

APNIC Routing I Workshop

Suva, Fiji

3 - 6 December 2013

Proudly Supported by:



APNIC



Presenter

Nurul Islam (Roman)

Senior Training Specialist, APNIC

Nurul maintains the APNIC training lab and is involved in delivering technical training for the APNIC community. He possesses specialized skills in designing and running IPv4/IPv6 routing and switching infrastructure for service provider and enterprise networks. Prior to his current role he looked after the IP and AS number allocations for the APNIC Members.

Areas of interests:

Internet Resource Management, IPv6, Routing and Switching, MPLS, BGP, Security, Internet Routing Registry and RPKI, ISP Services and Internetworking.

Contact:

Email: nurul@apnic.net



Presenter

Sheryl Hermoso (Shane)

Training Officer, APNIC

Sheryl has had various roles as a Network and Systems Administrator prior to joining APNIC. Starting her career as a Technical Support Assistant while studying at the University of the Philippines. Sheryl later finished her degree in Computer Engineering and continued to work in the same university as a Network Engineer, where she managed the DILNET network backbone and wireless infrastructure.

Areas of interests:

DNS/DNSSEC, IPv6, and Network Security

Contact:

Email: sheryl@apnic.net



Overview

Routing Workshop (3 Days)

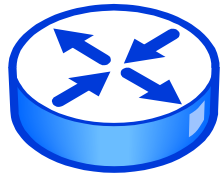
- Introduction to IP Routing
- Routing Protocol Basic
- IPv6 Address Structure
- Routing Lab Topology Overview
- Operation of OSPF Routing Protocol
- Lab Exercise on Basic Router and OSPF Dynamic Routing Configuration
- Basic BGP Operation
- BGP Attributes and Path Selection Process
- BGP Scaling Techniques
- Lab Exercise on iBGP, eBGP, RR, Peer group etc
- Internet Exchange [IX] Policy Overview and Configuration requirement

Overview

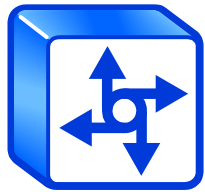
Routing Workshop (3 Days)

- **Introduction to IP Routing**
- Routing Protocol Basic
- IPv6 Address Structure
- Routing Lab Topology Overview
- Operation of OSPF Routing Protocol
- Lab Exercise on Basic Router and OSPF Dynamic Routing Configuration
- Basic BGP Operation
- BGP Attributes and Path Selection Process
- BGP Scaling Techniques
- Lab Exercise on iBGP, eBGP, RR, Peer group etc
- Internet Exchange [IX] Policy Overview and Configuration requirement

Graphics / Symbols Used



Router
(layer 3, IP datagram forwarding)



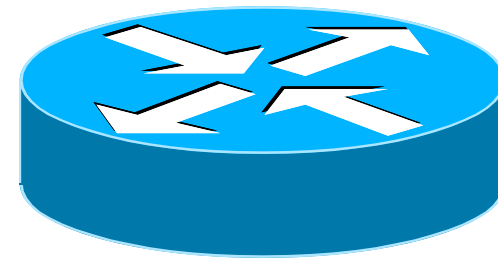
Network Access Server
(layer 3, IP datagram forwarding)



Ethernet switch
(layer 2, packet forwarding)

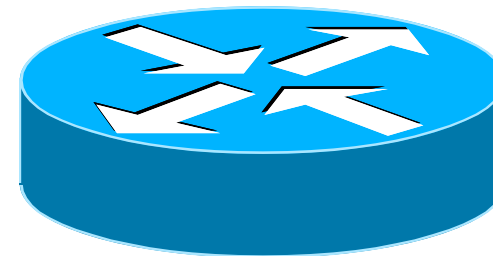
What does a router do?

- ?



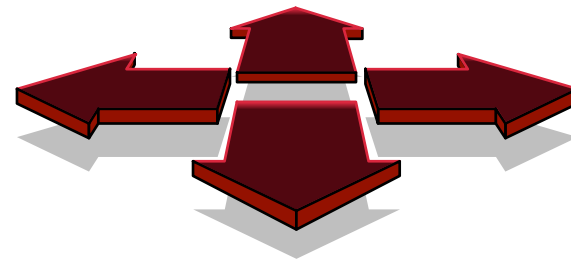
A day in a life of a router

- find path
- forward packet, forward packet, forward packet, forward packet...
- find alternate path
- forward packet, forward packet, forward packet, forward packet...
- repeat until powered off



Routing versus Forwarding

- Routing = building maps and giving directions
- Forwarding = moving packets between interfaces according to the “directions”



IP Routing – finding the path

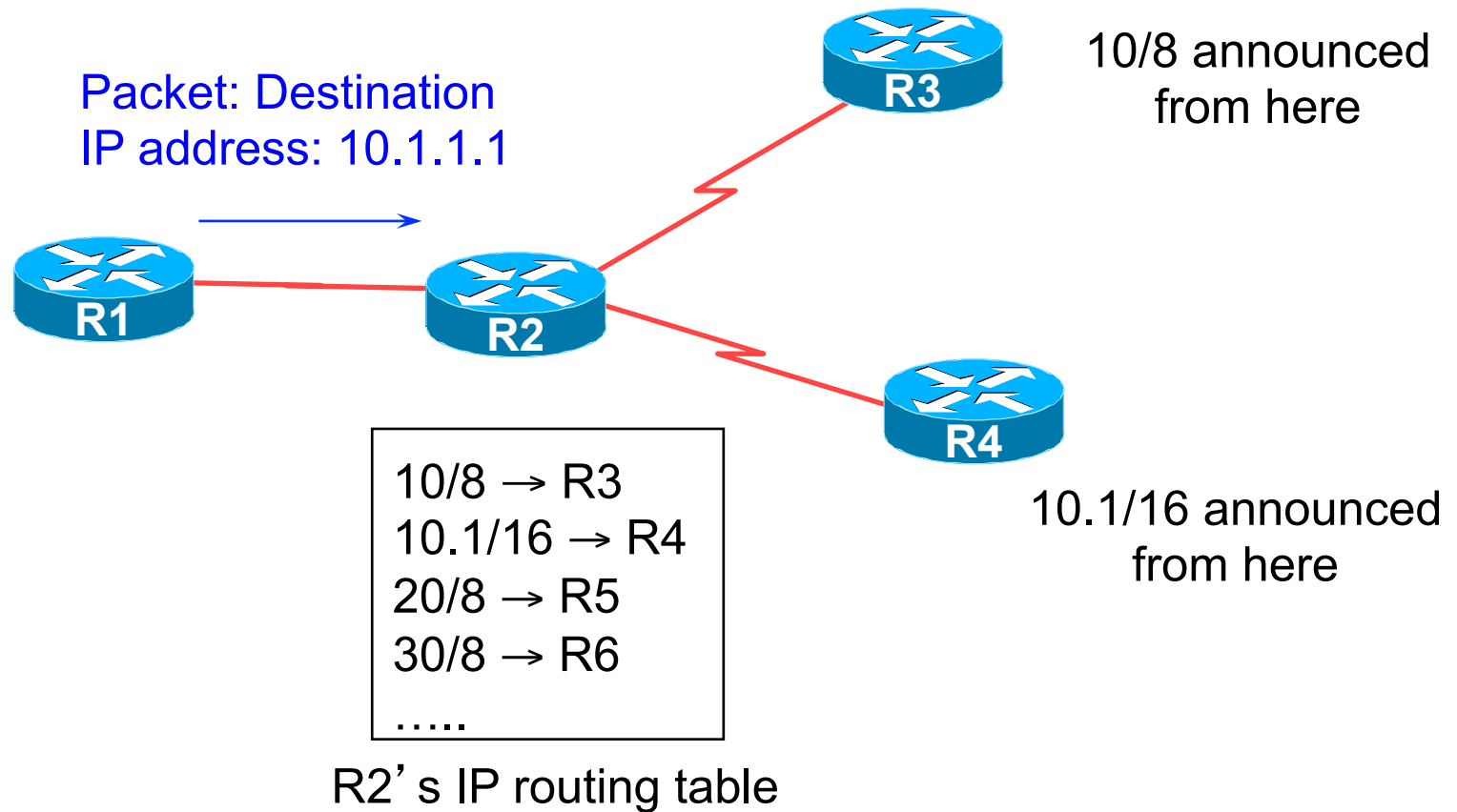
- Path derived from information received from a routing protocol
- Several alternative paths may exist
 - best path stored in forwarding table
- Decisions are updated periodically or as topology changes (event driven)
- Decisions are based on:
 - topology, policies and metrics (hop count, filtering, delay, bandwidth, etc.)

IP route lookup

- Based on destination IP address
- “longest match” routing
 - More specific prefix preferred over less specific prefix
 - **Example:** packet with destination of 10.1.1.1/32 is sent to the router announcing 10.1/16 rather than the router announcing 10/8.

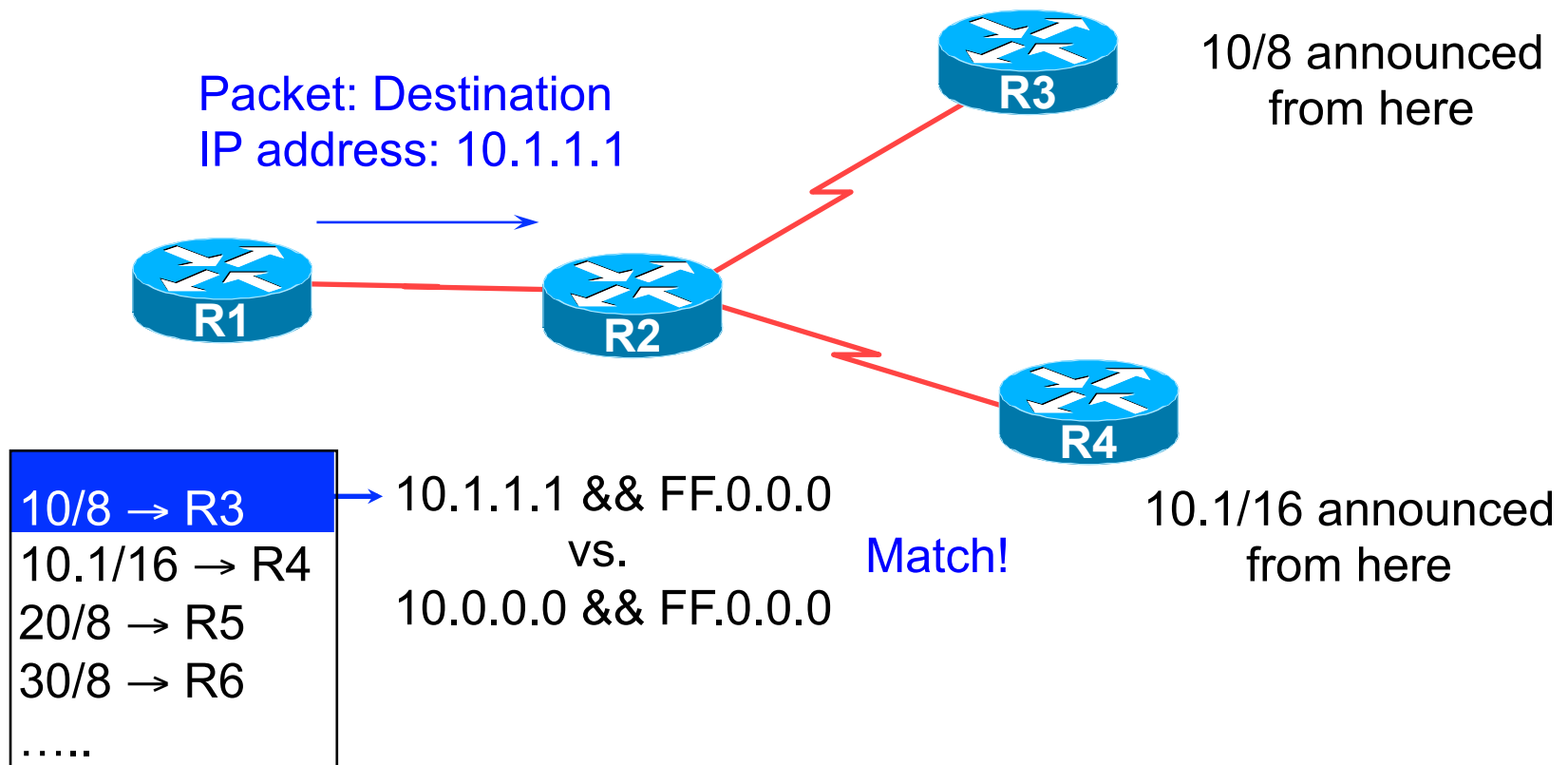
IP route lookup

- Based on destination IP address



IP route lookup: Longest match routing

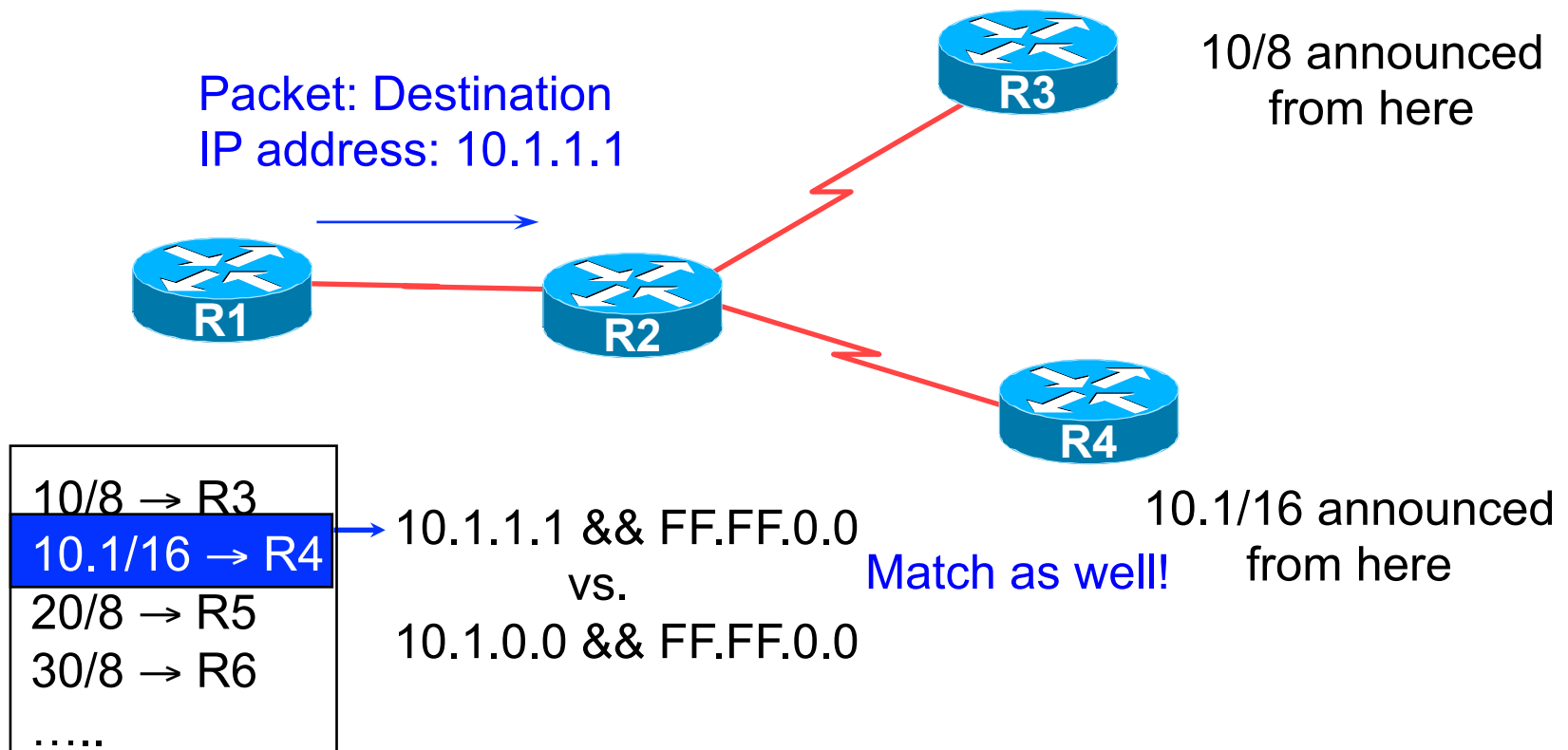
- Based on destination IP address



R2' s IP routing table

IP route lookup: Longest match routing

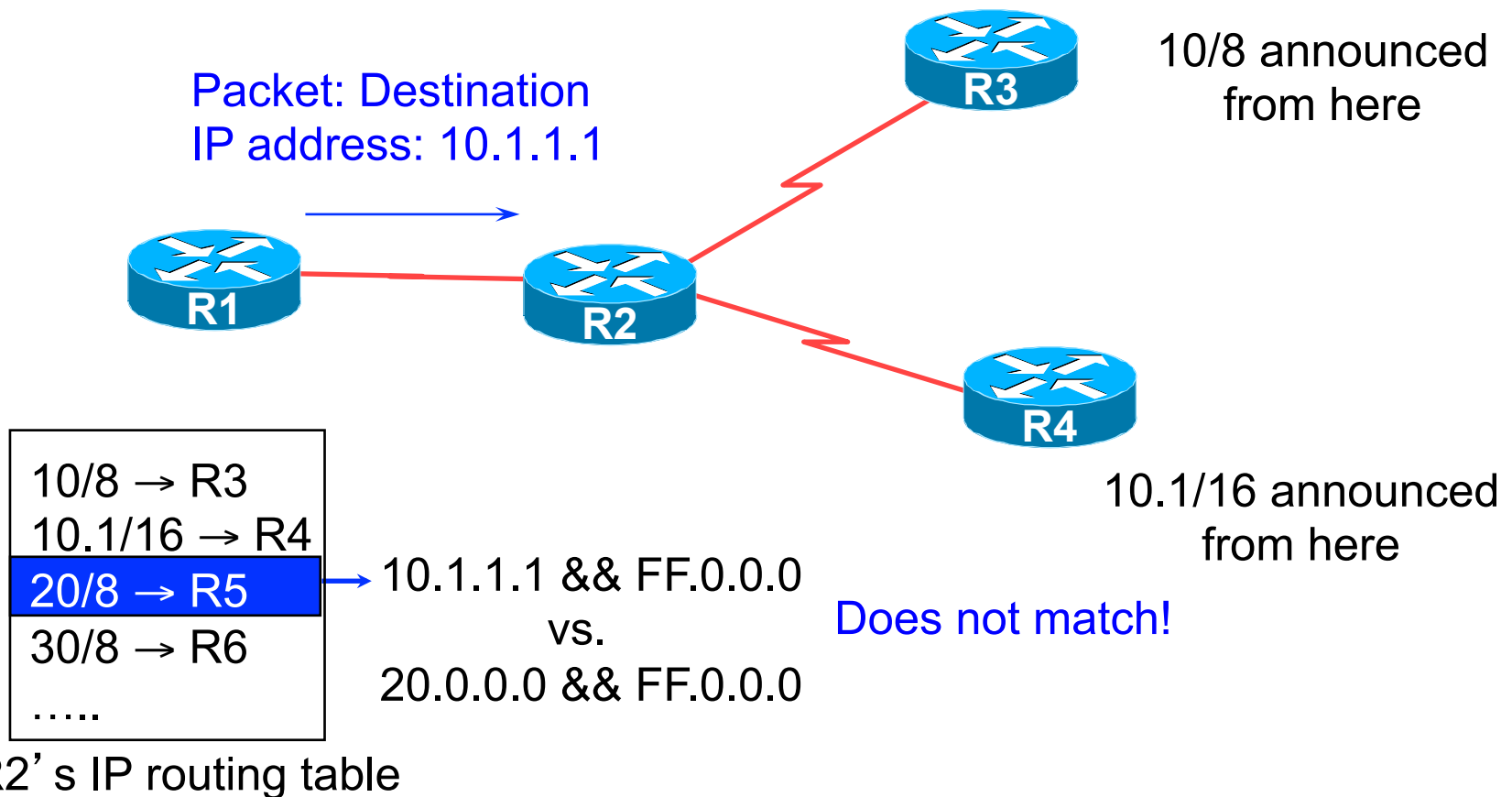
- Based on destination IP address



R2' s IP routing table

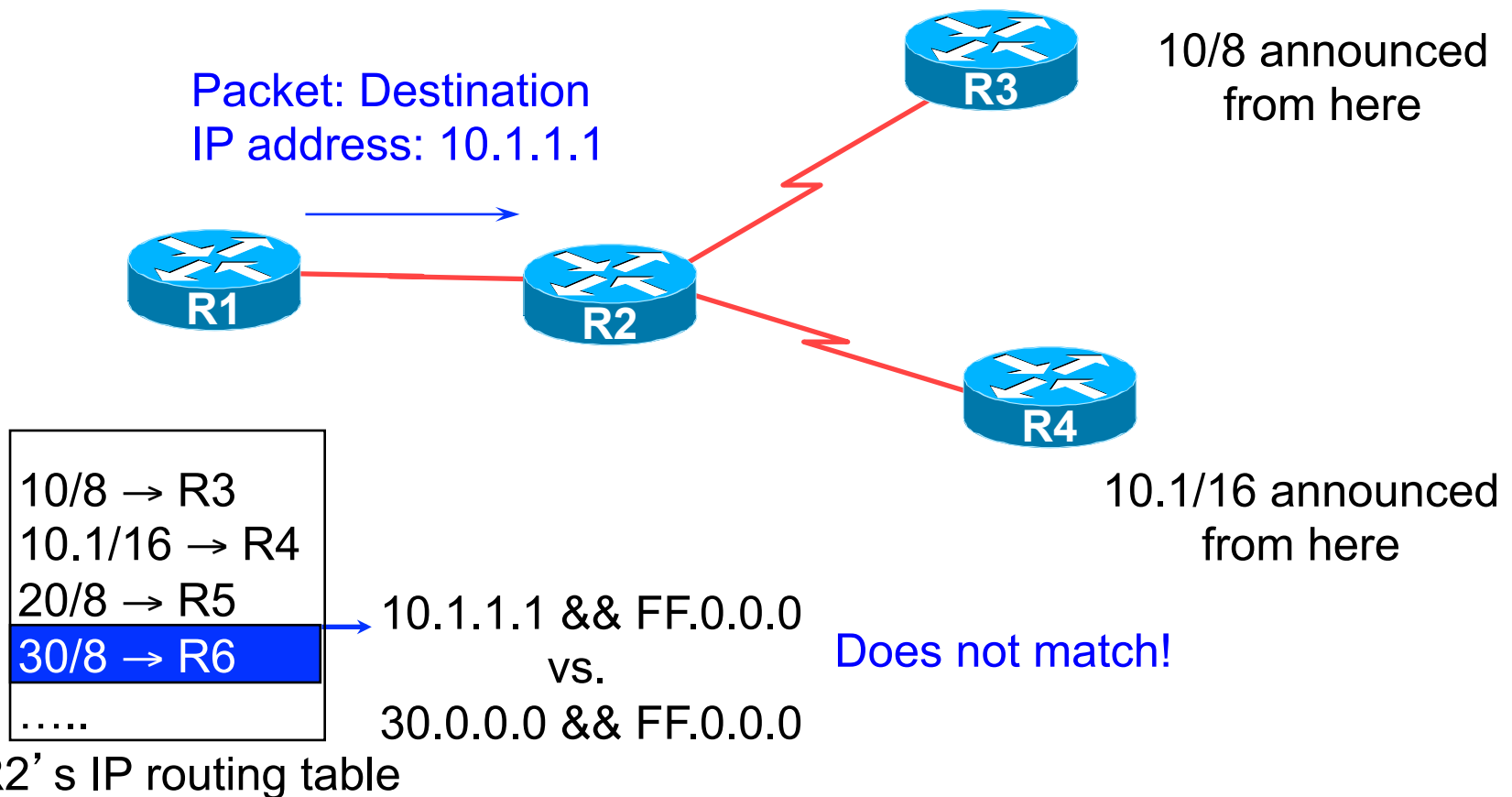
IP route lookup: Longest match routing

- Based on destination IP address



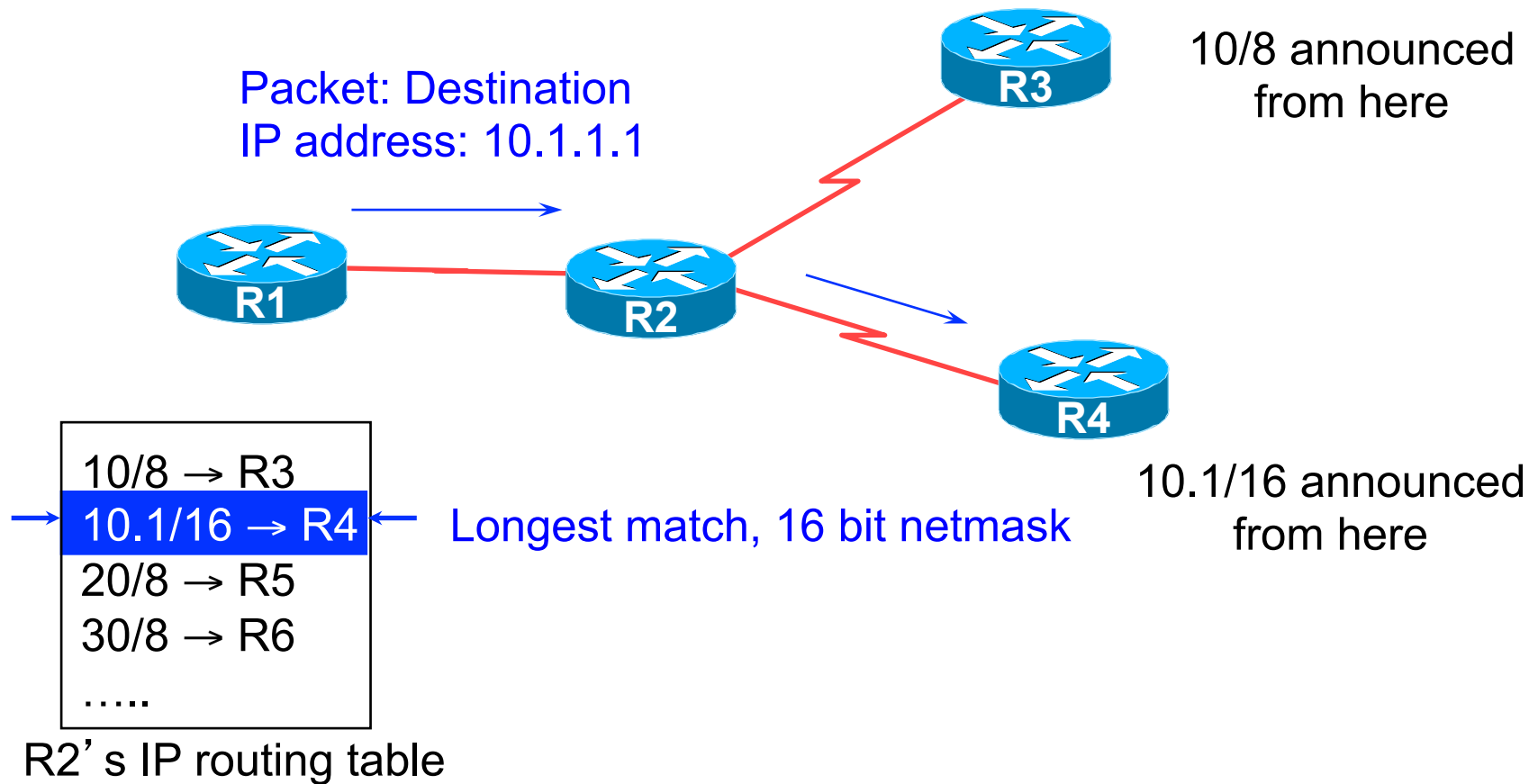
IP route lookup: Longest match routing

- Based on destination IP address



IP route lookup: Longest match routing

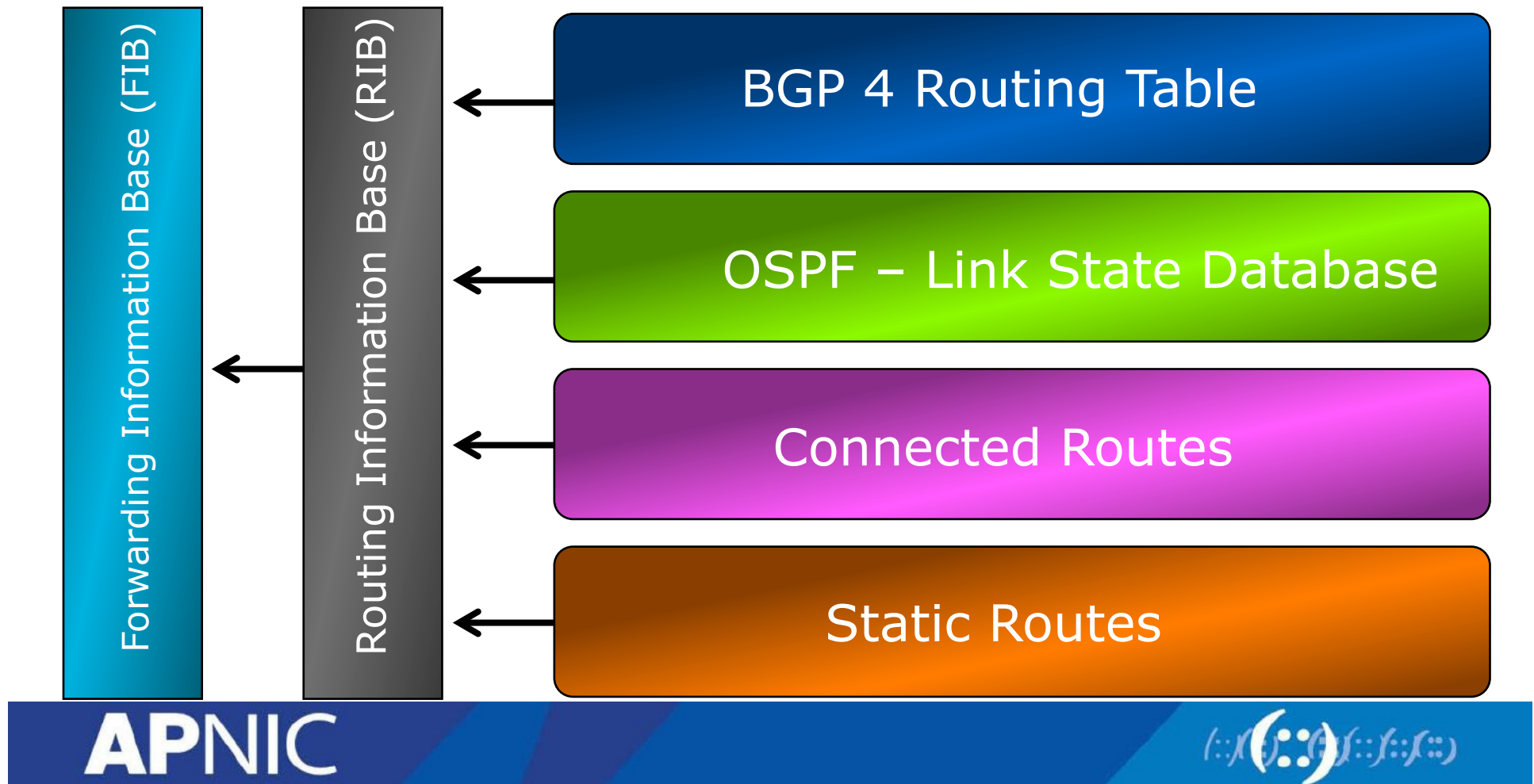
- Based on destination IP address



IP Forwarding

- Router decides which interface a packet is sent to
- Forwarding table populated by routing process
- Forwarding decisions:
 - destination address
 - class of service (fair queuing, precedence, others)
 - local requirements (packet filtering)
- Forwarding is usually aided by special hardware

Routing Tables Feed the Forwarding Table



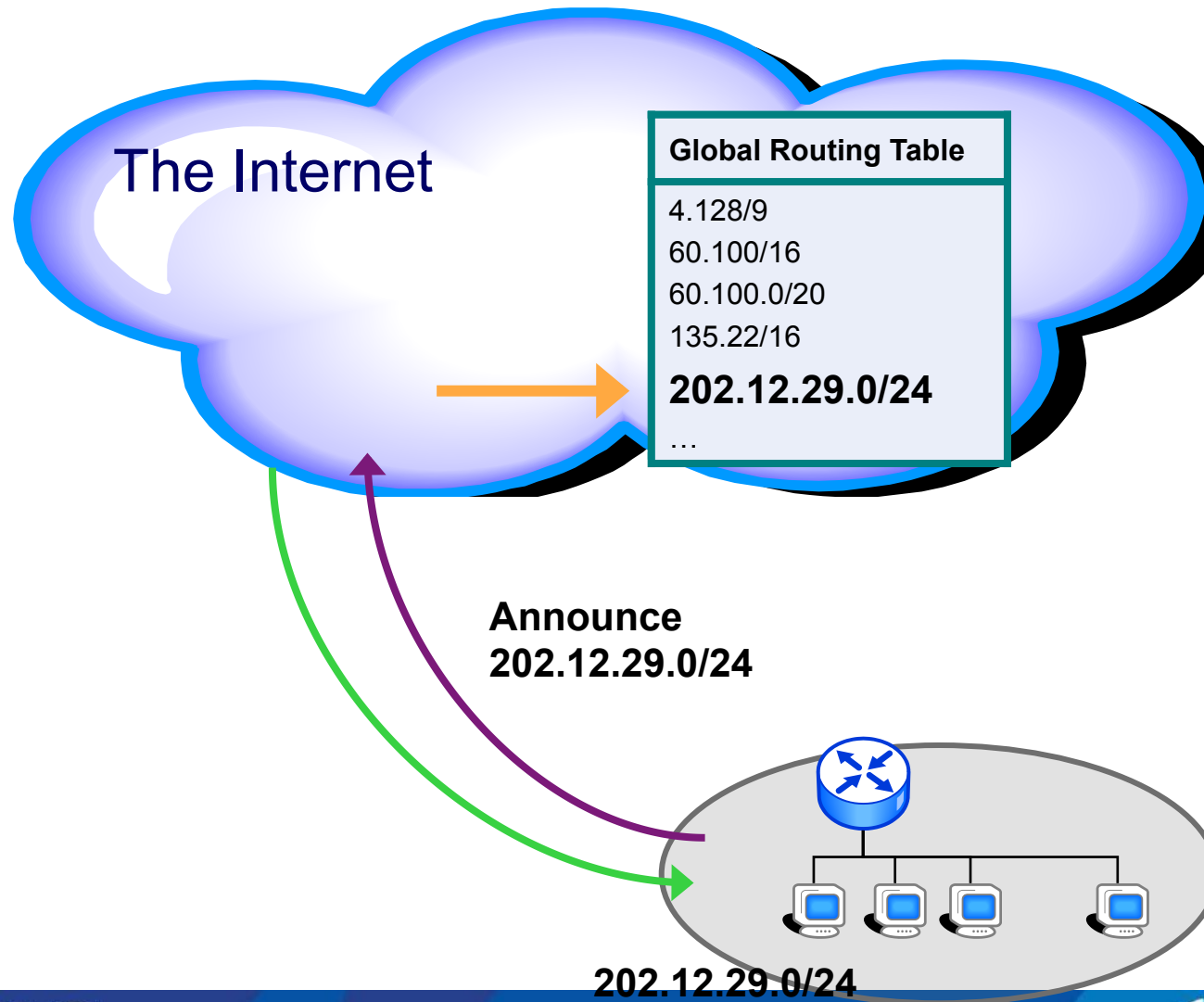
RIBs and FIBs

- FIB is the Forwarding Table
 - It contains destinations and the interfaces to get to those destinations
 - Used by the router to figure out where to send the packet
 - Careful! Some people still call this a route!
- RIB is the Routing Table
 - It contains a list of all the destinations and the various next hops used to get to those destinations – and lots of other information too!
 - One destination can have lots of possible next-hops – only the best next-hop goes into the FIB

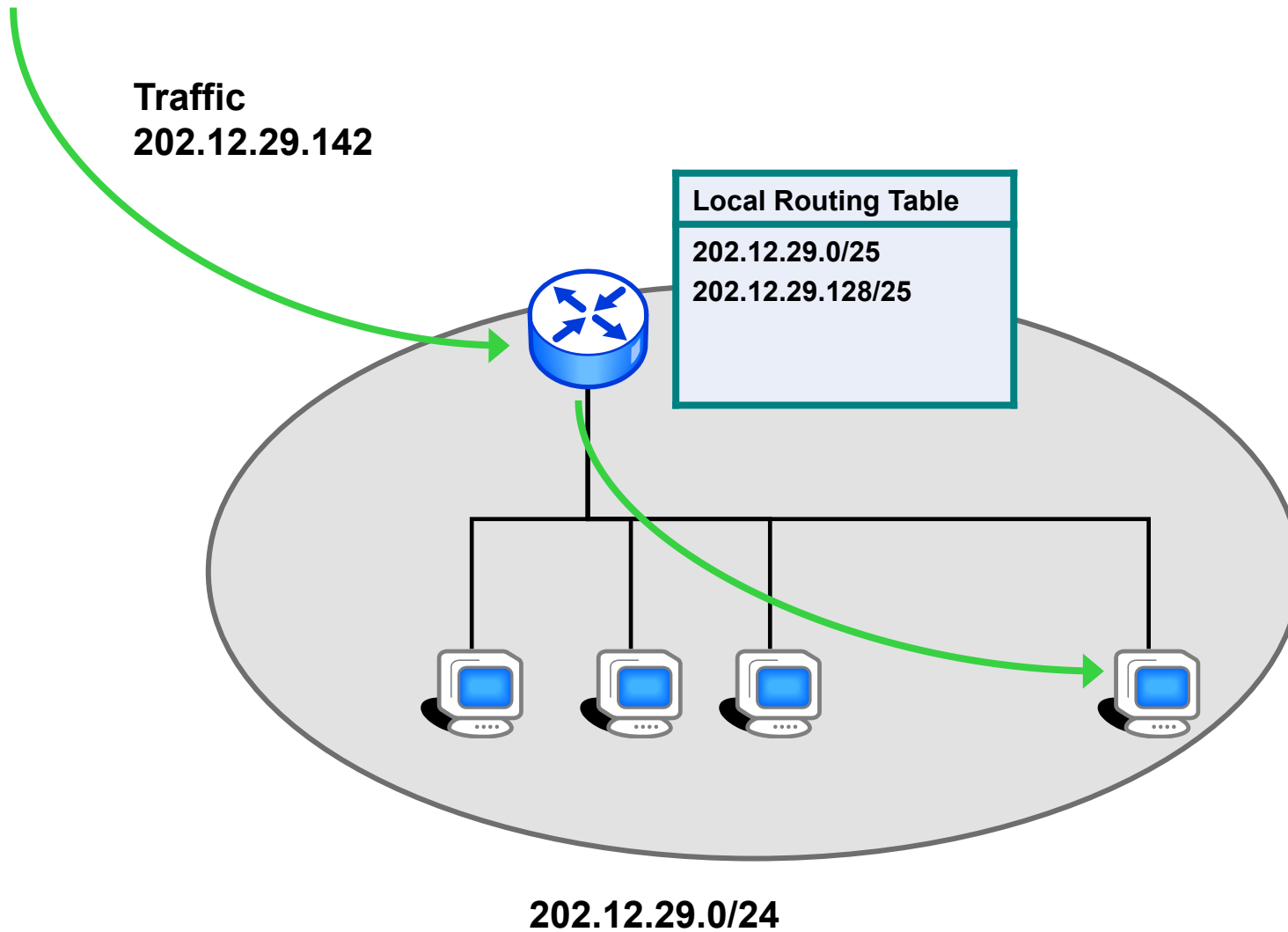
Explicit versus Default Routing

- Default:
 - simple, cheap (cycles, memory, bandwidth)
 - low granularity (metric games)
- Explicit (default free zone)
 - high overhead, complex, high cost, high granularity
- Hybrid
 - minimise overhead
 - provide useful granularity
 - requires some filtering knowledge

Internet Routing



Internet Routing



Egress Traffic

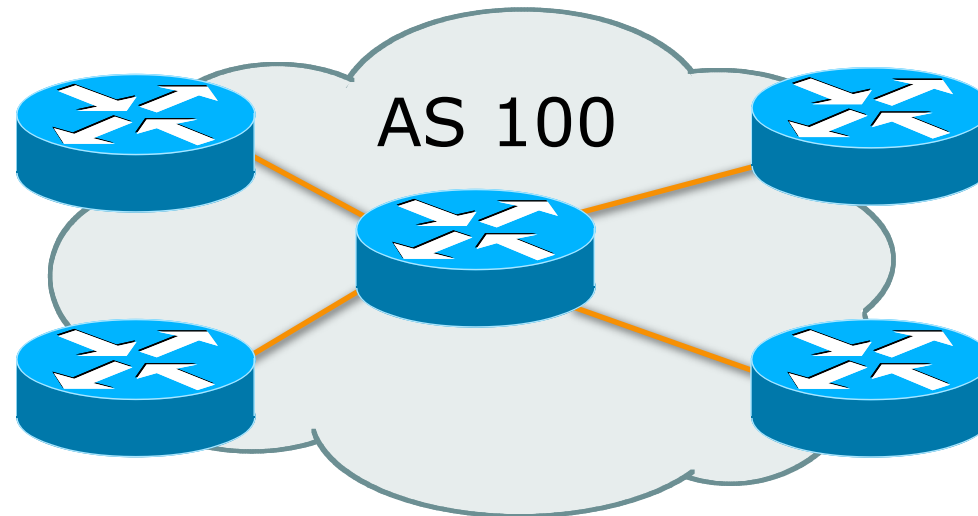
- How packets leave your network
- Egress traffic depends on:
 - route availability (what others send you)
 - route acceptance (what you accept from others)
 - policy and tuning (what you do with routes from others)
 - Peering and transit agreements

Ingress Traffic

- How packets get to your network and your customers' networks
- Ingress traffic depends on:
 - what information you send and to whom
 - based on your addressing and AS' s
 - based on others' policy (what they accept from you and what they do with it)

Autonomous System (AS)

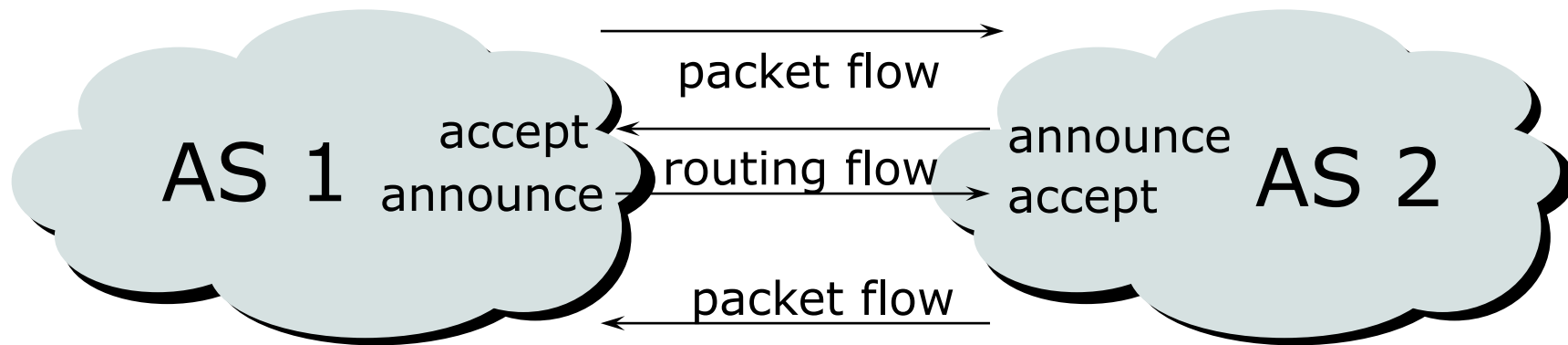
- Collection of networks with same routing policy
- Single routing protocol
- Usually under single ownership, trust and administrative control



Definition of terms

- **Neighbours**
 - AS's which directly exchange routing information
 - Routers which exchange routing information
- **Announce**
 - send routing information to a neighbour
- **Accept**
 - receive and use routing information sent by a neighbour
- **Originate**
 - insert routing information into external announcements (usually as a result of the IGP)
- **Peers**
 - routers in neighbouring AS' s or within one AS which exchange routing and policy information

Routing flow and packet flow



For networks in AS1 and AS2 to communicate:

- AS1 must announce to AS2

- AS2 must accept from AS1

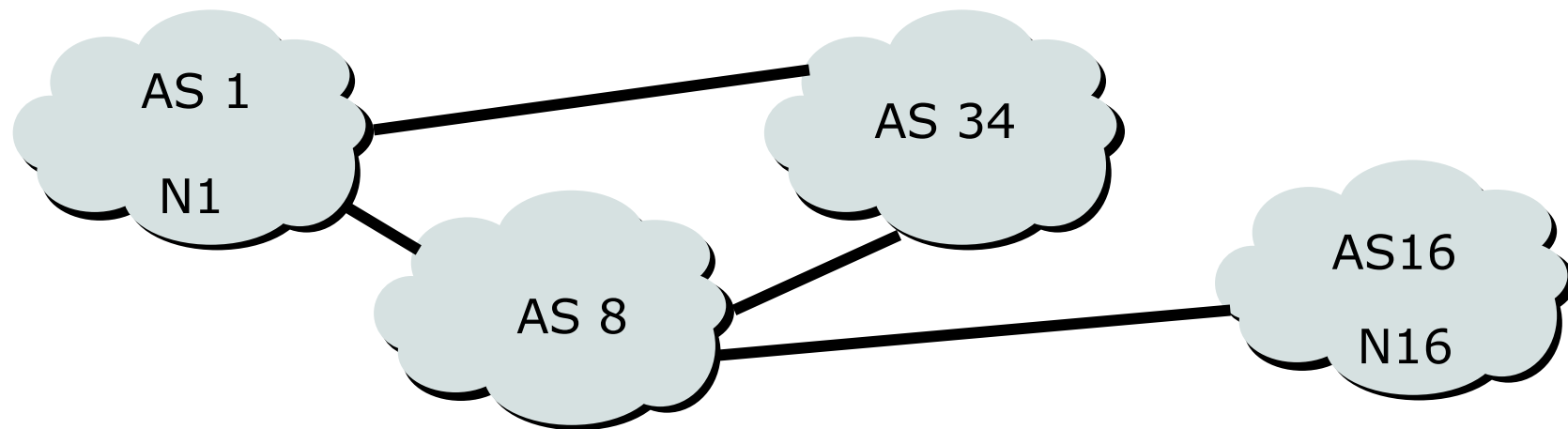
- AS2 must announce to AS1

- AS1 must accept from AS2

Routing flow and Traffic flow

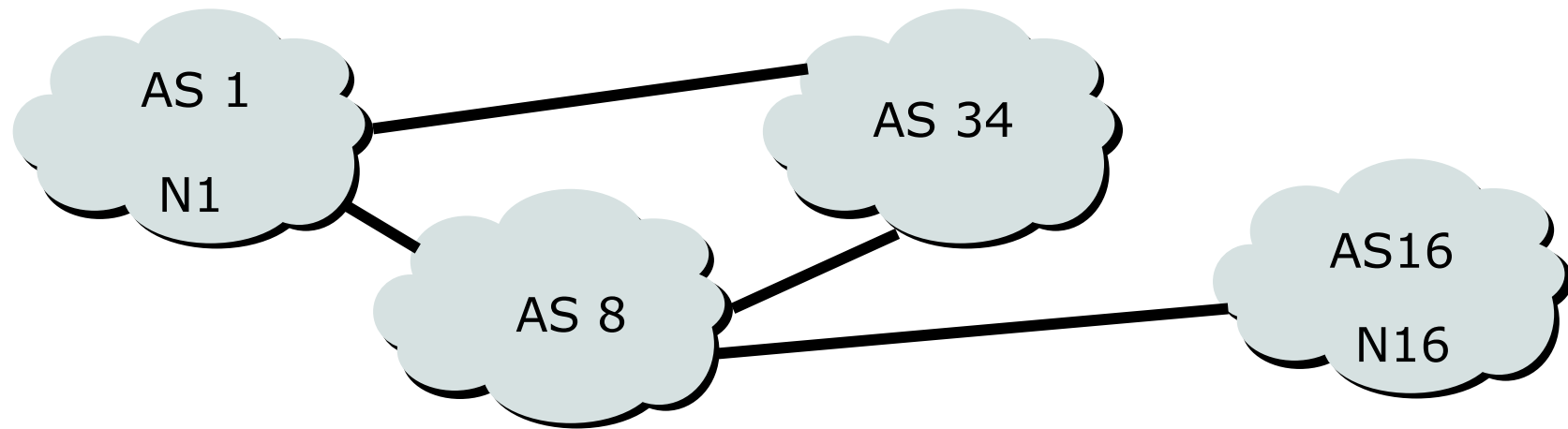
- Traffic flow is always in the opposite direction of the flow of Routing information
 - Filtering outgoing routing information inhibits traffic flow inbound
 - Filtering inbound routing information inhibits traffic flow outbound

Routing Flow/Packet Flow: With multiple ASes



- For net N1 in AS1 to send traffic to net N16 in AS16:
 - AS16 must originate and announce N16 to AS8.
 - AS8 must accept N16 from AS16.
 - AS8 must forward announcement of N16 to AS1 or AS34.
 - AS1 must accept N16 from AS8 or AS34.
- For two-way packet flow, similar policies must exist for N1

Routing Flow/Packet Flow: With multiple ASes

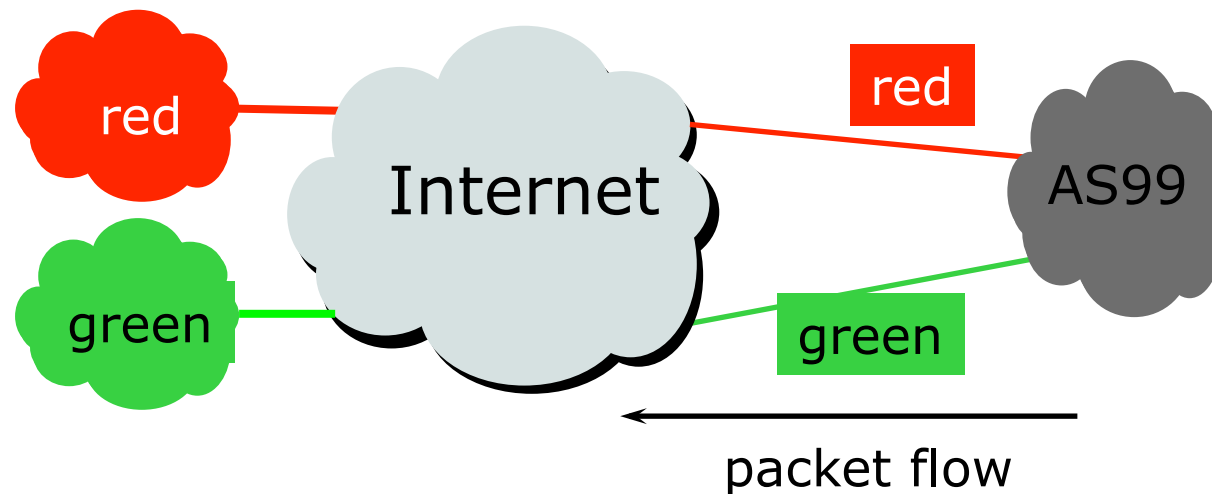


- As multiple paths between sites are implemented it is easy to see how policies can become quite complex.

Routing Policy

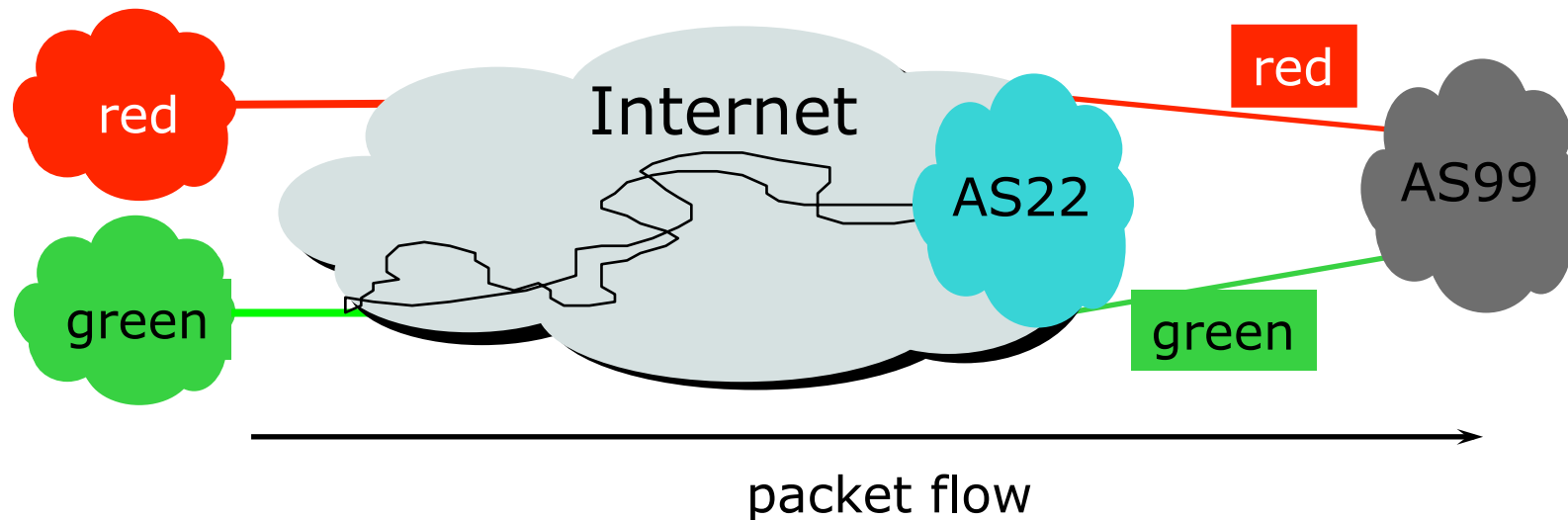
- Used to control traffic flow in and out of an ISP network
- ISP makes decisions on what routing information to accept and discard from its neighbours
 - Individual routes
 - Routes originated by specific ASes
 - Routes traversing specific ASes
 - Routes belonging to other groupings
 - Groupings which you define as you see fit

Routing Policy Limitations



- AS99 uses red link for traffic to the red AS and the green link for remaining traffic
- To implement this policy, AS99 has to:
 - Accept routes originating from the red AS on the red link
 - Accept all other routes on the green link

Routing Policy Limitations



- AS99 would like packets coming from the green AS to use the green link.
- But unless AS22 cooperates in pushing traffic from the green AS down the green link, there is very little that AS99 can do to achieve this aim

Routing Policy Issues

- As of 18 Oct 2013:
 - 482972 prefixes
 - Not realistic to set policy on all of them individually
 - 45591 origin AS' s
 - Too many to try and create individual policies for
- Routes tied to a specific AS or path may be unstable regardless of connectivity
- Solution: Groups of ASes are a natural abstraction for filtering purposes

Questions



Overview

Routing Workshop (3 Days)

- Introduction to IP Routing
- **Routing Protocol Basic**
- IPv6 Address Structure
- Routing Lab Topology Overview
- Operation of OSPF Routing Protocol
- Lab Exercise on Basic Router and OSPF Dynamic Routing Configuration
- Basic BGP Operation
- BGP Attributes and Path Selection Process
- BGP Scaling Techniques
- Lab Exercise on iBGP, eBGP, RR, Peer group etc
- Internet Exchange [IX] Policy Overview and Configuration requirement

1: How Does Routing Work?

- Internet is made up of the ISPs who connect to each other's networks
- How does an ISP in Kenya tell an ISP in Japan what customers they have?
- And how does that ISP send data packets to the customers of the ISP in Japan, and get responses back
 - After all, as in a local Ethernet, two way packet flow is needed for communication between two devices

2: How Does Routing Work?

- ISP in Kenya could buy a direct connection to the ISP in Japan
 - But this **doesn't scale** – thousands of ISPs, would need thousands of connections, and cost would be astronomical
- Instead, ISP in Kenya tells the neighbouring ISPs about his customers
 - The neighbour ISPs pass this information on to their neighbours, and so on
 - This process repeats until the information reaches the ISP in Japan

3: How Does Routing Work?

- This process is called “**Routing**”
- The mechanisms used are called “**Routing Protocols**”
- Routing and Routing Protocols ensure that the Internet can scale, that thousands of ISPs can provide connectivity to each other, giving us the Internet we see today

4: How Does Routing Work?

- ISP in Kenya doesn't actually tell his neighbouring ISPs the names of the customers
 - (network equipment does not understand names)
- Instead, he has received an IP address block as a member of the Regional Internet Registry serving Kenya
 - His customers have received address space from this address block as part of their “Internet service”
 - And he announces this address block to his neighbouring ISPs – this is called announcing a “route”

What is a Routing Protocol?

- A set of rules defined to facilitate the exchange of routing information between routers (Layer 3 device) inside networks
- Builds routing tables dynamically based on updates from its neighbours
 - Allows the router to find the best path in a network that has more than one path to a remote network.
- Maintains connectivity between devices within the network.

Distinction between *Routed* and *Routing* Protocols

- Routed protocols
 - Layer3 datagram that carry the information required in transporting the data across the network
- Routing protocols
 - Handles the updating requirement of the routers within the network for determining the path of the datagram across the network

Routing and Routed Protocols

Routed protocol	Routing protocol
AppleTalk	RTMP, AURP, EIGRP
IPX	RIP, NLSP, EIGRP
Vines	RTP
DecNet IV	DecNet
IP	RIPv2, OSPF, IS-IS, BGP and (Cisco Systems proprietary) EIGRP

Routing Information

A routing table entry must contain the following information:

- Network field
- Outgoing interface
- Metric field
- Next-hop field

Network Field

- Contains information of entries
 - Networks learned (destination logical network or subnets)
 - Manually (static or default routes)
 - Dynamically (learned from routing protocol as dynamic routes)
- Information recorded is the entry on where to forward traffic to its destination when the datagram is received.

Outgoing Interface Field

- Interface to where the router sends the datagram
- Informs the administrator of interface where the update came through

Metric Field

- Used to determine which path to use if there are multiple paths to the remote network
- Provides the value to select the best path
- But take note of the administrative distance selection process

Routing Protocol Metrics

Routing protocol	Metric
RIPv2	Hop count
EIGRP	Bandwidth, delay, load, reliability, MTU
OSPF	Cost (the higher the bandwidth indicates a lowest cost)
IS-IS	Cost

Next Hop Field

- Contains the destination address of the next forwarding router
 - Address of the next hop (outgoing interface) usually within the same subnet
 - iBGP (exemption to the rule)
- Identifies the next hop so that the router can create the Layer2 frame with the destination address

Routing Table Sample (Cisco)

```
Cisco-router#sh ip route
```

```
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B -  
BGP
```

```
D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
```

```
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
```

```
E1 - OSPF external type 1, E2 - OSPF external type 2, E - EGP
```

```
i - IS-IS, L1 - IS-IS level-1, L2 - IS-IS level-2, ia - IS-IS
```

```
inter area
```

```
* - candidate default, U - per-user static route, o - ODR
```

```
P - periodic downloaded static route
```

```
Gateway of last resort is not set ??????????????
```

```
202.41.143.0/24 is variably subnetted, 2 subnets, 2 masks
```

```
S 202.41.143.0/24 is directly connected, Null0
```

```
C 202.41.143.17/32 is directly connected, Loopback0
```

```
O E2 10.110.0.0 [110/5] via 10.119.254.6, 0:01:00, Ethernet2
```

```
O 10.67.10.0 [110/128] via 10.119.254.244, 0:02:22, Ethernet2
```

```
B 217.170.115.0/24 [20/0] via 12.123.29.249, 5d16h
```



Sample only

Types of Routing Protocols

- Routing protocols are essentially applications inside the router designed to ensure correct and timely exchange of information within the network
- The IP routing protocol has several distinctions which can be divided into different groups
 - The first in the group is how the routing protocol handles the subnet mask sent during the routing update
 - The early routing protocols don't support this but the newer ones like RIP2, OSPF, ISIS, BGP4 support it. The distinctions are called “**classful** and **classless**”

Classful Routing Protocol (**obsolete**)

- Periodic updates are done by the routing protocol
 - does not carry out subnet or routing masks because the assumption is always based on network bit boundaries
 - does not support VLSM which makes it inefficient for addressing the network
- This has been **obsolete** for a long time but for knowledge purposes there are two protocols designed for it.
 - RIPv1
 - IGRP
- It has created constraints to IP network design due to its limitations

Classless Routing Protocol

- Classless routing protocol was designed to overcome the constraints from classful routing
 - OSPF
 - IS-IS
 - EIGRP (Cisco)
 - BGP
 - lead also to the development of RIPv2

Distance Vector Routing Protocol

- Another distinction based on the behavior of learning the path to the destination network
 - Distance vector and
 - Link-State protocols
- Distance vector routing protocol
 - Early technology of IP routing protocols (RIPv2)
 - Concept design was for small networks

Link-State Routing Protocol

- Link-state routing protocol was designed to overcome the limitation of Distance Vector routing protocol. The main goal is to achieve the following:
 - Maintain a loop-free and accurate table
 - Utilises multicast address and make updated based on incremental
 - Fast convergence of the network
 - Reduce the network overhead during updates
 - Selection of best path based on link status
- Routing protocols that support link-state are:
 - OSPF
 - IS-IS

Routing Protocols

- Routers use “routing protocols” to exchange routing information with each other
 - **IGP** is used to refer to the process running on routers inside an ISP's network
 - **EGP** is used to refer to the process running between routers bordering directly connected ISP networks

What Is an IGP?

- Interior Gateway Protocol
- Within an Autonomous System
- Carries information about internal infrastructure prefixes
- Two widely used IGPs in service provider network:
 - OSPF
 - ISIS

Why Do We Need an IGP?

- ISP backbone scaling
 - Hierarchy
 - Limiting scope of failure
 - Only used for ISP's **infrastructure** addresses, not customers or anything else
 - Design goal is to **minimize** number of prefixes in IGP to aid scalability and rapid convergence

What Is an EGP?

- Exterior Gateway Protocol
- Used to convey routing information between Autonomous Systems
- De-coupled from the IGP
- Current EGP is BGP

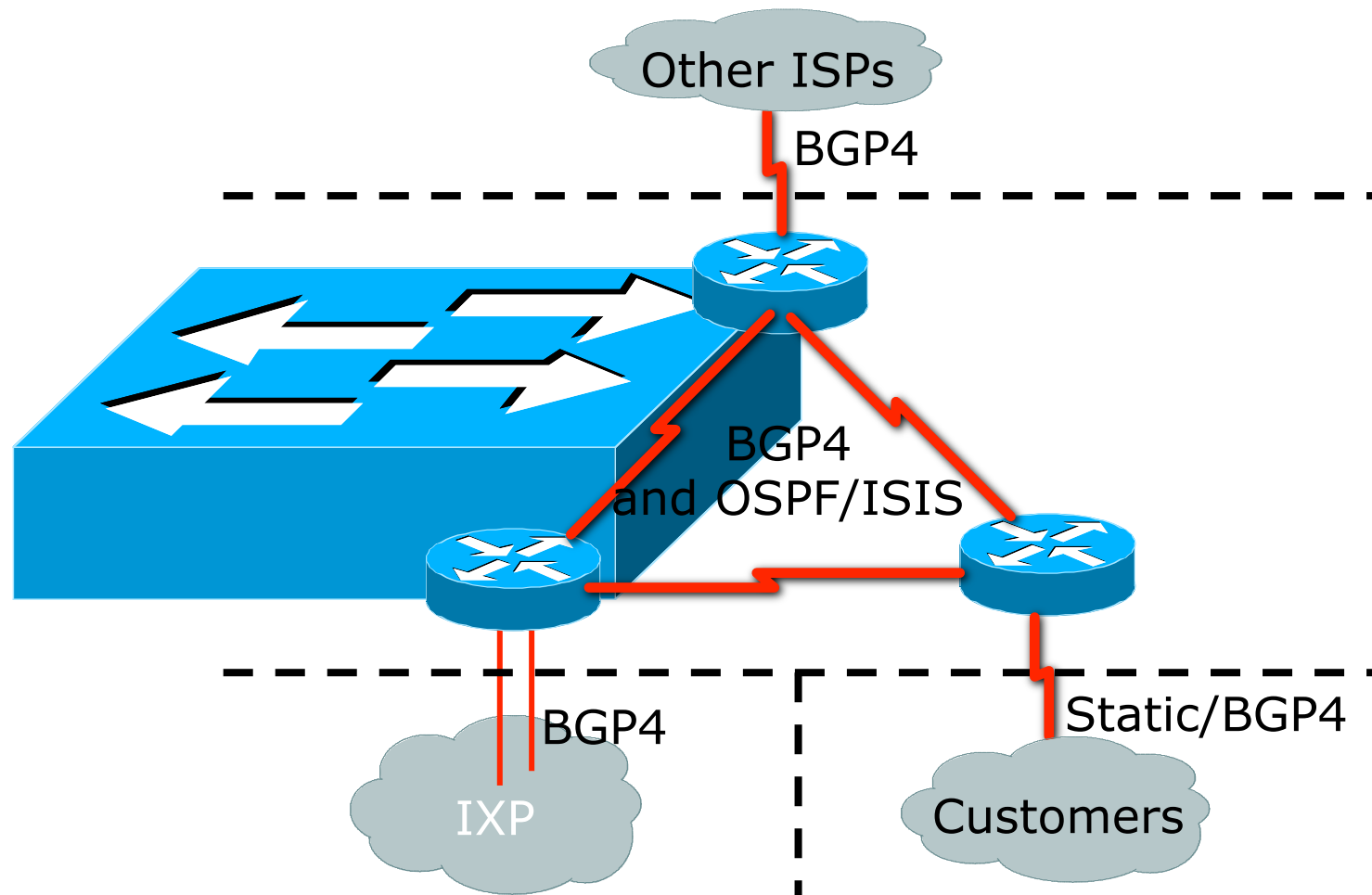
Why Do We Need an EGP?

- Scaling to large network
 - Hierarchy
 - Limit scope of failure
- Define Administrative Boundary
- Policy
 - Control reachability of prefixes
 - Merge separate organisations
 - Connect multiple IGPs

Interior versus Exterior Routing Protocols

- Interior
 - Automatic neighbour discovery
 - Generally trust your IGP routers
 - Prefixes go to all IGP routers
 - Binds routers in one AS together
 - Carries ISP infrastructure addresses only
 - ISPs aim to keep the IGP small for efficiency and scalability
- Exterior
 - Specifically configured peers
 - Connecting with outside networks
 - Set administrative boundaries
 - Binds AS's together
 - Carries customer prefixes
 - Carries Internet prefixes
 - EGPs are independent of ISP network topology

Hierarchy of Routing Protocols



Administrative Distance

- The method used for selection of route priority of IP routing protocol. The lowest administrative distance is preferred.
 - Manually entered routes are preferred over dynamically learned routes
 - Static routes
 - Default routes
 - Dynamically learned routes depend on the routing protocol metric calculation algorithm. For default metric values, the smallest metric value is preferred.

FYI: Cisco IOS Default Administrative Distances

Route Source	Default Distance
Connected Interface	0
Static Route	1
Enhanced IGRP Summary Route	5
External BGP	20
Internal Enhanced IGRP	90
IGRP	100
OSPF	110
IS-IS	115
RIP	120
EGP	140
External Enhanced IGRP	170
Internal BGP	200
Unknown	255

Overview

Routing Workshop (3 Days)

- Introduction to IP Routing
- Routing Protocol Basic
- **IPv6 Address Structure**
- Routing Lab Topology Overview
- Operation of OSPF Routing Protocol
- Lab Exercise on Basic Router and OSPF Dynamic Routing Configuration
- Basic BGP Operation
- BGP Attributes and Path Selection Process
- BGP Scaling Techniques
- Lab Exercise on iBGP, eBGP, RR, Peer group etc
- Internet Exchange [IX] Policy Overview and Configuration requirement

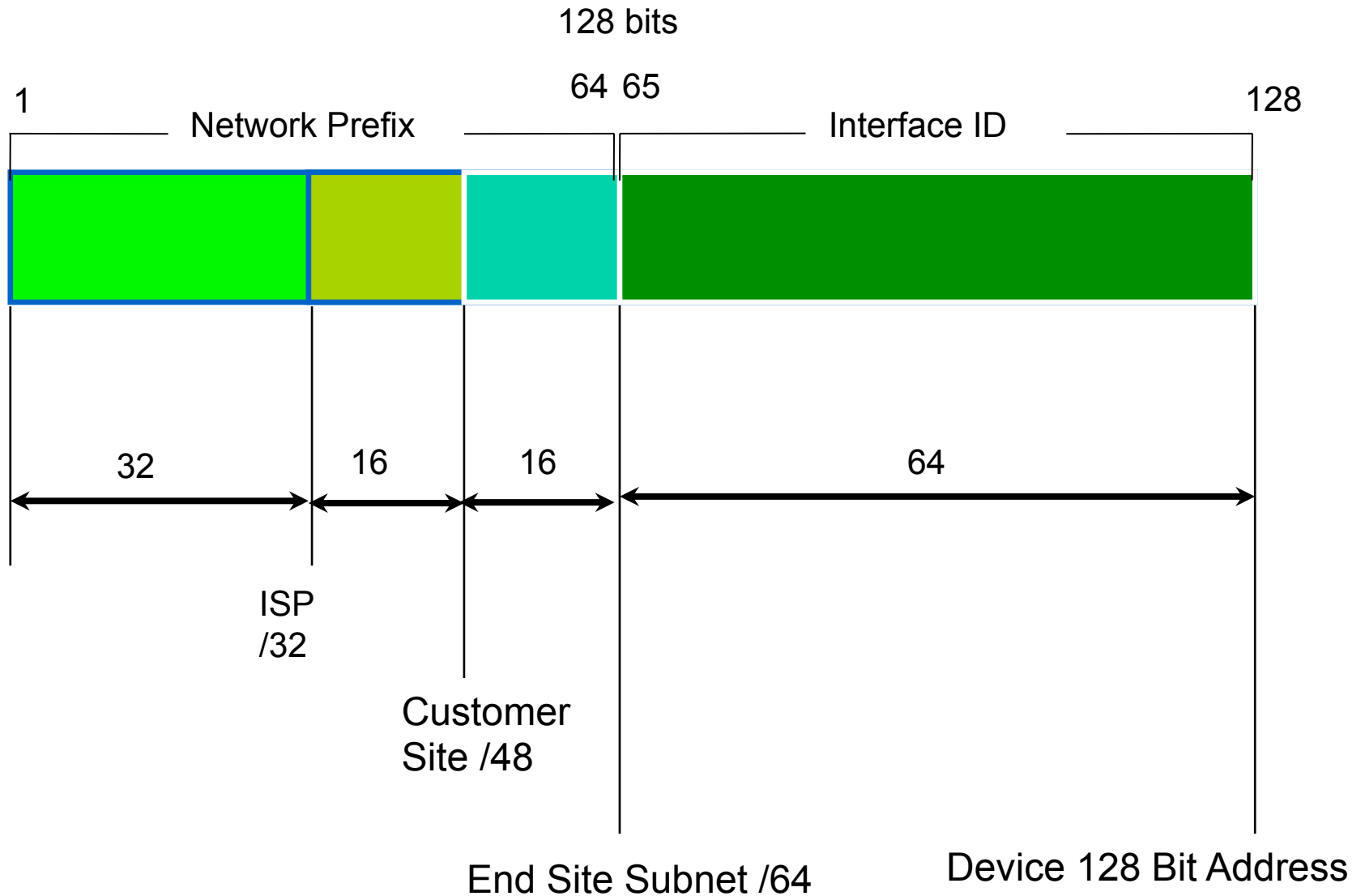
IPv6 Addressing

- An IPv6 address is 128 bits long
- So the number of addresses are 2^{128}
=340282366920938463463374607431768211455
(39 decimal digits)
=0xffffffffffffffffffffffffffffffff (32 hexadecimal digits)
- In hex 4 bit (nibble) is represented by a hex digit
- So 128 bit is reduced down to 32 hex digit

IPv6 Address Representation

- Hexadecimal values of eight 16 bit fields
 - X:X:X:X:X:X:X:X (X=16 bit number, ex: A2FE)
 - 16 bit number is converted to a 4 digit hexadecimal number
- Example:
 - FE38:DCE3:124C:C1A2:BA03:6735:EF1C:683D
 - Abbreviated form of address
 - 4EED:0023:0000:0000:0000:036E:1250:2B00
 - →4EED:23:0:0:0:36E:1250:2B00
 - →4EED:23::36E:1250:2B00
 - (Null value can be used only once)

IPv6 addressing structure



IPv6 addressing model

- **IPv6 Address type**

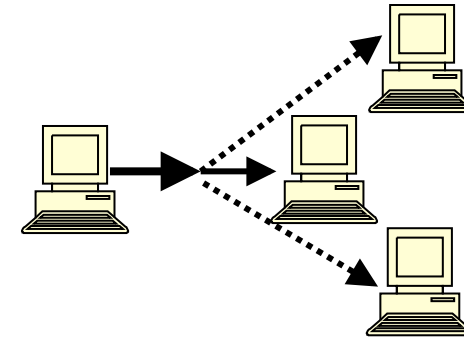
- Unicast

- An identifier for a single interface



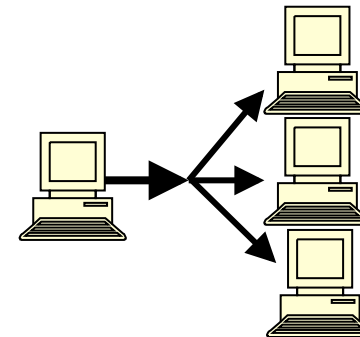
- Anycast

- An identifier for a set of interfaces



- Multicast

- An identifier for a group of nodes



Addresses Without a Network Prefix

- Localhost `::1/128`
- Unspecified Address `::/128`
- IPv4-mapped IPv6 address `::ffff/96 [a.b.c.d]`
- IPv4-compatible IPv6 address `::/96 [a.b.c.d]`

Local Addresses With Network Prefix

- Link Local Address
 - A special address used to communicate within the local link of an interface
 - i.e. anyone on the link as host or router
 - This address in packet destination that packet would never pass through a router
 - fe80::/10

Local Addresses With Network Prefix

- Unique Local IPv6 Unicast Address
 - Addresses similar to the RFC 1918 / private address like in IPv4 but will ensure uniqueness
 - A part of the prefix (40 bits) are generated using a pseudo-random algorithm and it's improbable that two generated ones are equal
 - fc00::/7
 - Example webtools to generate ULA prefix
 - <http://www.sixxs.net/tools/grh/ula/>
 - <http://www.goebel-consult.de/ipv6/createLULA>

Global Addresses With Network Prefix

- IPV6 Global Unicast Address

- Global Unicast Range: 0010 2000::/3

- 0011 3000::/3

- All five RIRs are given a /12 from the /3 to further distribute within the RIR region

- APNIC 2400:0000::/12
 - ARIN 2600:0000::/12
 - AfriNIC 2C00:0000::/12
 - LACNIC 2800:0000::/12
 - Ripe NCC 2A00:0000::/12

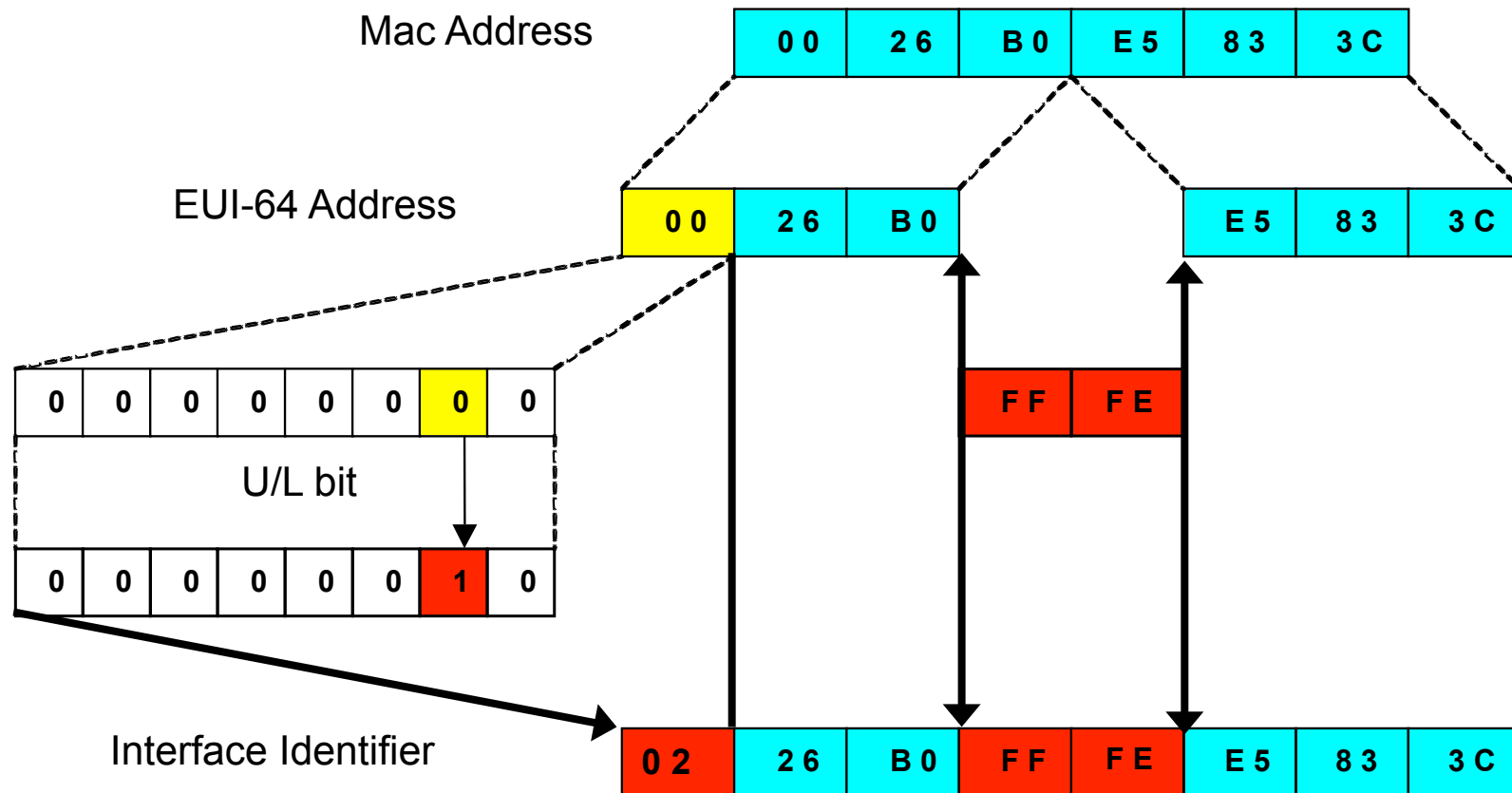
Examples and Documentation Prefix

- Two address ranges are reserved for examples and documentation purpose by RFC 3849
 - For example 3fff:ffff::/32
 - For documentation 2001:0DB8::/32

Interface ID

- The lowest-order 64-bit field addresses may be assigned in several different ways:
 - auto-configured from a 48-bit MAC address expanded into a 64-bit EUI-64
 - assigned via DHCP
 - manually configured
 - auto-generated pseudo-random number
 - possibly other methods in the future

EUI-64



Questions



Overview

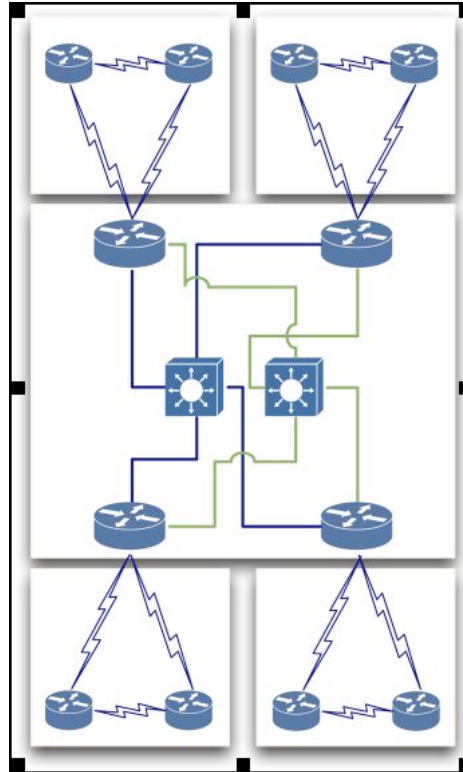
Routing Workshop (3 Days)

- Introduction to IP Routing
- Routing Protocol Basic
- IPv6 Address Structure
- **Routing Lab Topology Overview**
- Operation of OSPF Routing Protocol
- Lab Exercise on Basic Router and OSPF Dynamic Routing Configuration
- Basic BGP Operation
- BGP Attributes and Path Selection Process
- BGP Scaling Techniques
- Lab Exercise on iBGP, eBGP, RR, Peer group etc
- Internet Exchange [IX] Policy Overview and Configuration requirement

Training ISP Network Topology

- Scenario:
 - Training ISP has 4 main operating area or region
 - Each region has 2 small POP
 - Each region will have one datacenter to host content
 - Regional network are inter-connected with multiple link

Training ISP Network Topology



Training ISP Topology Diagram

Training ISP Network Topology

- Regional Network:
 - Each regional network will have 3 routers
 - 1 Core & 2 Edge Routers
 - 2 Point of Presence (POP) for every region
 - POP will use a router to terminate customer network i.e Edge Router
 - Each POP is an aggregation point of ISP customer

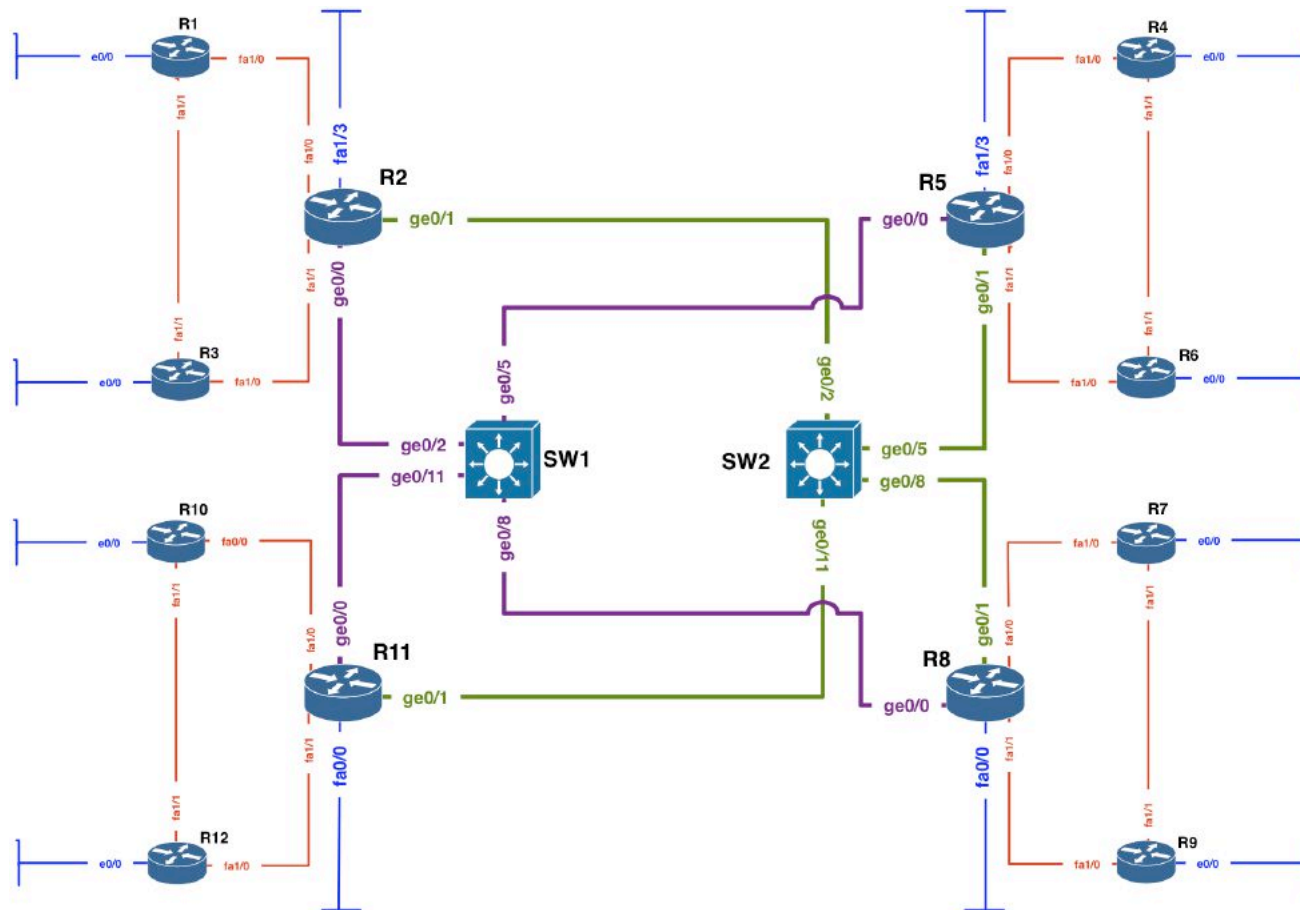
Training ISP Network Topology

- Access Network:
 - Connection between customer network & Edge router
 - Usually 10 to 100 MBPS link
 - Separate routing policy from most of ISP
 - Training ISP will connect them on edge router with separate customer IP prefix

Training ISP Network Topology

- Transport Link:
 - Inter-connection between regional core router
 - Higher data transmission capacity than access link
 - Training ISP has 2 transport link for link redundancy
 - 2 Transport link i.e Purple link & Green link are connected to two career grade switch

Training ISP Network Topology



Training ISP Core IP Backbone

Training ISP Network Topology

- Design Consideration:
 - Each regional network should have address summarization capability for customer block and CS link WAN.
 - Prefix planning should have scalability option for next couple of years for both customer block and infrastructure
 - No Summarization require for infrastructure WAN and loopback address

Training ISP Network Topology

- Design Consideration:
 - All WAN link should be ICMP reachable for link monitoring purpose (At least from designated host)
 - Conservation will get high preference for IPv4 address planning and aggregation will get high preference for IPv6 address planning.

Training ISP Network Topology

- Design Consideration:
 - OSPF is running in ISP network to carry infrastructure IP prefix
 - Each region is a separate OSPF area
 - Transport core is in OSPF area 0
 - Customer will connect on either static or eBGP (Not OSPF)
 - iBGP will carry external prefix within ISP core IP network

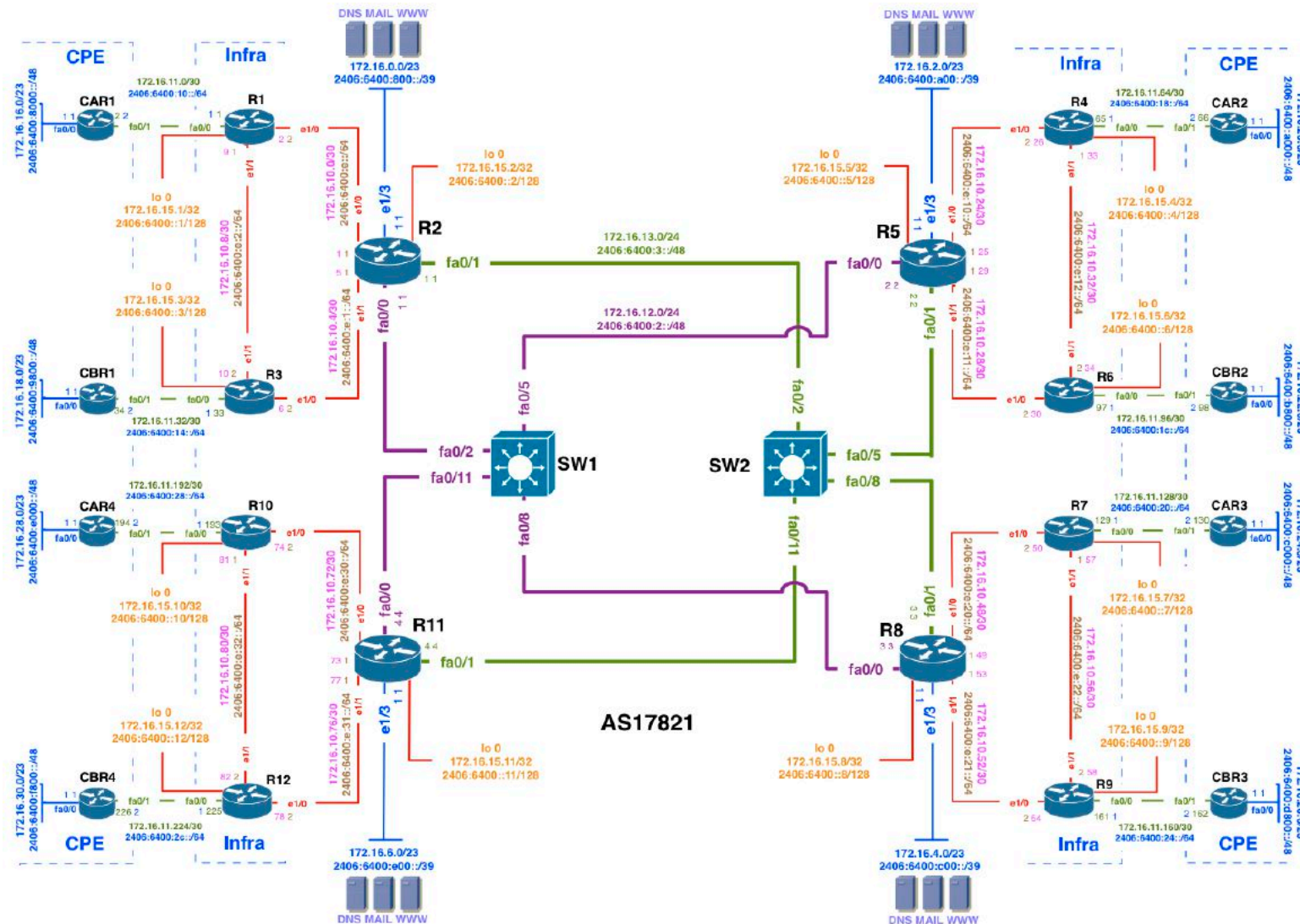
Training ISP IPV6 Addressing Plan

- IPv6 address plan consideration:
 - Big IPv6 address space can cause very very large routing table size
 - Most transit service provider apply IPv6 aggregation prefix filter (i.e. anything other than /48 & \leq /32 prefix size
 - Prefix announcement need to send to Internet should be either /32 or /48 bit boundary

Training ISP IPV6 Addressing Plan

- IPv6 address plan consideration (RFC3177):
 - WAN link can be used on /64 bit boundary
 - End site/Customer sub allocation can be made between /48~/64 bit boundary
 - APNIC Utilization/HD ratio will be calculated based on /56 end site assignment/sub-allocation

Training ISP IPV6 Addressing Plan



Addressing Plans – ISP Infrastructure

- What about LANs?
 - /64 per LAN
- What about Point-to-Point links?
 - Protocol design expectation is that /64 is used
 - /127 now recommended/standardised
 - <http://www.rfc-editor.org/rfc/rfc6164.txt>
 - (reserve /64 for the link, but address it as a /127)
 - Other options:
 - /126s are being used (mirrors IPv4 /30)
 - /112s are being used
 - Leaves final 16 bits free for node IDs
 - Some discussion about /80s, /96s and /120s too

Addressing Plans – ISP Infrastructure

- ISPs should receive /32 from their RIR
- Address block for router loop-back interfaces
 - Generally number all loopbacks out of one /48
 - /128 per loopback
- Address block for infrastructure
 - /48 allows 65k subnets
 - /48 per region (for the largest international networks)
 - /48 for whole backbone (for the majority of networks)
 - Summarise between sites if it makes sense

Addressing Plans – Customer

- Customers get one /48
 - Unless they have more than 65k subnets in which case they get a second /48 (and so on)
- In typical deployments today:
 - Several ISPs give small customers a /56 or single LAN end-sites a /64, e.g.:
 - /64 if end-site will only ever be a LAN
 - /56 for medium end-sites (e.g. small business)
 - /48 for large end-sites
 - (This is another very active discussion area)

Addressing Plans – Advice

- Customer address assignments should not be reserved or assigned on a per PoP basis
 - Same principle as for IPv4
 - ISP iBGP carries customer nets
 - Aggregation within the iBGP not required and usually not desirable
 - Aggregation in eBGP is very necessary
- Backbone infrastructure assignments:
 - Number out of a single /48
 - Operational simplicity and security
 - Aggregate to minimise size of the IGP

Addressing Plans Planning

- Registries will usually allocate the next block to be contiguous with the first allocation
 - Minimum allocation is /32
 - Very likely that subsequent allocation will make this up to a /31
 - So plan accordingly

Example Address Plan

- IPv6 Allocation From Registry is
 - 2406:6400::/32
- IPv4 Allocation From Registry is
 - 172.16.0.0/19

Training ISP IPV6 Addressing Plan

Table 1: Top level distribution infrastructure & customer

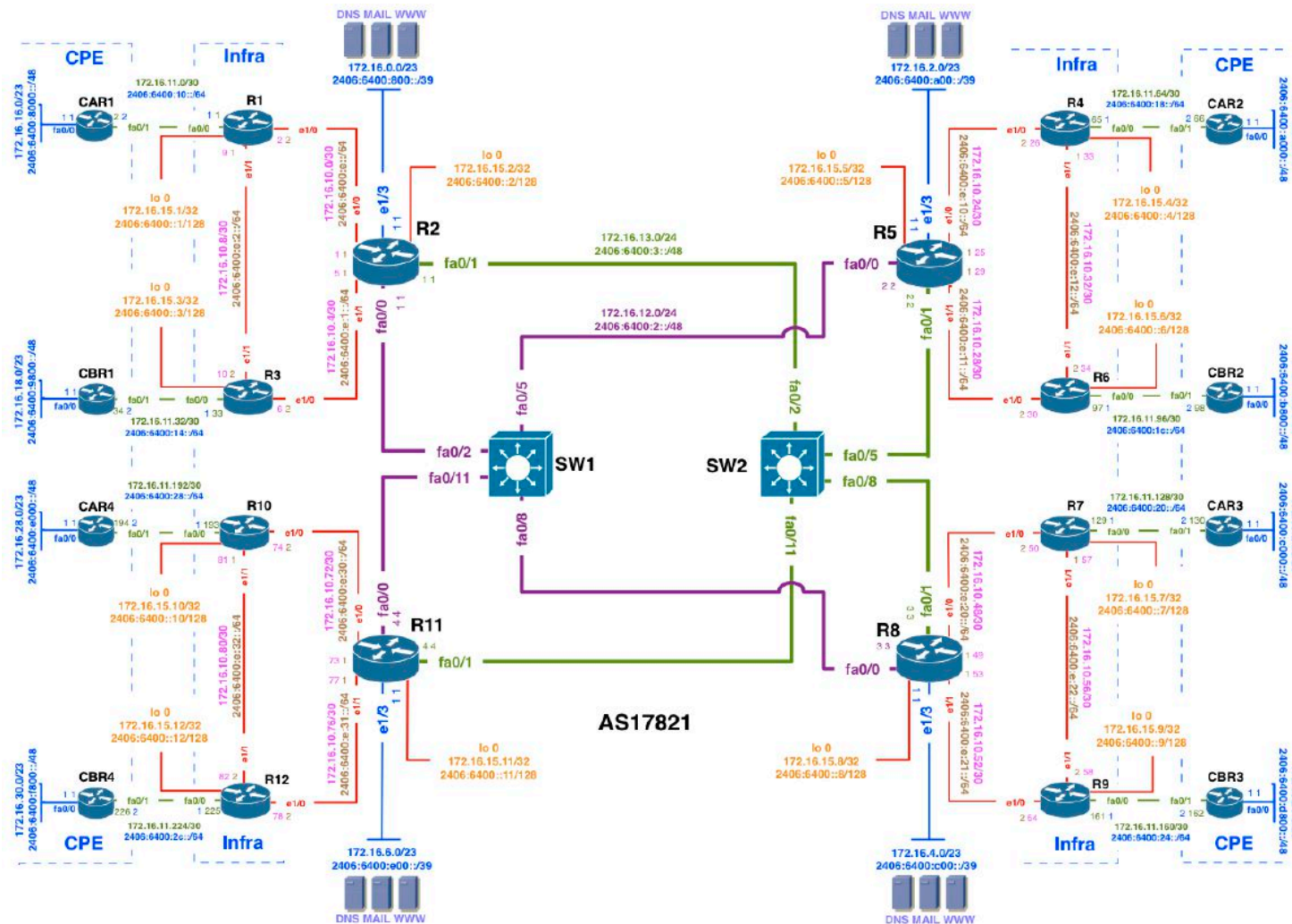
Block#	Prefix	Description	Reverse Domain	SOR	Registration
1	2406:6400::/32	Parent Block	0.0.4.6.6.0.4.2.ip6.arpa.	N/A	APNIC
2	2406:6400:0000:0000::/36	Infrastructure	0.0.0.4.6.6.0.4.2.ip6.arpa.	No	Optional
	2406:6400:1000:0000::/36				
	2406:6400:2000:0000::/36				
	2406:6400:3000:0000::/36				
	2406:6400:4000:0000::/36				
	2406:6400:5000:0000::/36				
	2406:6400:6000:0000::/36				
	2406:6400:7000:0000::/36				
3	2406:6400:8000:0000::/36	Customer network Region 1	8.0.0.4.6.6.0.4.2.ip6.arpa.	Not yet	Optional
	2406:6400:9000:0000::/36				
4	2406:6400:a000:0000::/36	Customer network Region 2	a.0.0.4.6.6.0.4.2.ip6.arpa.	Not yet	Optional
	2406:6400:b000:0000::/36				
5	2406:6400:c000:0000::/36	Customer network Region 3	c.0.0.4.6.6.0.4.2.ip6.arpa.	Not yet	Optional
	2406:6400:d000:0000::/36				
6	2406:6400:e000:0000::/36	Customer network Region 4	e.0.0.4.6.6.0.4.2.ip6.arpa.	Not yet	Optional
	2406:6400:f000:0000::/36				

Training ISP IPV6 Addressing Plan

Table 2: Top level summarization option infrastructure & customer

Block#	Prefix	Description	Reverse Domain
7	2406:6400:8000:0000::/35	CS net summary region1 [R2]	2x/36 arpa domain
8	2406:6400:a000:0000::/35	CS net summary region2 [R5]	2x/36 arpa domain
9	2406:6400:c000:0000::/35	CS net summary region3 [R8]	2x/36 arpa domain
10	2406:6400:e000:0000::/35	CS net summary region4 [R11]	2x/36 arpa domain

Training ISP IPV6 Addressing Plan



Training ISP IPV6 Addressing Plan

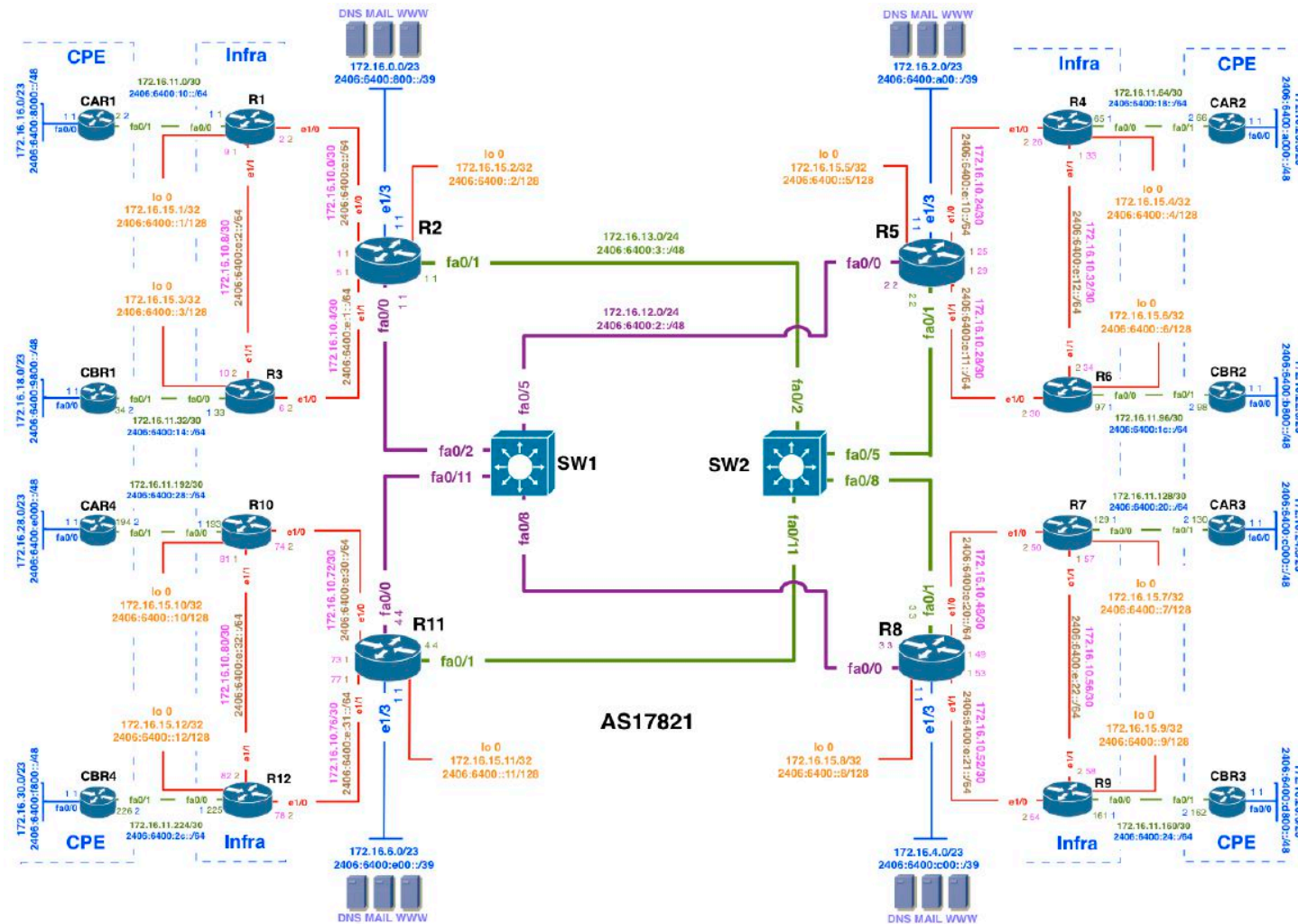
Table 3: Detail distribution infrastructure					
Block#	Prefix	Description	Reverse Domain	SOR	Registration
2	2406:6400:0000:0000::/36	Infrastructure	0.0.0.4.6.6.0.4.2.ip6.arpa.	No	Optional
11	2406:6400:0000:0000::/40	Loopback, Transport & WAN [Infra+CS]	0.0.0.0.4.6.6.0.4.2.ip6.arpa.	No	Optional
	2406:6400:0100:0000::/40				
	2406:6400:0200:0000::/40				
	2406:6400:0300:0000::/40				
	2406:6400:0400:0000::/40				
	2406:6400:0500:0000::/40				
	2406:6400:0600:0000::/40				
	2406:6400:0700:0000::/40				
16	2406:6400:0800:0000::/40	R2 DC	8.0.0.0.4.6.6.0.4.2.ip6.arpa.	No	Recommended
	2406:6400:0900:0000::/40				
17	2406:6400:0a00:0000::/40	R5 DC	a.0.0.0.4.6.6.0.4.2.ip6.arpa.	No	Recommended
	2406:6400:0b00:0000::/40				
18	2406:6400:0c00:0000::/40	R8 DC	c.0.0.0.4.6.6.0.4.2.ip6.arpa.	No	Recommended
	2406:6400:0d00:0000::/40				
19	2406:6400:0e00:0000::/40	R11 DC	e.0.0.0.4.6.6.0.4.2.ip6.arpa.	No	Recommended
	2406:6400:0f00:0000::/40				

Training ISP IPV6 Addressing Plan

Table 4: Datacenter prefix summarization options

Block#	Prefix	Description	Reverse Domain
12	2406:6400:0800:0000::/39	Region 1 DC Summary [R2]	
13	2406:6400:0a00:0000::/39	Region 2 DC Summary [R5]	
14	2406:6400:0c00:0000::/39	Region 3 DC Summary [R8]	
15	2406:6400:0e00:0000::/39	Region 4 DC Summary [R11]	

Training ISP IPV6 Addressing Plan



APNIC



Training ISP IPV6 Addressing Plan

Table 5: Further detail loopback, transport & infrastructure WAN					
Block#	Prefix	Description	Reverse Domain	SOR	Registration
11	2406:6400:0000:0000::/40	Loopback, Transport & Infra WAN	0.0.0.0.4.6.6.0.4.2.ip6.arpa.		
20	2406:6400:0000:0000::/48	Loopback		No	Recommended
	2406:6400:0001:0000::/48				
21	2406:6400:0002:0000::/48	Purple Transport		No	Recommended
22	2406:6400:0003:0000::/48	Green Transport		No	Recommended
	2406:6400:0004:0000::/48				
	2406:6400:0005:0000::/48				
	2406:6400:0006:0000::/48				
	2406:6400:0007:0000::/48				
	2406:6400:0008:0000::/48				
	2406:6400:0009:0000::/48				
	2406:6400:000A:0000::/48				
	2406:6400:000B:0000::/48				
	2406:6400:000C:0000::/48				
	2406:6400:000D:0000::/48				
23	2406:6400:000E:0000::/48	WAN Prefix Infra Link		No	Recommended
	2406:6400:000F:0000::/48				

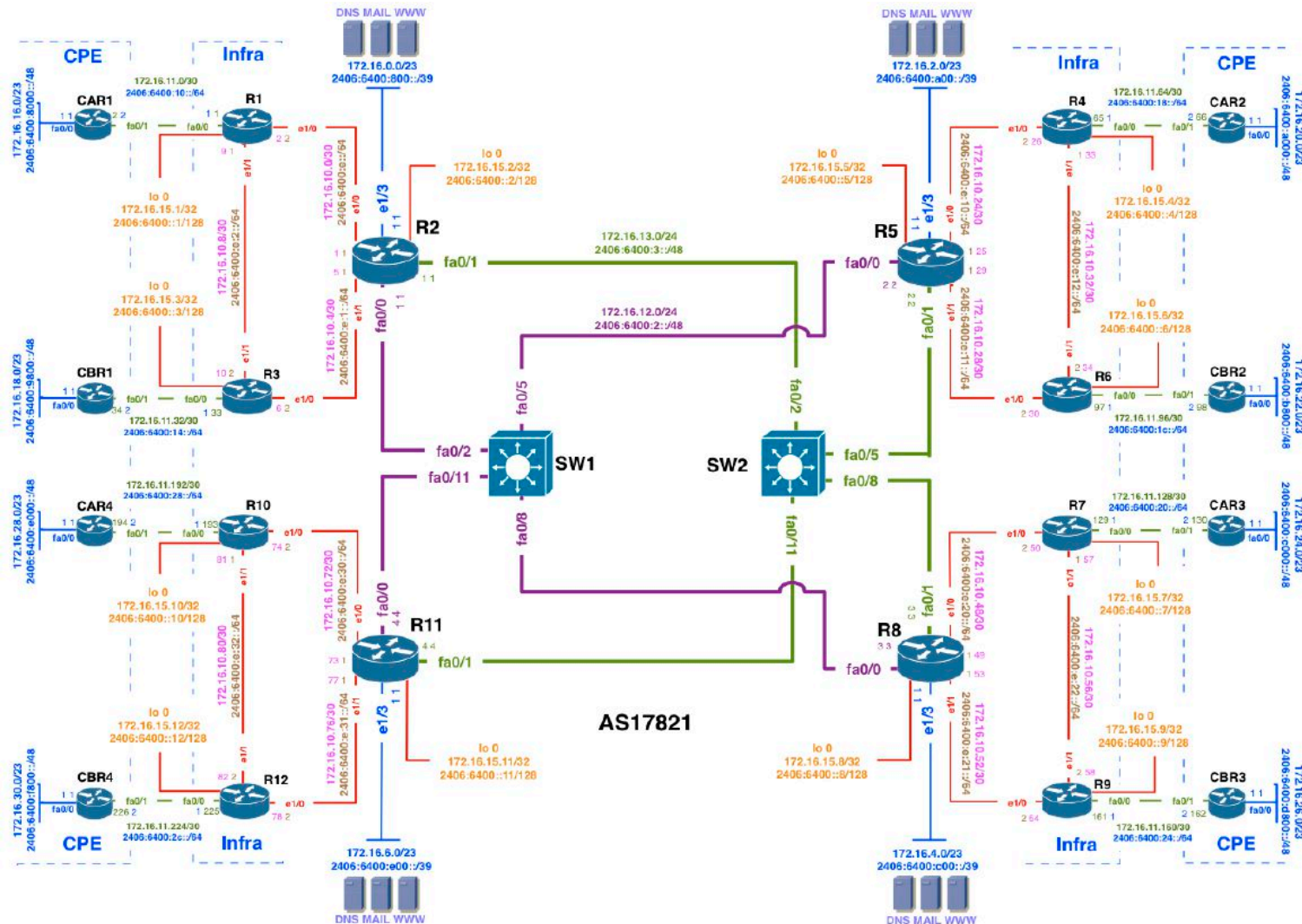
Training ISP IPV6 Addressing Plan

Table 6: Further detail CS link WAN					
Block#	Prefix	Description	Reverse Domain	SOR	Registration
27	2406:6400:0010:0000::/48	WAN Prefix CS Link R1 Region1		No	Recommended
	2406:6400:0011:0000::/48				
	2406:6400:0012:0000::/48				
	2406:6400:0013:0000::/48				
28	2406:6400:0014:0000::/48	WAN Prefix CS Link R3 Region1		No	Recommended
	2406:6400:0015:0000::/48				
	2406:6400:0016:0000::/48				
	2406:6400:0017:0000::/48				
32	2406:6400:0018:0000::/48	WAN Prefix CS Link R4 Region2		No	Recommended
	2406:6400:0019:0000::/48				
	2406:6400:001A:0000::/48				
	2406:6400:001B:0000::/48				
33	2406:6400:001C:0000::/48	WAN Prefix CS Link R6 Region2		No	Recommended
	2406:6400:001D:0000::/48				
	2406:6400:001E:0000::/48				
	2406:6400:001F:0000::/48				
37	2406:6400:0020:0000::/48	WAN Prefix CS Link R7 Region3		No	Recommended
	2406:6400:0021:0000::/48				
	2406:6400:0022:0000::/48				
	2406:6400:0023:0000::/48				
38	2406:6400:0024:0000::/48	WAN Prefix CS Link R9 Region3		No	Recommended
	2406:6400:0025:0000::/48				
	2406:6400:0026:0000::/48				
	2406:6400:0027:0000::/48				
42	2406:6400:0028:0000::/48	WAN Prefix CS Link R10 Region4		No	Recommended
	2406:6400:0029:0000::/48				
	2406:6400:002A:0000::/48				
	2406:6400:002B:0000::/48				
43	2406:6400:002C:0000::/48	WAN Prefix CS Link R12 Region4		No	Recommended
	2406:6400:002D:0000::/48				
	2406:6400:002E:0000::/48				
	2406:6400:002F:0000::/48				

Training ISP IPV6 Addressing Plan

Table 7: CS link WAN summarization options			
Block#	Prefix	Description	Reverse Domain
24	2406:6400:0010:0000::/45	WAN CS Link Region1 Summary [R2]	
25	2406:6400:0010:0000::/46	WAN CS Link Region1 POP1 Summary [R1]	
26	2406:6400:0014:0000::/46	WAN CS Link Region1 POP2 Summary [R3]	
Block#	Prefix	Description	Reverse Domain
29	2406:6400:0018:0000::/45	WAN Prefix CS Link Region2 Summary [R5]	
30	2406:6400:0018:0000::/46	WAN CS Link Region2 POP1 Summary [R4]	
31	2406:6400:001C:0000::/46	WAN CS Link Region2 POP2 Summary [R6]	
Block#	Prefix	Description	Reverse Domain
34	2406:6400:0020:0000::/45	WAN Prefix CS Link Region3 Summary [R8]	
35	2406:6400:0020:0000::/46	WAN CS Link Region3 POP1 Summary [R7]	
36	2406:6400:0024:0000::/46	WAN CS Link Region3 POP2 Summary [R9]	
Block#	Prefix	Description	Reverse Domain
39	2406:6400:0028:0000::/45	WAN Prefix CS Link Region4 Summary [R11]	
40	2406:6400:0028:0000::/46	WAN CS Link Region4 POP1 Summary [R10]	
41	2406:6400:002C:0000::/46	WAN CS Link Region4 POP2 Summary [R12]	

Training ISP IPV6 Addressing Plan



Training ISP IPV6 Addressing Plan

Table 8: Further detail loopback

Block#	Prefix	Description	PTR Record	SOR	Registration
20	2406:6400:0000:0000::/48	Loopback		No	Recommended
			YES		
43	2406:6400:0000:0000::1/128	Router1 loopback 0	YES	No	No
44	2406:6400:0000:0000::2/128	Router2 loopback 0	YES	No	No
45	2406:6400:0000:0000::3/128	Router3 loopback 0	YES	No	No
46	2406:6400:0000:0000::4/128	Router4 loopback 0	YES	No	No
47	2406:6400:0000:0000::5/128	Router5 loopback 0	YES	No	No
48	2406:6400:0000:0000::6/128	Router6 loopback 0	YES	No	No
49	2406:6400:0000:0000::7/128	Router7 loopback 0	YES	No	No
50	2406:6400:0000:0000::8/128	Router8 loopback 0	YES	No	No
51	2406:6400:0000:0000::9/128	Router9 loopback 0	YES	No	No
52	2406:6400:0000:0000::10/128	Router10 loopback 0	YES	No	No
53	2406:6400:0000:0000::11/128	Router11 loopback 0	YES	No	No
54	2406:6400:0000:0000::12/128	Router12 loopback 0	YES	No	No

Training ISP IPV6 Addressing Plan

Table 9: Further detail transport					
Block#	Prefix	Description	PTR Record	SOR	Registration
21	2406:6400:0002:0000::/48	Purple Transport		No	Recommended
	2406:6400:0002:0000::1/48	Router2 fa0/0	YES	No	No
	2406:6400:0002:0000::2/48	Router5 fa0/0	YES	No	No
	2406:6400:0002:0000::3/48	Router8 fa0/0	YES	No	No
	2406:6400:0002:0000::4/48	Router11 fa0/0	YES	No	No
Block#	Prefix	Description	PTR Record	SOR	Registration
22	2406:6400:0003:0000::/48	Green Transport		No	Recommended
	2406:6400:0003:0000::1/48	Router2 fa0/1	YES	No	No
	2406:6400:0003:0000::2/48	Router5 fa0/1	YES	No	No
	2406:6400:0003:0000::3/48	Router8 fa0/1	YES	No	No
	2406:6400:0003:0000::4/48	Router11 fa0/1	YES	No	No

Training ISP IPV6 Addressing Plan

Table 10: Further detail Infra WAN					
Block#	Prefix	Description	PTR Record	SOR	Registration
23	2406:6400:000E:0000::/48	WAN Prefix Infra Link		No	Recommended
55	2406:6400:000E:0000::/64	R2[::1]-R1[::2]	YES	No	No
56	2406:6400:000E:0001::/64	R2[::1]-R3[::2]	YES	No	No
57	2406:6400:000E:0002::/64	R1[::1]-R3[::2]	YES	No	No
	2406:6400:000E:0003::/64				
	2406:6400:000E:0004::/64				
	2406:6400:000E:0005::/64				
	2406:6400:000E:0006::/64				
	2406:6400:000E:0007::/64				
	2406:6400:000E:0008::/64				
	2406:6400:000E:0009::/64				
	2406:6400:000E:000A::/64				
	2406:6400:000E:000B::/64				
	2406:6400:000E:000C::/64				
	2406:6400:000E:000D::/64				
	2406:6400:000E:000E::/64				
	2406:6400:000E:000F::/64				
58	2406:6400:000E:0010::/64	R5[::1]-R4[::2]	YES	No	No
59	2406:6400:000E:0011::/64	R5[::1]-R6[::2]	YES	No	No
60	2406:6400:000E:0012::/64	R4[::1]-R6[::2]	YES	No	No
	2406:6400:000E:0013::/64				
	2406:6400:000E:0014::/64				
	2406:6400:000E:0015::/64				
	2406:6400:000E:0016::/64				
	2406:6400:000E:0017::/64				
	2406:6400:000E:0018::/64				
	2406:6400:000E:0019::/64				
	2406:6400:000E:001A::/64				
	2406:6400:000E:001B::/64				
	2406:6400:000E:001C::/64				
	2406:6400:000E:001D::/64				
	2406:6400:000E:001E::/64				
	2406:6400:000E:001F::/64				
61	2406:6400:000E:0020::/64	R8[::1]-R7[::2]	YES	No	No
62	2406:6400:000E:0021::/64	R8[::1]-R9[::2]	YES	No	No
63	2406:6400:000E:0022::/64	R7[::1]-R9[::2]	YES	No	No
	2406:6400:000E:0023::/64				
	2406:6400:000E:0024::/64				
	2406:6400:000E:0025::/64				
	2406:6400:000E:0026::/64				
	2406:6400:000E:0027::/64				
	2406:6400:000E:0028::/64				
	2406:6400:000E:0029::/64				
	2406:6400:000E:002A::/64				
	2406:6400:000E:002B::/64				
	2406:6400:000E:002C::/64				
	2406:6400:000E:002D::/64				
	2406:6400:000E:002E::/64				
	2406:6400:000E:002F::/64				
64	2406:6400:000E:0030::/64	R11[::1]-R10[::2]	YES	No	No
65	2406:6400:000E:0031::/64	R11[::1]-R12[::2]	YES	No	No
66	2406:6400:000E:0032::/64	R10[::1]-R12[::2]	YES	No	No
	2406:6400:000E:0033::/64				
	2406:6400:000E:0034::/64				
	2406:6400:000E:0035::/64				
	2406:6400:000E:0036::/64				
	2406:6400:000E:0037::/64				
	2406:6400:000E:0038::/64				
	2406:6400:000E:0039::/64				
	2406:6400:000E:003A::/64				

Training ISP IPV6 Addressing Plan

Table 11: Detail CS link WAN Region 1					
Block#	Prefix	Description	PTR Record	SOR	Registration
27	2406:6400:0010:0000::/48	WAN Prefix CS Link R1 Region1		No	Recommended
	2406:6400:0010:0000::/64	R1[::1]-CAR1[::2]	Yes	No	No
	2406:6400:0010:0001::/64		Yes	No	No
	2406:6400:0010:0002::/64		Yes	No	No
	2406:6400:0010:0003::/64		Yes	No	No
	2406:6400:0010:0004::/64		Yes	No	No
	2406:6400:0010:0005::/64		Yes	No	No
	2406:6400:0010:0006::/64		Yes	No	No
	2406:6400:0010:0007::/64		Yes	No	No
	2406:6400:0010:0008::/64		Yes	No	No
	2406:6400:0010:0009::/64		Yes	No	No
	2406:6400:0010:000A::/64		Yes	No	No
	2406:6400:0010:000B::/64		Yes	No	No
	2406:6400:0010:000C::/64		Yes	No	No
	2406:6400:0010:000D::/64		Yes	No	No
	2406:6400:0010:000E::/64		Yes	No	No
	2406:6400:0010:000F::/64		Yes	No	No
Block#	Prefix	Description	PTR Record	SOR	Registration
28	2406:6400:0014:0000::/48	WAN Prefix CS Link R3 Region1		No	Recommended
	2406:6400:0014:0000::/64	R3[::1]-CBR1[::2]	Yes	No	No
	2406:6400:0014:0001::/64		Yes	No	No
	2406:6400:0014:0002::/64		Yes	No	No
	2406:6400:0014:0003::/64		Yes	No	No
	2406:6400:0014:0004::/64		Yes	No	No
	2406:6400:0014:0005::/64		Yes	No	No
	2406:6400:0014:0006::/64		Yes	No	No
	2406:6400:0014:0007::/64		Yes	No	No
	2406:6400:0014:0008::/64		Yes	No	No
	2406:6400:0014:0009::/64		Yes	No	No
	2406:6400:0014:000A::/64		Yes	No	No
	2406:6400:0014:000B::/64		Yes	No	No
	2406:6400:0014:000C::/64		Yes	No	No
	2406:6400:0014:000D::/64		Yes	No	No
	2406:6400:0014:000E::/64		Yes	No	No
	2406:6400:0014:000F::/64		Yes	No	No

Training ISP IPV6 Addressing Plan

Table 12: Detail CS link WAN Region 2					
Block#	Prefix	Description	PTR Record	SOR	Registration
32	2406:6400:0018:0000::/48	WAN Prefix CS Link R4 Region2		No	Recommended
	2406:6400:0018:0000::/64	R4[::1]-CAR2[::2]	Yes	No	No
	2406:6400:0018:0001::/64		Yes	No	No
	2406:6400:0018:0002::/64		Yes	No	No
	2406:6400:0018:0003::/64		Yes	No	No
	2406:6400:0018:0004::/64		Yes	No	No
	2406:6400:0018:0005::/64		Yes	No	No
	2406:6400:0018:0006::/64		Yes	No	No
	2406:6400:0018:0007::/64		Yes	No	No
	2406:6400:0018:0008::/64		Yes	No	No
	2406:6400:0018:0009::/64		Yes	No	No
	2406:6400:0018:000A::/64		Yes	No	No
	2406:6400:0018:000B::/64		Yes	No	No
	2406:6400:0018:000C::/64		Yes	No	No
	2406:6400:0018:000D::/64		Yes	No	No
	2406:6400:0018:000E::/64		Yes	No	No
	2406:6400:0018:000F::/64		Yes	No	No
Block#	Prefix	Description	PTR Record	SOR	Registration
33	2406:6400:001C:0000::/48	WAN Prefix CS Link R6 Region2		No	Recommended
	2406:6400:001C:0000::/64	R6[::1]-CBR2[::2]	Yes	No	No
	2406:6400:001C:0001::/64		Yes	No	No
	2406:6400:001C:0002::/64		Yes	No	No
	2406:6400:001C:0003::/64		Yes	No	No
	2406:6400:001C:0004::/64		Yes	No	No
	2406:6400:001C:0005::/64		Yes	No	No
	2406:6400:001C:0006::/64		Yes	No	No
	2406:6400:001C:0007::/64		Yes	No	No
	2406:6400:001C:0008::/64		Yes	No	No
	2406:6400:001C:0009::/64		Yes	No	No
	2406:6400:001C:000A::/64		Yes	No	No
	2406:6400:001C:000B::/64		Yes	No	No
	2406:6400:001C:000C::/64		Yes	No	No
	2406:6400:001C:000D::/64		Yes	No	No
	2406:6400:001C:000E::/64		Yes	No	No
	2406:6400:001C:000F::/64		Yes	No	No

Training ISP IPV6 Addressing Plan

Table 13: Detail CS link WAN Region3					
Block#	Prefix	Description	PTR Record	SOR	Registration
37	2406:6400:0020:0000::/48	WAN Prefix CS Link R7 Region3		No	Recommended
	2406:6400:0020:0000::/64	R7[::1]-CAR3[::2]	Yes	No	No
	2406:6400:0020:0001::/64		Yes	No	No
	2406:6400:0020:0002::/64		Yes	No	No
	2406:6400:0020:0003::/64		Yes	No	No
	2406:6400:0020:0004::/64		Yes	No	No
	2406:6400:0020:0005::/64		Yes	No	No
	2406:6400:0020:0006::/64		Yes	No	No
	2406:6400:0020:0007::/64		Yes	No	No
	2406:6400:0020:0008::/64		Yes	No	No
	2406:6400:0020:0009::/64		Yes	No	No
	2406:6400:0020:000A::/64		Yes	No	No
	2406:6400:0020:000B::/64		Yes	No	No
	2406:6400:0020:000C::/64		Yes	No	No
	2406:6400:0020:000D::/64		Yes	No	No
	2406:6400:0020:000E::/64		Yes	No	No
	2406:6400:0020:000F::/64		Yes	No	No
Block#	Prefix	Description	PTR Record	SOR	Registration
38	2406:6400:0024:0000::/48	WAN Prefix CS Link R9 Region3		No	Recommended
	2406:6400:0024:0000::/64	R9[::1]-CBR3[::2]	Yes	No	No
	2406:6400:0024:0001::/64		Yes	No	No
	2406:6400:0024:0002::/64		Yes	No	No
	2406:6400:0024:0003::/64		Yes	No	No
	2406:6400:0024:0004::/64		Yes	No	No
	2406:6400:0024:0005::/64		Yes	No	No
	2406:6400:0024:0006::/64		Yes	No	No
	2406:6400:0024:0007::/64		Yes	No	No
	2406:6400:0024:0008::/64		Yes	No	No
	2406:6400:0024:0009::/64		Yes	No	No
	2406:6400:0024:000A::/64		Yes	No	No
	2406:6400:0024:000B::/64		Yes	No	No
	2406:6400:0024:000C::/64		Yes	No	No
	2406:6400:0024:000D::/64		Yes	No	No
	2406:6400:0024:000E::/64		Yes	No	No
	2406:6400:0024:000F::/64		Yes	No	No

Training ISP IPV6 Addressing Plan

Table 14: Detail CS link WAN Region 4

Block#	Prefix	Description	PTR Record	SOR	Registration
42	2406:6400:0028:0000::/48	WAN Prefix CS Link R10 Region4		No	Recommended
	2406:6400:0028:0000::/64	R10[::1]-CAR4[::2]	Yes	No	No
	2406:6400:0028:0001::/64		Yes	No	No
	2406:6400:0028:0002::/64		Yes	No	No
	2406:6400:0028:0003::/64		Yes	No	No
	2406:6400:0028:0004::/64		Yes	No	No
	2406:6400:0028:0005::/64		Yes	No	No
	2406:6400:0028:0006::/64		Yes	No	No
	2406:6400:0028:0007::/64		Yes	No	No
	2406:6400:0028:0008::/64		Yes	No	No
	2406:6400:0028:0009::/64		Yes	No	No
	2406:6400:0028:000A::/64		Yes	No	No
	2406:6400:0028:000B::/64		Yes	No	No
	2406:6400:0028:000C::/64		Yes	No	No
	2406:6400:0028:000D::/64		Yes	No	No
	2406:6400:0028:000E::/64		Yes	No	No
	2406:6400:0028:000F::/64		Yes	No	No
Block#	Prefix	Description	PTR Record	SOR	Registration
43	2406:6400:002C:0000::/48	WAN Prefix CS Link R12 Region4		No	Recommended
	2406:6400:002C:0000::/64	R12[::1]-CBR4[::2]	Yes	No	No
	2406:6400:002C:0001::/64		Yes	No	No
	2406:6400:002C:0002::/64		Yes	No	No
	2406:6400:002C:0003::/64		Yes	No	No
	2406:6400:002C:0004::/64		Yes	No	No
	2406:6400:002C:0005::/64		Yes	No	No
	2406:6400:002C:0006::/64		Yes	No	No
	2406:6400:002C:0007::/64		Yes	No	No
	2406:6400:002C:0008::/64		Yes	No	No
	2406:6400:002C:0009::/64		Yes	No	No
	2406:6400:002C:000A::/64		Yes	No	No
	2406:6400:002C:000B::/64		Yes	No	No
	2406:6400:002C:000C::/64		Yes	No	No
	2406:6400:002C:000D::/64		Yes	No	No
	2406:6400:002C:000E::/64		Yes	No	No
	2406:6400:002C:000F::/64		Yes	No	No

Training ISP IPV6 Addressing Plan

Table 15: Customer block Region 1					
Block#	Prefix	Description	Reverse DNS	SOR	Registration
7	2406:6400:8000:0000::/35	Customer block Region 1			
	2406:6400:8000:0000::/40	Customer block POP1 [R1]		>= /48 Yes	Yes
	2406:6400:8100:0000::/40				
	2406:6400:8200:0000::/40				
	2406:6400:8300:0000::/40				
	2406:6400:8400:0000::/40				
	2406:6400:8500:0000::/40				
	2406:6400:8600:0000::/40				
	2406:6400:8700:0000::/40				
	2406:6400:8800:0000::/40	Customer block future use/POP		>= /48 Yes	Yes
	2406:6400:8900:0000::/40				
	2406:6400:8A00:0000::/40				
	2406:6400:8B00:0000::/40				
	2406:6400:8C00:0000::/40				
	2406:6400:8D00:0000::/40				
	2406:6400:8E00:0000::/40				
	2406:6400:8F00:0000::/40				
	2406:6400:9000:0000::/40	Customer block future use/POP		>= /48 Yes	Yes
	2406:6400:9100:0000::/40				
	2406:6400:9200:0000::/40				
	2406:6400:9300:0000::/40				
	2406:6400:9400:0000::/40				
	2406:6400:9500:0000::/40				
	2406:6400:9600:0000::/40				
	2406:6400:9700:0000::/40				
	2406:6400:9800:0000::/40	Customer block POP2 [R3]		>= /48 Yes	Yes
	2406:6400:9900:0000::/40				
	2406:6400:9A00:0000::/40				
	2406:6400:9B00:0000::/40				
	2406:6400:9C00:0000::/40				
	2406:6400:9D00:0000::/40				
	2406:6400:9E00:0000::/40				
	2406:6400:9F00:0000::/40				

Training ISP IPV6 Addressing Plan

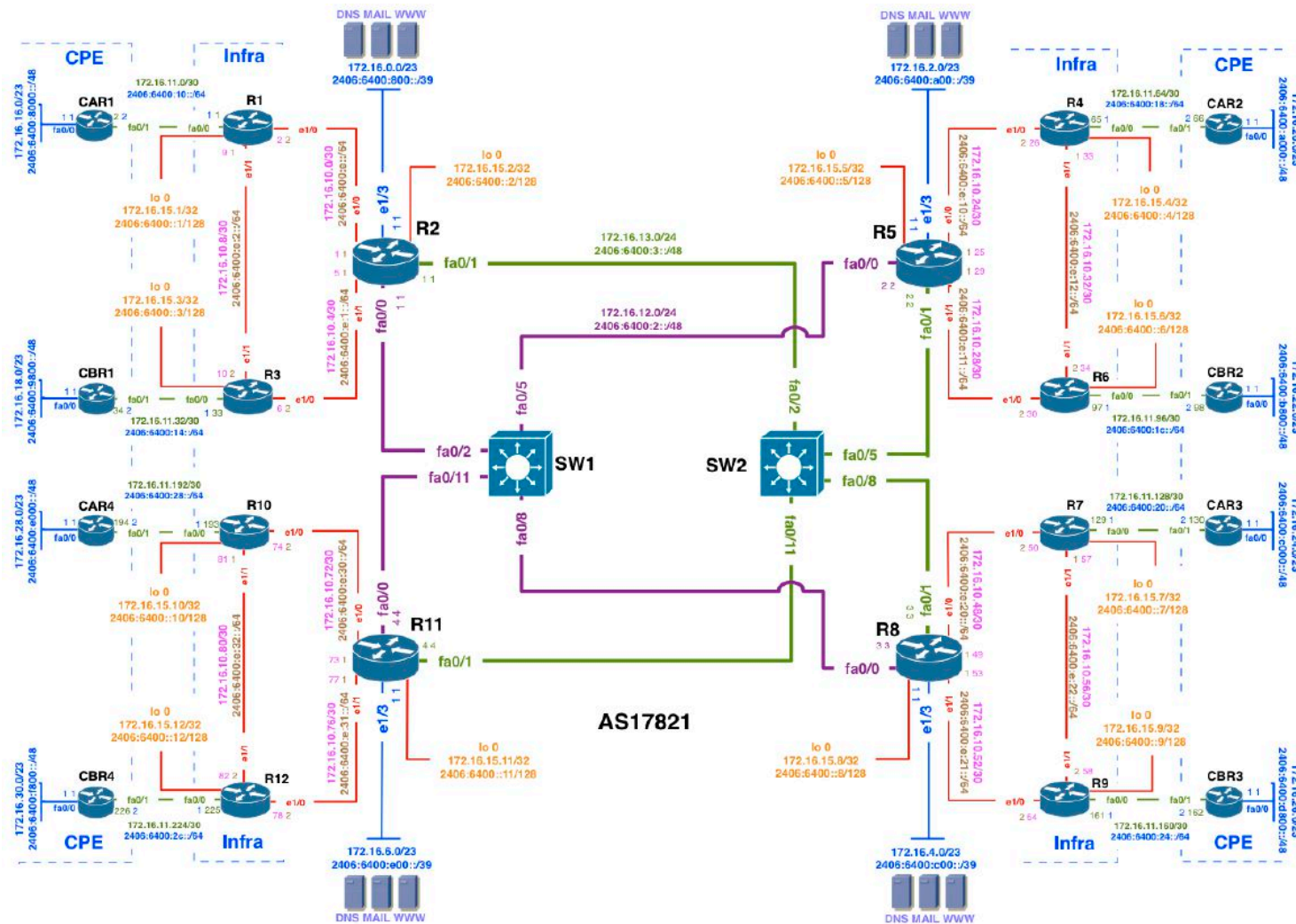
Table 16: Summarization oprions customer block Region 1			
Block#	Prefix	Description	Reverse Domain
	2406:6400:8000:0000::/35	Customer block Region 1 [R2]	
	2406:6400:8000:0000::/37	Customer block POP1 [R1]	
	2406:6400:8800:0000::/37	Customer block future use/POP	
	2406:6400:9000:0000::/37	Customer block future use/POP	
	2406:6400:9800:0000::/37	Customer block POP2 [R3]	

Training ISP IPV6 Addressing Plan

Table 17: Detail customer block Region 1

Block#	Prefix	Description	Reverse DNS	SOR	Registration
	2406:6400:8000:0000::/40	1st Customer block POP1 [R1]			
	2406:6400:8000:0000::/48	1st Customer prefix POP1 [R1]		Yes	Yes
	2406:6400:8001:0000::/48				
	2406:6400:8002:0000::/48				
	2406:6400:8003:0000::/48				
	2406:6400:8004:0000::/48				
	2406:6400:8005:0000::/48				
	2406:6400:8006:0000::/48				
	2406:6400:8007:0000::/48				
	2406:6400:9800:0000::/40	1st Customer block POP2 [R3]			
	2406:6400:9800:0000::/48	1st Customer prefix POP2 [R3]		Yes	Yes
	2406:6400:9801:0000::/48				
	2406:6400:9802:0000::/48				
	2406:6400:9803:0000::/48				
	2406:6400:9804:0000::/48				
	2406:6400:9805:0000::/48				
	2406:6400:9806:0000::/48				
	2406:6400:9807:0000::/48				

Training ISP IPV6 Addressing Plan



Training ISP IPV6 Addressing Plan

Table 18: Customer block Region 2					
Block#	Prefix	Description	Reverse DNS	SOR	Registration
8	2406:6400:a000:0000::/35	Customer block Region 2			
	2406:6400:A000:0000::/40	Customer block POP1 [R4]		>= /48 Yes	Yes
	2406:6400:A100:0000::/40				
	2406:6400:A200:0000::/40				
	2406:6400:A300:0000::/40				
	2406:6400:A400:0000::/40				
	2406:6400:A500:0000::/40				
	2406:6400:A600:0000::/40				
	2406:6400:A700:0000::/40				
	2406:6400:A800:0000::/40	Customer block future use/POP		>= /48 Yes	Yes
	2406:6400:A900:0000::/40				
	2406:6400:AA00:0000::/40				
	2406:6400:AB00:0000::/40				
	2406:6400:AC00:0000::/40				
	2406:6400:AD00:0000::/40				
	2406:6400:AE00:0000::/40				
	2406:6400:AF00:0000::/40				
	2406:6400:B000:0000::/40	Customer block future use/POP		>= /48 Yes	Yes
	2406:6400:B100:0000::/40				
	2406:6400:B200:0000::/40				
	2406:6400:B300:0000::/40				
	2406:6400:B400:0000::/40				
	2406:6400:B500:0000::/40				
	2406:6400:B600:0000::/40				
	2406:6400:B700:0000::/40				
	2406:6400:B800:0000::/40	Customer block POP2 [R6]		>= /48 Yes	Yes
	2406:6400:B900:0000::/40				
	2406:6400:BA00:0000::/40				
	2406:6400:BB00:0000::/40				
	2406:6400:BC00:0000::/40				
	2406:6400:BD00:0000::/40				
	2406:6400:BE00:0000::/40				
	2406:6400:BF00:0000::/40				

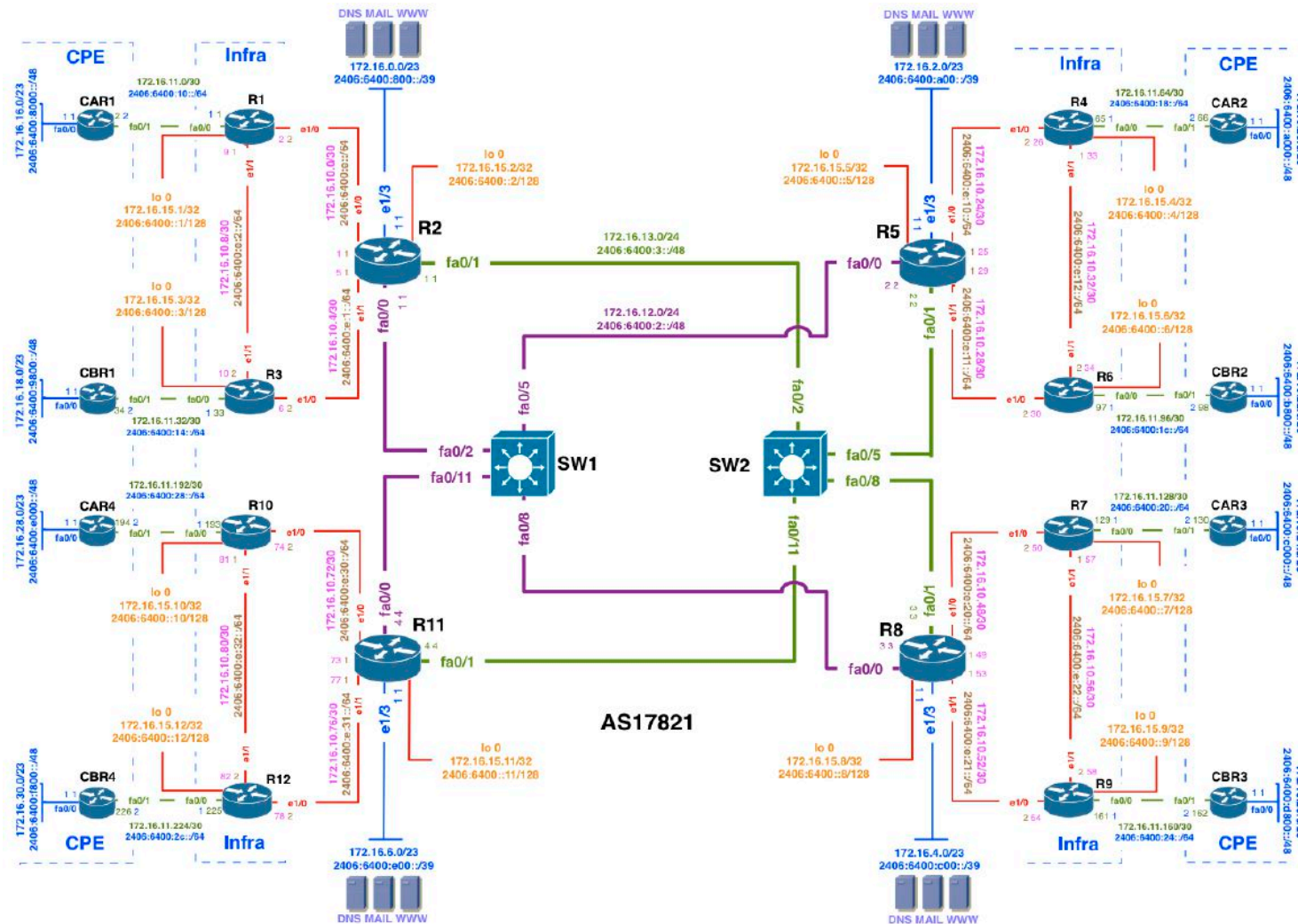
Training ISP IPV6 Addressing Plan

Table 19: Summarization oprions customer block Region 2			
Block#	Prefix	Description	Reverse Domain
	2406:6400:A000:0000::/35	Customer block Region 2 [R5]	
	2406:6400:A000:0000::/37	Customer block POP1 [R4]	
	2406:6400:A800:0000::/37	Customer block future use/POP	
	2406:6400:B000:0000::/37	Customer block future use/POP	
	2406:6400:B800:0000::/37	Customer block POP2 [R6]	

Training ISP IPV6 Addressing Plan

Table 20: Detail customer block Region 2					
Block#	Prefix	Description	Reverse DNS	SOR	Registration
	2406:6400:A000:0000::/40	1st Customer block POP1 [R4]			
	2406:6400:A000:0000::/48	1st Customer prefix POP1 [R4]		Yes	Yes
	2406:6400:A001:0000::/48				
	2406:6400:A002:0000::/48				
	2406:6400:A003:0000::/48				
	2406:6400:A004:0000::/48				
	2406:6400:A005:0000::/48				
	2406:6400:A006:0000::/48				
	2406:6400:A007:0000::/48				
	2406:6400:B800:0000::/40	1st Customer block POP2 [R6]			
	2406:6400:B800:0000::/48	1st Customer prefix POP2 [R6]		Yes	Yes
	2406:6400:B801:0000::/48				
	2406:6400:B802:0000::/48				
	2406:6400:B803:0000::/48				
	2406:6400:B804:0000::/48				
	2406:6400:B805:0000::/48				
	2406:6400:B806:0000::/48				
	2406:6400:B807:0000::/48				

Training ISP IPV6 Addressing Plan



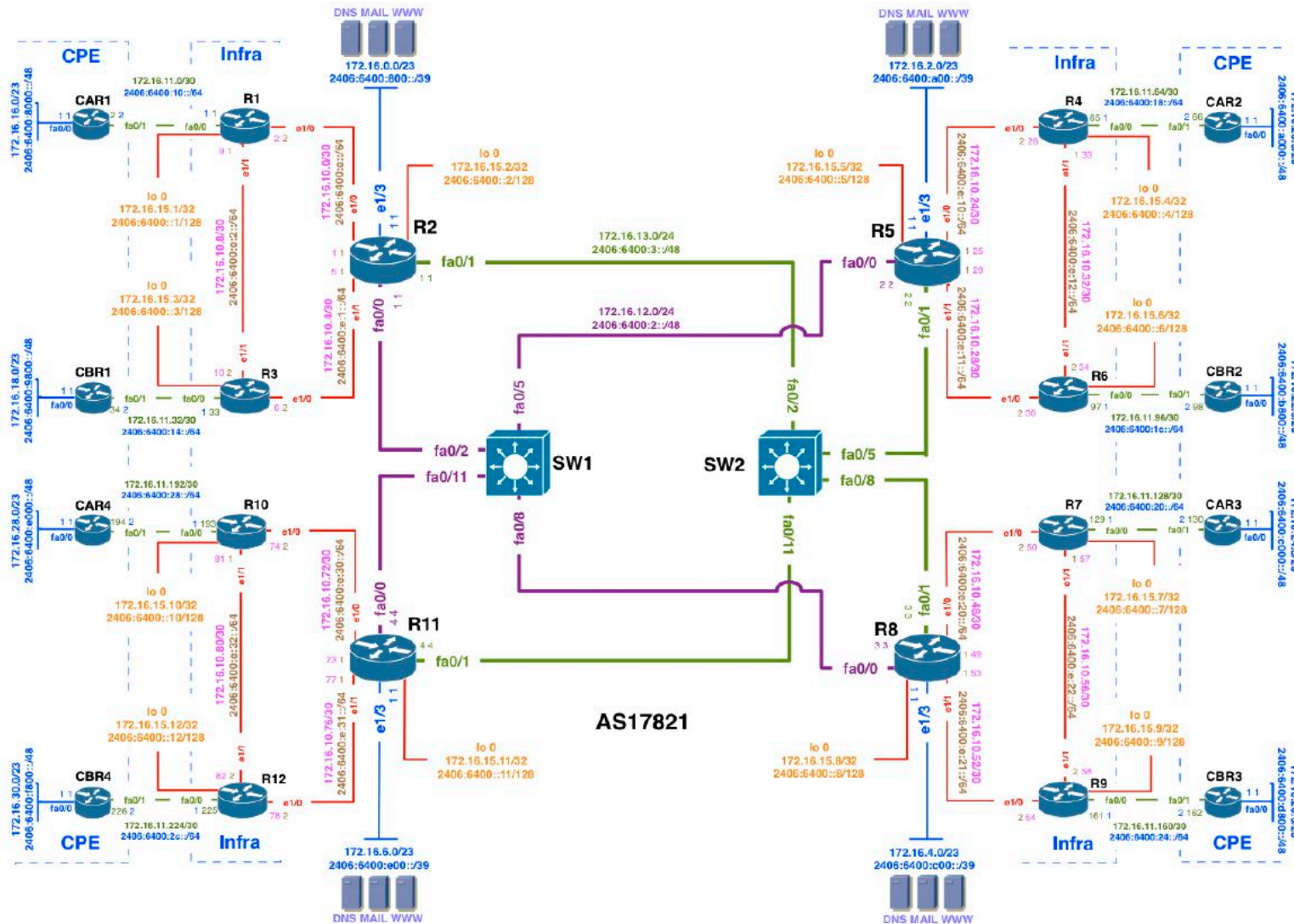
Training ISP IPV6 Addressing Plan

Table 21: Customer block Region 3					
Block#	Prefix	Description	Reverse DNS	SOR	Registration
9	2406:6400:c000:0000::/35	Customer block Region 3			
	2406:6400:C000:0000::/40	Customer block POP1 [R7]		>= /48 Yes	Yes
	2406:6400:C100:0000::/40				
	2406:6400:C200:0000::/40				
	2406:6400:C300:0000::/40				
	2406:6400:C400:0000::/40				
	2406:6400:C500:0000::/40				
	2406:6400:C600:0000::/40				
	2406:6400:C700:0000::/40				
	2406:6400:C800:0000::/40	Customer block future use/POP		>= /48 Yes	Yes
	2406:6400:C900:0000::/40				
	2406:6400:CA00:0000::/40				
	2406:6400:CB00:0000::/40				
	2406:6400:CC00:0000::/40				
	2406:6400:CD00:0000::/40				
	2406:6400:CE00:0000::/40				
	2406:6400:CF00:0000::/40				
	2406:6400:D000:0000::/40	Customer block future use/POP		>= /48 Yes	Yes
	2406:6400:D100:0000::/40				
	2406:6400:D200:0000::/40				
	2406:6400:D300:0000::/40				
	2406:6400:D400:0000::/40				
	2406:6400:D500:0000::/40				
	2406:6400:D600:0000::/40				
	2406:6400:D700:0000::/40				
	2406:6400:D800:0000::/40	Customer block POP2 [R9]		>= /48 Yes	Yes
	2406:6400:D900:0000::/40				
	2406:6400:DA00:0000::/40				
	2406:6400:DB00:0000::/40				
	2406:6400:DC00:0000::/40				
	2406:6400:DD00:0000::/40				
	2406:6400:DE00:0000::/40				
	2406:6400:DF00:0000::/40				

Training ISP IPV6 Addressing Plan

Table 22: Summarization oprions customer block Region 3			
Block#	Prefix	Description	Reverse Domain
	2406:6400:c000:0000::/35	Customer block Region 3 [R8]	
	2406:6400:C000:0000::/37	Customer block POP1 [R7]	
	2406:6400:C800:0000::/37	Customer block future use/POP	
	2406:6400:D000:0000::/37	Customer block future use/POP	
	2406:6400:D800:0000::/37	Customer block POP2 [R9]	

Training ISP IPV6 Addressing Plan



Training ISP IPV6 Addressing Plan

Table 24: Customer block Region 4					
Block#	Prefix	Description	Reverse DNS	SOR	Registration
10	2406:6400:e000:0000::/35	Customer block Region 4			
	2406:6400:E000:0000::/40	Customer block POP1 [R10]		>= /48 Yes	Yes
	2406:6400:E100:0000::/40				
	2406:6400:E200:0000::/40				
	2406:6400:E300:0000::/40				
	2406:6400:E400:0000::/40				
	2406:6400:E500:0000::/40				
	2406:6400:E600:0000::/40				
	2406:6400:E700:0000::/40				
	2406:6400:E800:0000::/40	Customer block future use/POP		>= /48 Yes	Yes
	2406:6400:E900:0000::/40				
	2406:6400:EA00:0000::/40				
	2406:6400:EB00:0000::/40				
	2406:6400:EC00:0000::/40				
	2406:6400:ED00:0000::/40				
	2406:6400:EE00:0000::/40				
	2406:6400:EF00:0000::/40				
	2406:6400:F000:0000::/40	Customer block future use/POP		>= /48 Yes	Yes
	2406:6400:F100:0000::/40				
	2406:6400:F200:0000::/40				
	2406:6400:F300:0000::/40				
	2406:6400:F400:0000::/40				
	2406:6400:F500:0000::/40				
	2406:6400:F600:0000::/40				
	2406:6400:F700:0000::/40				
	2406:6400:F800:0000::/40	Customer block POP2 [R12]		>= /48 Yes	Yes
	2406:6400:F900:0000::/40				
	2406:6400:FA00:0000::/40				
	2406:6400:FB00:0000::/40				
	2406:6400:FC00:0000::/40				
	2406:6400:FD00:0000::/40				
	2406:6400:FE00:0000::/40				
	2406:6400:FF00:0000::/40				

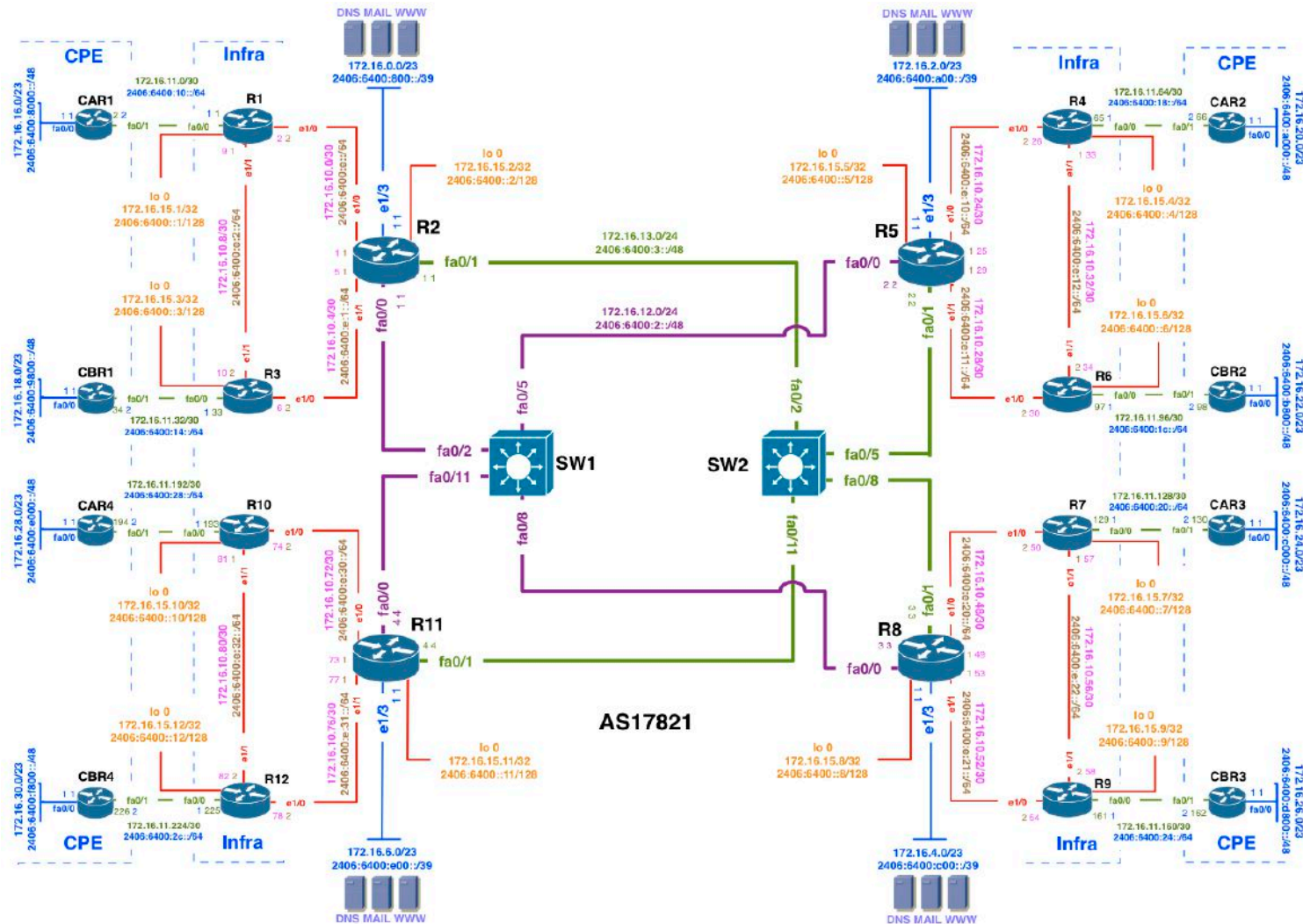
Training ISP IPV6 Addressing Plan

Table 25: Summarization oprions customer block Region 4			
Block#	Prefix	Description	Reverse Domain
	2406:6400:e000:0000::/35	Customer block Region 4 [R11]	
	2406:6400:E000:0000::/37	Customer block POP1 [R10]	
	2406:6400:E800:0000::/37	Customer block future use/POP	
	2406:6400:F000:0000::/37	Customer block future use/POP	
	2406:6400:F800:0000::/37	Customer block POP2 [R12]	

Training ISP IPV6 Addressing Plan

Table 26: Detail customer block Region 4					
Block#	Prefix	Description	Reverse DNS	SOR	Registration
	2406:6400:E000:0000::/40	1st Customer block POP1 [R10]			
	2406:6400:E000:0000::/48	1st Customer prefix POP1 [R10]		Yes	Yes
	2406:6400:E001:0000::/48				
	2406:6400:E002:0000::/48				
	2406:6400:E003:0000::/48				
	2406:6400:E004:0000::/48				
	2406:6400:E005:0000::/48				
	2406:6400:E006:0000::/48				
	2406:6400:E007:0000::/48				
	2406:6400:F800:0000::/40	1st Customer block POP2 [R10]			
	2406:6400:F800:0000::/48	1st Customer prefix POP2 [R10]		Yes	Yes
	2406:6400:F801:0000::/48				
	2406:6400:F802:0000::/48				
	2406:6400:F803:0000::/48				
	2406:6400:F804:0000::/48				
	2406:6400:F805:0000::/48				
	2406:6400:F806:0000::/48				
	2406:6400:F807:0000::/48				

Training ISP IPV6 Addressing Plan



Training ISP IPV4 Addressing Plan

Summary parent block IPV4

Block#	Prefix	Size	Description
1	172.16.0.0	/19	Parent block
2	172.16.0.0	/20	Infrastructure
3	172.16.16.0	/20	Customer network

Training ISP IPV4 Addressing Plan

Detail DC infrastructure block IPV4

Block#	Prefix	Size	Description	SOR	Register
2	172.16.0.0	/20	Infrastructure		
4	172.16.0.0	/23	Router2 DC summary net		
5	172.16.0.0	/24	Router2 DC	No	Recommended
6	172.16.2.0	/23	Router5 DC summary net		
7	172.16.2.0	/24	Router5 DC	No	Recommended
8	172.16.4.0	/23	Router8 DC summary net		
9	172.16.4.0	/24	Router8 DC	No	Recommended
10	172.16.6.0	/23	Router11 DC summary net		
11	172.16.6.0	/24	Router11 DC	No	Recommended

Training ISP IPV4 Addressing Plan

Detail infrastructure WAN block IPV4

12	172.16.10.0	/24	WAN prefix		Optional
13	172.16.10.0	/30	Router2-1 WAN	No	
14	172.16.10.4	/30	Router2-3 WAN	No	
15	172.16.10.8	/30	Router1-3 WAN	No	
16	172.16.10.24	/30	Router5-4 WAN	No	
17	172.16.10.28	/30	Router5-6 WAN	No	
18	172.16.10.32	/30	Router4-6 WAN	No	
19	172.16.10.48	/30	Router8-7 WAN	No	
20	172.16.10.52	/30	Router8-9 WAN	No	
21	172.16.10.56	/30	Router7-9 WAN	No	
22	172.16.10.72	/30	Router11-10 WAN	No	
23	172.16.10.76	/30	Router11-12 WAN	No	
24	172.16.10.80	/30	Router10-12 WAN	No	

Training ISP IPV4 Addressing Plan

Detail customer link WAN block

Block#	Prefix	Size	Description	SOR	Register
	172.16.11.0	/26	WAN CS Link Region1		
	172.16.11.0	/27	WAN CS Link POP1 [R1]		
	172.16.11.0	/30	R1[::1]-CAR1[::2]	No	No
	172.16.11.4	/30			
	172.16.11.32	/27	WAN CS Link POP2 [R3]		
	172.16.11.32	/30	R3[::33]-CBR1[::34]	No	No
	172.16.11.36	/30			
	172.16.11.64	/26	WAN CS Link Region2		
	172.16.11.64	/27	WAN CS Link POP1 [R4]		
	172.16.11.64	/30	R4[::65]-CAR2[::66]	No	No
	172.16.11.68	/30			
	172.16.11.96	/27	WAN CS Link POP2 [R6]		
	172.16.11.96	/30	R6[::97]-CBR2[::98]	No	No
	172.16.11.100	/30			
	172.16.11.128	/26	WAN CS Link Region3		
	172.16.11.128	/27	WAN CS Link POP1 [R7]		
	172.16.11.128	/30	R7[::129]-CAR3[::130]	No	No
	172.16.11.132	/30			
	172.16.11.160	/27	WAN CS Link POP2 [R9]		
	172.16.11.160	/30	R9[::161]-CBR3[::162]	No	No
	172.16.11.164	/30			
	172.16.11.192	/26	WAN CS Link Region4		
	172.16.11.192	/27	WAN CS Link POP1 [R10]		
	172.16.11.192	/30	R10[::193]-CAR4[::194]	No	No
	172.16.11.196	/30			
	172.16.11.224	/27	WAN CS Link POP2 [R12]		
	172.16.11.224	/30	R12[::225]-CBR4[::226]	No	No
	172.16.11.228	/30			

Training ISP IPV4 Addressing Plan

Detail infrastructure block Transport & Loopback IPV4

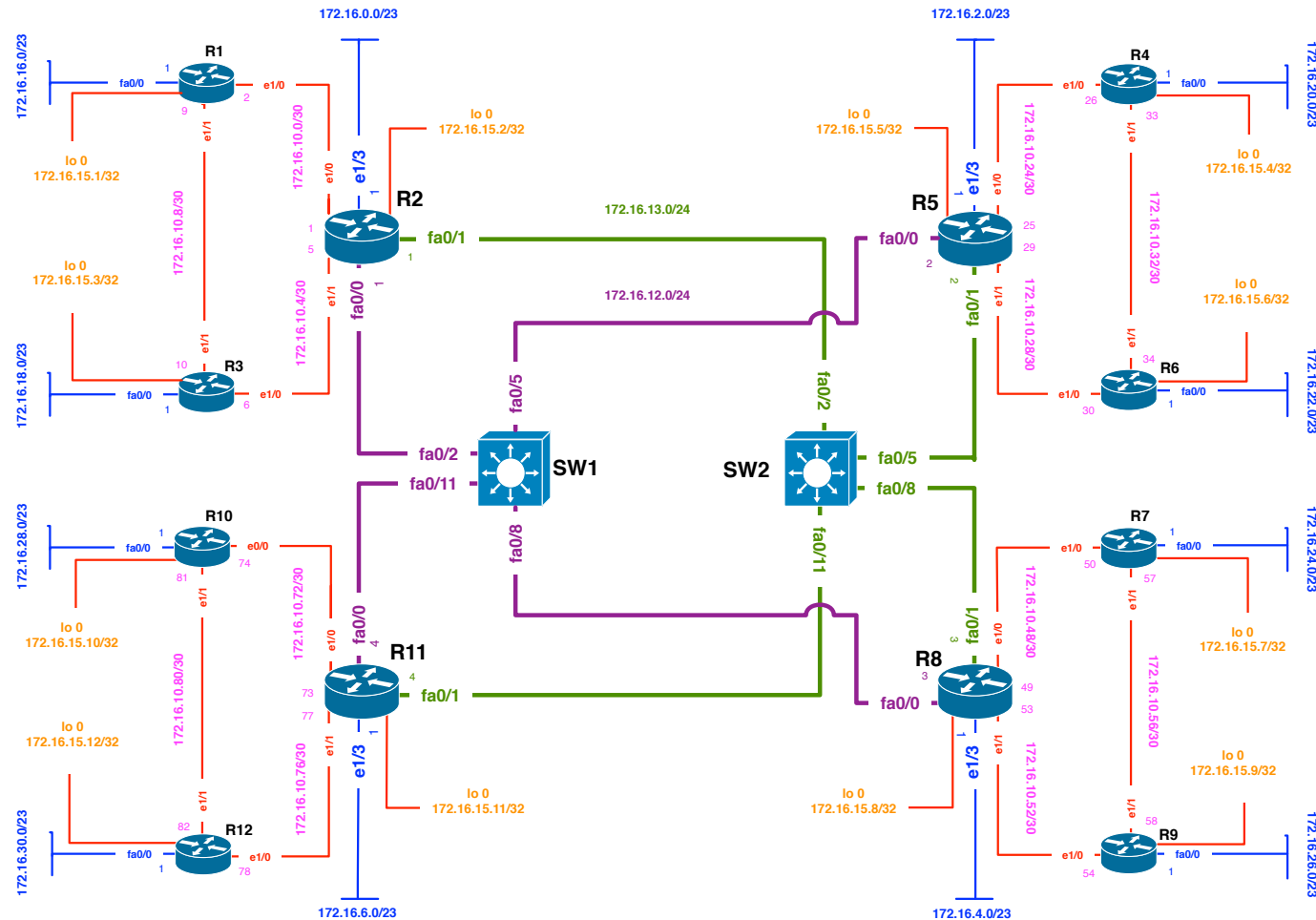
25	172.16.12.0	/24	Transport link PURPLE	No	
26	172.16.13.0	/24	Transport link GREEN	No	
27	172.16.15.0	/24	Loopback	No	

Training ISP IPV4 Addressing Plan

Detail customer block

Block#	Prefix	Size	Description	SOR	Register
28	172.16.6.0	/20	Customer network		
29	172.16.16.0	/22	Router2 summary net		
30	172.16.16.0	/23	Router1 CS network	Yes	Must
31	172.16.18.0	/23	Router3 CS network	Yes	Must
32	172.16.20.0	/22	Router5 summary net		
33	172.16.20.0	/23	Router4 CS network	Yes	Must
34	172.16.22.0	/23	Router6 CS network	Yes	Must
35	172.16.24.0	/22	Router8 summary net		
36	172.16.24.0	/23	Router7 CS network	Yes	Must
37	172.16.26.0	/23	Router9 CS network	Yes	Must
38	172.16.28.0	/22	Router11 summary net		
39	172.16.28.0	/23	Router10 CS network	Yes	Must
40	172.16.30.0	/23	Router12 CS network	Yes	Must

Training ISP IPV4 Addressing Plan



Training ISP IPv4 Address Plan

Questions



Overview

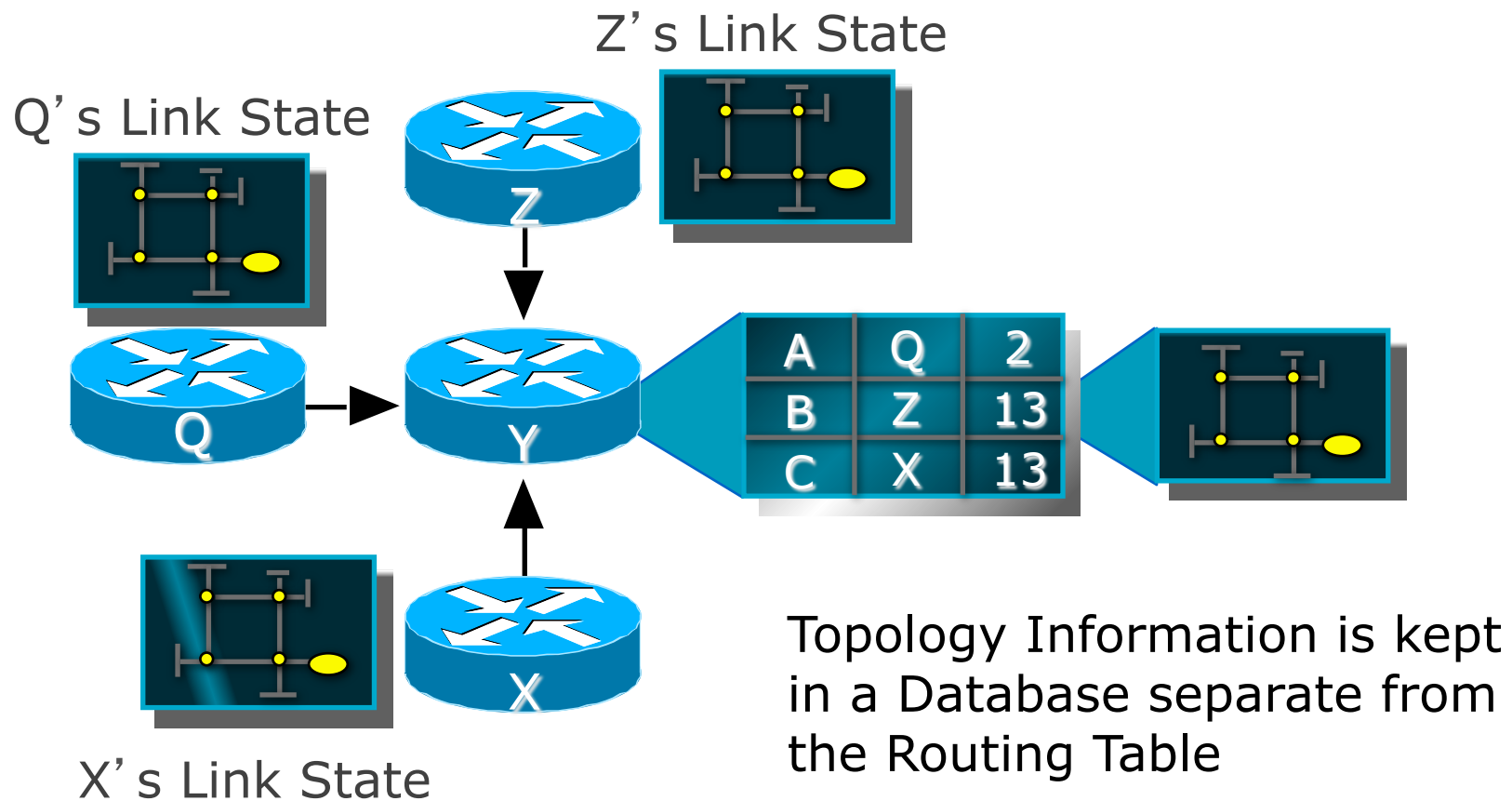
Routing Workshop (3 Days)

- Introduction to IP Routing
- Routing Protocol Basic
- IPv6 Address Structure
- Routing Lab Topology Overview
- **Operation of OSPF Routing Protocol**
- Lab Exercise on Basic Router and OSPF Dynamic Routing Configuration
- Basic BGP Operation
- BGP Attributes and Path Selection Process
- BGP Scaling Techniques
- Lab Exercise on iBGP, eBGP, RR, Peer group etc
- Internet Exchange [IX] Policy Overview and Configuration requirement

OSPF

- Open Shortest Path First
- Link state or SPF technology
- Developed by OSPF working group of IETF (RFC 1247)
- OSPFv2 (IPv4) standard described in RFC2328
- OSPFv3 (IPv6) standard described in RFC2740
- Designed for:
 - TCP/IP environment
 - Fast convergence
 - Variable-length subnet masks
 - Discontiguous subnets
 - Incremental updates
 - Route authentication
- Runs on IP, Protocol 89

Link State Routing Protocol



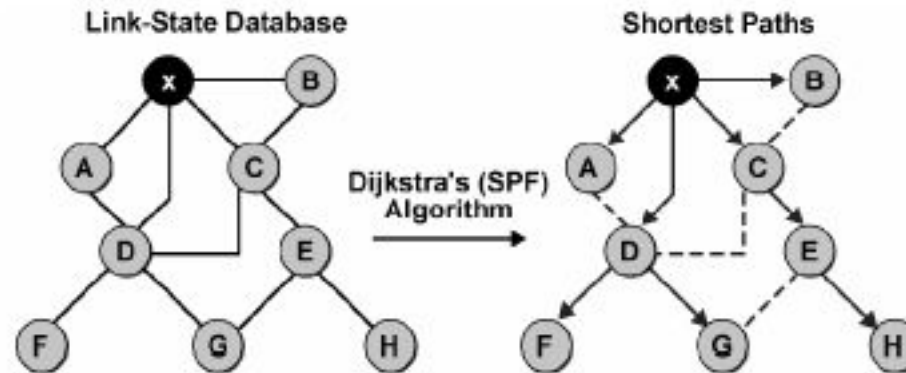
What is Link State Routing

- Do not send full routing table on periodic interval
- Maintain three tables to collect routing information
 - Neighbor table
 - Topology Table
 - Routing table
- Use Shortest Path First (SPF) algorithm to select best path from topology table
- Send very small periodic (Hello) message to maintain link condition
- Send triggered update instantly when network change occur

Link State Data Structure

- Neighbor Table
 - List of all recognized neighboring router to whom routing information will be interchanged
- Topology Table
 - Also called LSDB which maintain list of routers and their link information i.e network destination, prefix length, link cost etc
- Routing table
 - Also called forwarding table contain only the best path to forward data traffic

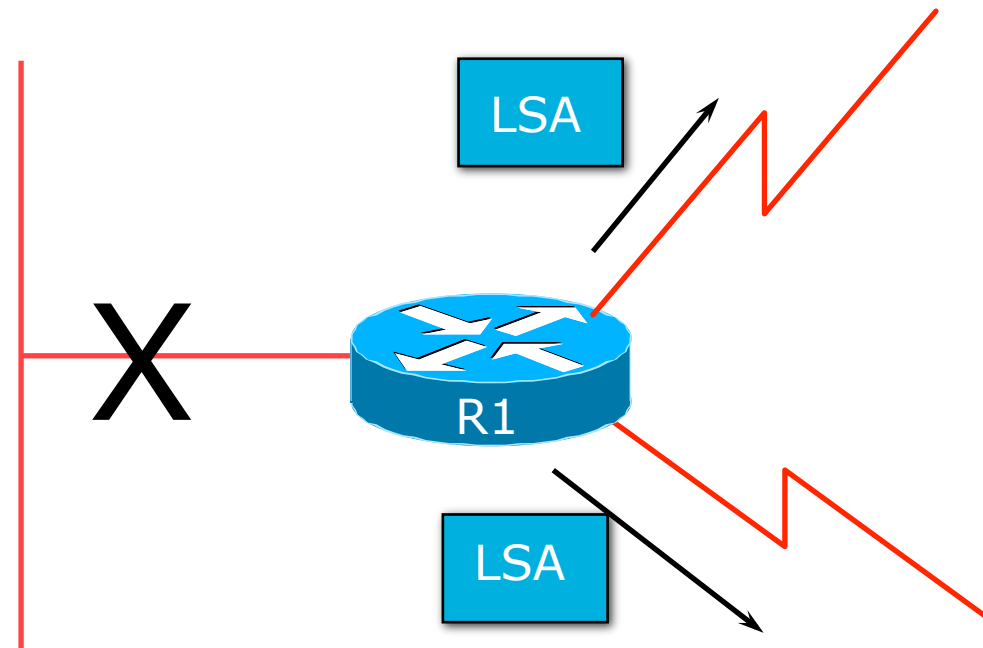
Shortest Path First (SPF) Tree



• Assume all links are Ethernet, with an OSPF cost of 10

- Every router in an OSPF network maintain an identical topology database
- Router place itself at the root of SPF tree when calculate the best path

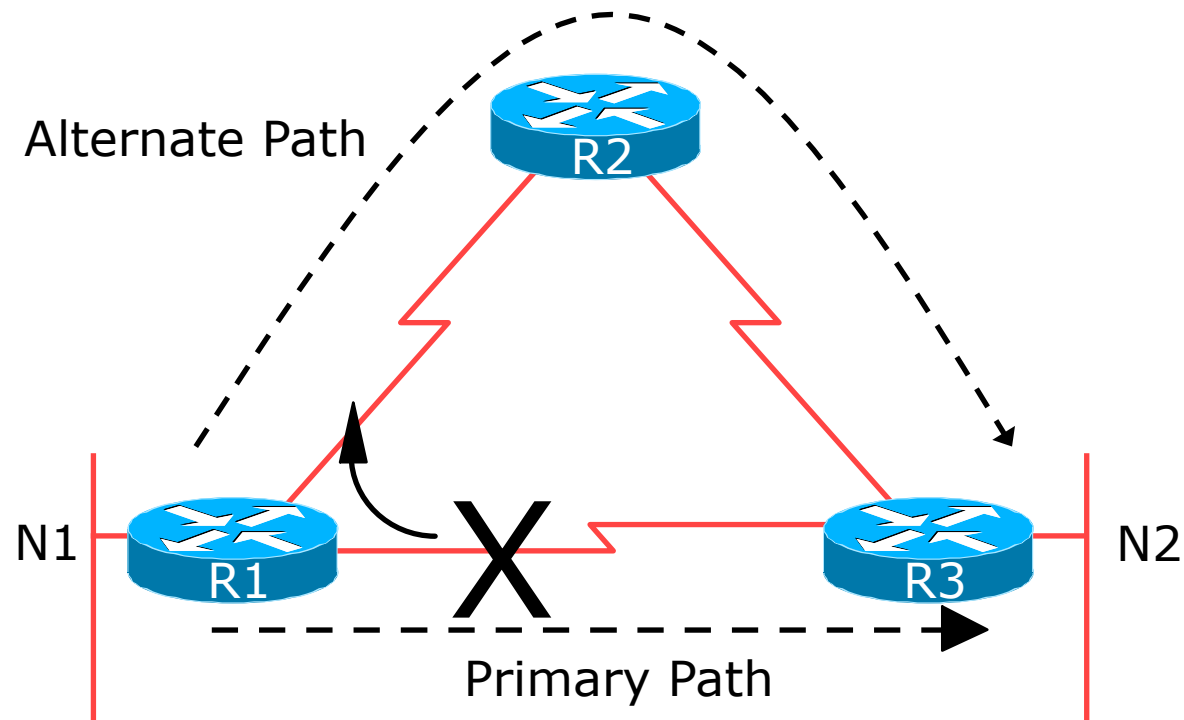
Low Bandwidth Utilisation



- Only changes propagated
- Uses multicast on multi-access broadcast networks

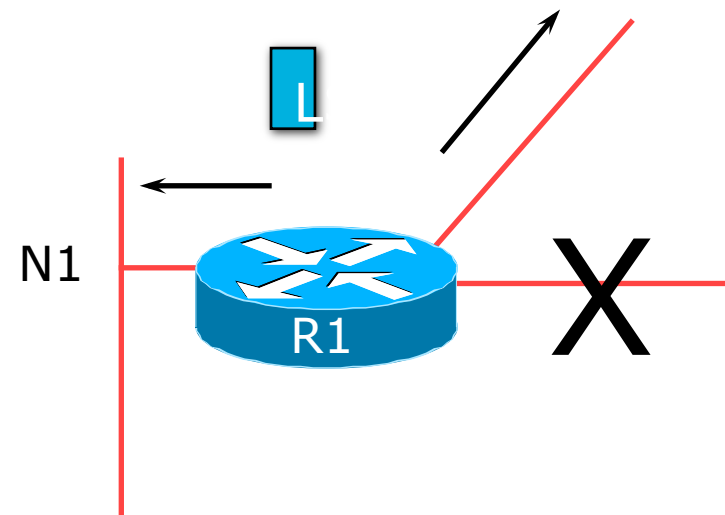
Fast Convergence

- Detection Plus LSA/SPF
 - Known as the Dijkstra Algorithm



Fast Convergence

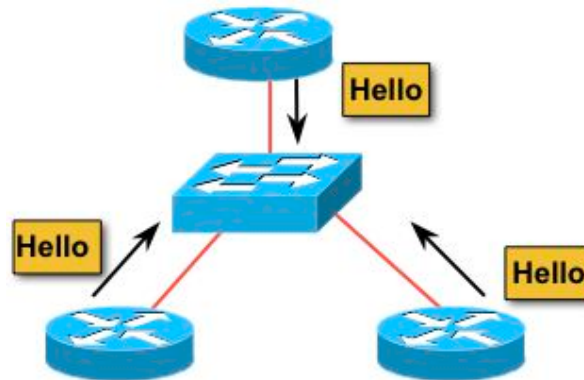
- Finding a new route
 - LSA flooded throughout area
 - Acknowledgement based
 - Topology database synchronised
 - Each router derives routing table to destination network



Basic OSPF Operation

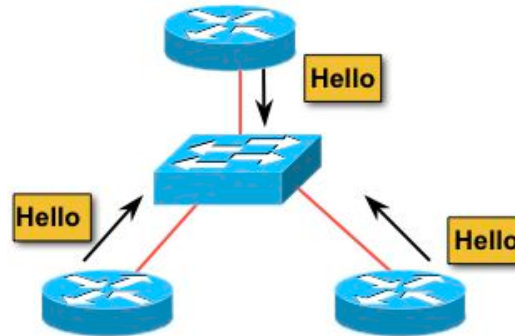
- Neighbor discovery
 - Send L3 multicast message (hello) to discover neighbors
- Exchanging topology table (LSDB)
 - Send L3 multicast message (DBD packets)
- Use SPF algorithm to select best path
 - Each router independently calculates best path from an identical topology database of an OSPF network or area
- Building up routing table
 - All the SPF selected best paths are installed in routing table for the traffic to be forwarded

OSPF Neighbor Discovery Process



- Use IP packet to send hello message. At start routers are at **OSPF Down State**
- Use multicast address 224.0.0.5/FF02::5 to make sure single IP packet will be forwarded to every router within OSPF network. Router now at **OSPF Init State**

OSPF Neighbor Discovery Process



- All neighboring router with OSPF enabled receive the hello packet
- Checks contents of the hello message and if certain information match it reply (Unicast) to that hello with sending its router ID in the neighbor list.
- This is OSPF **Two-way State**

Contents Of A Hello Packet

- Required information to build up adjacency:
 - Router ID of sending router
 - Hello and dead interval time *
 - List of neighbors
 - Network mask
 - Router priority
 - Area ID *
 - DR & BDR IP
 - Authentication information (If any) *

* Need to match to create neighbor relationship

Discovering Network Information

- After creating 2-way neighbor relationship neighboring routers will start exchanging network related information
- At this stage they will decide who will send network information first. Router with the highest router ID will start sending first. This stage is called OSPF **Exstart Stage**
- Then they will start exchanging link state database. This stage is **Exchange Stage**

Adding Network Information

- When router receive the LSDB it perform following action:
 - Acknowledge the receipt of DBD by sending Ack packet (LSAck)
 - Compare the information it received with the existing DB (if any)
 - If the new DB is more up to date the router send link state request (LSR) for detail information of that link. This is **Loading Stage**
- When all LSR have been satisfied and all routers has an identical LSDB this stage is OSPF **Full Stage**

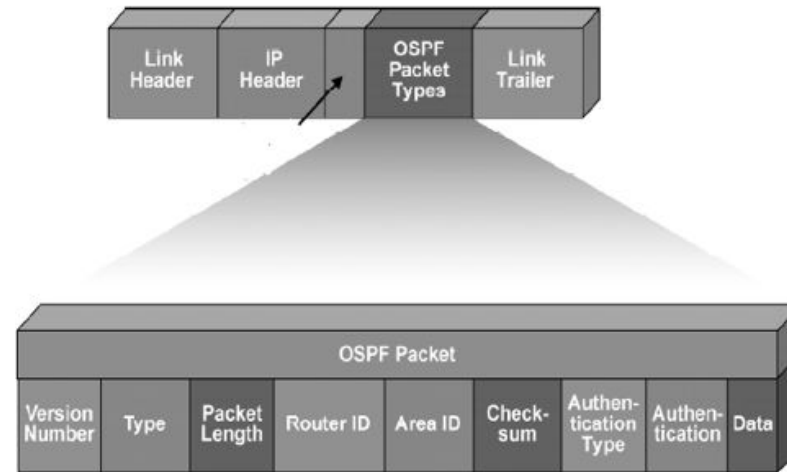
Maintaining Routing Information

- Send periodic updates (Hello) to all neighbors to make sure link with the neighbor is active. I.e 10 sec for LAN
- Send triggered (Instant) update if any network information changed
- Maintain link state sequence number to make sure all information are up-to-date
- Sequence number is 4-byte number that begins with 0x80000001 to 0x7fffffff

OSPF Packet Types

- OSPF use following five packet types to flow routing information between routers:
 - 1: hello [every 10 sec]
 - Hello Builds adjacencies between neighbors
 - 2: DBD [Database Descriptor Packet]
 - DBD for database synchronization between routers
 - 3: LSR [Link State Request Packet]
 - Requests specific link-state records from router to router
 - 4: LSU [Link State Update Packet]
 - Sends specifically requested link-state records
 - 5: LSAck [Link State Ack Packet]
 - Acknowledges the above packet types

Format of OSPF Packet



- All five OSPF packets encapsulated in IP payload (Not TCP)
- To ensure reliable deliver using IP packet OSPF use its own Ack packet (Type 5)

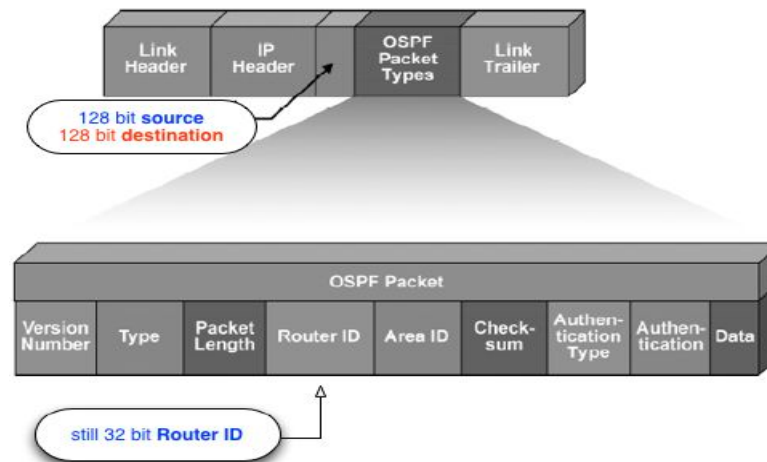
Format of OSPF Packet Header Field

- Version number
 - Either OSPF version 2 (IPv4) or version 3 (IPv6)
- Packet type
 - Differentiates the five OSPF packet types [Type 1 to Type 5]
- Packet length
 - Length of OSPF packet in bytes
- Router ID
 - Defines which router is the source of the packet [Not always source address of IP header]
- Area ID
 - Defines the area where the packet originated
- Checksum
 - Used for packet-header error-detection to ensure that the OSPF packet was not corrupted during transmission
- Authentication type
 - An option in OSPF that describes either clear-text passwords or encrypted Message Digest 5 (MD5) formats for router authentication

Content of OSPF Packet Data

- Data (for hello packet):
 - Contains a list of known neighbors
- Data (for DBD packet):
 - Contains a summary of the LSDB, which includes all known router IDs and their last sequence number, among a number of other fields
- Data (for LSR packet):
 - Contains the type of LSU needed and the router ID of the needed LSU
- Data (for LSU packet):
 - Contains the full LSA entry. Multiple LSA entries can fit in one OSPF update packet
- Data (for LSAck packet):
 - Is empty

Difference is OSPFv3 for IPv6



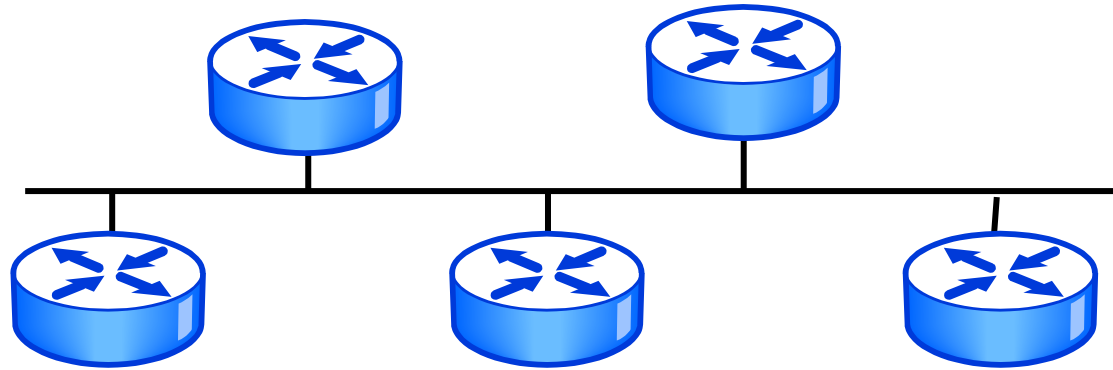
- OSPFv3 still use 32 bit number as router ID
- So OSPFv3 operation and packet types are same as OSPFv2
- Change will be in IP header where source address will be interface address and destination will be FF02::5 which is 128 bit address.
- Change will be in DBD [t2] and LSU packet [t4] to carry 128 bit prefix

OSPF Network Topology

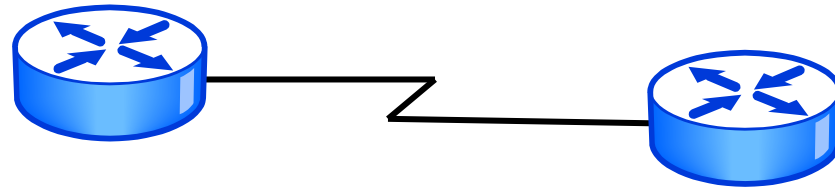
- OSPF network can be made up of different types of network links
- Neighbor relationship behavior will also be different for each network type
- It is important for OSPF to be configured correctly based on its network types to function properly
- Some network type create neighbor relationship automatically, while some need to create it manually

OSPF Network Topology

**Broadcast
Multi-access**



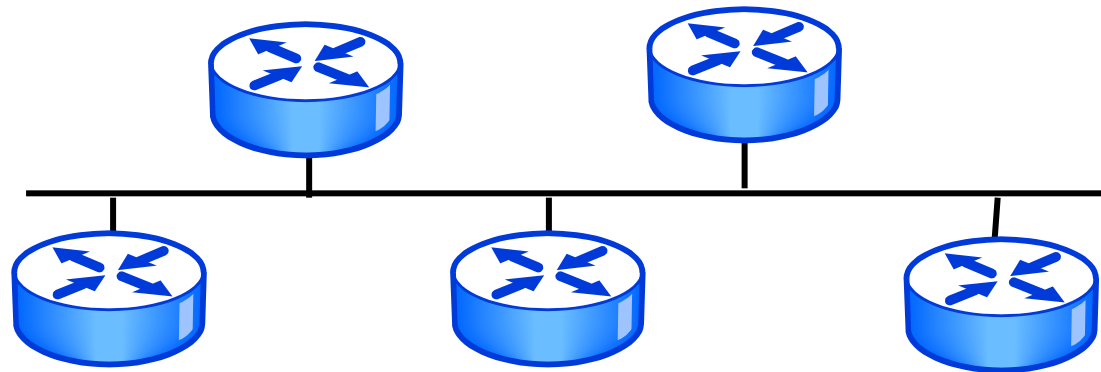
Point-to-Point



**Non Broadcast
Multi-access (NBMA)**



Broadcast Multi-access Network



- Generally LAN type of technologies like Ethernet or Token Ring
- Neighbor relationship are created automatically
- DB/BDR election is required
- Default OSPF hello is 10 sec dead interval is 40 sec

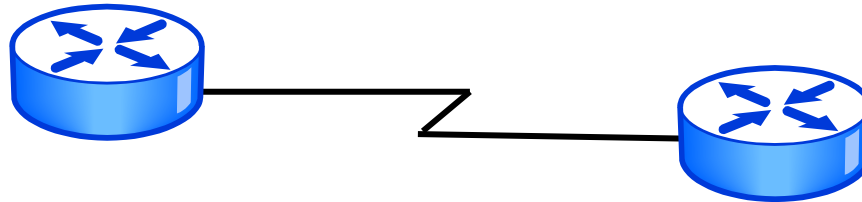
Broadcast Multi-access Network

- Broadcast network use flooding process to send routing update
- Broadcast network use DR/BDR concept to reduce routing update traffic in the LAN
- Packet sent to DR/BDR use 224.0.0.6/FF02::6 multicast address
- Packets from DR to all other routers use 224.0.0.5/FF02::5 multicast address
- All neighbor routers form full adjacencies relation with the DR and BDR only

DB/BDR Election Process

- Router with the highest priority value is the DR, Second highest is BDR
- In the event of tie router with the highest IP address on an interface become DR and second highest is BDR
- DR/BDR election can be manipulated by using router-ID command.
- In practice loopback IP address is used as router ID and the highest IP address will become DR, Second highest is BDR
- The DR/BDR election is non-preemptive
- Generates network link advertisements
- Assists in database synchronization

Point-to-Point Network



- Usually a serial interface running either PPP or HDLC
- Neighbor relationship are created automatically
- No DR or BDR election required
- Default OSPF hello is 10 sec and dead interval is 40 sec

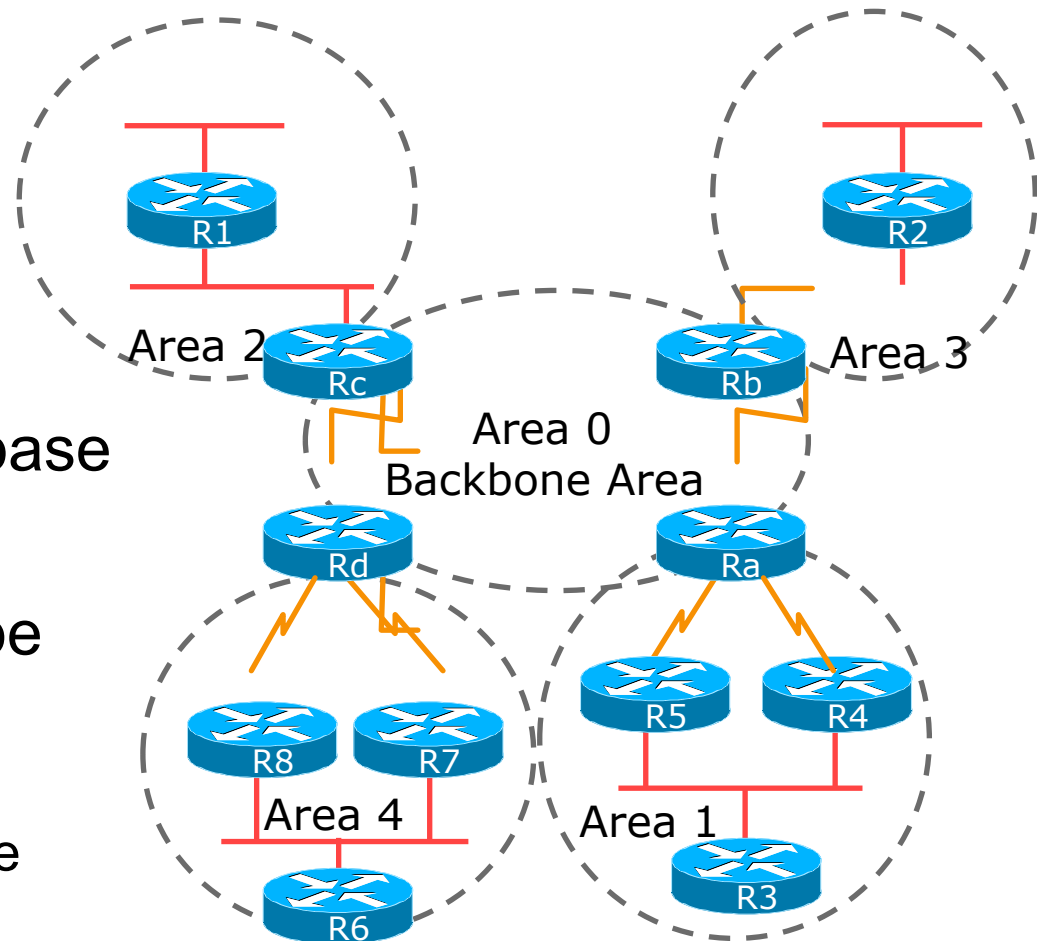
Non Broadcast Multi-access Network



- A single interface interconnects multiple sites like Frame Relay/ATM/X.25
- NBMA topologies support multiple routers, but without broadcasting capabilities
- OSPF neighbor relation need to create manually, DR/BDR will be elected
- Default OSPF hello is 30 sec and dead interval is 120 sec

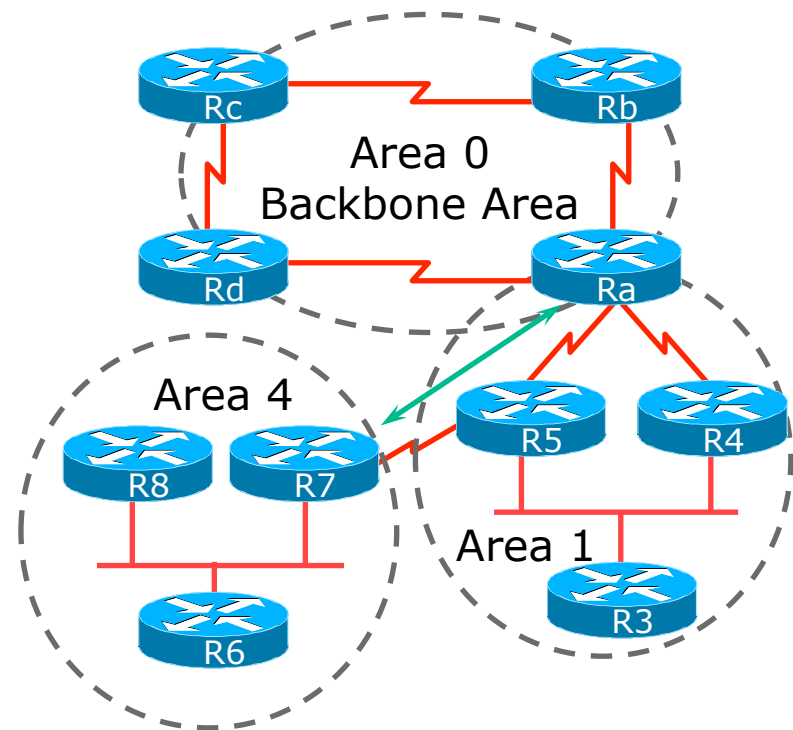
OSPF Areas

- Area is a group of contiguous hosts and networks
 - Reduces routing traffic
- Per area topology database
 - Invisible outside the area
- Backbone area **MUST** be contiguous
 - All other areas must be connected to the backbone

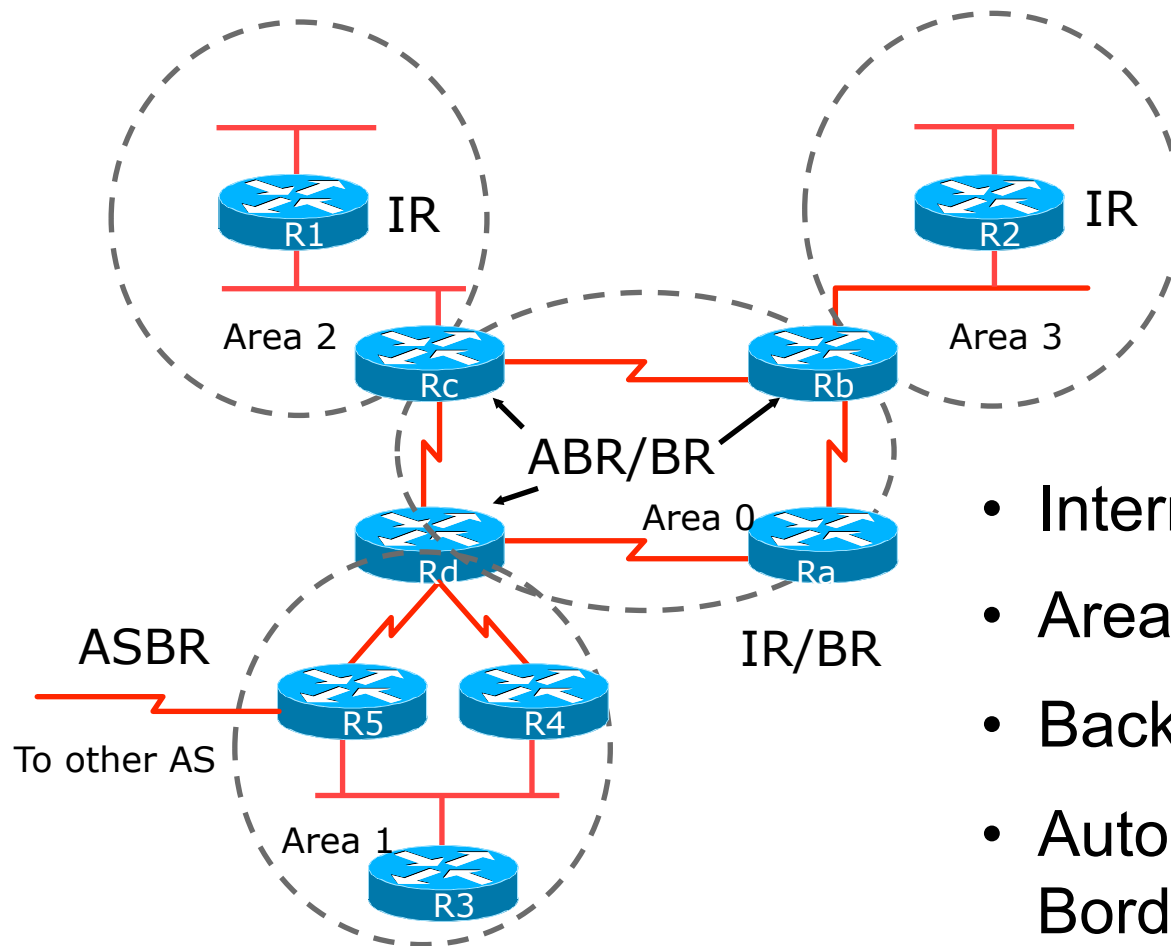


Virtual Links between OSPF Areas

- Virtual Link is used when it is not possible to physically connect the area to the backbone
- ISPs avoid designs which require virtual links
 - Increases complexity
 - Decreases reliability and scalability

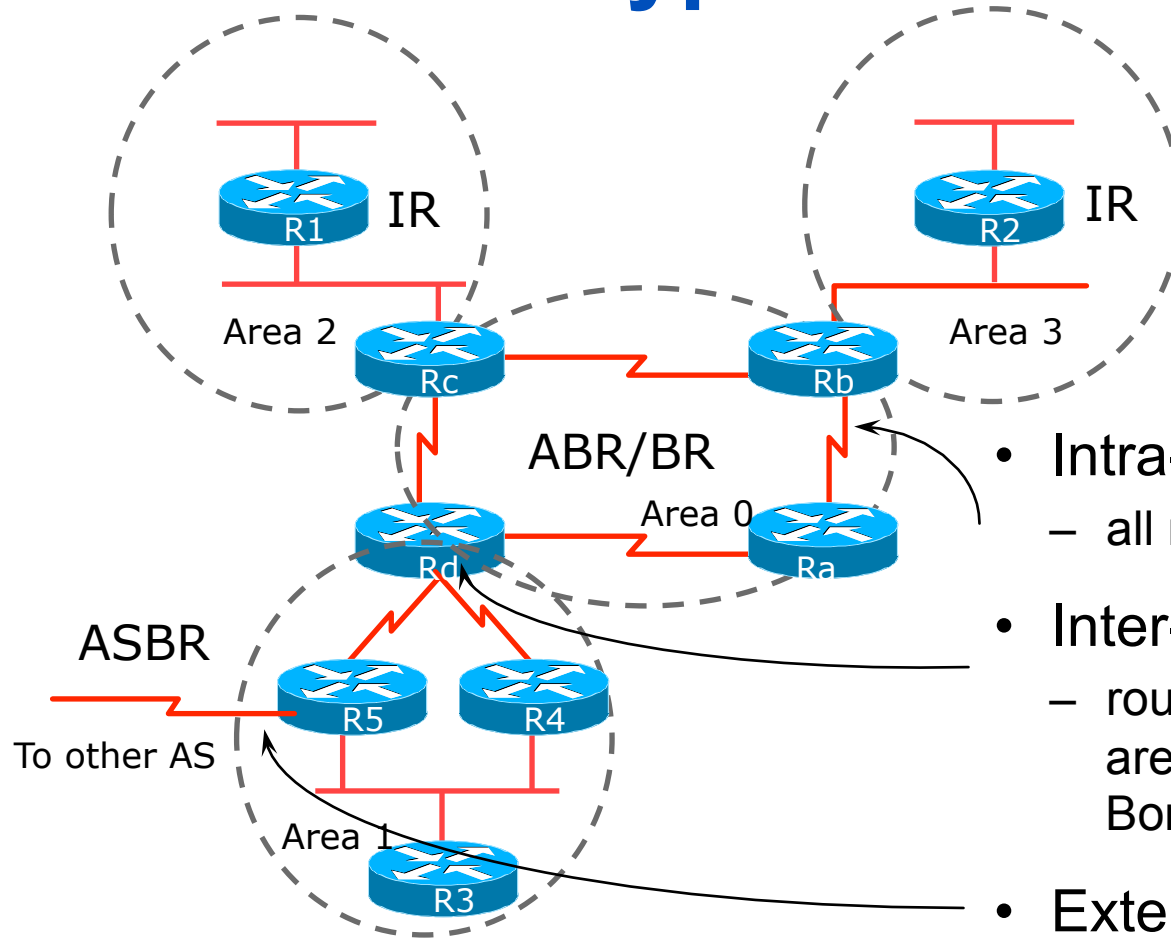


Classification of Routers



- Internal Router (IR)
- Area Border Router (ABR)
- Backbone Router (BR)
- Autonomous System Border Router (ASBR)

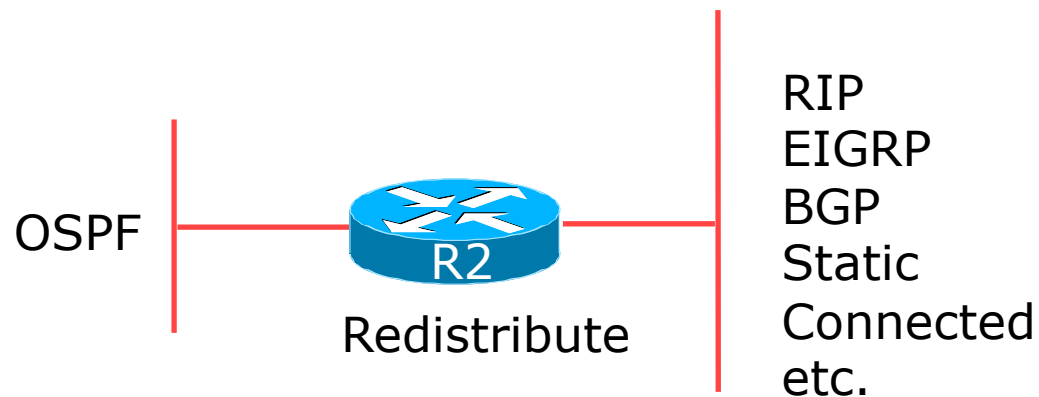
OSPF Route Types



- Intra-area Route
 - all routes inside an area
- Inter-area Route
 - routes advertised from one area to another by an Area Border Router
- External Route
 - routes imported into OSPF from other protocol or static routes

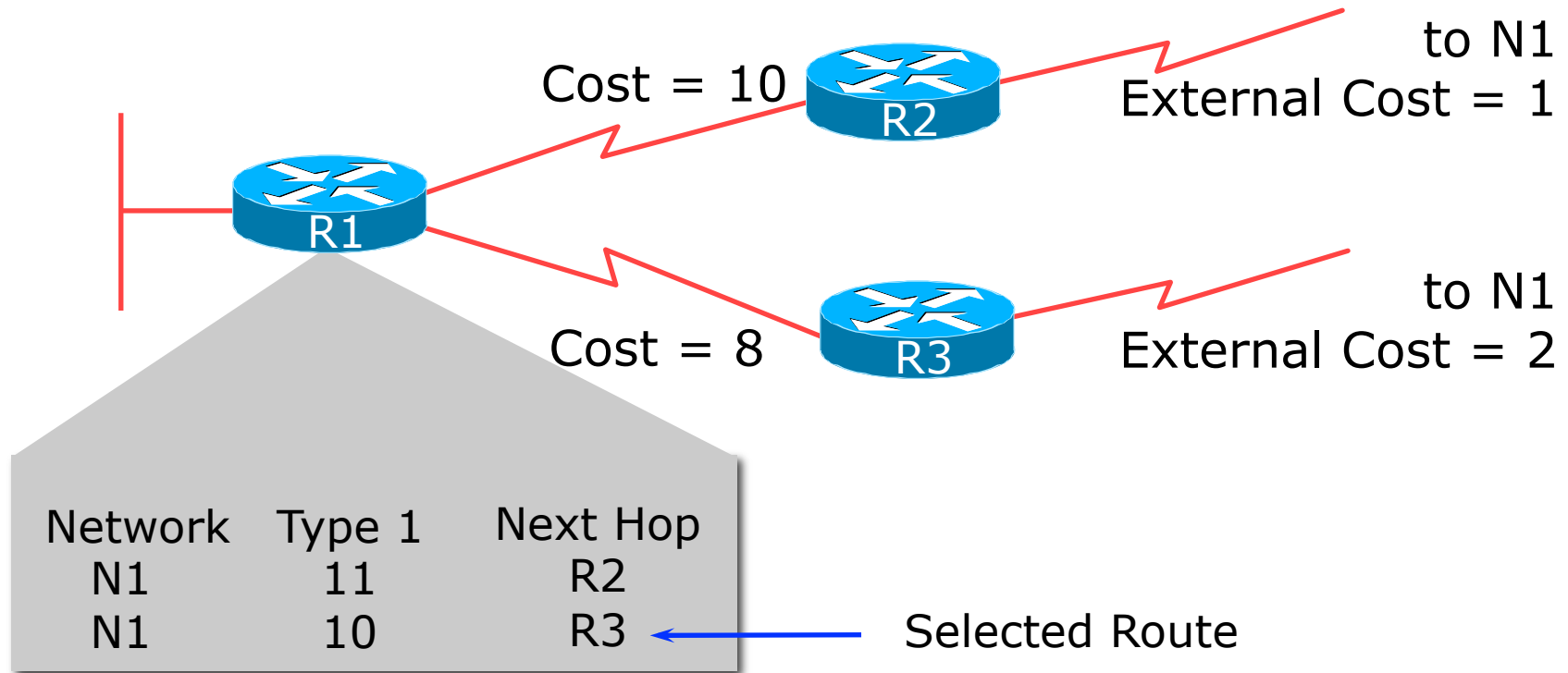
External Routes

- Prefixes which are redistributed into OSPF from other protocols
- Flooded unaltered throughout the AS
 - **Recommendation: Avoid redistribution!!**
- OSPF supports two types of external metrics
 - Type 1 external metrics
 - Type 2 external metrics (Cisco IOS default)



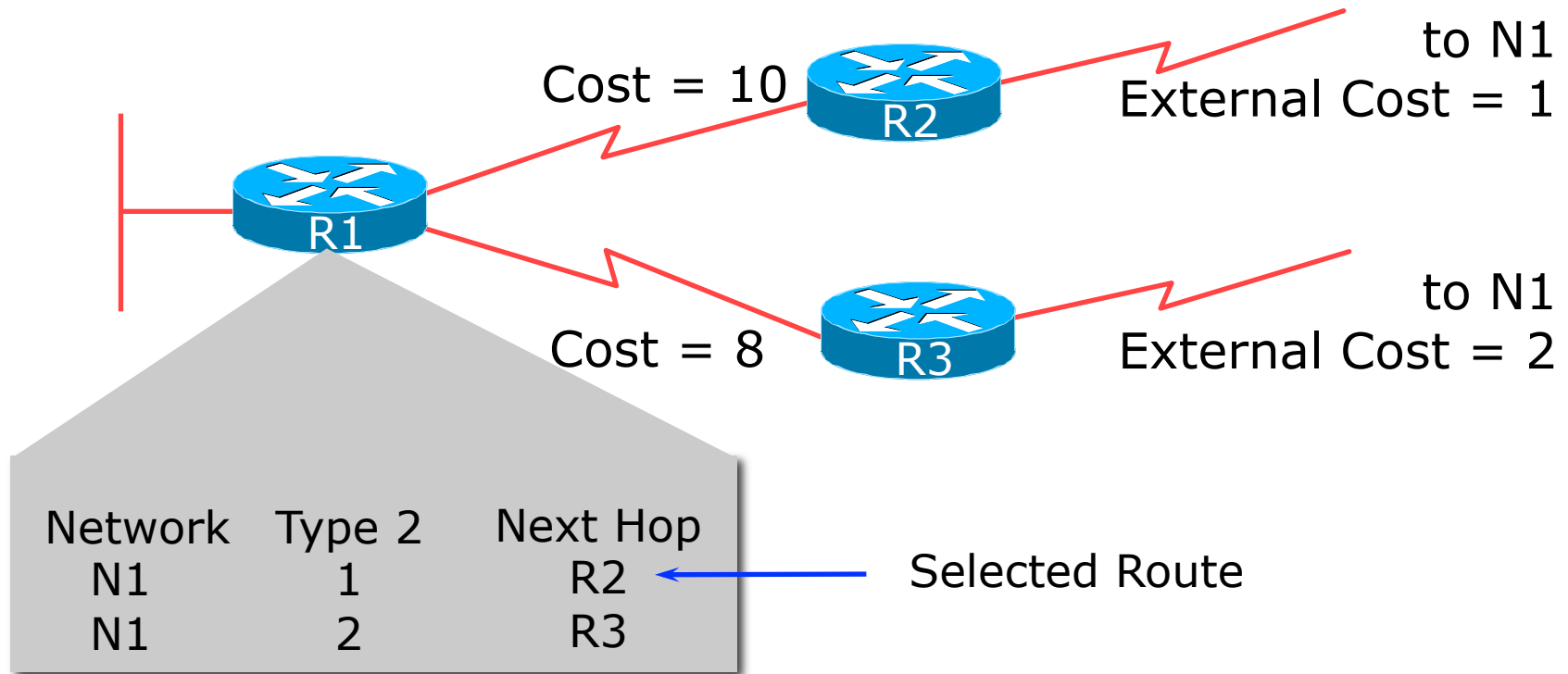
External Routes

- Type 1 external metric: metrics are added to the summarised internal link cost



External Routes

- Type 2 external metric: metrics are compared without adding to the internal link cost



Topology/Link State Database

- A router has a separate LS database for each area to which it belongs
- All routers belonging to the same area have identical database
- SPF calculation is performed separately for each area
- LSA flooding is bounded by area
- Recommendation:
 - Limit the number of areas a router participates in!!
 - 1 to 3 is fine (typical ISP design)
 - >3 can overload the CPU depending on the area topology complexity

Different Types of LSAs

- Six distinct type of LSAs
 - Type 1 : Router LSA
 - Type 2 : Network LSA
 - Type 3 & 4: Summary LSA
 - Type 5 & 7: External LSA (Type 7 is for NSSA)
 - Type 6: Group membership LSA
 - Type 9, 10 & 11: Opaque LSA (9: Link-Local, 10: Area)

Router LSA (Type 1)

- Describes the state and cost of the router's links to the area
- All of the router's links in an area must be described in a single LSA
- Flooded throughout the particular area and no more
- Router indicates whether it is an ASBR, ABR, or end point of virtual link

Network LSA (Type 2)

- Generated for every transit broadcast and NBMA network
- Describes all the routers attached to the network
- Only the designated router originates this LSA
- Flooded throughout the area and no more

Summary LSA (Type 3 and 4)

- Describes the destination outside the area but still in the AS
- Flooded throughout a single area
- Originated by an ABR
- Only inter-area routes are advertised into the backbone
- Type 4 is the information about the ASBR

External LSA (Type 5 and 7)

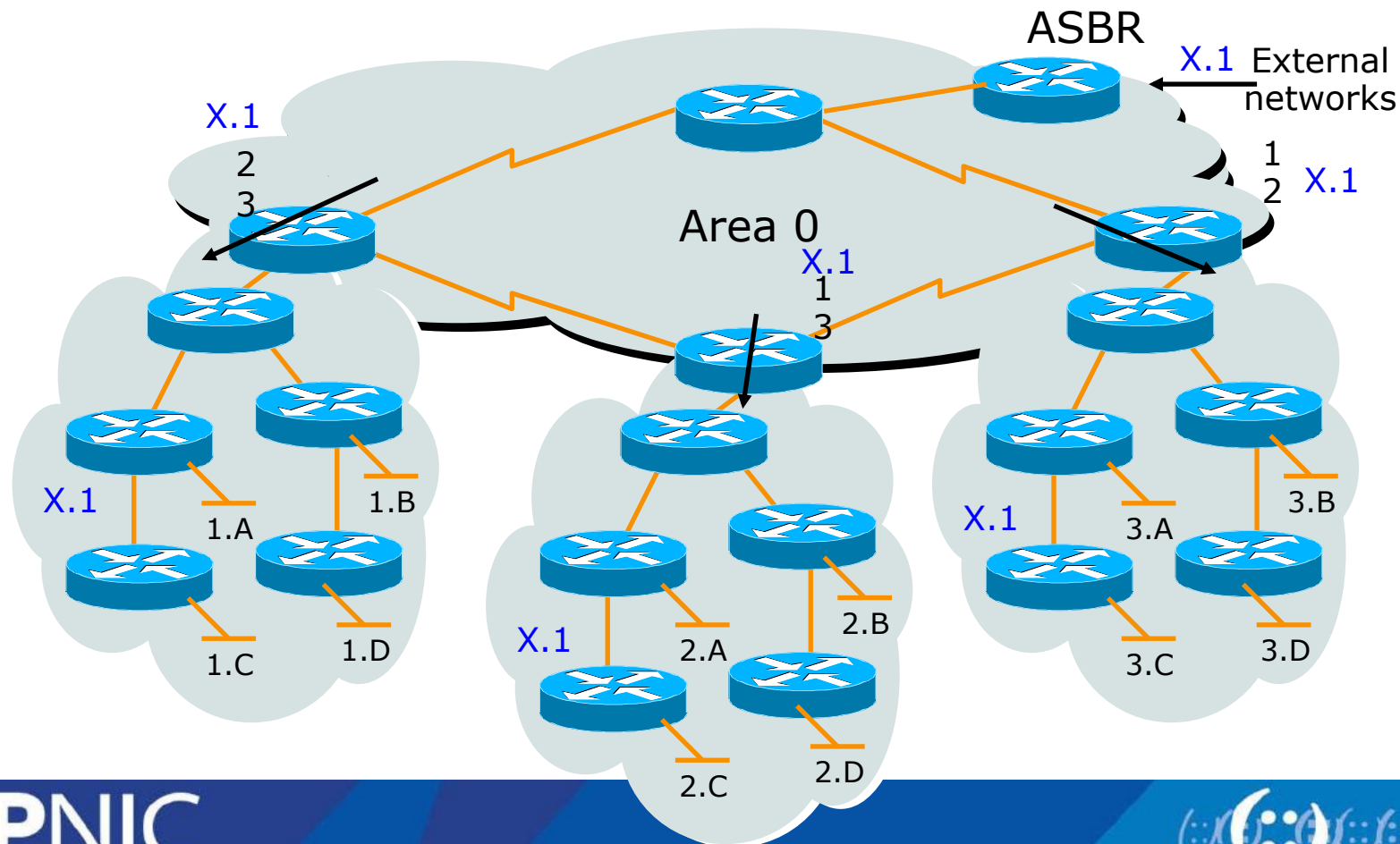
- Defines routes to destination external to the AS
- Default route is also sent as external
- Two types of external LSA:
 - E1: Consider the total cost up to the external destination
 - E2: Considers only the cost of the outgoing interface to the external destination
- (Type 7 LSAs used to describe external LSA for one specific OSPF area type)

Types of Areas

- Regular
- Stub
- Totally Stubby
- Not-So-Stubby
- Only “regular” areas are useful for ISPs
 - Other area types handle redistribution of other routing protocols into OSPF – ISPs don’t redistribute anything into OSPF
- The next slides describing the different area types are provided for information only

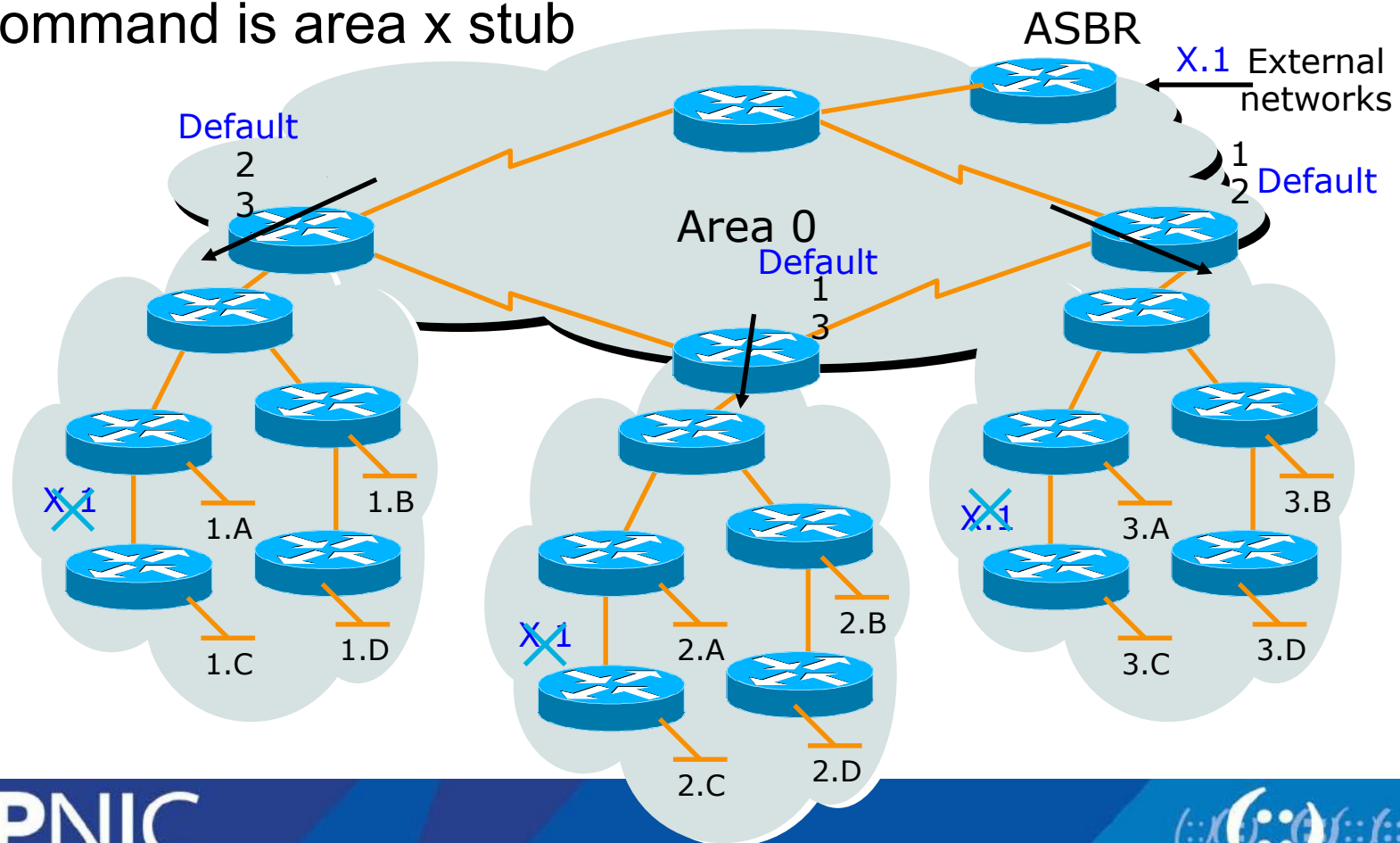
Regular Area (Not a Stub)

- From Area 1's point of view, summary networks from other areas are injected, as are external networks such as X.1



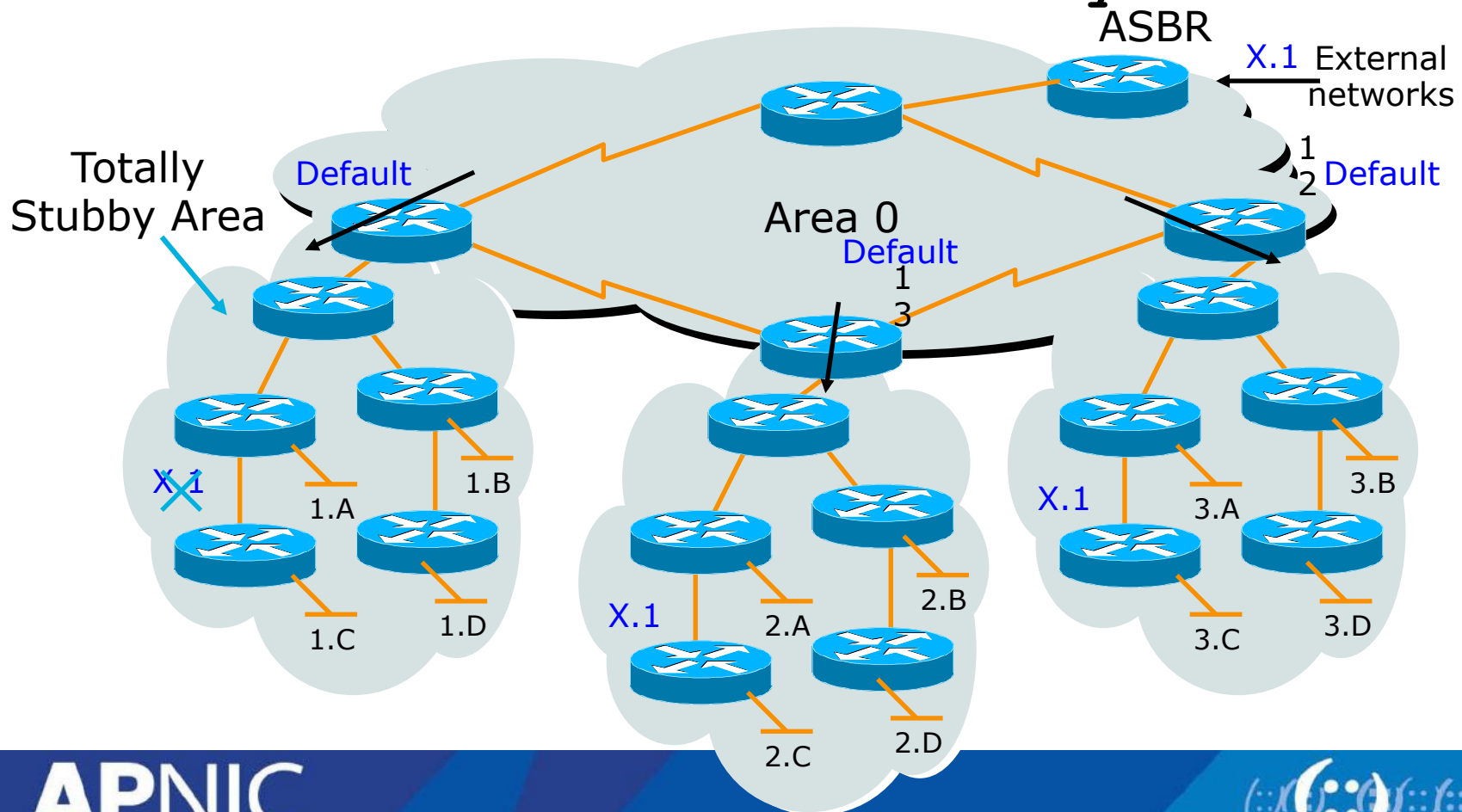
Normal Stub Area

- Summary networks, default route injected
- Command is area x stub



Totally Stubby Area

- Only a default route injected
 - Default path to closest area border router
- Command is **area x stub no-summary**



ISP Use of Areas

- ISP networks use:
 - Backbone area
 - Regular area
- Regular area
 - Summarisation of point to point link addresses used within areas
 - Loopback addresses allowed out of regular areas without summarisation

Questions



Overview

Routing Workshop (3 Days)

- Introduction to IP Routing
- Routing Protocol Basic
- IPv6 Address Structure
- Routing Lab Topology Overview
- Operation of OSPF Routing Protocol
- **Lab Exercise on Basic Router and OSPF Dynamic Routing Configuration**
- Basic BGP Operation
- BGP Attributes and Path Selection Process
- BGP Scaling Techniques
- Lab Exercise on iBGP, eBGP, RR, Peer group etc
- Internet Exchange [IX] Policy Overview and Configuration requirement

Configuration of OSPF as IGP

Scenario:

- Training ISP need to configure OSPF as IGP for both IPv4 and IPv6
- Dual stack mechanism will be used to ensure both IPv4 and IPv6 operation
- OSPFv3 supports IPv6 routed protocol
- IGP is used to carry next hop only for BGP

Configuration of OSPF as IGP

Minimum Router OS require for OSPF3:

- Cisco IOS
 - 12.2(15)T or later (For OSPFv3)
 - 12.2(2)T or later (For IPv6 support)
- Jun OS
 - JUNOS 8.4 or later

Configuration of OSPF as IGP

- Before enabling OSPFv3 on an Interface following steps must be done on a Router:
 - Enable IPv6 unicast routing
 - Enable IPv6 CEF (Optional)

```
config t  
ipv6 unicast-routing  
ipv6 cef (distributed cef)
```


Configuration of OSPF as IGP

Configure interface for both IPv4 and IPv6:

```
interface e1/0
description WAN R1-R2
no ip redirects
no ip directed-broadcast
no ip unreachable
ip address 172.16.10.2 255.255.255.252
no shutdown
```

```
interface e1/0
ipv6 address 2406:6400:000F:0000::2/64
ipv6 enable
```

Configuration of OSPF as IGP

Verify Interface configuration:

```
sh ip interface e0/0
```

```
ping 172.16.10.1
```

```
sh ipv6 interface e0/0
```

```
ping 2406:6400:000F:0000::2
```

Configuration of OSPF as IGP

IPv4 Interface configuration for Router1:

```
interface loopback 0
description Router1 Loopback
no ip redirects
no ip directed-broadcast
no ip unreachables
ip address 172.16.15.1 255.255.255.255
no shutdown

interface e1/0
description WAN R1-R2
no ip redirects
no ip directed-broadcast
no ip unreachables
ip address 172.16.10.2 255.255.255.252
no shutdown
```

Configuration of OSPF as IGP

IPv4 Interface configuration for Router1:

```
interface e1/1
description WAN R1-R3
no ip redirects
no ip directed-broadcast
no ip unreachable
ip address 172.16.10.9 255.255.255.252
no shutdown

interface fa0/0
description Router1 customer network
no ip redirects
no ip directed-broadcast
no ip unreachable
no cdp enable
ip address 172.16.16.1 255.255.255.0
no shutdown
```

Configuration of OSPF as IGP

IPv6 Interface configuration for Router1:

```
interface loopback 0
ipv6 address 2406:6400:0000:0000::1/128
ipv6 enable
interface e1/0
ipv6 address 2406:6400:000F:0000::2/64
ipv6 enable
interface e1/1
ipv6 address 2406:6400:000F:0002::1/64
ipv6 enable
interface fa0/0
ipv6 address 2406:6400:0100:0000::1/48
ipv6 enable
```

Configuration of OSPF as IGP

- OSPF Configuration for IPv4:
 - OSPF for IPv4 can be configured from global configuration mode
 - Interface mode configuration will also activate OSPF process on your running config

Configuration of OSPF as IGP

- OSPF Configuration for IPv6:
 - OSPF for IPv6 need to configure from Interface configuration mode
 - Interface mode configuration will automatically activate OSPF process on your running config

Configuration of OSPF as IGP

OSPF for IPv6 Configuration Command:

```
router ospf 17821
log-adjacency-changes
passive-interface default
network 172.16.15.1 0.0.0.0 area 1
no passive-interface e1/0
network 172.16.10.0 0.0.0.3 area 1
no passive-interface e1/1
network 172.16.10.8 0.0.0.3 area 1
```


Configuration of OSPF as IGP

OSPF for IPv6 Configuration Command:

```
interface loopback 0
```

```
ipv6 ospf 17821 area 1
```

```
interface e1/0
```

```
ipv6 ospf 17821 area 1
```

```
interface e1/1
```

```
ipv6 ospf 17821 area 1
```

Configuration of OSPF as IGP

Verify OSPF configuration:

```
sh run
```

```
!
```

```
interface Ethernet1/0
```

```
description WAN R1-R2
```

```
ip address 172.16.10.2 255.255.255.252
```

```
no ip redirects
```

```
no ip unreachable
```

```
half-duplex
```

```
ipv6 address 2406:6400:F::2/64
```

```
ipv6 enable
```

```
ipv6 ospf 17821 area 1
```

Configuration of OSPF as IGP

Example OSPF configuration for Router1:

```
router ospf 17821
log-adjacency-changes
passive-interface default
network 172.16.15.1 0.0.0.0 area 1
no passive-interface e1/0
network 172.16.10.0 0.0.0.3 area 1
no passive-interface e1/1
network 172.16.10.8 0.0.0.3 area 1
```

Configuration of OSPF as IGP

Example OSPF configuration for Router1:

```
interface loopback 0
ipv6 ospf 17821 area 1
interface e1/0
ipv6 ospf 17821 area 1
interface e1/1
ipv6 ospf 17821 area 1
```

OSPF Packet Type

Five OSPF Packet Type:

t: Specifies the OSPF packet type:

- | | |
|----------|------------------------------|
| 1: hello | [every 10 sec] |
| 2: DBD | [Database Descriptor Packet] |
| 3: LSR | [Link State Request Packet] |
| 4: LSU | [Link State Update Packet] |
| 5: LSAck | [Link State Ack Packet] |

debug ip ospf packet

debug ipv6 ospf packet

Questions



Overview

Routing Workshop (3 Days)

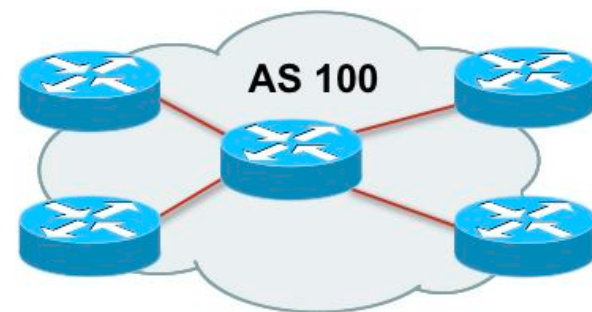
- Introduction to IP Routing
- Routing Protocol Basic
- IPv6 Address Structure
- Routing Lab Topology Overview
- Operation of OSPF Routing Protocol
- Lab Exercise on Basic Router and OSPF Dynamic Routing Configuration
- **Basic BGP Operation**
- BGP Attributes and Path Selection Process
- BGP Scaling Techniques
- Lab Exercise on iBGP, eBGP, RR, Peer group etc
- Internet Exchange [IX] Policy Overview and Configuration requirement

What is Border Gateway Protocol?

- BGP:
 - A path vector routing protocol to exchange routing information between different Autonomous System (AS)
 - ASes are the building block of BGP operational unites
 - AS is a collection of routers with a common routing policy
 - Specification is defined in RFC4271

What is an Autonomous System (AS)

- An AS is a collection of networks with same routing policy
- Usually under a single administrative control unit
- A public AS is identified by a unique number called AS number
- Around 32000 ASes are visible on the Internet now



BGP Features

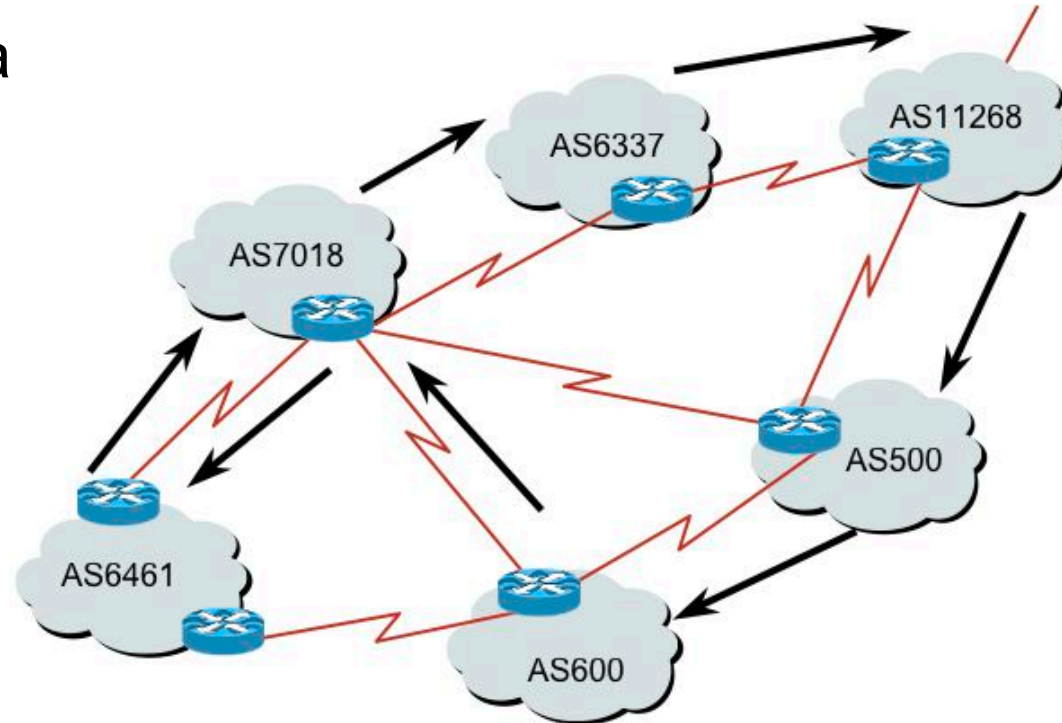
- Path Vector Routing Protocol
- Send incremental updates to peers
- Runs over TCP –Port 179
- Select path based on routing policy/ organization' s business requirement
- Support Classless Inter Domain Routing (CIDR) concept
- Widely used in today' s Internet Backbone
- Current BGP version is MP-BGP

What is Path Vector Routing Protocol

- A path vector routing protocol is used to span different autonomous systems
- It defines a route as a collection of a number of AS that it passes through from source AS to destination AS
- This list of ASes are called AS path and used to avoid routing loop
- AS path is also used to select path to destination

What is AS path?

- An AS path example



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

AS Path

BGP Traffic Arrangement Definition

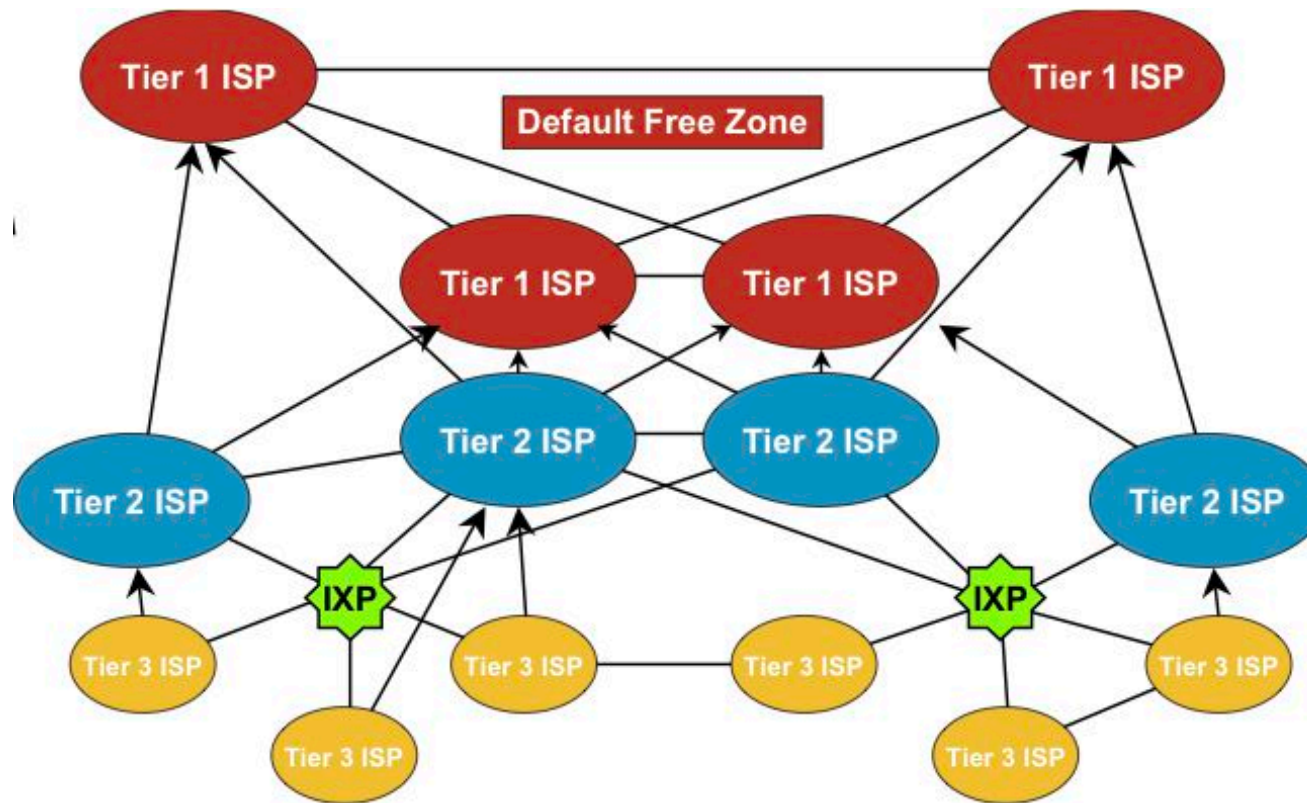
- Transit
 - Forwarding traffic through the network usually for a fee
 - I.e Internet service from upstream ISP
- Peering
 - Exchanging traffic without any fee
 - I.e Connection in an IXP
- Default
 - Where to send traffic if there no explicit route match in the routing table

What is Default Free Zone?

- Default free zone is made up of Tier One ISP routers which have explicit routing information about every part of the Global Internet
- So there is no need of default route
- If there is no destination network match, then that prefix is still not announced/ used by any ISP yet

ISP Hierarchical Connection

- Connectivity Diagram:



BGP General Operation

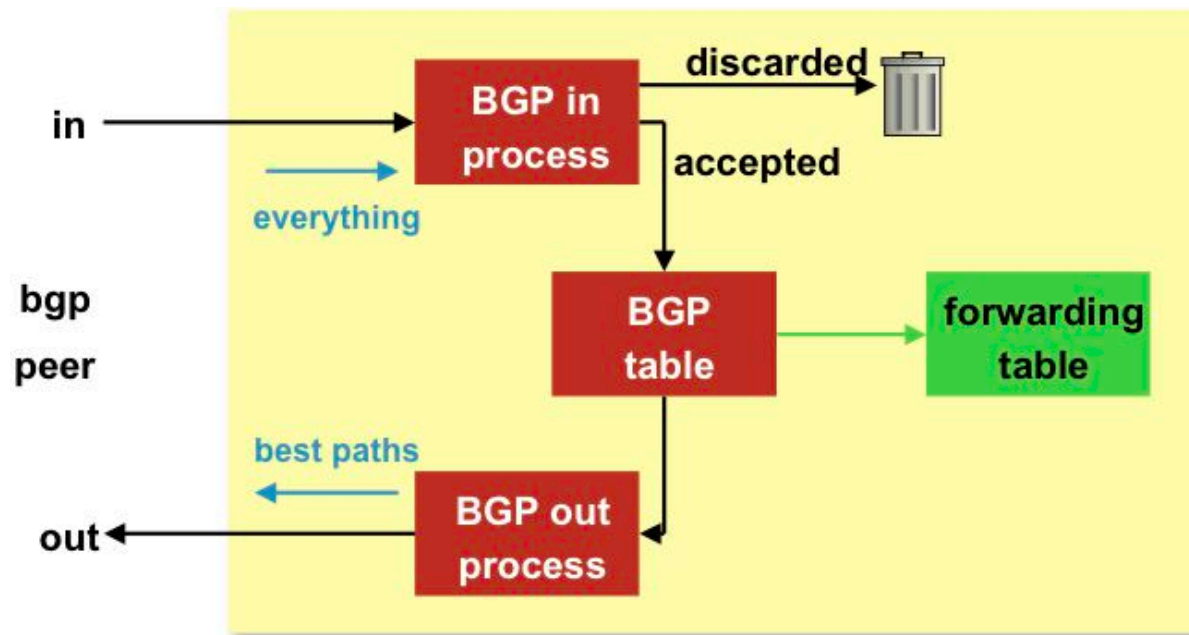
- BGP maintain 3 database i.e Neighbor Table, BGP Table and Forwarding Table
- Learns multiple paths via internal and external BGP speakers
- Picks the best path and installs them on the forwarding tables
- Best path is sent to external BGP neighbors
- Policies are applied by influencing the best path selection

Constructing the Forwarding Table

- BGP “In” process
 - Receives path information from peers
 - Results of BGP path selection placed in the BGP table “best path” flagged
- BGP “Out” process
 - Announce “best path” information to peers
- Best path installed in forwarding table if:
 - Prefix and prefix length are equal
 - Lowest protocol distance

Constructing the Forwarding Table

- Flowchart:



BGP Terminology

- Neighbor
 - Any two routers that have formed a TCP connection to exchange BGP routing information are called peers or neighbors
- iBGP
 - iBGP refers to the BGP neighbor relationship within the same AS.
 - The neighbors do not have to be directly connected.
- eBGP
 - When BGP neighbor relationship are formed between two peers belongs to different AS are called eBGP.
 - EBGP neighbors by default need to be directly connected.

Building Neighbor Relationship

- After adding BGP neighbor:
 - Both router establish a TCP connection and send open message
 - If open message is accepted then both send keepalive message to each other to confirm open message
 - After both confirm open message by sending keepalive message they establish BGP neighbor relationship and exchange routing information

BGP message type

- Open Message
 - To establish BGP neighbor relationship
- Keepalive message
 - Only contain message header to maintain neighbor relationship. Sent every periodic interval
- Update message
 - Contain path information. One update message contain one path information. Multiple path need multiple update message to be sent
- Notification message
 - Sent when an error condition occur and BGP connection closed immediately

BGP Open message

- Open message contain:
 - BGP Version number
 - AS number of the local router
 - BGP holdtime in second to elapse between the successive keepalive message
 - BGP router ID which is a 32 bit number. Usually an IPv4 address is used as router ID
 - Optional parameters i.e types, length and value encoded. An example optional parameter is session authentication info

BGP Keepalive Message

- Send between BGP peers after every periodic interval (60 sec)
- It refresh hold timer from expiration (180sec)
- A keepalive message contain only the message header

BGP Update Message

- An update message contain:
 - Withdrawn routes: a list contain address prefix that are withdrawn from service
 - Path attributes: includes AS path, origin code, local pref etc
 - Network-layer reachability information: includes a list of address prefix reachable by this path

BGP Notification message

- Only sent when an error condition occur and detected in a network and BGP connection is closed immediately
- Notification message contain an error code, an error subcode, and data that are related to that error

BGP Neighbor Relationship States

- BGP neighbor goes through following steps:
 - Idle: Router is searching its routing table to reach the neighbor
 - Connect: Router found route and completed TCP three-way handshake
 - Open Sent: Open message sent with the parameter for BGP session
 - Open Confirm: Router receive agreement on the parameter to establish BGP session
 - Established: Peering is established and routing information exchange began

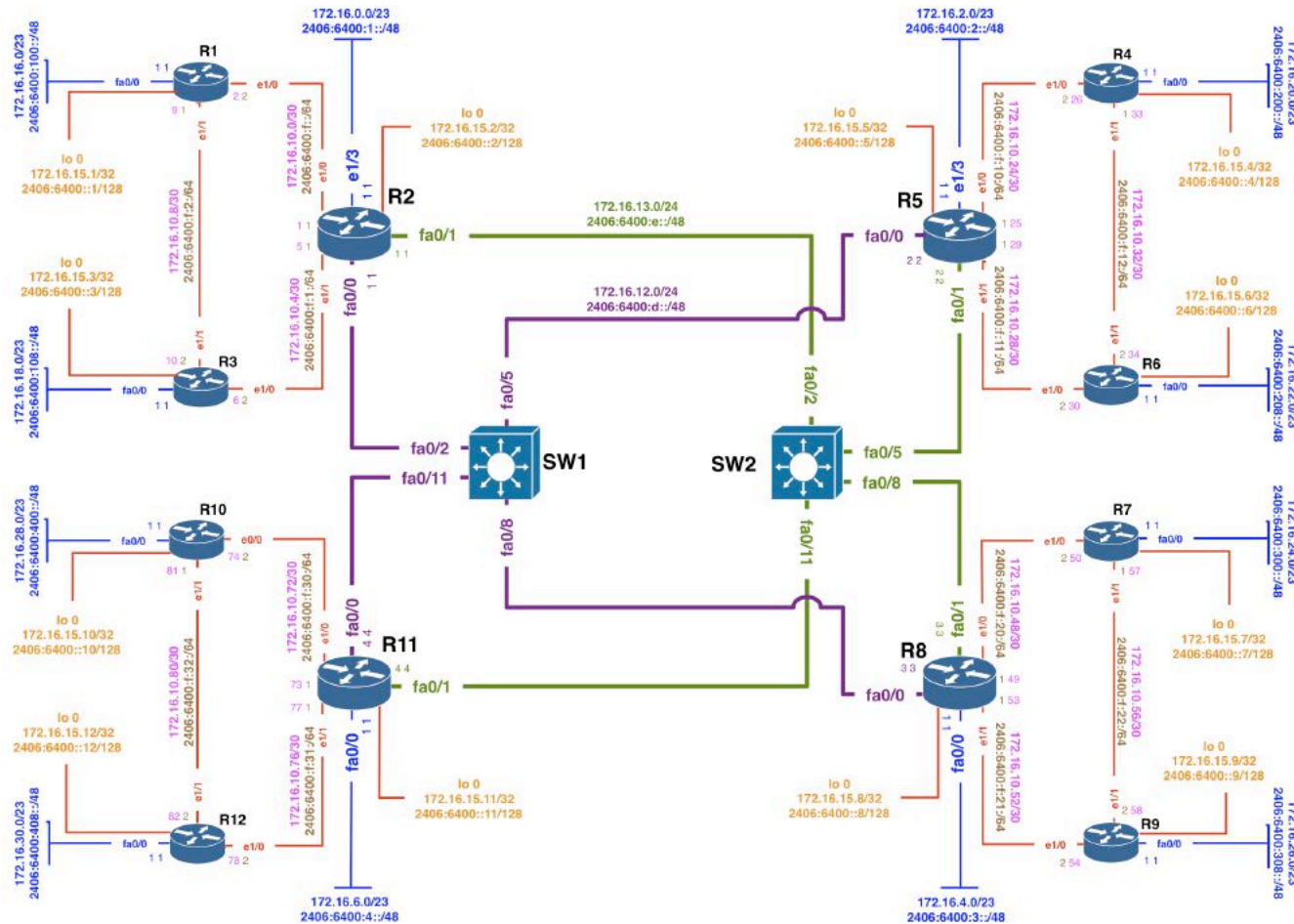
Troubleshoot BGP Neighbor Relation

- Idle:
 - The router can not find address of the neighbor in its routing table
- Active:
 - Router found address of the neighbor in its routing table sent open message and waiting for the response from the neighbor
- Cycle between Active/Idle
 - Neighbor might peer with wrong address
 - Does not have neighbor statement on the other side
 - BGP open message source IP address does not match with remote side neighbor statement or no route to source IP address

iBGP Peering

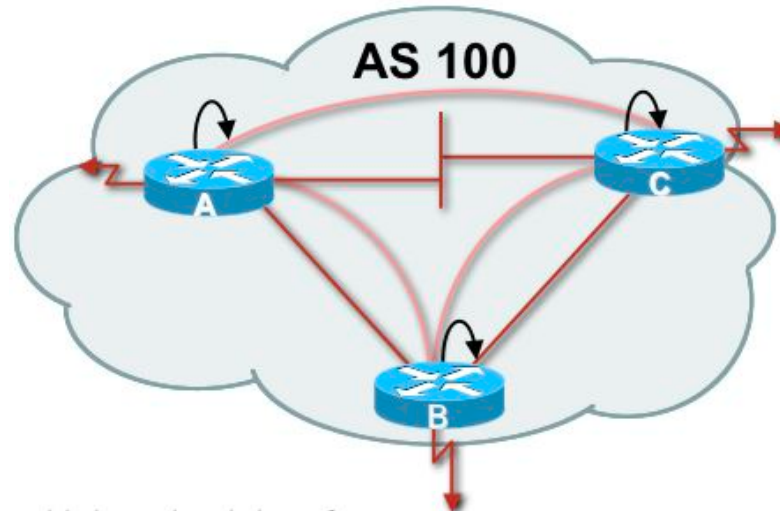
- BGP peer within the same AS
- Not required to be directly connected
- iBGP peering require full mesh peering
 - Within an AS all iBGP speaker must peer with other iBGP speaker
 - They originate connected network
 - Pass on prefixes learned from outside AS
 - They do not forward prefixes learned from other iBGP peer

Training ISP IPV6 Addressing Plan



Training ISP IPv6 Address Plan

iBGP Peering with Loopback Interface

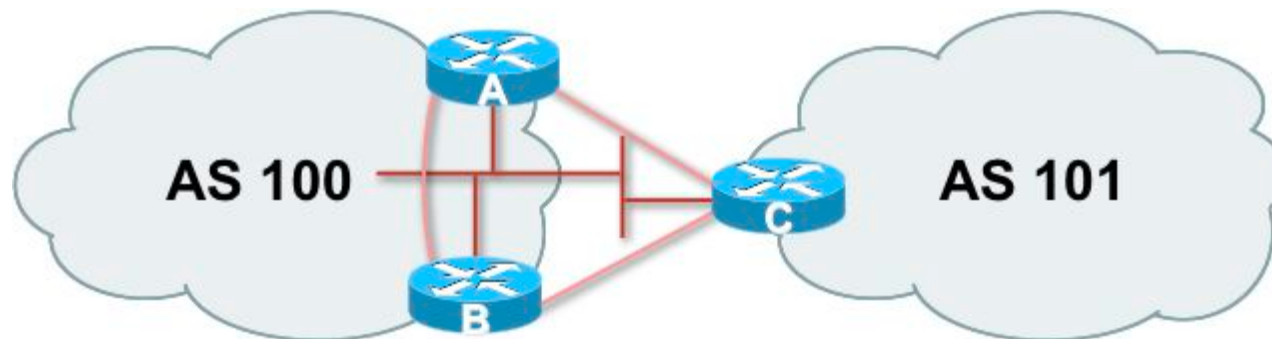


- If iBGP speakers has multiple connection then it is advisable to peer with loopback
- Connected network can go down which might loose iBGP peering
- Loopback interface will never go down

iBGP Neighbor Update Source

- This command allows the BGP process to use the IP address of a specified interface as the source IP address of all BGP updates to that neighbor
- A loopback interface is usually used as it will never goes down as long as the router is operational
- All BGP message will use the referenced interface as source of the messages

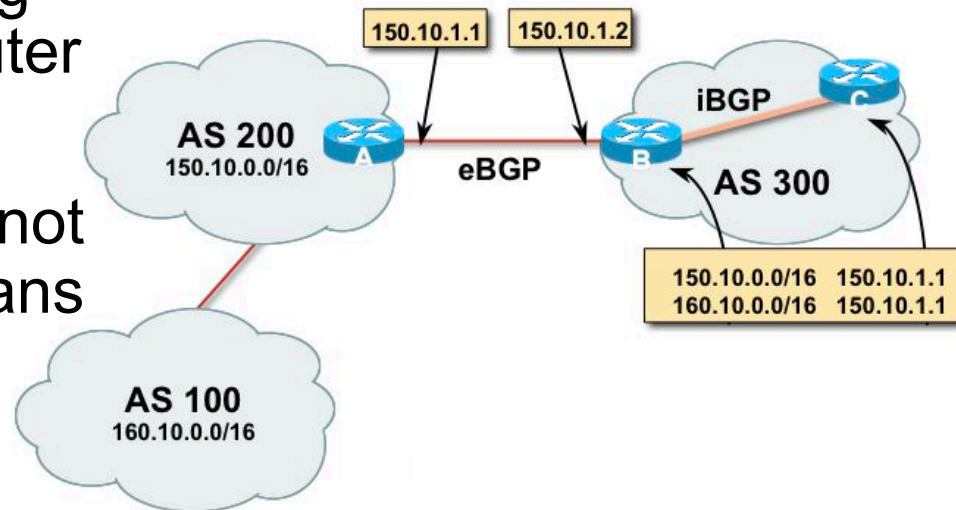
eBGP Peering



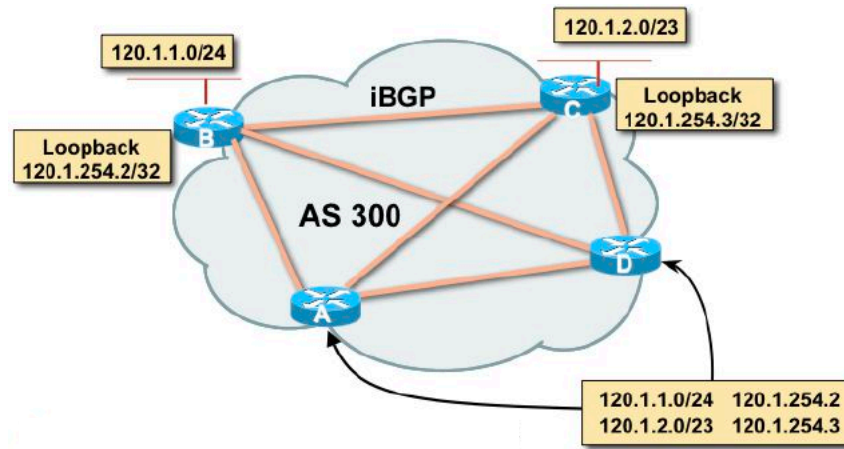
- Peering with BGP speaker in different AS
- Peers should be directly connected and share same WAN link
- eBGP neighbors are usually routed through connected network

BGP Next Hop Behavior

- BGP is an AS-by-AS routing protocol not a router-by-router routing protocol.
- In BGP, the next hop does not mean the next router it means the IP address to reach the next AS
 - I.e Router A advertise 150.10.0.0/16 and 160.10.0.0/16 to router B in eBGP with next hop 150.10.1.1
 - Router B will update Router C in iBGP keeping the next hop unchanged



iBGP Next Hop



- Next hop is iBGP router loopback address
- Recursive route look-up
- Loopback address need to announce through IGP (OSPF)

BGP Synchronous Rule

- BGP do not use or advertise any route to an external neighbor learned by iBGP until a matching route has been learned from an IGP i.e OSPF or static
- It ensure consistency of information throughout the AS
- Avoid black hole route within an AS
- It is safe to turn off if all routers with in the AS run full-mesh iBGP
- Advisable to disable this feature (BCP)

Questions



Overview

Routing Workshop (3 Days)

- Introduction to IP Routing
- Routing Protocol Basic
- IPv6 Address Structure
- Routing Lab Topology Overview
- Operation of OSPF Routing Protocol
- Lab Exercise on Basic Router and OSPF Dynamic Routing Configuration
- Basic BGP Operation
- **BGP Attributes and Path Selection Process**
- BGP Scaling Techniques
- Lab Exercise on iBGP, eBGP, RR, Peer group etc
- Internet Exchange [IX] Policy Overview and Configuration requirement

BGP Attributes

BGP metrics are called path attributes. Here is the classifications BGP attributes:

Well-known mandatory

- AS-Path
- Next-hop
- Origin

Well-known discretionary

- Local preference
- Atomic aggregate

Optional transitive

- Community
- Aggregator

Optional non-transitive

- Multi-exit-discriminator (MED)

Well-Known Attributes

- Must be recognized by all compliant BGP implementations
- Are propagated to other neighbors

Well-Known Mandatory Attributes

- Must be present in all update messages
- ***AS Path***
- ***Next-hop***
- ***Origin***

Well-Known Discretionary Attributes

- May be present in update messages
- ***Local preference***
- ***Atomic aggregate***

Optional Attributes

- Recognized by some implementations (could be private) expected not to be recognized by everyone
- Recognized optional attributes are propagated to other neighbors based on their meaning

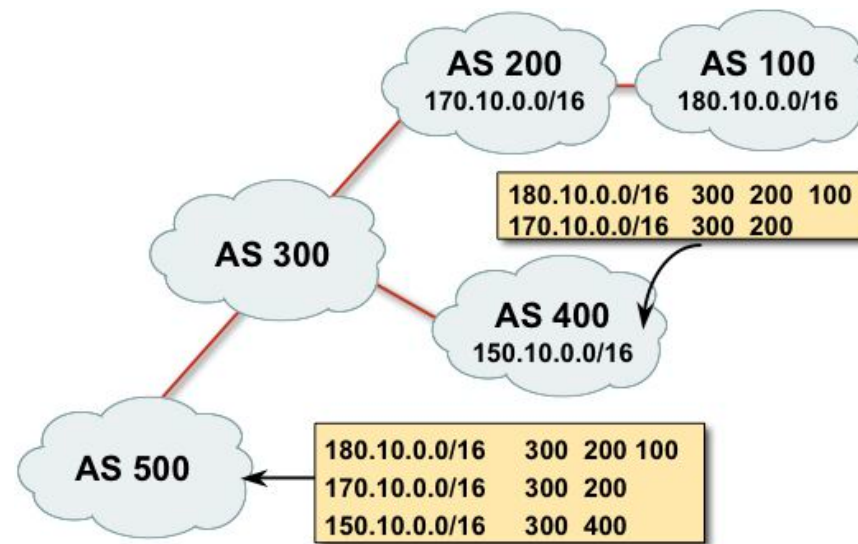
Optional Transitive Attributes

- If not recognized, are marked as partial and propagated to other neighbors
- **Community**
- **Aggregator**

Optional Non Transitive attributes

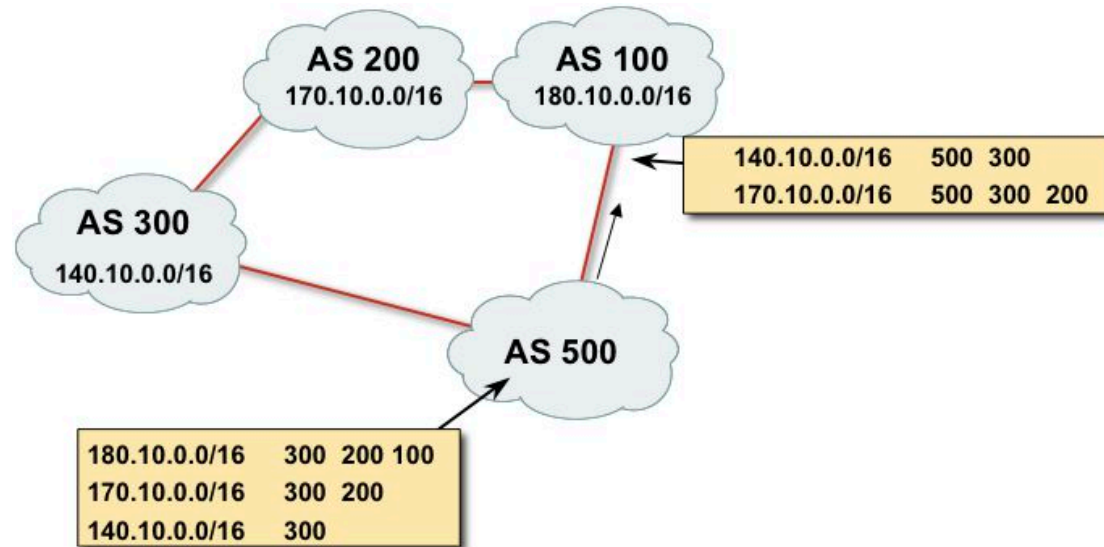
- Discarded if not recognized
- **Multi Exit Discriminator (MED)**

AS Path Attribute



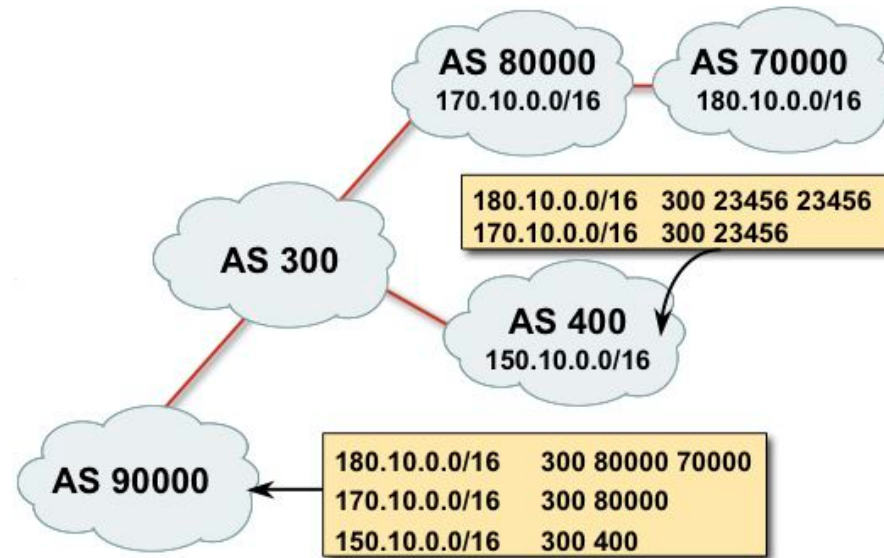
- Sequence of ASes a route has traversed
- Used for
 - Loop detection
 - Path metrics where the length of the AS Path is used as in path selection

AS Path Loop Detection



- 180.10.0.0/16 is not accepted by AS100 as the prefix has AS100 in its AS-PATH
- This is loop detection in action

AS Path Attribute (2 byte and 4 byte)



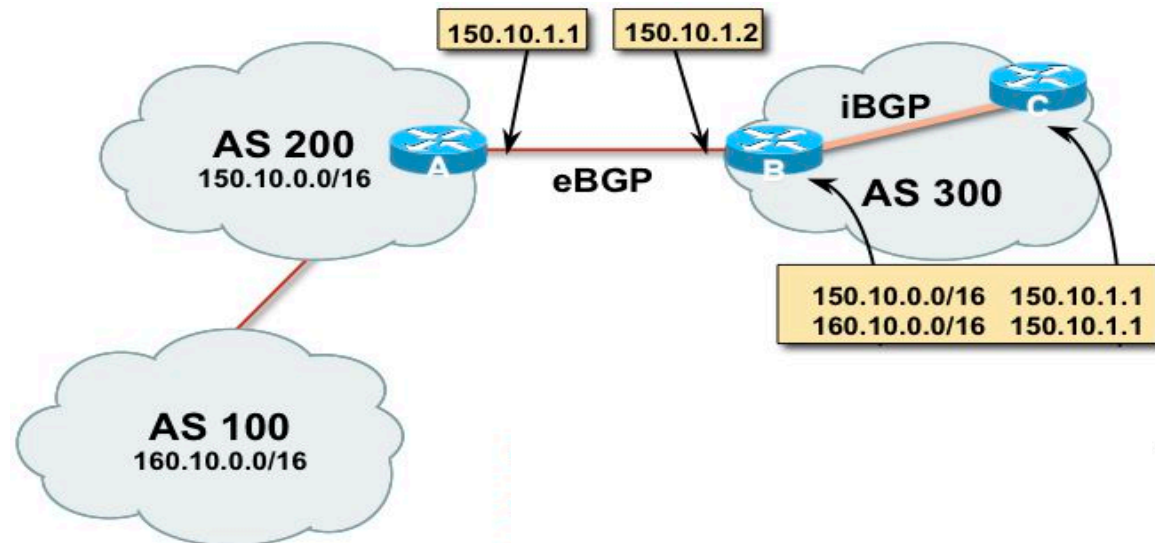
- Internet with 16-bit and 32-bit ASNs
 - 32-bit ASNs are 65536 and above
 - AS-PATH length maintained

AS Path and AS4 Path Example

Router5:

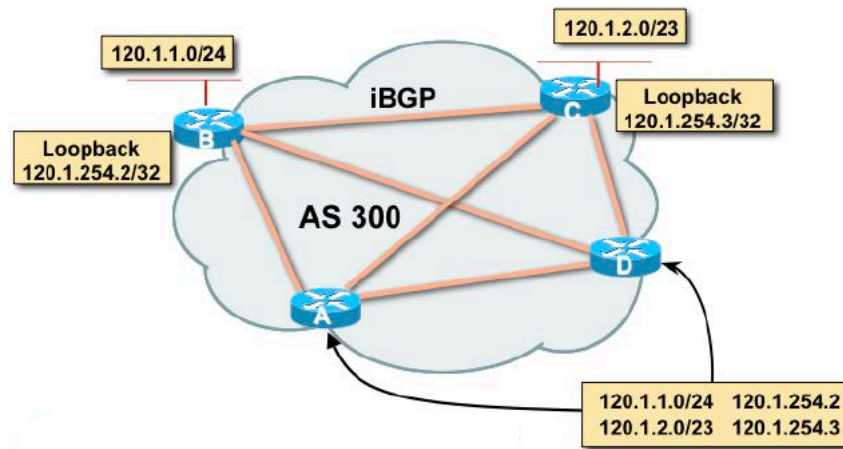
Network	Next Hop	Metric	LocPrf	Weight	Path
*> 2001::/32	2406:6400:F:41::1				
			0	23456	38610 6939 I
* i	2406:6400:D::5	0	100	0	45192 4608 4826 6939 i
*> 2001:200::/32	2406:6400:F:41::1				
			0	23456	38610 6939 2500 i
* i	2406:6400:D::5	0	100	0	45192 4608 4826 6939 2500 i

eBGP Next Hop



- The IP address to reach the next AS
 - Router A advertise 150.10.0.0/16 and 160.10.0.0/16 to router B in eBGP with next hop 150.10.1.1 (Change it to own IP)
 - Router B will update Router C in iBGP keeping the next hop unchanged
- Well known mandatory attribute

iBGP Next Hop



- Next hop is iBGP router loopback address
- Recursive route look-up
- Loopback address need to announce through IGP (OSPF)
- iBGP send update next-hop unchanged

Next Hop Best Practice

- IOS default is for external next-hop to be propagated unchanged to iBGP peers
 - This means that IGP has to carry external next-hops
 - Forgetting means external network is invisible
 - With many eBGP peers, it is unnecessary extra load on IGP
- ISP Best Practice is to change external next-hop to be that of the local router
 - neighbor x.x.x.x next-hop-self

Next Hop Self Configuration

- Next hop default behavior can be changed by using next-hop-self command
- Forces all updates for this neighbor to be advertised with this router as the next hop
- The IP address used for next-hop-self will be the same as the source IP address of the BGP packet

BGP Origin Attribute

- The origin attribute informs all autonomous systems how the prefix introduced into BGP
- Well known mandatory attribute
- Three values: IGP, EGP, incomplete
 - IGP generated by BGP network statement
 - EGP generated by EGP
 - Incomplete redistributed from another routing protocol

BGP Origin Attribute Example

Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,

r RIB-failure, S Stale

Origin codes: i - IGP, e - EGP, ? – incomplete

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 2001::/32	2406:6400:F:41::1	0	23456	38610	6939 i
* i	2406:6400:D::5	0	100	0	45192 4608 4826 6939 i

Aggregator

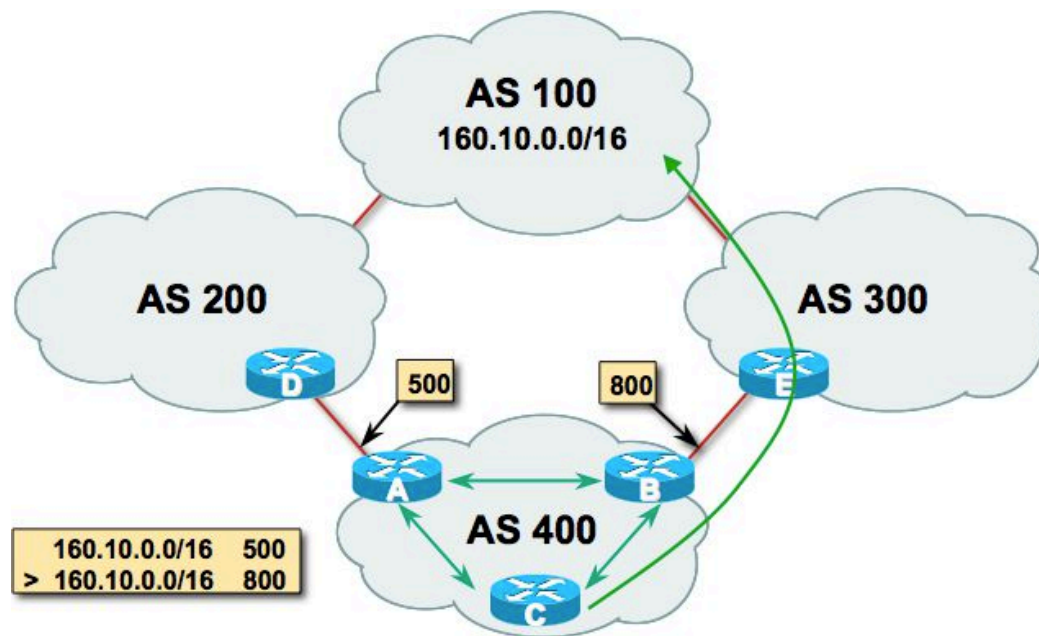
- Conveys the IP address of the router or BGP speaker generating the aggregate route
- Optional & transitive attribute
- Useful for debugging purposes
- Does not influence best path selection
- Creating aggregate using “aggregate-address” sets the aggregator attribute:

```
router bgp 100  
  aggregate-address 100.1.0.0 255.255.0.0
```

BGP Local Preference Attribute

- Local preference is used to advertise to IBGP neighbors only about how to leave their AS (Outbound Traffic).
- Paths with highest preference value are most desirable
- Local preference attribute is well-known and discretionary and is passed only within the AS
- Cisco Default Local Pref is 100

BGP Local Preference Attribute



- For destination 160.10.0.0/16 Router A advertise local pref 500 and Router B advertise local pref 800 in iBGP
- 800 will win best path (Router B)

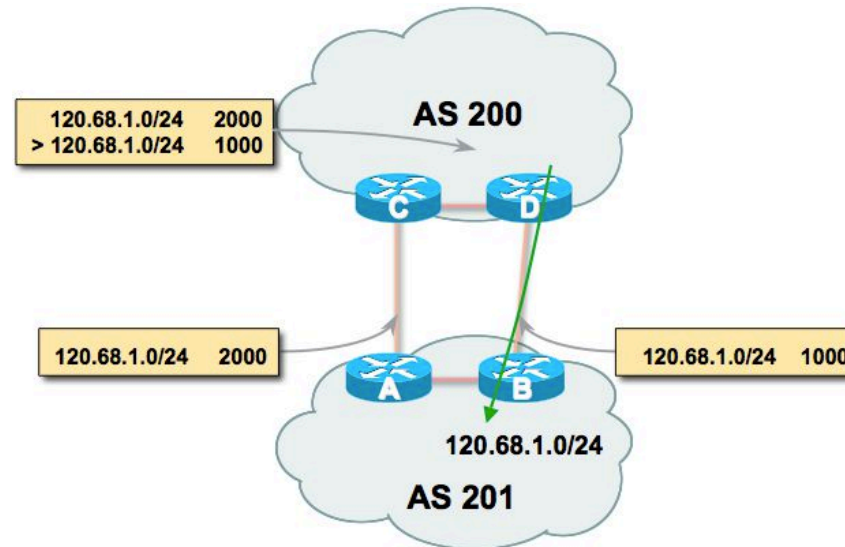
BGP Local Pref Attribute Example

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 2001::/32	2406:6400:F:41::1				
		0	23456	38610	6939 i
* i	2406:6400:D::5	0	100	0	45192 4608 4826
		6939	i		
*> 2001:200::/32	2406:6400:F:41::1				
		0	23456	38610	6939 2500 i
* i	2406:6400:D::5	0	100	0	45192 4608 4826
		6939	2500	i	

BGP MED Attribute

- MED is used to advertise to EBGP neighbors about how to exit their AS to reach networks owned by this AS (Incoming traffic).
- MED is sent to EBGP neighbors only.
- The paths with the lowest MED value are the most desirable
- The MED attribute is optional and non transitive

BGP MED Attribute



- For prefix 120.68.1.0/24 Router B send MED 1000 and router A send MED 2000 to eBGP neighbor
- Incoming traffic from AS200 will choose Router B since lowest MED will win

BGP MED Example

Network Next Hop Metric LocPrf Weight Path

*> 2001::/32 2406:6400:F:41::1

0 23456 38610 6939 i

* i 2406:6400:D::5 0 100 0 45192 4608 4826 6939 i

*> 2001:200::/32 2406:6400:F:41::1

0 23456 38610 6939 2500 i

* i 2406:6400:D::5 0 100 0 45192 4608 4826 6939 2500 i

BGP Community Attribute

- Community is a tagging technique to mark a set of routes
- Upstream service provider routers can then use these flags to apply specific routing policies (i.e local preference etc) within their network
- Represented as two 16 bit integers (RFC1998)
- Common format is <local-ASN>:xx
- I.e 0:0 to 0:65535 and 65535:0 to 65535:65535 are reserved
- Very useful in applying policies within and between ASes
- Optional & transitive attribute

BGP Route Selection Process

- Step 1: Prefer highest weight (local to router)
- Step 2: Prefer highest local preference (global within AS)
- Step 3: Prefer route originated by the local router
- Step 4: Prefer shortest AS path
- Step 5: Prefer lowest origin code (IGP < EGP < incomplete)
- Step 6: Prefer lowest MED (from other AS)
- Step 7: Prefer EBGP path over IBGP path
- Step 8: Prefer the path through the closest IGP neighbor
- Step 9: Prefer oldest route for EBGP paths
- Step 10: Prefer the path with the lowest neighbor BGP router ID

Questions



Overview

Routing Workshop (3 Days)

- Introduction to IP Routing
- Routing Protocol Basic
- IPv6 Address Structure
- Routing Lab Topology Overview
- Operation of OSPF Routing Protocol
- Lab Exercise on Basic Router and OSPF Dynamic Routing Configuration
- Basic BGP Operation
- BGP Attributes and Path Selection Process
- **BGP Scaling Techniques**
- Lab Exercise on iBGP, eBGP, RR, Peer group etc
- Internet Exchange [IX] Policy Overview and Configuration requirement

BGP Scaling Techniques

- Original BGP specification and implementation was fine for the Internet of the early 1990s
 - But didn't scale
- Issues as the Internet grew included:
 - Scaling the iBGP mesh beyond a few peers?
 - Implement new policy without causing flaps and route churning?
 - Keep the network stable, scalable, as well as simple?

BGP Scaling Techniques

- Current Best Practice Scaling Techniques
 - Route Refresh
 - Peer-groups
 - Route Reflectors (and Confederations)
- Deprecated Scaling Techniques
 - Soft Reconfiguration
 - Route Flap Damping

Route Refresh

- Policy Changes:
 - Hard BGP peer reset required after every policy change because the router does not store prefixes that are rejected by policy
- Hard BGP peer reset:
 - Tears down BGP peering
 - Consumes CPU
 - Severely disrupts connectivity for all networks
- Solution:
 - Route Refresh

Route Refresh Capability

- Facilitates non-disruptive policy changes
- No configuration is needed
 - Automatically negotiated at peer establishment
- No additional memory is used
- Requires peering routers to support “route refresh capability” – RFC2918
- Tell peer to resend full BGP announcement
 - `clear ip bgp x.x.x.x [soft] in`
- Resend full BGP announcement to peer
 - `clear ip bgp x.x.x.x [soft] out`

BGP Peer Group

- Problem – how to scale iBGP
 - Large iBGP mesh slow to build
 - iBGP neighbors receive the same update
 - Router CPU wasted on repeat calculations
- Solution – peer-groups
 - Group peers with the same outbound policy
 - Updates are generated once per group

BGP Peer Group -Advantages

- Makes configuration easier
- Makes configuration less prone to error
- Makes configuration more readable
- Lower router CPU load
- iBGP mesh builds more quickly
- Members can have different inbound policy
- Can be used for eBGP neighbors too!

BGP Peer Group -BCP

- Always configure peer-groups for iBGP
 - Even if there are only a few iBGP peers
 - Easier to scale network in the future
- Consider using peer-groups for eBGP
 - Especially useful for multiple BGP customers using same AS (RFC2270)
 - Also useful at Exchange Points where ISP policy is generally the same to each peer

Configuring a Peer Group

```
router bgp 100

  neighbor ibgp-peer peer-group
  neighbor ibgp-peer remote-as 100
  neighbor ibgp-peer update-source loopback 0
  neighbor ibgp-peer send-community
  neighbor ibgp-peer route-map outfilter out
  neighbor 1.1.1.1 peer-group ibgp-peer
  neighbor 2.2.2.2 peer-group ibgp-peer
  neighbor 2.2.2.2 route-map infilter in
  neighbor 3.3.3.3 peer-group ibgp-peer
```

! note how 2.2.2.2 has different inbound filter from peer-group !

Peer Groups

- Always configure peer-groups for iBGP
 - Even if there are only a few iBGP peers
 - Easier to scale network in the future
- Consider using peer-groups for eBGP
 - Especially useful for multiple BGP customers using same AS (RFC2270)
 - Also useful at Exchange Points where ISP policy is generally the same to each peer
- Peer-groups are essentially obsoleted
 - But are still widely considered best practice
 - Replaced by update-groups (internal coding – not configurable)
 - Enhanced by peer-templates (allowing more complex constructs)

BGP Route Reflector

- In a transit AS all router in the core need to know the complete routing table coming from Internet
- Global routing table size is above 300k prefix
- Practically impossible to redistribute these route in IGP i.e OSPF
- Solution is to forward these large routing information by iBGP

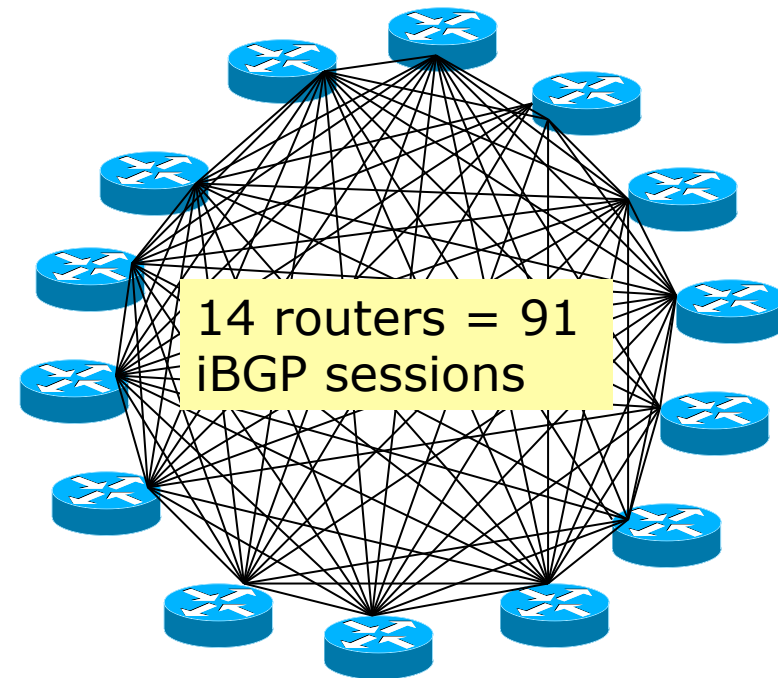
BGP Route Reflector

- iBGP use TCP to make reliable delivery of these large routing information across all core router in a transit ISP
- TCP can not broadcast so need to make individual delivery to all iBGP speaker
- To protect routing loop iBGP use split horizon rule so can not send routing information to neighbor learn via iBGP
- So iBGP need full mesh peering with other core router in a transit ISP

Scaling iBGP mesh

- Avoid $\frac{1}{2}n(n-1)$ iBGP mesh

$n=1000 \Rightarrow$ nearly
half a million
ibgp sessions!

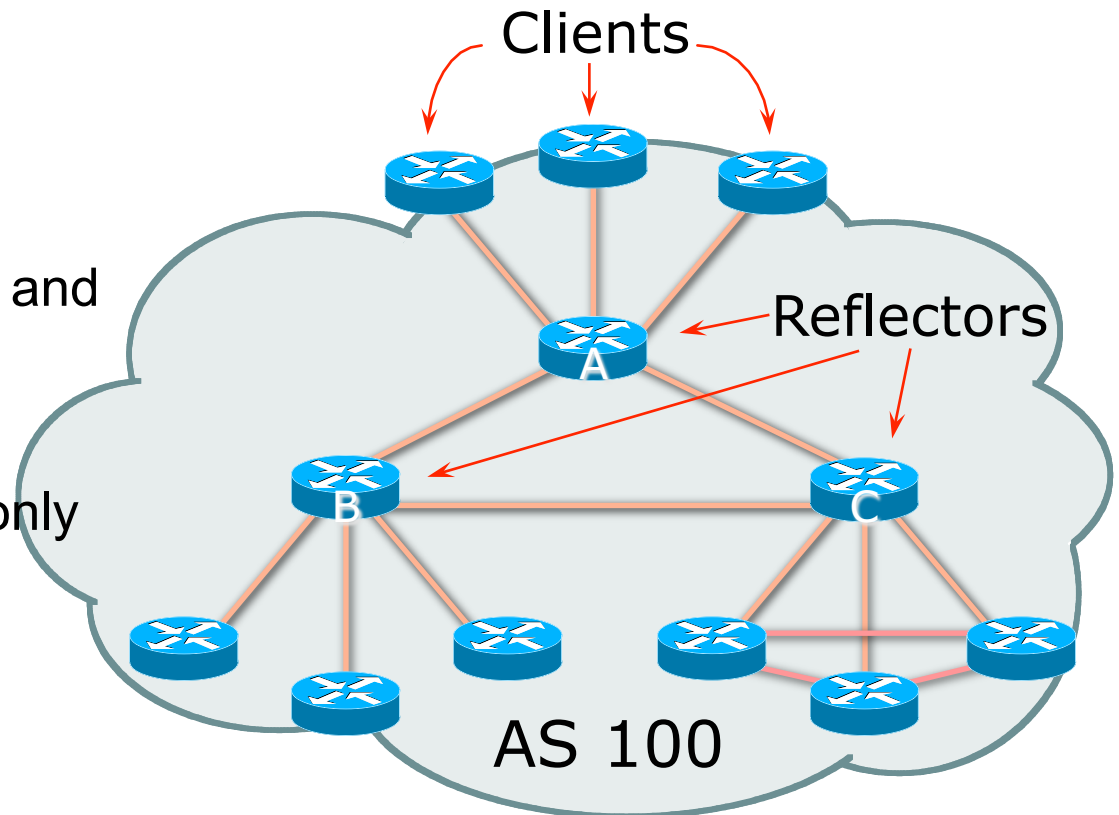


Two solutions

- Route reflector – simpler to deploy and run
- Confederation – more complex, has corner case advantages

Route Reflector

- Reflector receives path from clients and non-clients
- Selects best path
- If best path is from client, reflect to other clients and non-clients
- If best path is from non-client, reflect to clients only
- Non-meshed clients
- Described in RFC4456



BGP Scaling Techniques

- These 3 techniques should be core requirements on all ISP networks
 - Route Refresh (or Soft Reconfiguration)
 - Peer groups
 - Route Reflectors

Questions



Overview

Routing Workshop (3 Days)

- Introduction to IP Routing
- Routing Protocol Basic
- IPv6 Address Structure
- Routing Lab Topology Overview
- Operation of OSPF Routing Protocol
- Lab Exercise on Basic Router and OSPF Dynamic Routing Configuration
- Basic BGP Operation
- BGP Attributes and Path Selection Process
- BGP Scaling Techniques
- **Lab Exercise on iBGP, eBGP, RR, Peer group etc**
- Internet Exchange [IX] Policy Overview and Configuration requirement

Case study- Deployment IPv6 in EGP

- Scenario:
 - BGP4 is used in Training ISP network
 - iBGP is used between internal routers in Training ISP to carry external prefixes (i.e Customer & Global Internet Prefixes)
 - Route Reflector is used to resolve iBGP full mesh scalability issue.

Case study- Deployment IPv6 in EGP

- Scenario:
 - Transit service with upstream ASes is configured with eBGP
 - Customer network from downstream can also be configured with eBGP or static
 - Training ISP is having one native IPv6 transit and one tunnel IPv6 transit with AS45192 & AS131107 (2.35 as dot)

Case study- Deployment IPv6 in EGP

- Basic BGP Configuration:

```
router bgp 17821  
address-family ipv6  
no synchronization
```


Case study- Deployment IPv6 in EGP

Adding iBGP Neighbor:

```
router bgp 17821
```

```
address-family ipv6
```

```
!
```

```
neighbor 2406:6400:0000:0000::2 remote-as 17821
```

```
neighbor 2406:6400:0000:0000::2 update-source loopback 0
```

```
neighbor 2406:6400:0000:0000::2 activate
```

iBGP neighbor is always recommended with loopback interface

Case study- Deployment IPv6 in EGP

Announcing IPv6 Prefix:

```
router bgp 17821
address-family ipv6
!
neighbor 2406:6400:0000:0000::2 remote-as 17821
neighbor 2406:6400:0000:0000::2 update-source loopback 0
neighbor 2406:6400:0000:0000::2 activate
!
network 2406:6400:0100:0000::/48
```

Case study- Deployment IPv6 in EGP

Add Pull-up route if needed:

```
router bgp 17821
```

```
address-family ipv6
```

```
!
```

```
neighbor 2406:6400:0000:0000::2 remote-as 17821
```

```
neighbor 2406:6400:0000:0000::2 update-source loopback 0
```

```
neighbor 2406:6400:0000:0000::2 activate
```

```
!
```

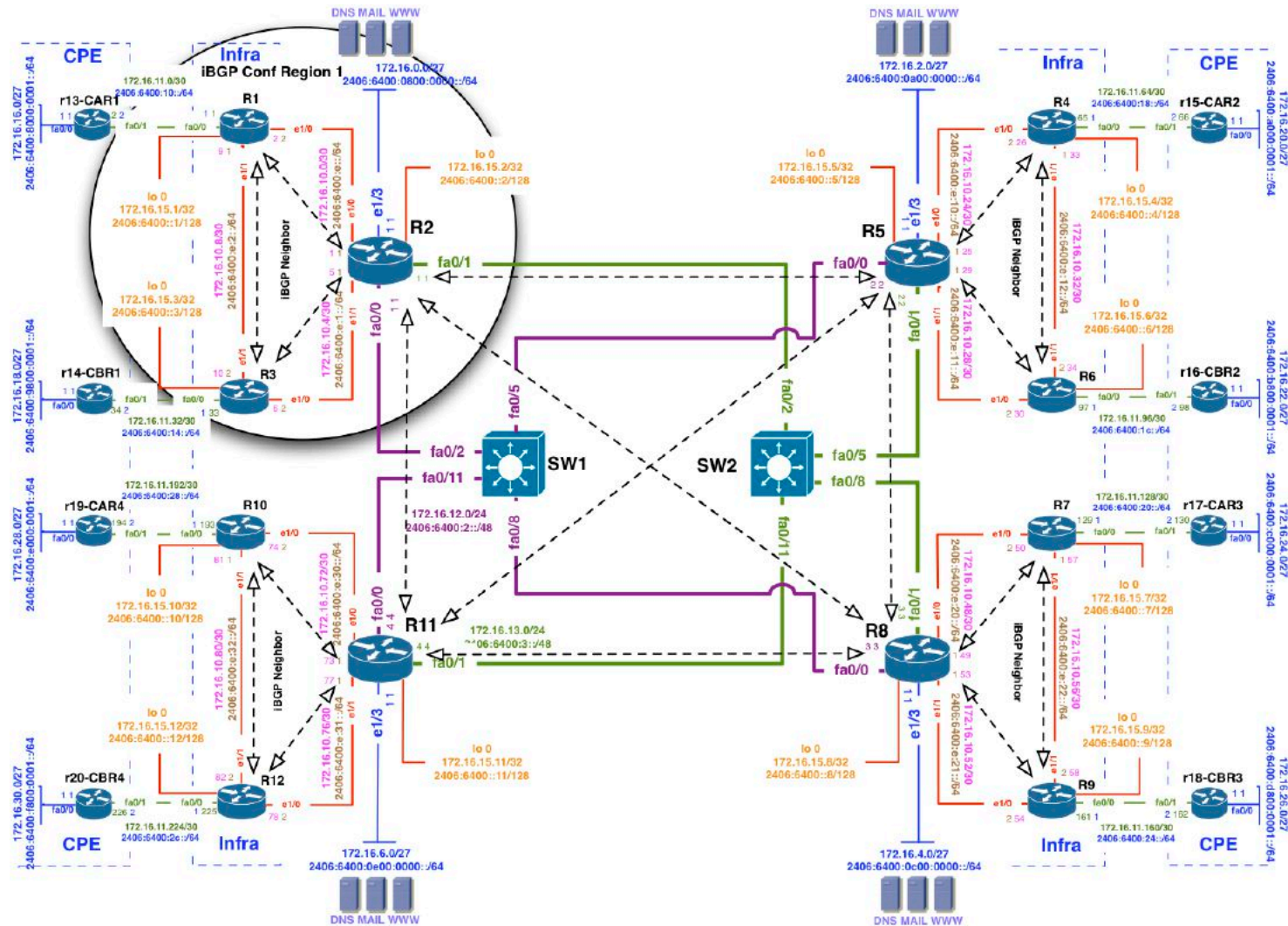
```
network 2406:6400:0100:0000::/48
```

```
exit
```

```
exit
```

```
ipv6 route 2406:6400:0100:0000::/48 null 0
```

iBGP Peering For Region 1



IPv4 iBGP Conf POP Router

Router1

```
config t
router bgp 17821
address-family ipv4
no auto-summary
no synchronization
neighbor 172.16.15.2 remote-as 17821
neighbor 172.16.15.2 update-source loopback 0
neighbor 172.16.15.2 activate
neighbor 172.16.15.3 remote-as 17821
neighbor 172.16.15.3 update-source loopback 0
neighbor 172.16.15.3 activate
network 172.16.16.0 mask 255.255.254.0
exit
exit
ip route 172.16.16.0 255.255.254.0 null 0 permanent
exit
wr
```

IPv4 iBGP Configuration Verification

POP Router

```
sh bgp ipv4 unicast summary
```

```
sh bgp ipv4 unicast
```

```
sh ip route bgp
```

```
sh bgp ipv4 unicast neighbors [router 1.....router12  
loopback] advertised-routes
```

```
sh bgp ipv4 unicast neighbors [router 1.....router12  
loopback] received-routes
```

```
sh ip route [R2, R5, R8, R11 datacenter prefix]
```

IPv6 iBGP Conf POP Router

Router1

```
config t
router bgp 17821
address-family ipv6
no synchronization
neighbor 2406:6400:0000:0000::2 remote-as 17821
neighbor 2406:6400:0000:0000::2 update-source loopback 0
neighbor 2406:6400:0000:0000::2 activate
neighbor 2406:6400:0000:0000::3 remote-as 17821
neighbor 2406:6400:0000:0000::3 update-source loopback 0
neighbor 2406:6400:0000:0000::3 activate
network 2406:6400:0100:0000::/45
exit
exit
ipv6 route 2406:6400:0100:0000::/45 null 0
exit
wr
```

IPv6 iBGP Configuration Verification

POP Router

```
sh bgp ipv6 unicast summary
```

```
sh bgp ipv6 unicast
```

```
sh ipv6 route bgp
```

```
sh bgp ipv6 unicast neighbors [router 1.....router12  
loopback] advertised-routes
```

```
sh bgp ipv6 unicast neighbors [router 1.....router12  
loopback] received-routes
```

```
sh ipv6 route [R2, R5, R8, R11 datacenter prefix]
```


IPv4 iBGP Conf Core Router

Router2 Configuration

```
config t
router bgp 17821
address-family ipv4
no auto-summary
no synchronization
neighbor 172.16.15.1 remote-as 17821
neighbor 172.16.15.1 update-source loopback 0
neighbor 172.16.15.1 activate
neighbor 172.16.15.3 remote-as 17821
neighbor 172.16.15.3 update-source loopback 0
neighbor 172.16.15.3 activate
neighbor 172.16.15.5 remote-as 17821
neighbor 172.16.15.5 update-source loopback 0
neighbor 172.16.15.5 activate
neighbor 172.16.15.8 remote-as 17821
neighbor 172.16.15.8 update-source loopback 0
neighbor 172.16.15.8 activate
neighbor 172.16.15.11 remote-as 17821
neighbor 172.16.15.11 update-source loopback 0
neighbor 172.16.15.11 activate
network 172.16.0.0 mask 255.255.254.0
exit
exit
ip route 172.16.0.0 255.255.254.0 null 0 permanent
exit
Wr
```

IPv4 iBGP Conf Core Router

Router2 Configuration

```
config t
router bgp 17821
address-family ipv4
no auto-summary
no synchronization
neighbor 172.16.15.1 remote-as 17821
neighbor 172.16.15.1 update-source loopback 0
neighbor 172.16.15.1 activate
neighbor 172.16.15.3 remote-as 17821
neighbor 172.16.15.3 update-source loopback 0
neighbor 172.16.15.3 activate
neighbor 172.16.15.5 remote-as 17821
neighbor 172.16.15.5 update-source loopback 0
neighbor 172.16.15.5 activate
neighbor 172.16.15.8 remote-as 17821
neighbor 172.16.15.8 update-source loopback 0
neighbor 172.16.15.8 activate
neighbor 172.16.15.11 remote-as 17821
neighbor 172.16.15.11 update-source loopback 0
neighbor 172.16.15.11 activate
network 172.16.0.0 mask 255.255.254.0
exit
exit
ip route 172.16.0.0 255.255.254.0 null 0 permanent
exit
Wr
```

IPv4 iBGP Configuration Verification

Core Router

```
sh bgp ipv4 unicast summary
```

```
sh bgp ipv4 unicast
```

```
sh ip route bgp
```

```
sh bgp ipv4 unicast neighbors [router 1.....router12  
loopback] advertised-routes
```

```
sh bgp ipv4 unicast neighbors [router 1.....router12  
loopback] received-routes
```

```
sh ip route [R2, R5, R8, R11 datacenter prefix]
```

IPv6 iBGP Conf Core Router

Router2 Configuration

```
config t
router bgp 17821
address-family ipv6
no synchronization
neighbor 2406:6400:0000:0000::1 remote-as 17821
neighbor 2406:6400:0000:0000::1 update-source loopback 0
neighbor 2406:6400:0000:0000::1 activate
neighbor 2406:6400:0000:0000::3 remote-as 17821
neighbor 2406:6400:0000:0000::3 update-source loopback 0
neighbor 2406:6400:0000:0000::3 activate
neighbor 2406:6400:0000:0000::5 remote-as 17821
neighbor 2406:6400:0000:0000::5 update-source loopback 0
neighbor 2406:6400:0000:0000::5 activate
neighbor 2406:6400:0000:0000::8 remote-as 17821
neighbor 2406:6400:0000:0000::8 update-source loopback 0
neighbor 2406:6400:0000:0000::8 activate
neighbor 2406:6400:0000:0000::11 remote-as 17821
neighbor 2406:6400:0000:0000::11 update-source loopback 0
neighbor 2406:6400:0000:0000::11 activate
network 2406:6400:0001:0000::/48
exit
exit
ipv6 route 2406:6400:0001:0000::/48 null 0
exit
wr
```

IPv6 iBGP Configuration Verification

- Core Router

```
sh bgp ipv6 unicast summary
```

```
sh bgp ipv6 unicast
```

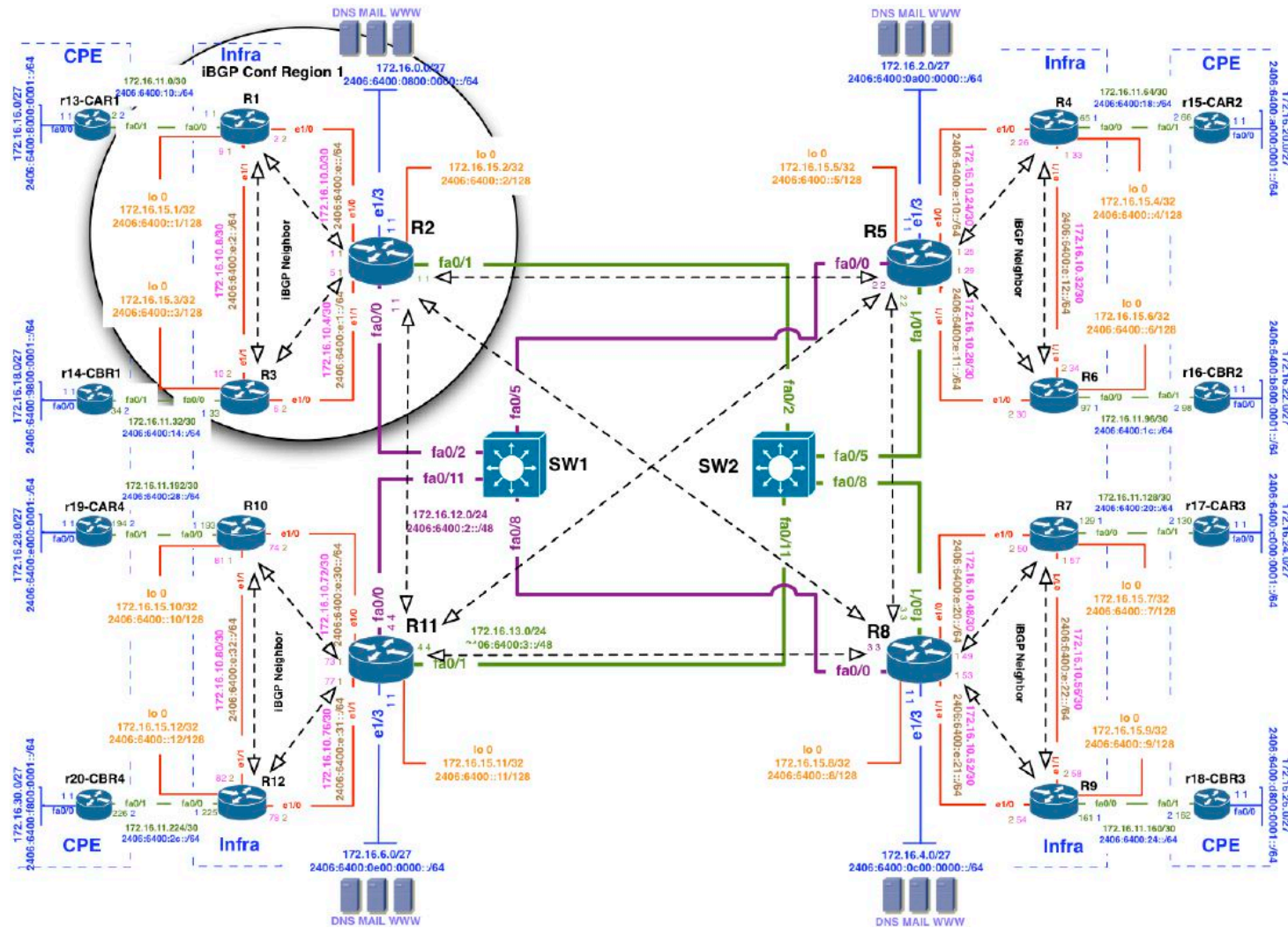
```
sh ipv6 route bgp
```

```
sh bgp ipv6 unicast neighbors [router 1.....router12  
loopback] advertised-routes
```

```
sh bgp ipv6 unicast neighbors [router 1.....router12  
loopback] received-routes
```

```
sh ipv6 route [R2, R5, R8, R11 datacenter prefix]
```

iBGP Full Mesh Issue



iBGP Full Mesh Issue

Route reflector configuration:

```
router bgp 17821
```

```
address-family ipv6
```

```
!
```

```
neighbor 2406:6400:0000:0000::1 remote-as 17821
```

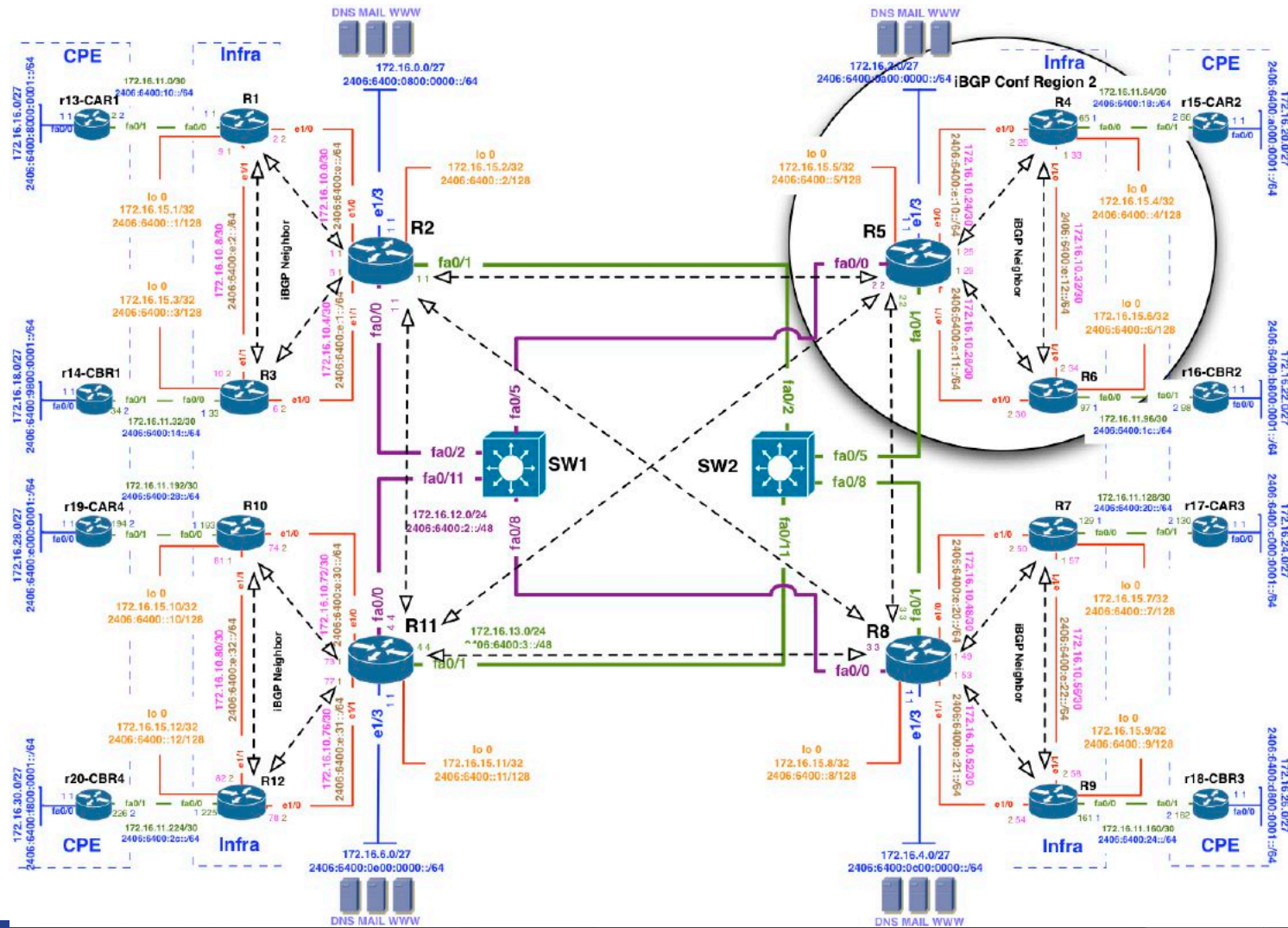
```
neighbor 2406:6400:0000:0000::1 update-source loopback 0
```

```
neighbor 2406:6400:0000:0000::1 activate
```

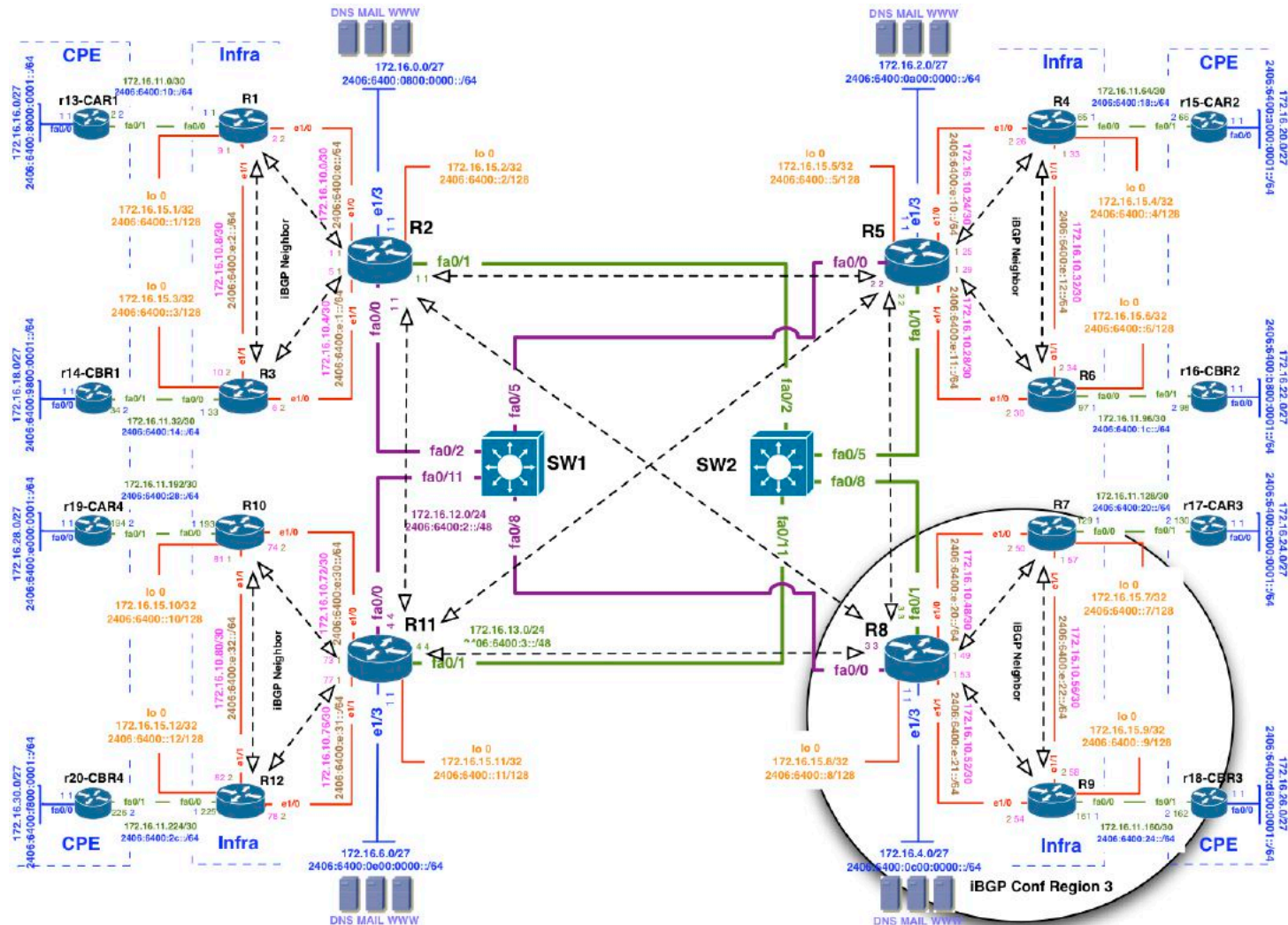
```
!
```

```
neighbor 2406:6400:0000:0000::1 route-reflector-client
```

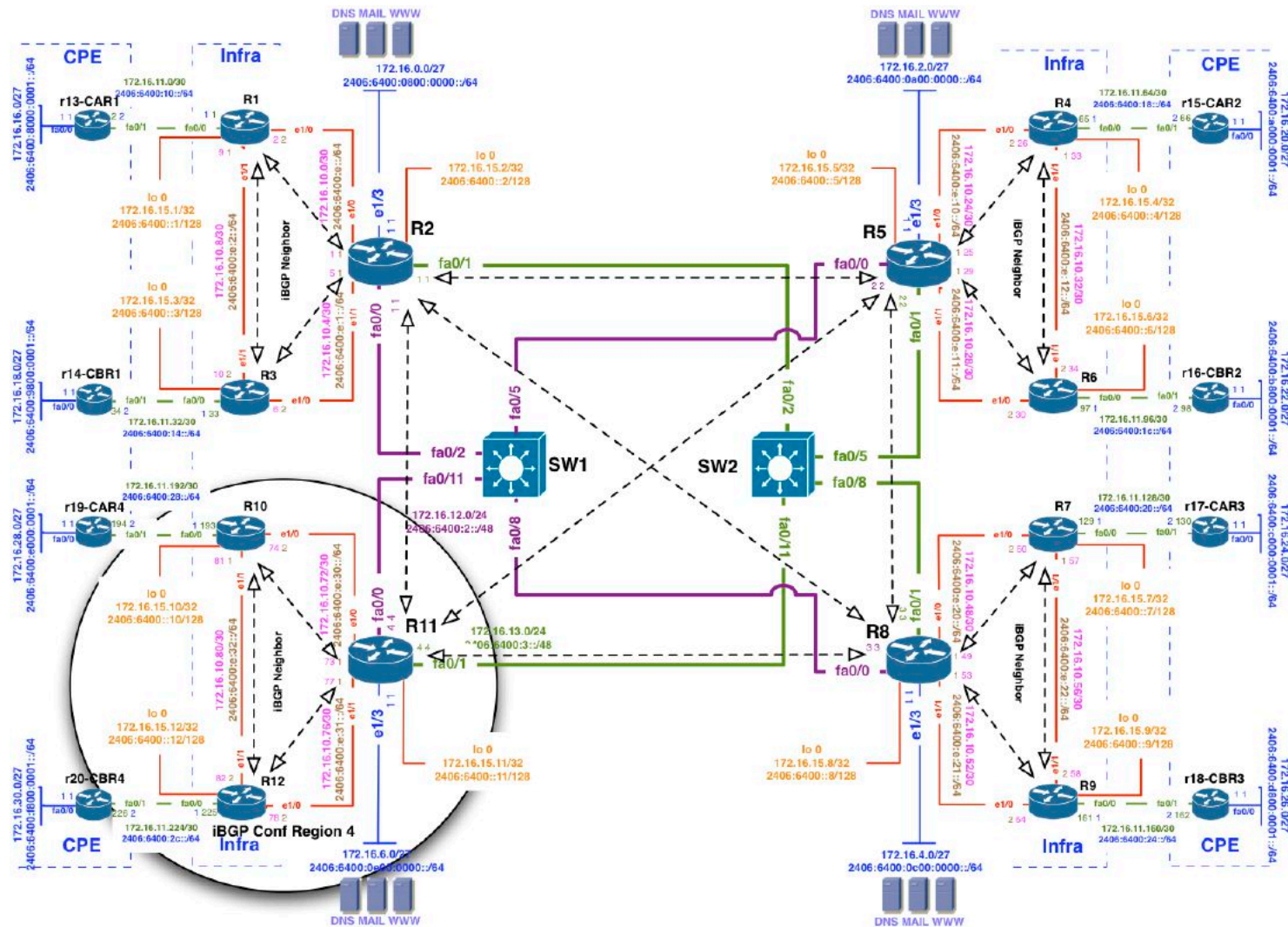

iBGP Peering For Region 2



iBGP Peering For Region 3



iBGP Peering For Region 4



Questions



Overview

Routing Workshop (3 Days)

- Introduction to IP Routing
- Routing Protocol Basic
- IPv6 Address Structure
- Routing Lab Topology Overview
- Operation of OSPF Routing Protocol
- Lab Exercise on Basic Router and OSPF Dynamic Routing Configuration
- Basic BGP Operation
- BGP Attributes and Path Selection Process
- BGP Scaling Techniques
- Lab Exercise on iBGP, eBGP, RR, Peer group etc
- **Internet Exchange [IX] Policy Overview and Configuration requirement**

Case Study- IXP Configuration

- Two type of traffic exchange between ISPs
- Transit
 - Where ISP will pay to send/receive traffic
 - Downstream ISP will pay upstream ISP for transit service
- Peering
 - ISPs will not pay each other to interchange traffic
 - Works well if win win for both
 - Reduce cost on expensive transit link

IX Peering Model

- BLPA (Bi-Lateral Peering Agreement)
 - IX will only provide layer two connection/switch port to ISPs
 - Every ISPs will arrange necessary peering arrangement with others by their mutual business understanding.
- MLPA (Multi-Lateral Peering Agreement)
 - IX will provide layer two connection/switch port to ISPs
 - Each ISP will peer with a route server on the IX.
 - Route server will collect and distribute directly connected routes to every peers.

IXP Peering Policy

- BLPA is applicable where different categories of ISPs are connected in an IX
 - Large ISPs can choose to peer with large ISPs (base on their traffic volume)
 - Small ISPs will arrange peering with small ISPs
- Would be preferable for large ISPs
 - They will peer with selected large ISPs (Equal traffic interchange)
 - Will not loose business by peering with small ISP

IXP Peering Policy

- MLPA model works well to widen the IX scope of operation (i.e national IX).
- Easy to manage peering
 - Peer with the route server and get all available local routes.
 - Do not need to arrange peering with every ISPs connected to the IX.
- Unequal traffic condition can create not intersected situation to peer with route server

IXP Peering Policy

- Both peering model can be available in an IX.
- Member will select peering model i.e either BLPA or MLPA (Route Server Peering)
- IX will provide switch port
- Mandatory MLPA model some time not preferred by large ISP (Business Interest)
 - Can create not interested situation to connect to an IX

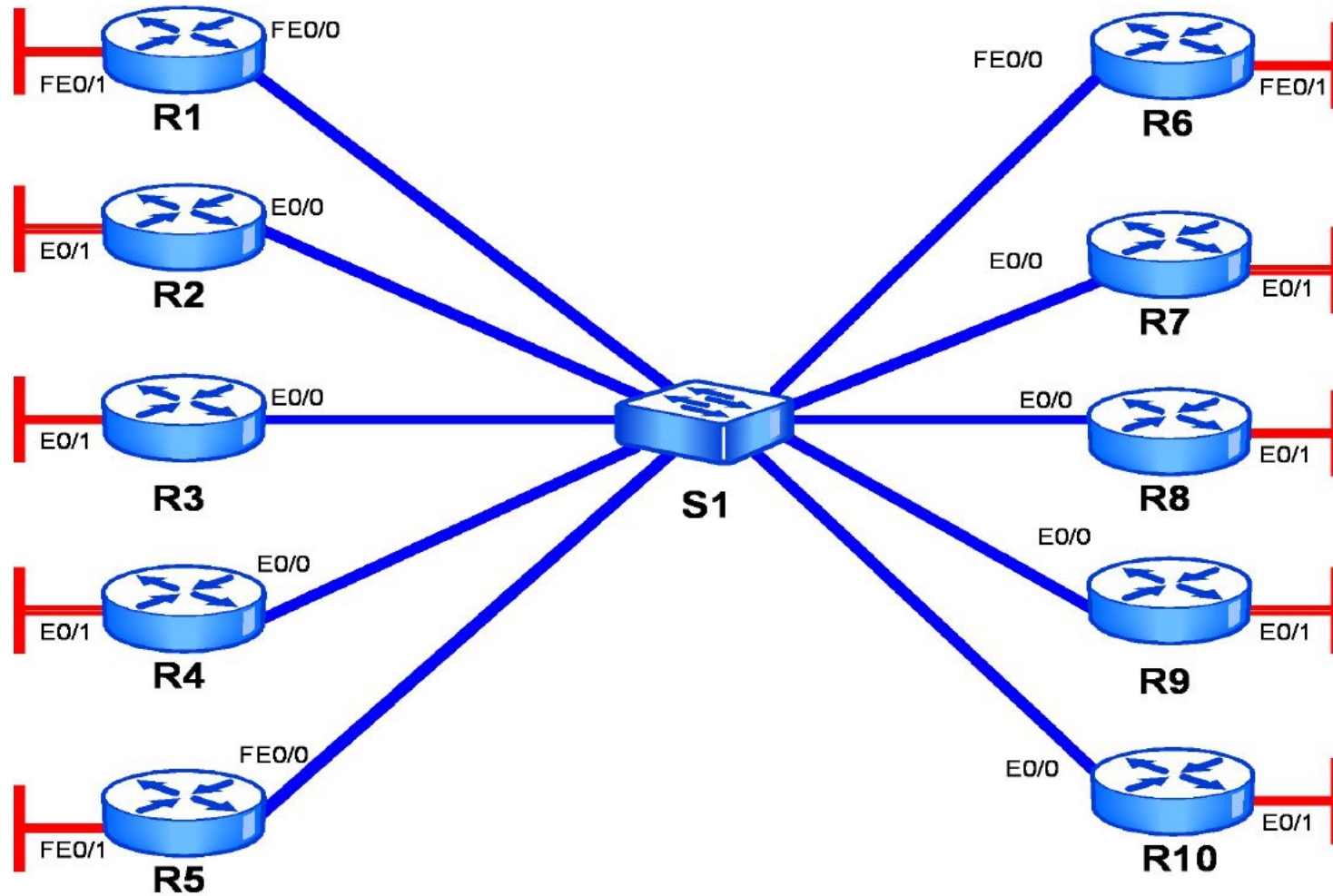
IXP Operating Cost

- Access link
- Link maintenance
- Utility
- Administration

IXP Cost Model

- Not for profit
- Cost sharing
- Membership based
- Commercial IX

IXP Network Diagram



Case Study- IXP Configuration

- Required **global & interface** commands to enable IPv6

```
Router(Config)#ipv6 unicast-routing  
Router(Config)#ipv6 cef (optional)
```

- Configure IPv6 address on interface

```
Router(Config-if)#ipv6 address 2001:0df0:00aa::1/64  
Router(Config-if)#ipv6 enable
```

- Verify IPv6 configuration

```
Router#sh ipv6 interface fa0/0
```

- Verify connectivity

```
Router#ping 2001:0df0:00aa::1
```

Case Study- IXP Configuration

- Required **BGP** commands to enable IPv6 routing

```
Router(config)# router bgp 1
```

```
Router2(config-router)#bgp router-id 10.0.0.1 (if no 32 bit  
address on any interface)
```

```
Router(config-router)# address-family ipv6
```

```
Router(config-router-af)# no synchronization
```

```
Router(config-router-af)#neighbor 2001:0df0:00aa::1  
remote-as 2 (EBGP)
```

```
Router(config-router-af)#neighbor 2001:0df0:00aa::1  
activate
```

```
Router(config-router-af)#network 2001:0df0:00aa::/48
```

- Verify BGP IPv6 configuration

```
Router#sh bgp ipv6 unicast summary (summarized  
neighbor list)
```

```
Router#sh bgp ipv6 unicast (BGP database)
```

```
Router#sh ipv6 route bgp (BGP routing table)
```

Case Study- IXP Configuration

Required command to add IX prefix filter

- Create prefix filter in global mode

```
Router(config)#ipv6 prefix-list AS1 seq 2 permit 2001:0df0:aa::/48
```

- Apply prefix filter in BGP router configuration mode

```
Router(config-router)# address-family ipv6
```

```
Router(config-router-af)#neighbor 2001:0df0:aa::1  
prefix-list AS1 in
```

```
Router(config-router-af)#neighbor 2001:0df0:aa::1  
prefix-list AS1 out
```

Case Study- IXP Configuration

- Controlling routing update traffic (Not data traffic)
- Incoming routing update (Will control outgoing data traffic)
- Outgoing routing update (Will control incoming data traffic)

Questions



Four-byte AS Numbers

AS Numbers

- Two Ranges:
 - [0 – 65535] are the original 16 bit
 - [65536 – 4294967295] are the new 32 bit
- Usages
 - 0 and 65535 Reserved
 - 1 to 64495 Public Internet
 - 64496 to 64511 Documentation –RFC5398
 - 64512 to 65534 Private use
 - 23456 represent 32 Bit range in 16 bit world
 - 65536 to 65551 Documentation – RFC 5398
 - 65552 to 4294967295 Public Internet

32 bit AS Number Representation

- AS DOT
 - Based upon 2-Byte AS representation
 - <Higher2bytes in decimal> . <Lower2bytes in decimal>
 - For example: AS 65546 is represented as 1.10
 - Easy to read, however hard for regular expressions
 - There is a meta character “.” in regular expression
 - i.e For example, a.c matches "abc", etc., but [a.c] matches only "a", ".", or "c".

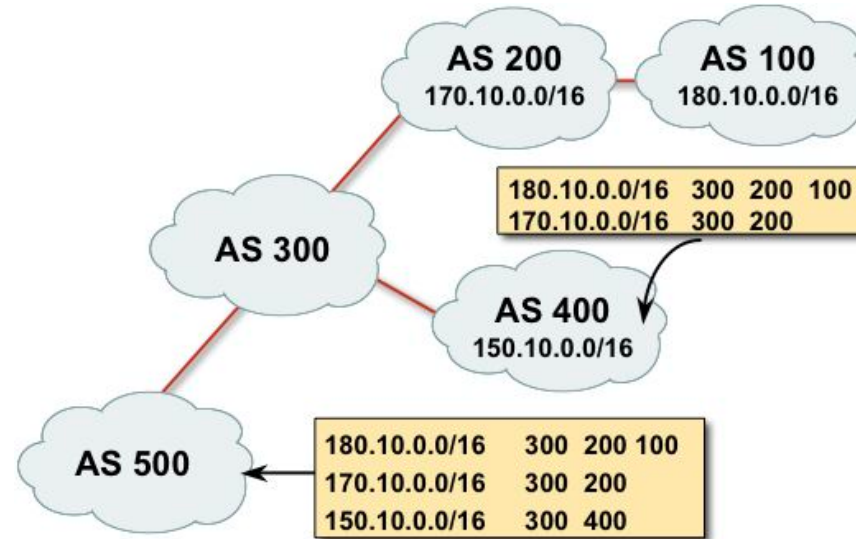
32 bit AS Number Representation

- AS PLAIN
 - ASPLAIN IETF preferred notation
 - Continuation on how a 2-Byte AS number has been represented historically
 - Notation: The 32 bit binary AS number is translated into a Single decimal value Example: AS 65546
 - Total AS Plain range (0 – 65535 - 65,536 - 4,294,967,295)

4 Byte AS Numbers

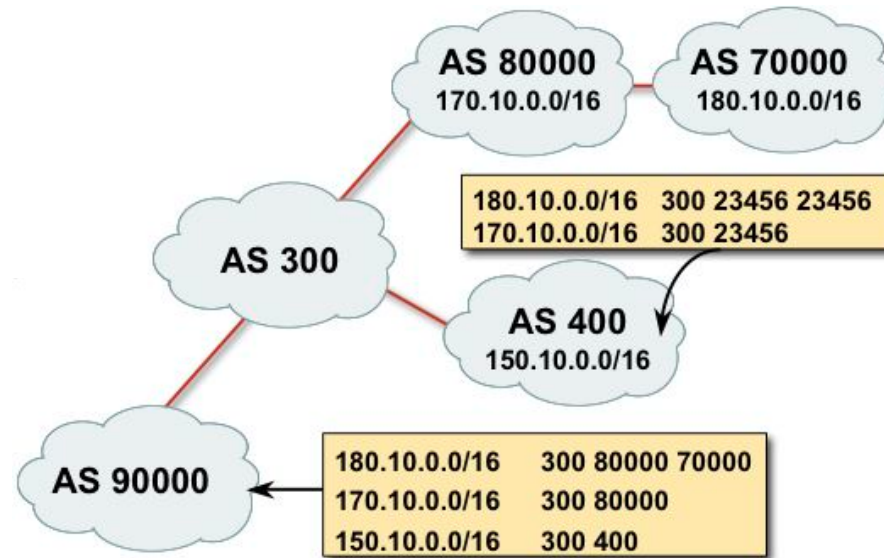
- 32 Bit range representation specified in RFC5396
- APNIC resource range:
 - In AS DOT: 2.0 ~ 2.1023
 - In AS PLAIN: 131072 ~ 132095
- AS number converter
- <http://submit.apnic.net/cgi-bin/convert-asn.pl>

AS Path Attribute



- Sequence of ASes a route has traversed
- Used for
 - Loop detection
 - Path metrics where the length of the AS Path is used as in path selection

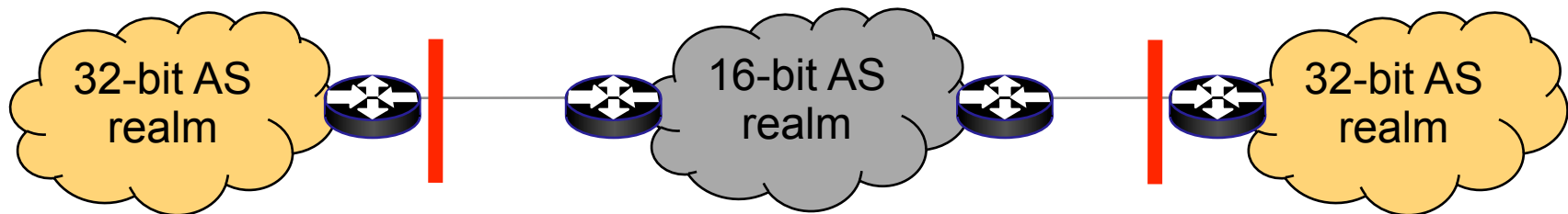
AS Path Attribute (2 byte and 4 byte)



- Internet with 16-bit and 32-bit ASNs
 - 32-bit ASNs are 65536 and above
 - AS-PATH length maintained

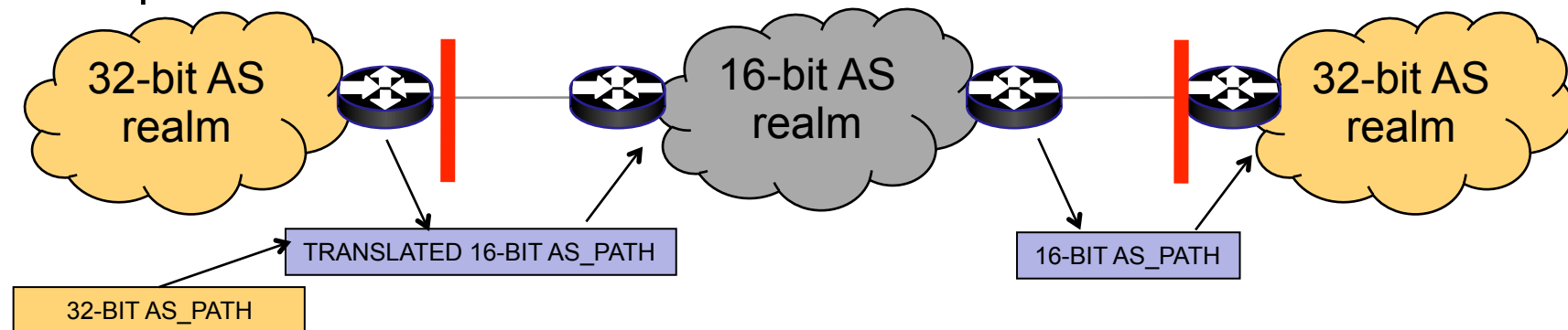
32-bit AS Transition

- Think about this space as a set of NEW / OLD boundaries
- Define the NEW / OLD and the OLD / NEW transitions
- Preserve all BGP information at the transition interfaces
 - Translate 32-bit AS Path information into a 16-bit representation
 - Tunnel 32-bit AS Path information through 16-bit AS domain as an update attribute



32-bit AS Transition

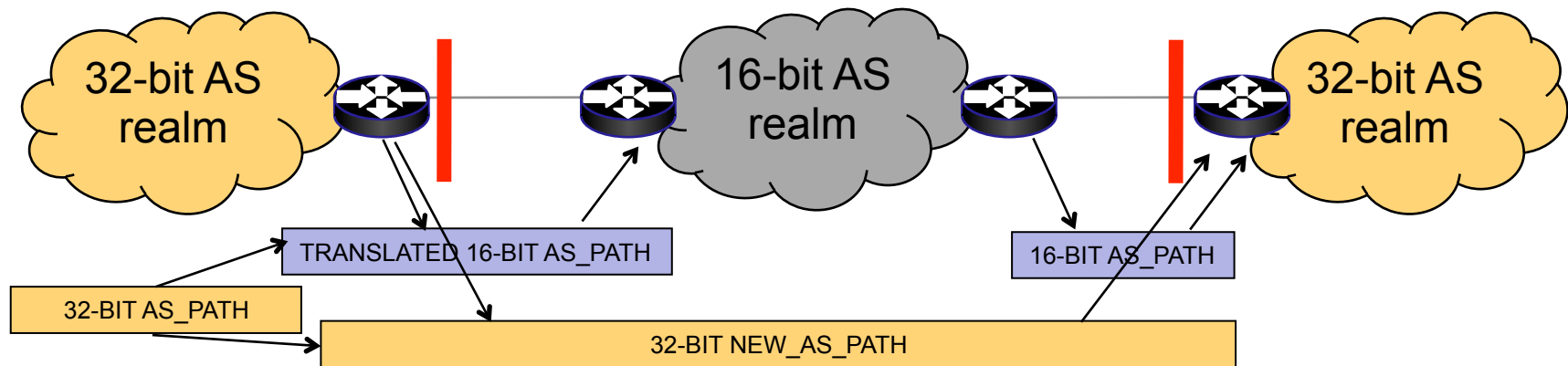
- Think about this space as a set of NEW / OLD boundaries
- Define the NEW / OLD and the OLD / NEW transitions
- Preserve all BGP information at the transition interfaces
 - Translate 32-bit AS Path information into a 16-bit representation
 - Tunnel 32-bit AS Path information through 16-bit AS domain as an update attribute



TRANSLATE all 32-bit-only AS numbers to **AS23456**

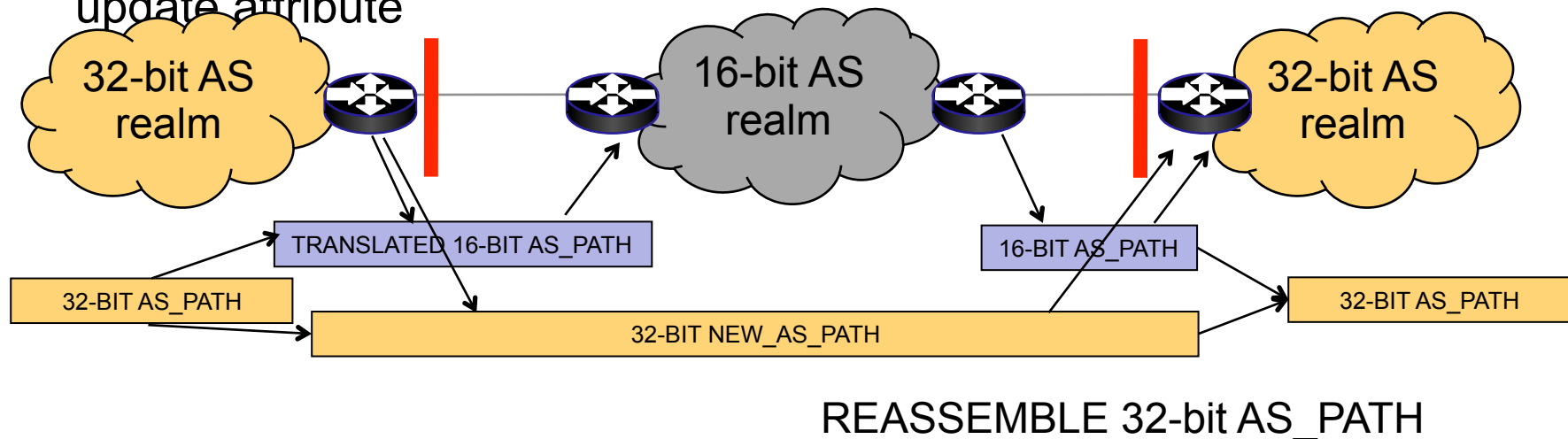
32-bit AS Transition

- Think about this space as a set of NEW / OLD boundaries
- Define the NEW / OLD and the OLD / NEW transitions
- Preserve all BGP information at the transition interfaces
 - Translate 32-bit AS Path information into a 16-bit representation
 - Tunnel 32-bit AS Path information through 16-bit AS domain as an update attribute

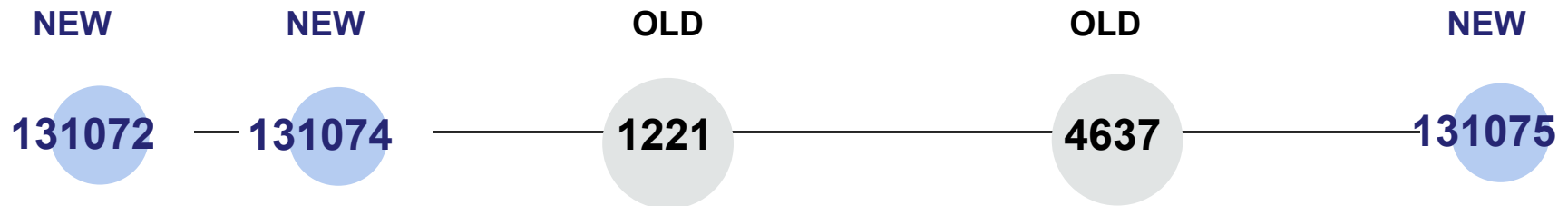


32-bit AS Transition

- Think about this space as a set of NEW / OLD boundaries
- Define the NEW / OLD and the OLD / NEW transitions
- Preserve all BGP information at the transition interfaces
 - Translate 32-bit AS Path information into a 16-bit representation
 - Tunnel 32-bit AS Path information through 16-bit AS domain as an update attribute



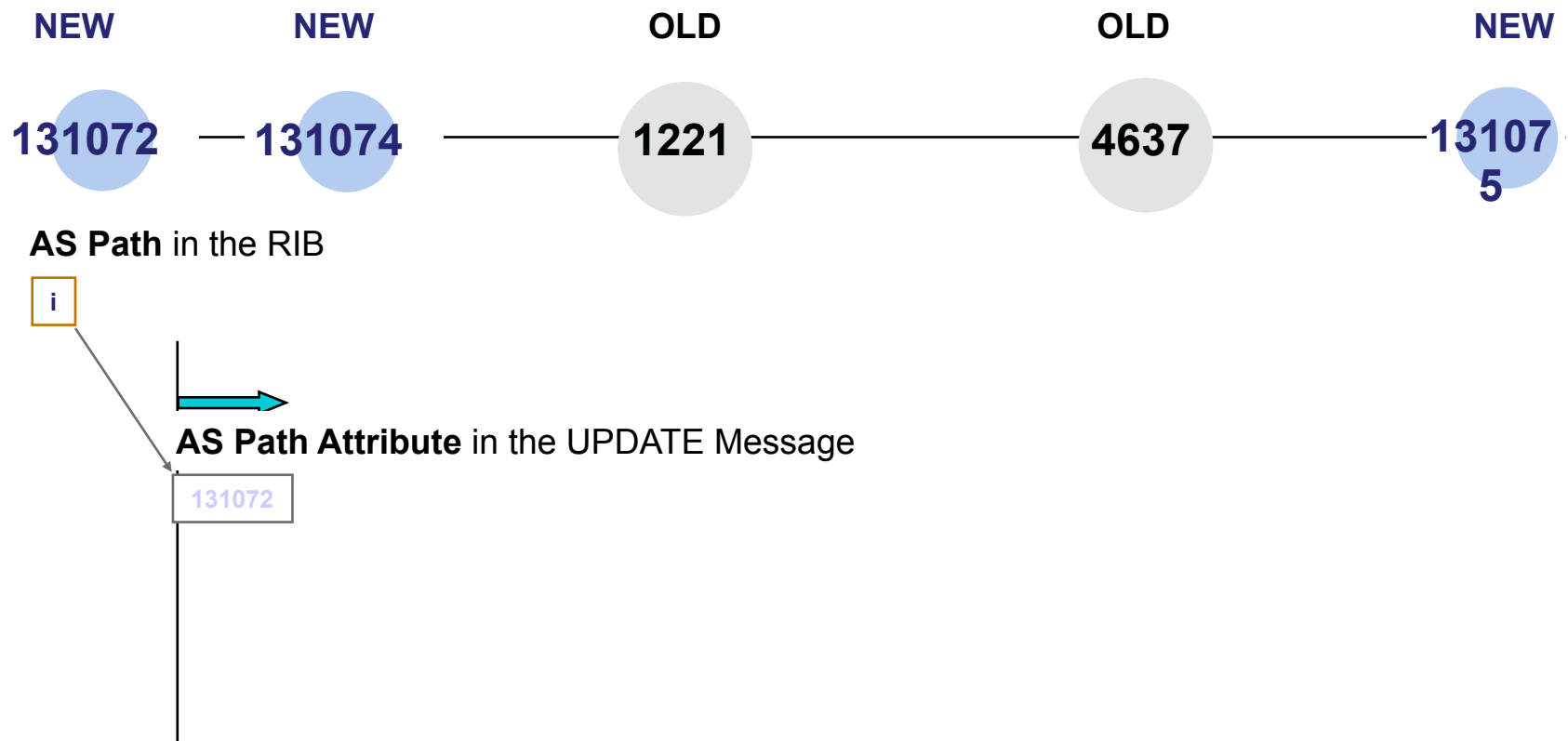
32-bit / 16-bit BGP Example



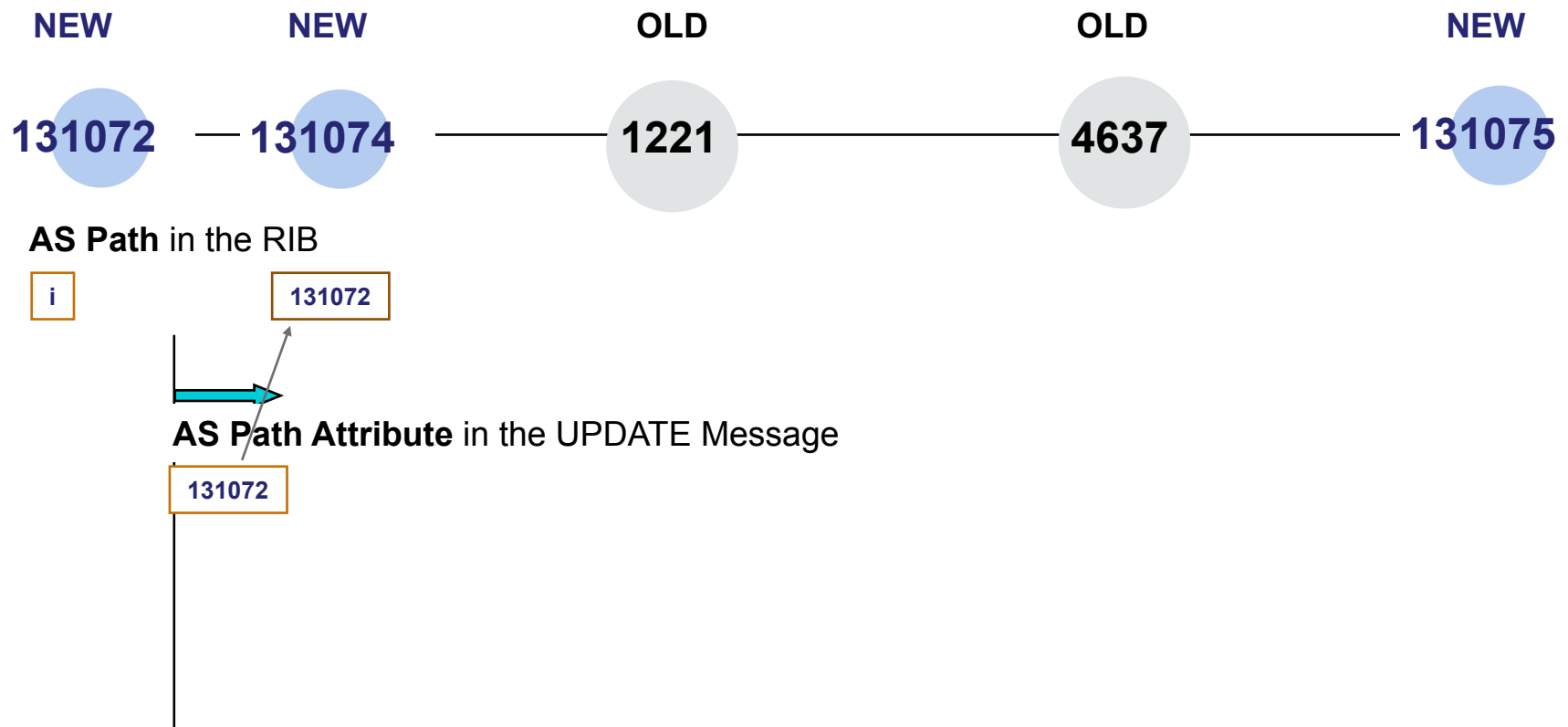
AS Path in the RIB

i

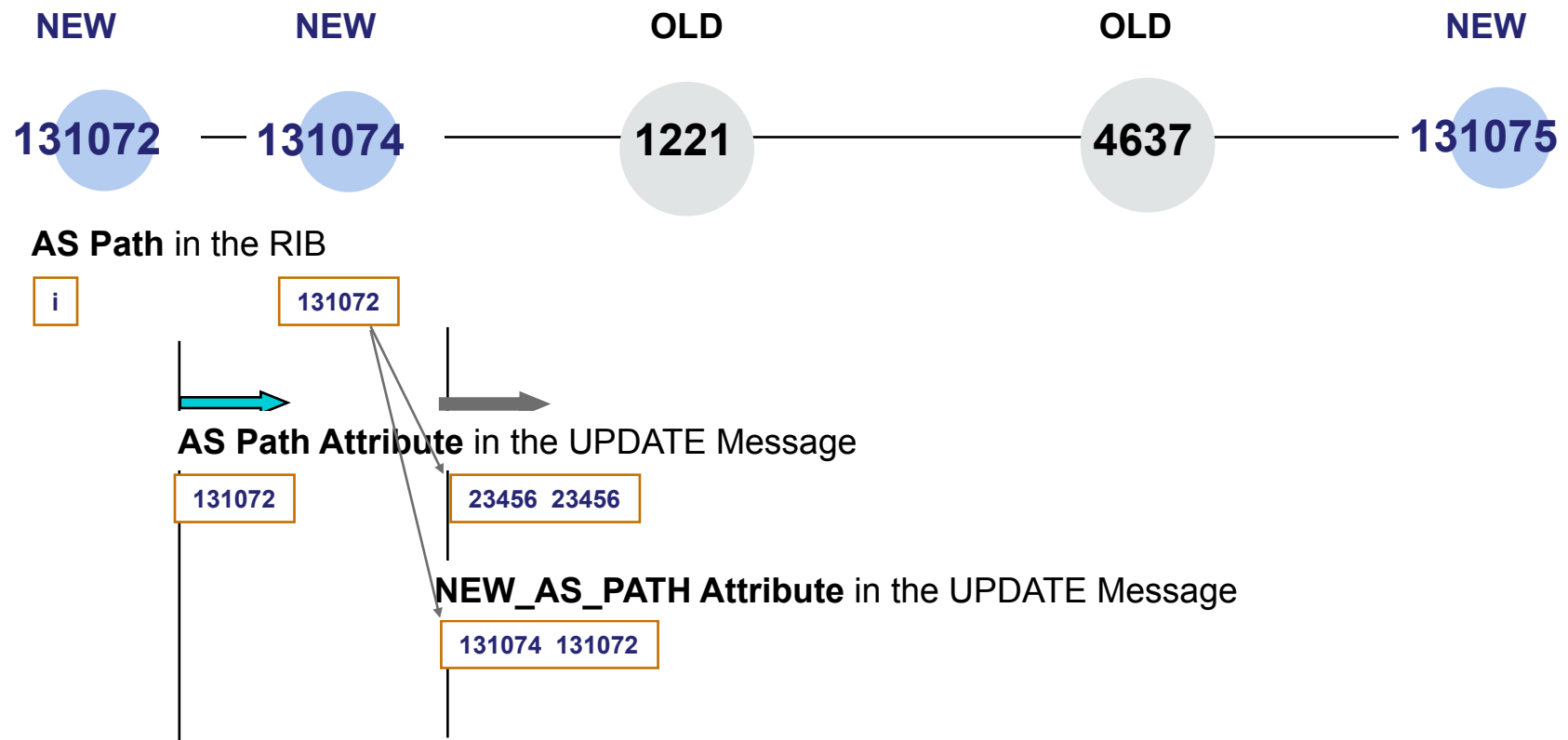
32-bit / 16-bit BGP Example



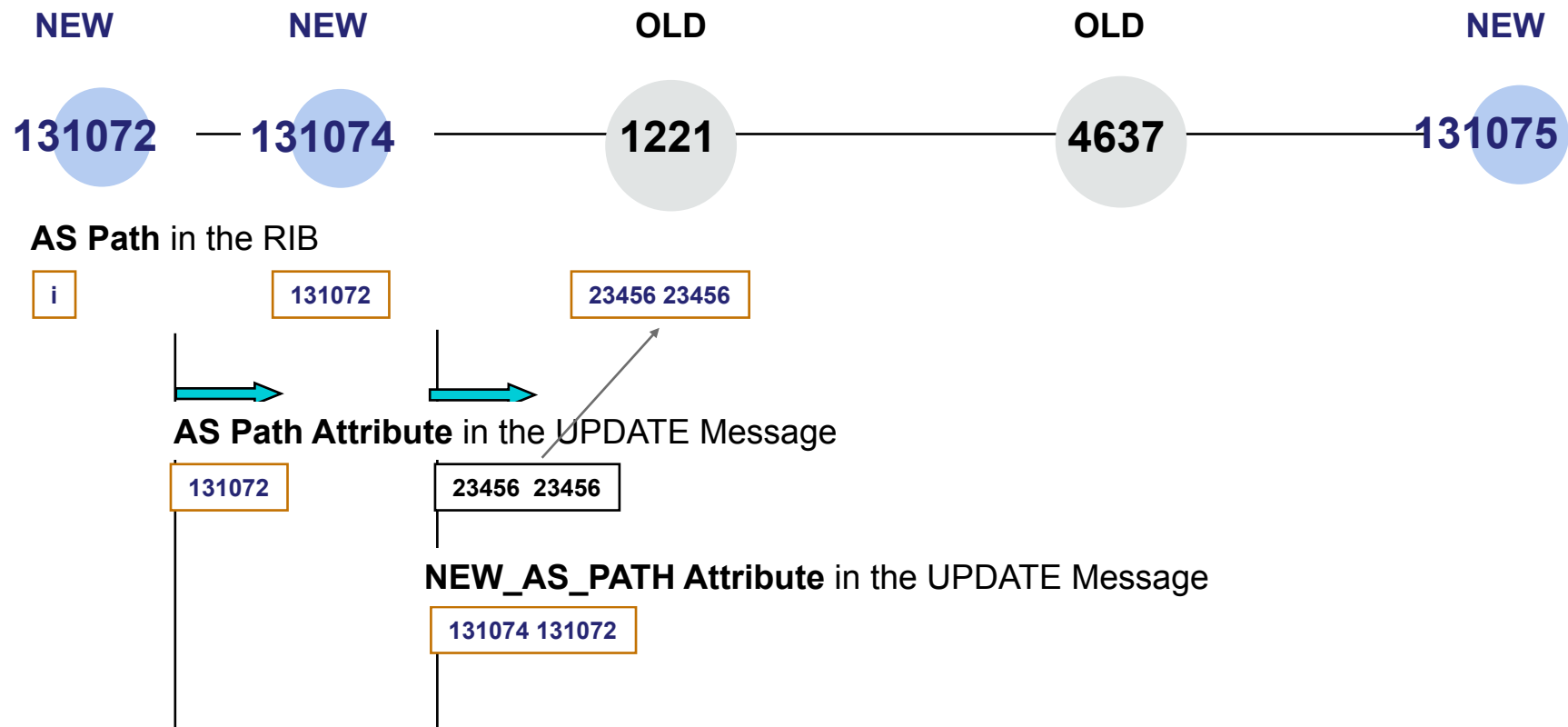
32-bit / 16-bit BGP Example



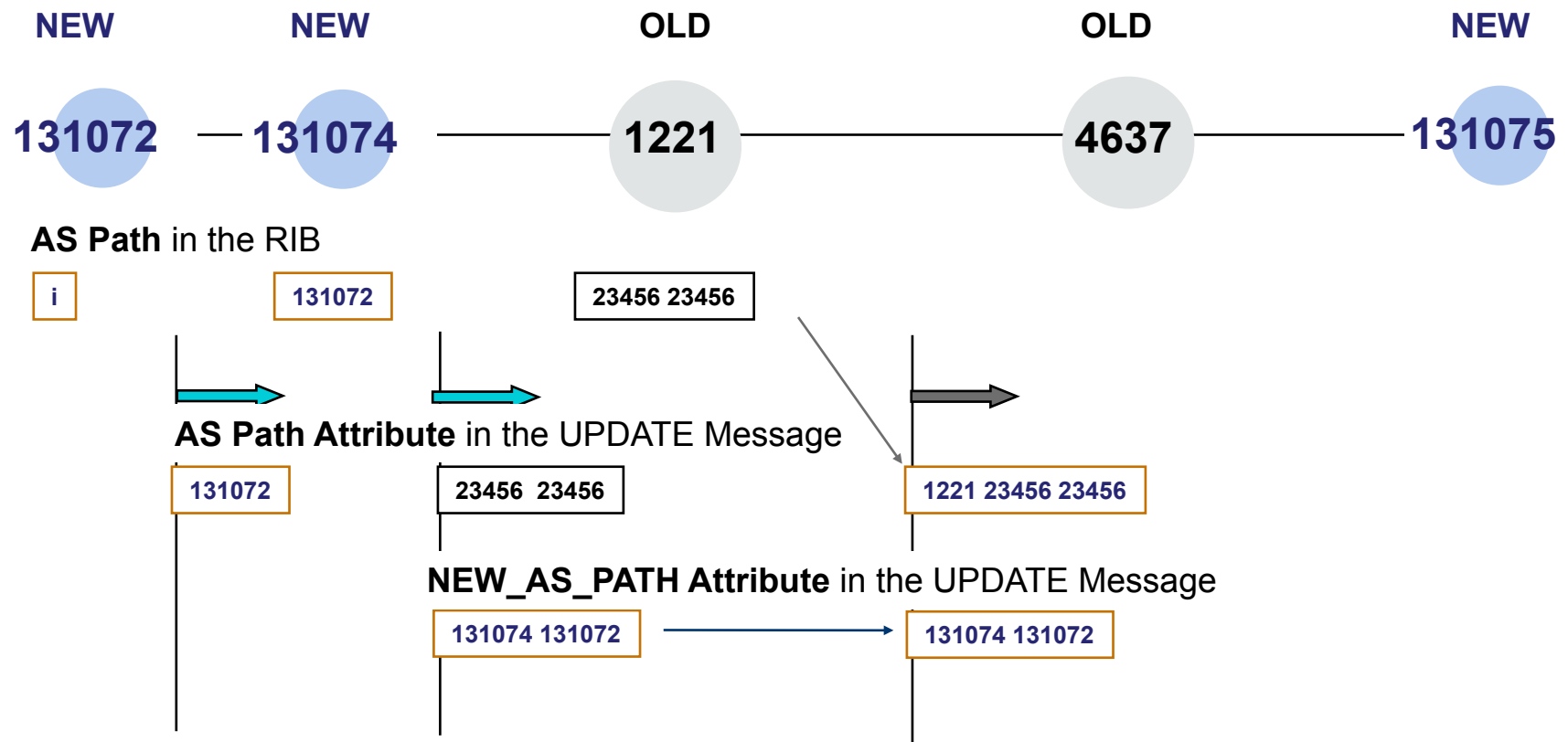
32-bit / 16-bit BGP Example



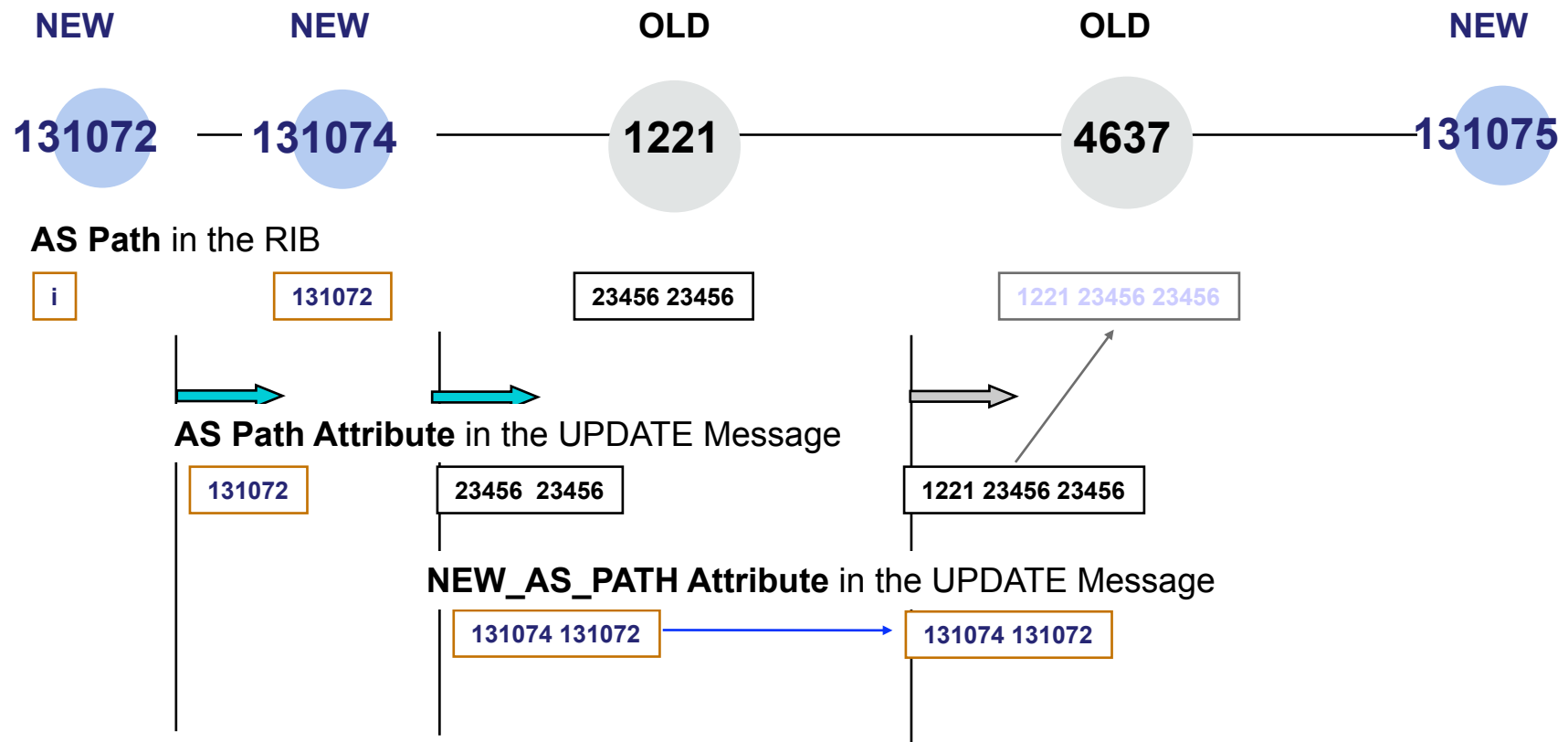
32-bit / 16-bit BGP Example



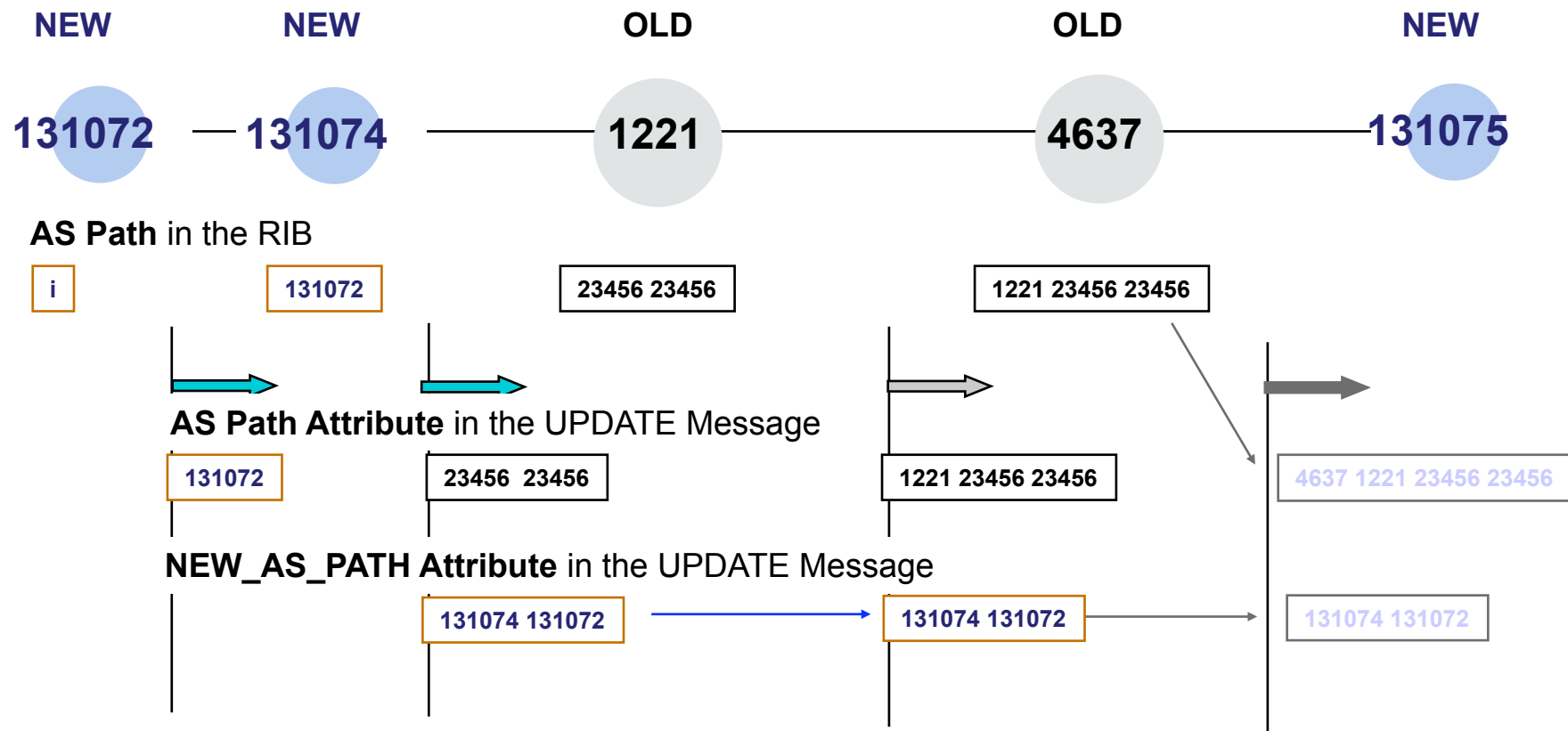
32-bit / 16-bit BGP Example



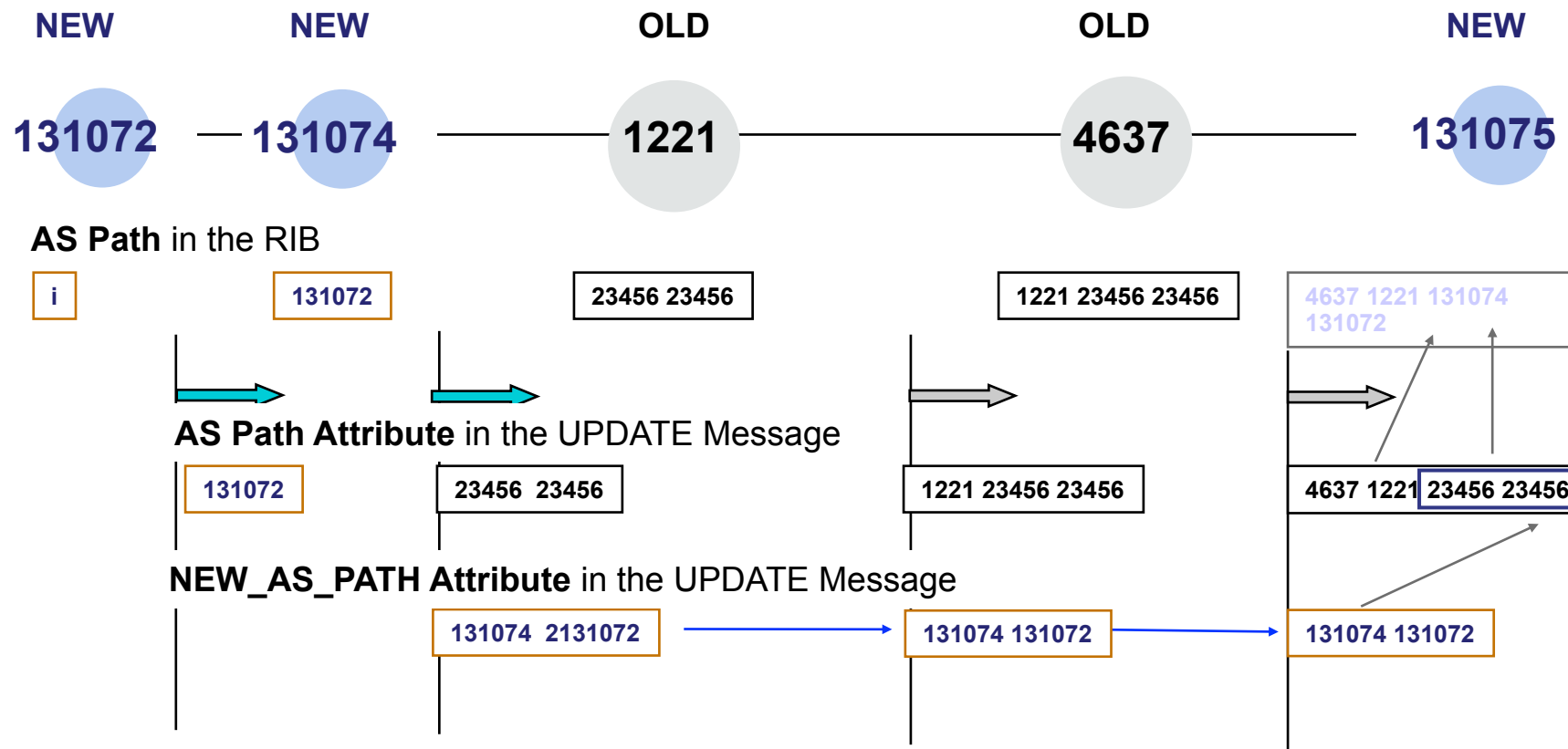
32-bit / 16-bit BGP Example



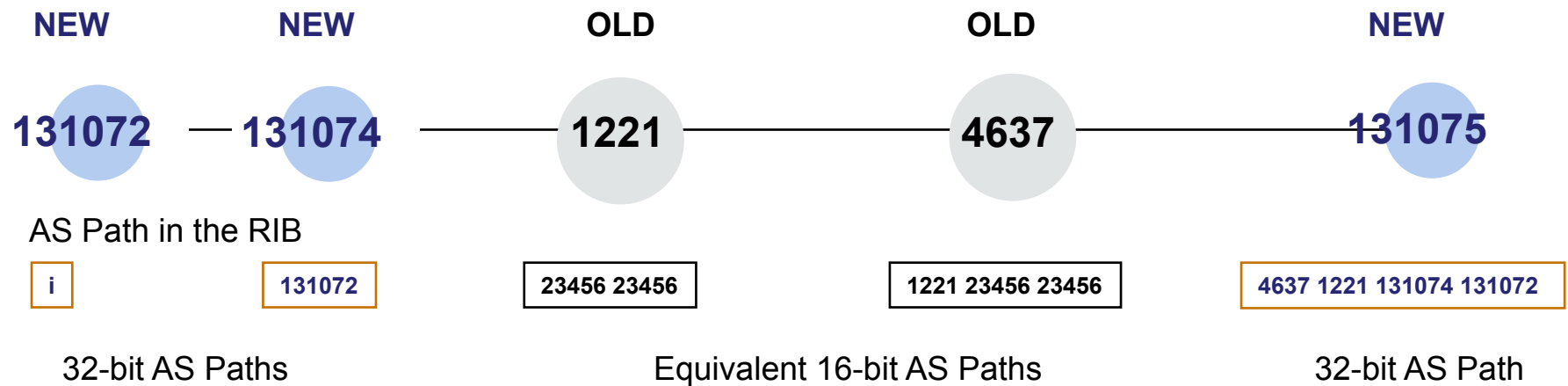
32-bit / 16-bit BGP Example



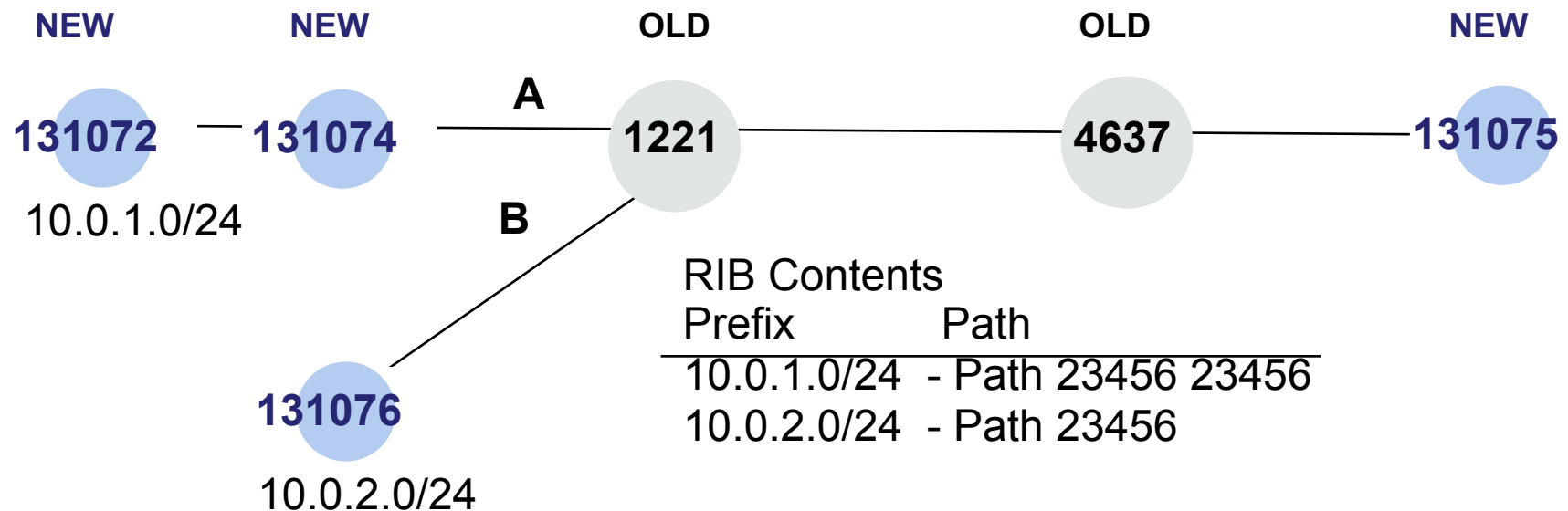
32-bit / 16-bit BGP Example



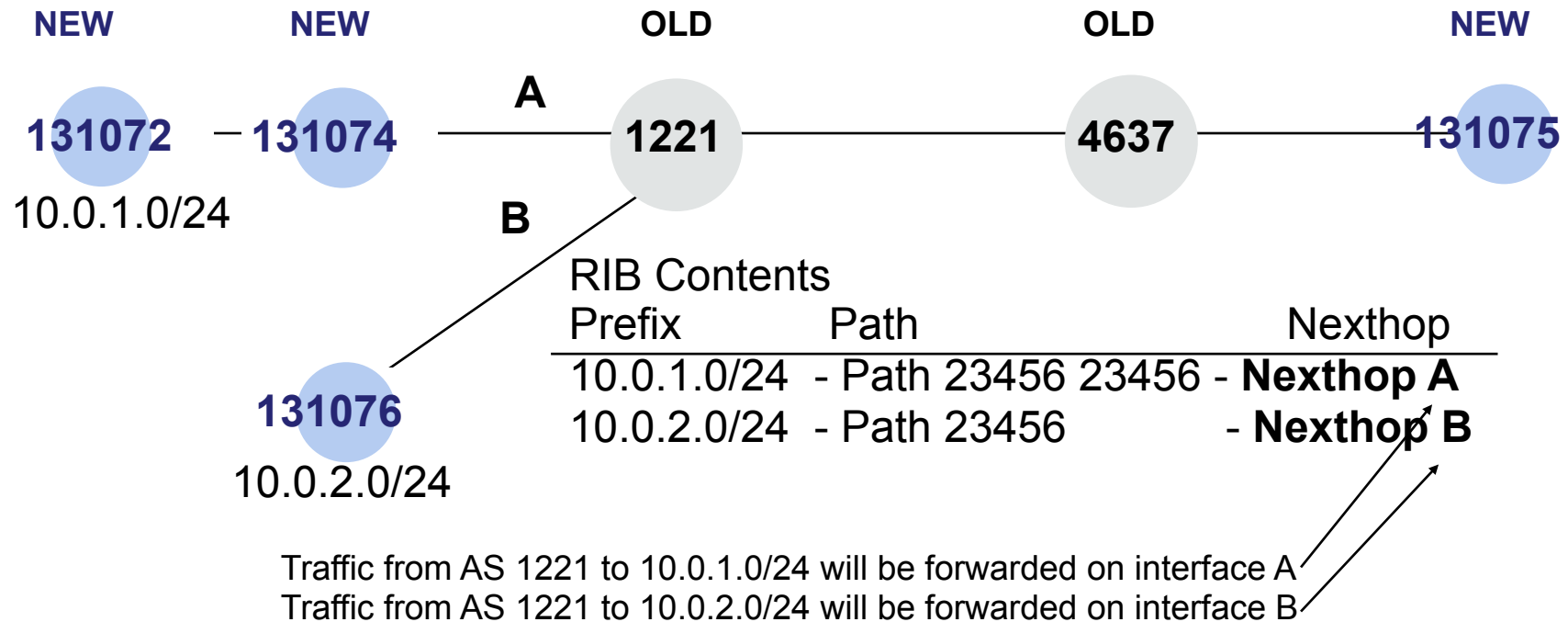
32-bit / 16-bit BGP Example



Can old-BGP get Confused?



NO! BGP Nexthop is the key!



This is standard BGP behaviour – nothing changes here for BGP as it is used today

NEW_AS_PATH Attribute

- BGP speakers in 16-bit AS domains should support NEW_AS_PATH as a transitive optional attribute in UPDATE messages
 - because that's where the 32-bit path is hiding
 - That's a "SHOULD" not a "MUST",
 - Its better if you do, but nothing fatally breaks if you don't
 - Mixed 2 / 4 Byte loops will get detected in the 16-bit world as a fallback
- Default BGP configurations will do the right thing here

NEW_AGGREGATOR Attribute

- BGP speakers in 16-bit AS domains should support NEW_AGGREGATOR as a transitive optional attribute in UPDATE messages
 - because that's where the 32-bit Aggregator AS is hiding
 - That's a "SHOULD" not a "MUST", by the way
 - Its better if you do, but nothing fatally breaks if you don't
- Default BGP configurations should do the right thing here
 -

AS 23456

- AS 23456 is going to appear in many 16-bit AS paths – both origin and transit
- This is not an error – it's a 16-bit token holder for a 32-bit AS number

AS Path and AS4 Path Example

- Router5:

Network Next Hop Metric LocPrf Weight Path

*> 2001::/32 2406:6400:F:41::1

0 23456 38610 6939 i

* i 2406:6400:D::5 0 100 0 45192 4608 4826 6939 i

*> 2001:200::/32 2406:6400:F:41::1

0 23456 38610 6939 2500 i

* i 2406:6400:D::5 0 100 0 45192 4608 4826 6939 2500 i

Questions



Thank you!

End of Session

APNIC

