EIGRP	90
IGRP	100
OSPF	110
Intermediate System-to-Intermediate System (IS-IS)	115
Routing Information Protocol (RIP)	120
Exterior Gateway Protocol (EGP)	140
On Demand Routing (ODR)	160
External EIGRP	170
Internal BGP	200
Unknown*	255

http://www.cisco.com/c/en/us/support/docs/ip/border-gateway-protocol-bgp/15986-admin-distance.html

Huawei VRP Default Preferences

Routing Protocol or Route Type	Default Preference
Direct	0
OSPF	10
IS-IS	15
Static	60
RIP	100
OSPF ASE	150
OSPF NSSA	150
IBGP	255
EBGP	255

Junos Default Route Preference Values

How Route Is Learned	Default Preference
Directly connected network	0
System routes	4
OSPF internal route	10
IS-IS Level 1 internal route	15
IS-IS Level 2 internal route	18
Redirects	30
RIP	100
RIPng	100
Aggregate	130
OSPF AS external routes	150
IS-IS Level 1 external route	160
IS-IS Level 2 external route	165
BGP	170

https://www.juniper.net/documentation/en_US/junos/topics/reference/general/routing-protocols-default-route-preference-values.html

Thank You!

APNIC