IPv4/IPv6 Advanced Routing Workshop

npNOG2

15-17 June, 2017

Kathmandu, Nepal

APNIC

Instructor

Ram Krishna Pariyar

Senior Network Engineer, SUBISU

Ram is a ICANN SSR2 Review Team Member. He has been working at Subisu Cablenet Pvt Ltd as Senior Network Engineer R&D Ip Engineering since last Nine yours. Before joining Subisu, he was working at Prime College as Cisco Academy Instructor and System administrator for close to 6 Years.

Ram has more than fifteen years of operational and Academic experience & knowledge in Routing, Switching and Network Security in multi-vendor networks like Cisco, Juniper, Huawei and Brocade etc.

Areas of Interest:

IPV4/IPv6 Advanced Routing, MPLS Network Design and Deployment, Network Security, Virtualization

Contact:

Email: pariyar.ramkrishna@gmail.com

Instructor

Tashi Phuntsho

Senior Network Analyst/Technical Trainer, APNIC

Tashi has experience in network design, operation, and maintenance having worked for more than 13 years as a core network engineer in service provider network. He has been involved in capacity development in the APNIC community by providing training in number of technical areas, such as Routing & Switching, Network Design and Architecture, Network Security, IPv6, DNSSEC, and so on.

Tashi completed his undergraduate studies in Electrical and Electronics engineering from India, complemented by postgraduate studies in Network Security from Japan and Network Systems from Australia.

Areas of Interest:

BGP, IS-IS/OSPF, IPv6, Blockchain Technology, Securing Internet Routing (RPKI), DWDM, Network Security, Next Generation Networks (SDN, Internet of Things), DNS and DNSSEC.

Contact:

Email: tashi@apnic.net

Instructor

Shaqayeq Saleh

Network Engineer, APNIC

Shaqayeq has over a decade experience in computer networking, network design, operation and maintenance in service provider network.

Shaqayeq completed her undergraduate degree in Computer Science, Computer/Software from Iran, followed by postgraduate study in Information Technology (Network Management) from Australia.

Areas of Interest:

BGP, IOSPF, EIGRP, IPv6, MPLS, CWDM/DWDM, Network Security

Contact:

Email: shaqayeq@apnic.net

Sessions

- Session 1:
  - 9:30 AM – 11:00 AM
- Session 2:
  - 11:30 AM – 13:00 PM
- Session 3:
  - 14:00 PM – 15:30 PM
- Session 4:
  - 16:00 PM – 17:30 PM

Link to workshop session:

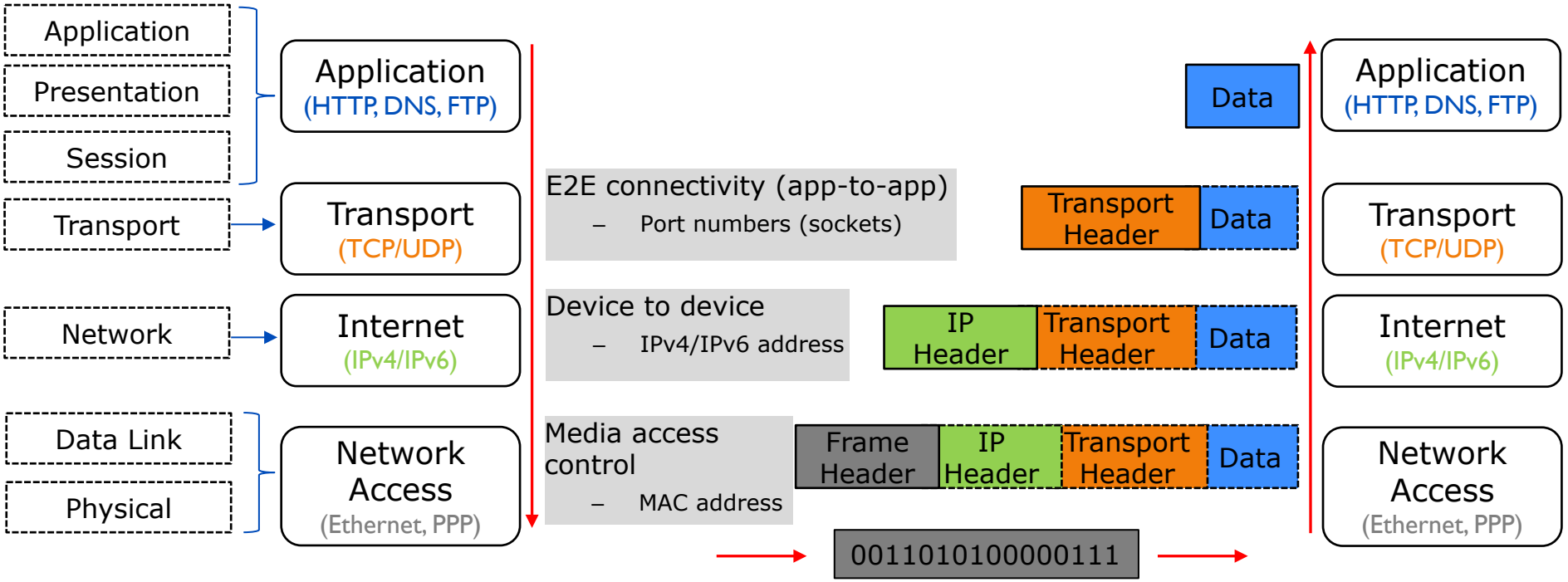
- SSID: npnog2
- PW: workshop

- Course Material: https://wiki.apnictraining.net

Agenda

- **TCP/IP communication and IP Routing**
- Internet Routing and Routing Protocols
- IPv6 Address Structure
- Training Lab Topology
- OSPF Operation
- Lab exercise: Basic router and OSPF
- BGP – operation and attributes
- BGP Scaling Techniques
- Lab: iBGP, eBGP, Peer-group, RR

TCP/IP Communication



Addressing is the key**!**

Internet/Network Layer

- Host to host communication across networks
  - Addressing
    - unique and hierarchical network-wide address
  - Routing
    - the best path to the destination
- Current protocols
  - IPv4 and IPv6

L3 Device/Router

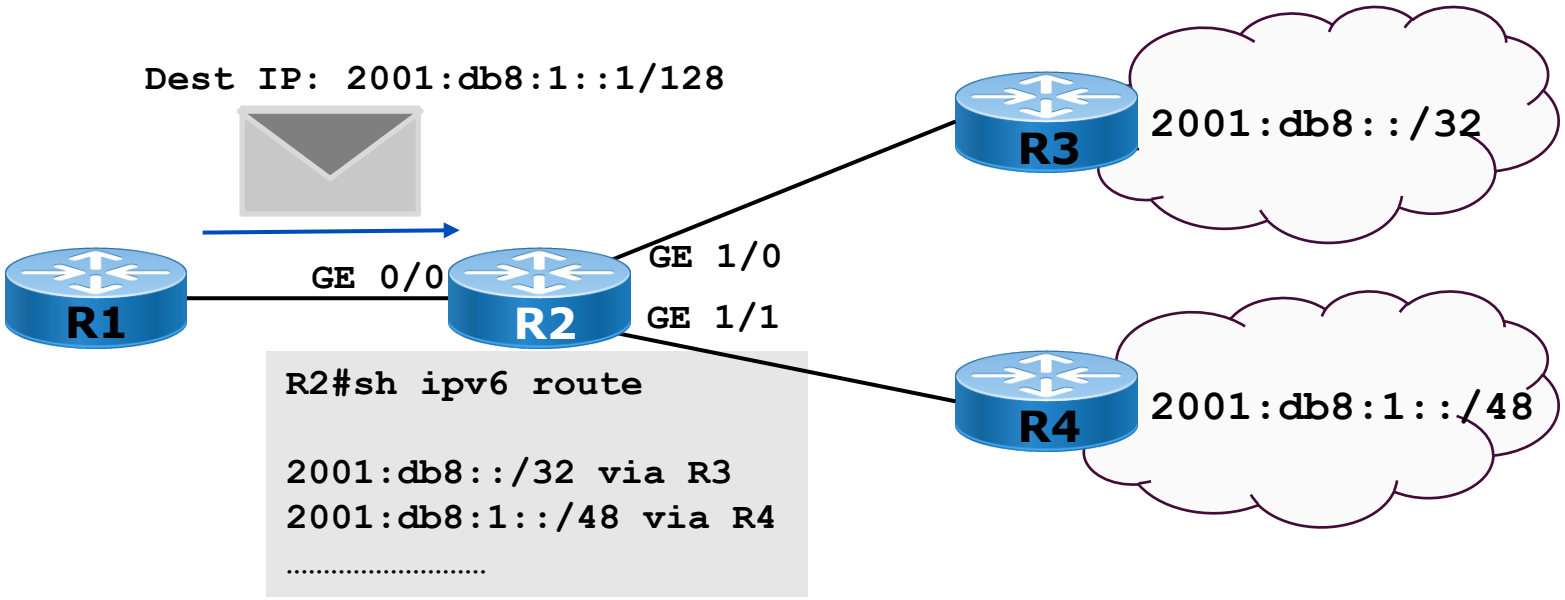
- L3 device gets the packet one step closer
  - The next hop to reach the destination!

- Router
  - Finds the best path to the destination, and
  - Forwards the packet to the next hop (a step closer) to reach the destination

Best path lookup – Routing Decision

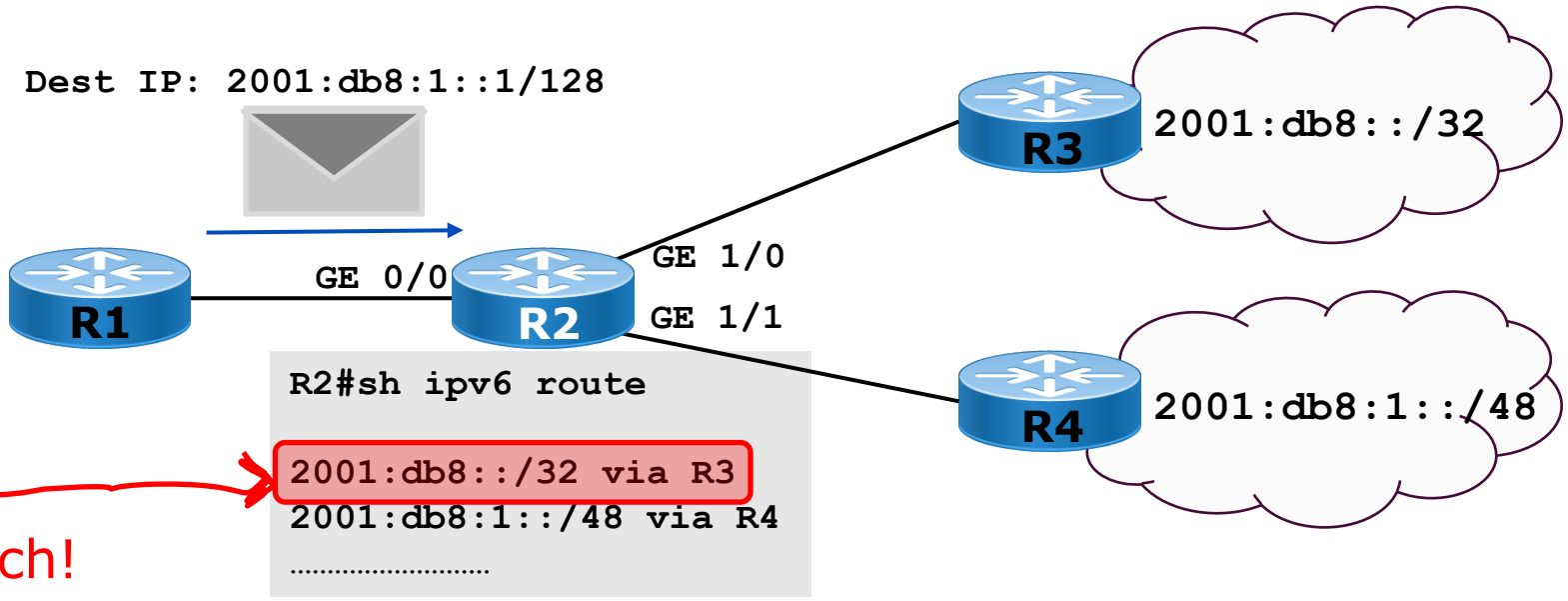
- Inspects the destination address of the packet
  - Network portion
- Looks up its routing table for a "best match"
  - Longest matching left-most bits
- If no match, checks for default route
  - If no default route, drop the packet!

Best path (route) lookup



2001:db8::/32	0010000000000001:1101101110000000::
2001:db8:1::/48	0010000000000001:1101101110000000:0000000000000001::

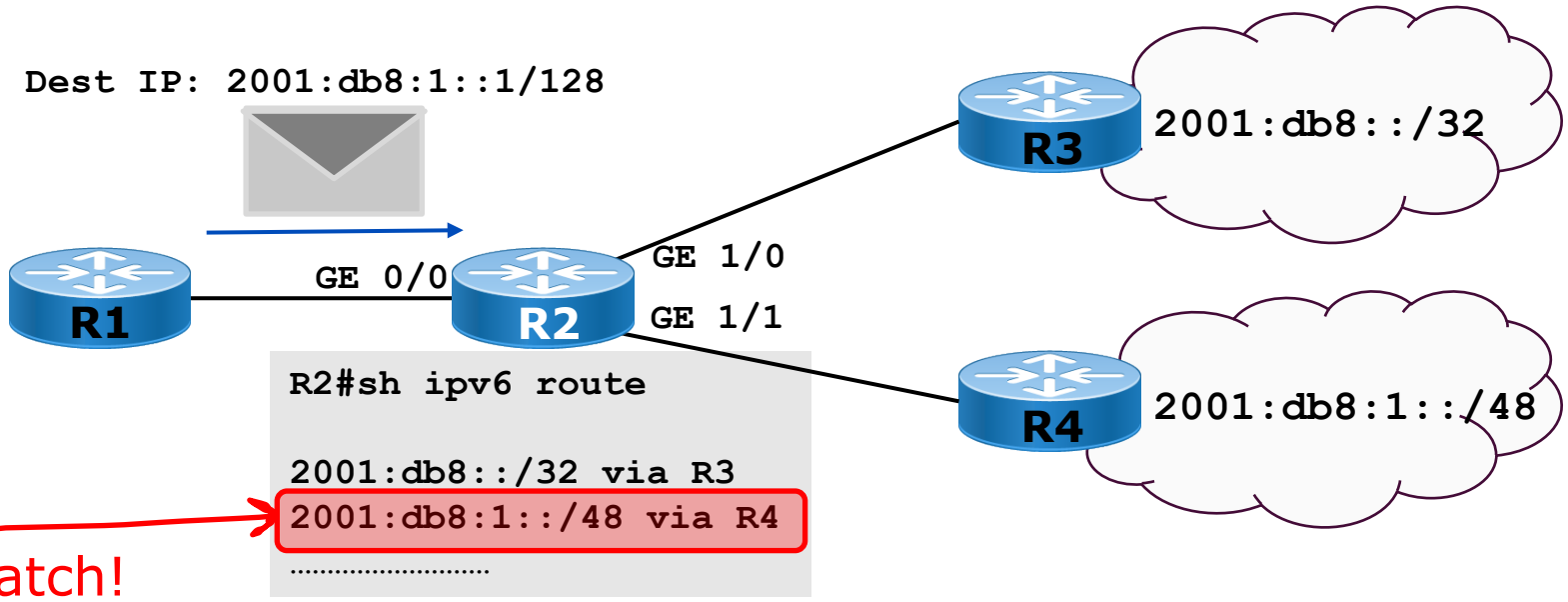
Best path (route) lookup



2001:db8:1::1	0010000000000001:1101101110000000:0000000000000001:0:0:0:0:0000000000000001
FFFF:FFFF:: (/32)	1111111111111111:1111111111111111:0000000000000000:0:0:0:0:0000000000000000
2001:db8::	0010000000000001:1101101110000000::

AND

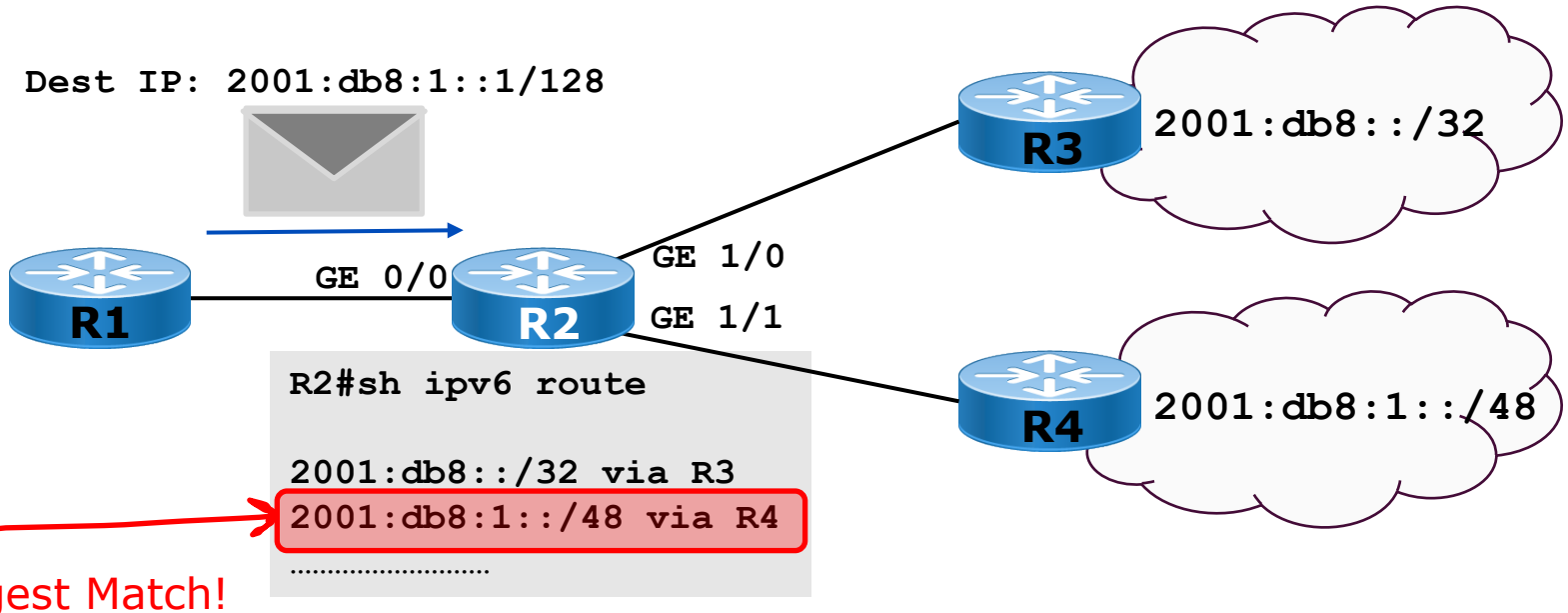
Best path (route) lookup



2001:db8:1::1	0010000000000001:1101101110000000:0000000000000001:0:0:0:0:0000000000000001
FFFF:FFFF:FFFF::(/48)	1111111111111111:1111111111111111:1111111111111111:0:0:0:0:0000000000000000
2001:db8:1::	0010000000000001:1101101110000000:0000000000000001::

AND

Best path (route) lookup

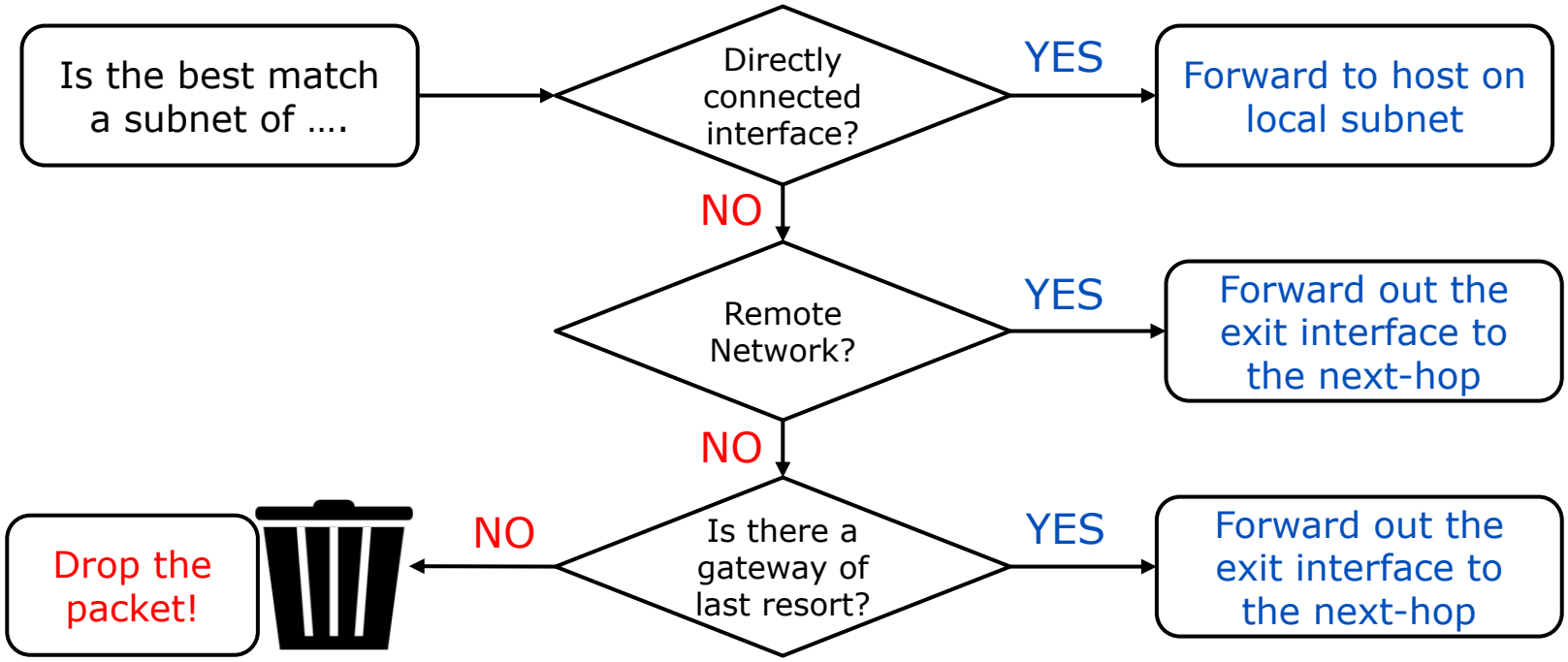


2001:db8:1::1	0010000000000001:1101101110000000:0000000000000001:0:0:0:0:0000000000000001
FFFF:FFFF:FFFF::(/48)	1111111111111111:1111111111111111:1111111111111111:0:0:0:0:0000000000000000
2001:db8:1::	0010000000000001:1101101110000000:0000000000000001::

AND

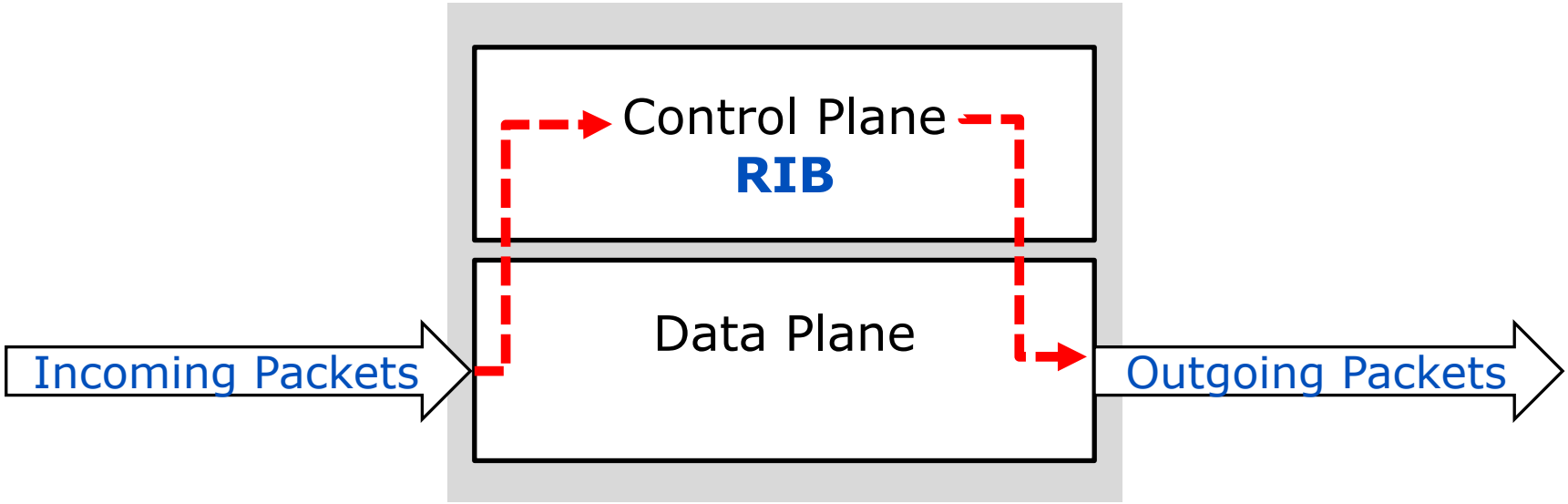
Forwarding Decision

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



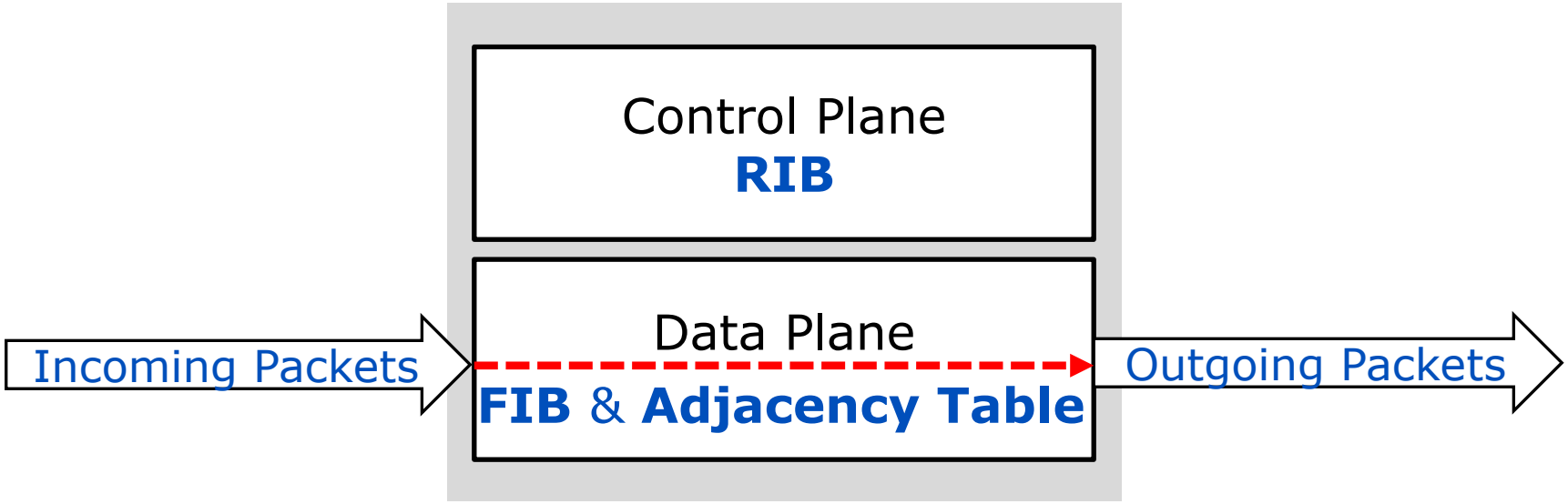
Process vs Hardware Switching

- Incoming packet forwarded to the control plane (CPU)
  - routing table (RIB) lookup, frame re-write (next-hop MAC), and forwarded to the exit interface



Process vs Hardware Switching

- Instead two hardware based tables
  - FIB derived from the RIB- all destinations and next-hops
  - Adjacency table from the ARP table- L2 header info for each next-hop in the FIB



Questions

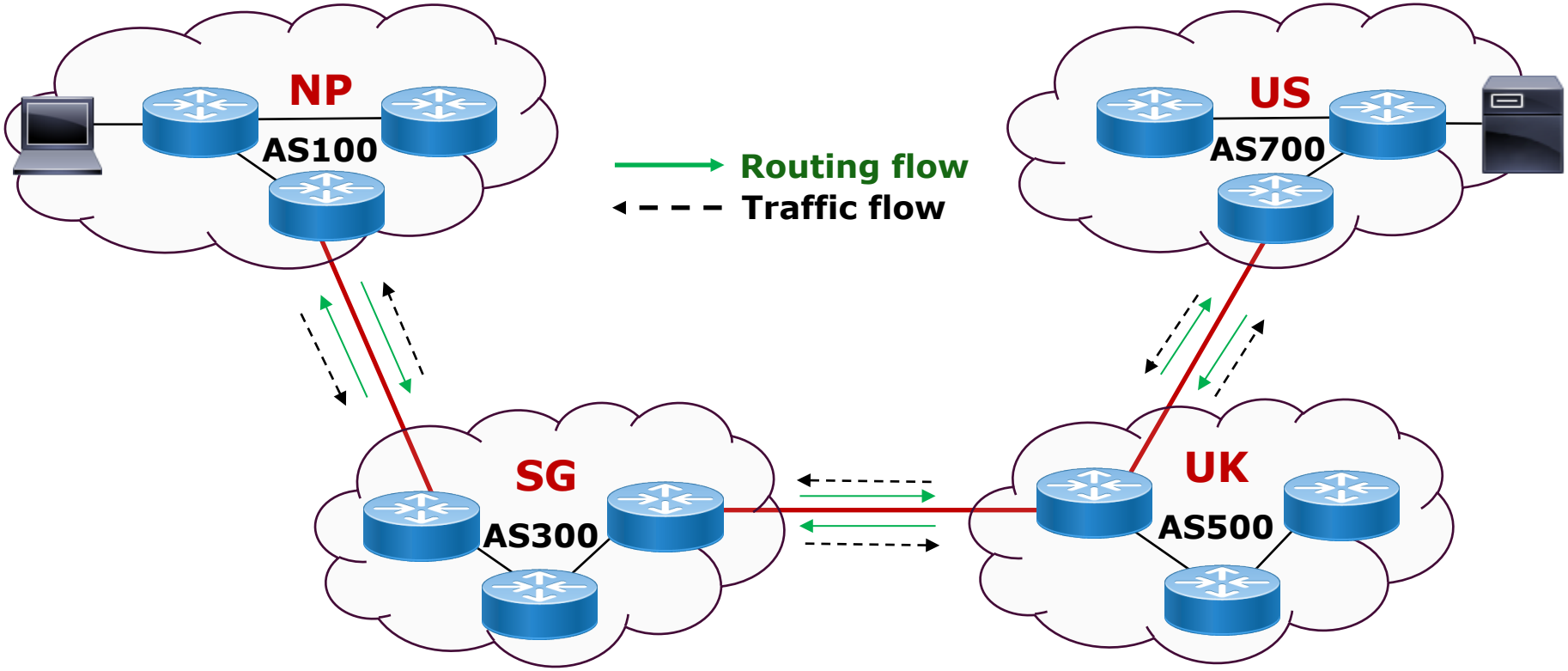
Agenda

- TCP/IP communication and IP Routing
- **Internet Routing and Routing Protocols**
- IPv6 Address Structure
- Training Lab Topology
- OSPF Operation
- Lab exercise: Basic router and OSPF
- BGP – operation and attributes
- BGP Scaling Techniques
- Lab: iBGP, eBGP, Peer-group, RR

Internet Routing

- How does a user in PH access a service hosted in the US?
  - The ISP in PH could directly connect to the ISP in US
    - Neither scalable nor economical
  - Instead, the PH ISP shares its network information with its neighbor ISPs
  - The ISP in US does the same with its own neighbors
    - Neighbor ISPs propagate the information to their neighbors, and so on…
    - Eventually, they both learn about each other's network!

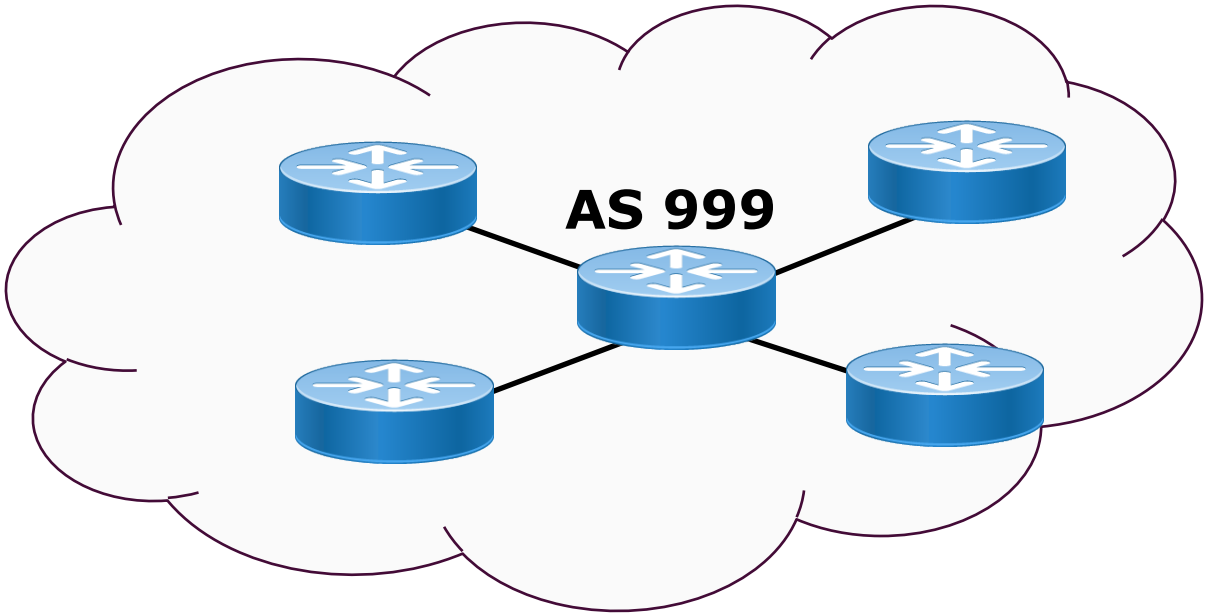
Internet Routing



Exchange of network information – Routing
Networks (ASes) connected together – Internet

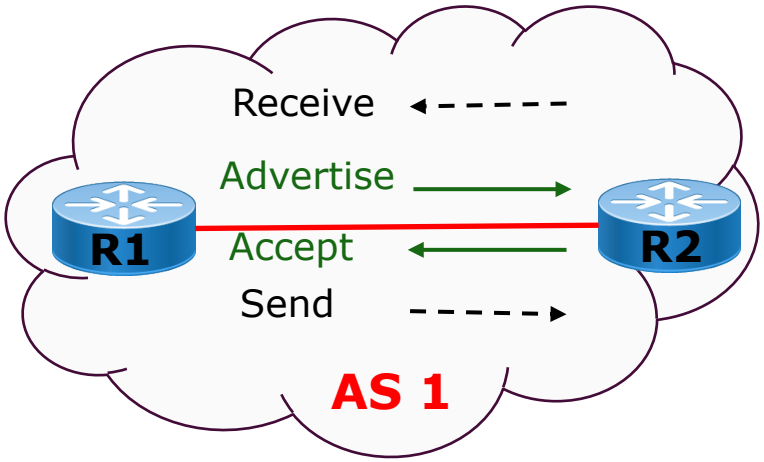
Autonomous System (AS)

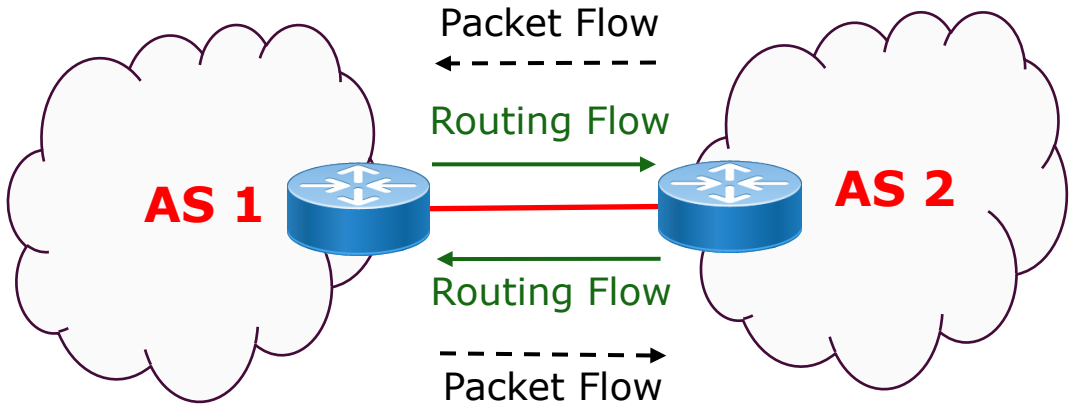
- A group of networks with the same routing policy
  - Usually under single administrative control



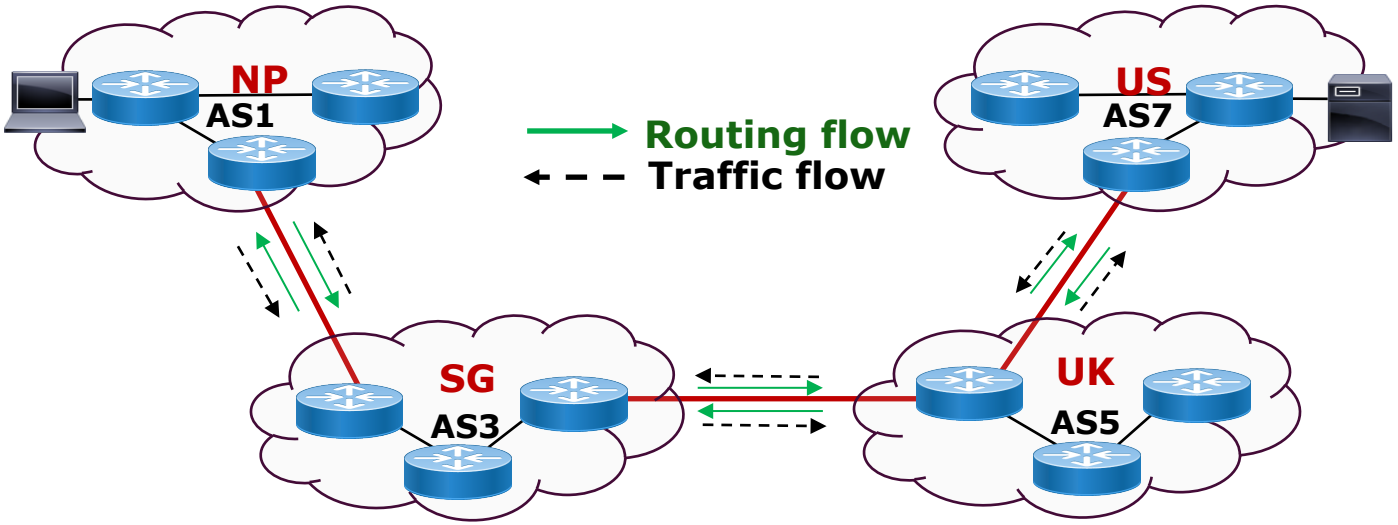
Routing Flow & Traffic Flow

- Traffic and network info always flow in opposite direction!
  - network info exchanged in both directions for bi-directional traffic flow
  - manipulate inbound/outbound routing info to influence outgoing/incoming traffic



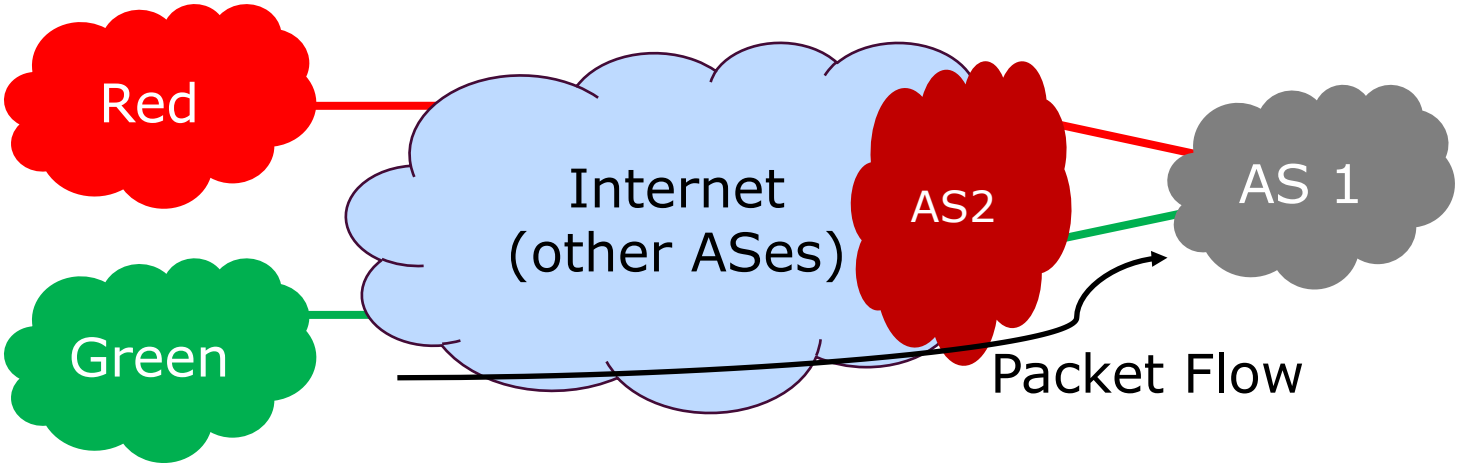


Routing & Traffic Flow: Internet



- For user (N1) in AS1 to send traffic to user (N7) in AS7:
  - AS7 must originate and announce N7 to AS5.
  - AS5 must accept N7 from AS7, and advertise to AS3.
  - AS3 must accept and forward N7 to AS1
  - AS1 must accept N7 from AS3

Routing Policy Limitations



- For the above policy, AS1
  - Needs to accept routes originating from Red AS over Red link, and Green AS over Green link
- But any intermediate AS (AS2) needs to cooperate

Routing Protocols

- How do routers exchange network information with each other?
  - Routing Protocols!
  - IGP & EGP

Interior Gateway Protocol (IGP)

- To exchange network info within an AS
  - To carry infrastructure info (loopbacks & ptp)
    - No customer routes!
  - Allows all routers within an AS to learn about each other

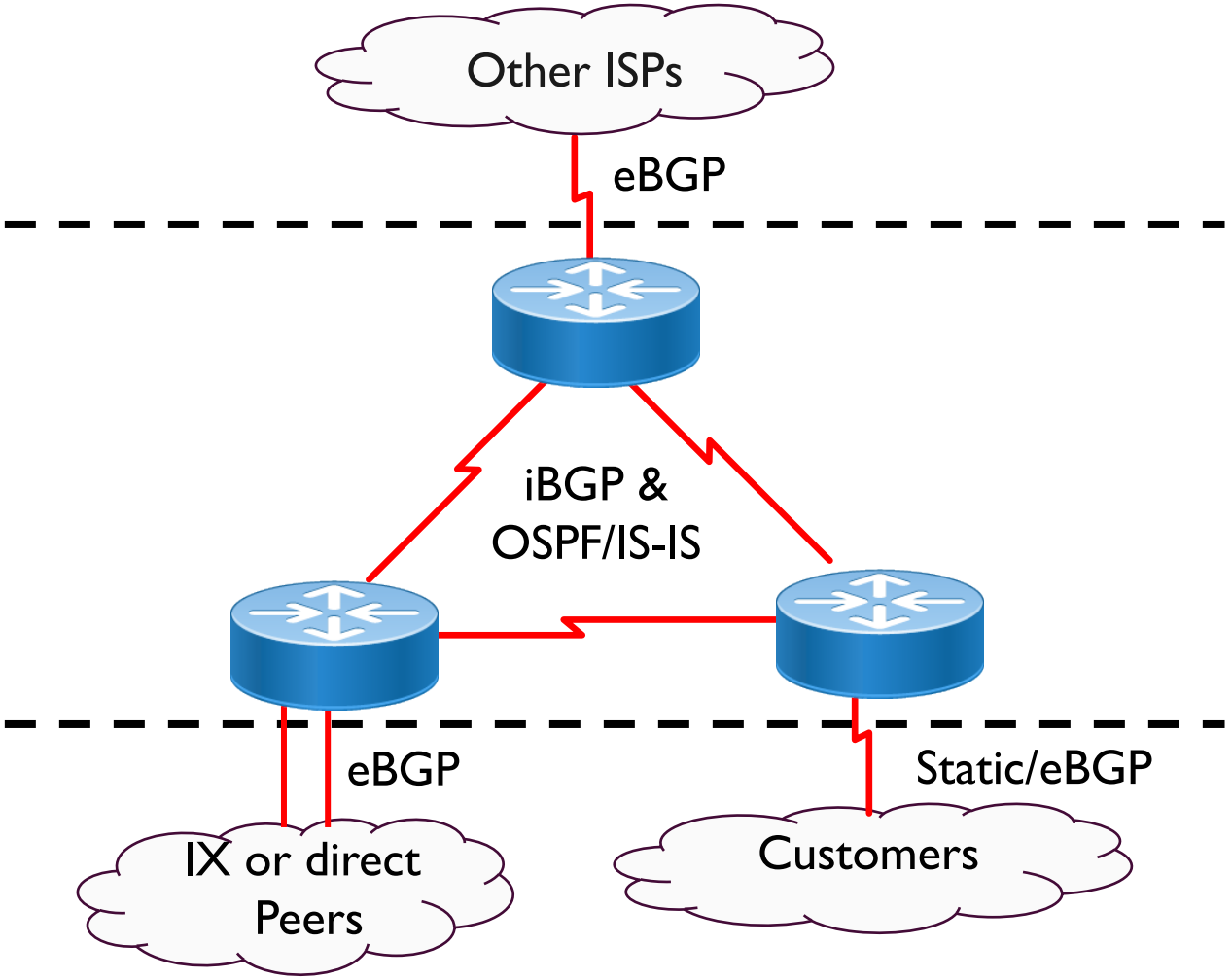
- Two most widely used IGPs in operator networks
  - OSPF & IS-IS
    - Uses the SPF algorithm
    - Best path selection based on lowest cost/metric
    - Supports hierarchical routing – scalability!

Exterior Gateway Protocol (EGP - BGP)

- To exchange network info between ASes
  - Implement routing policies (manipulate traffic path)
  - Define administrative boundary

- BGP is the de facto EGP!

Routing Protocols Hierarchy



Cisco Default AD

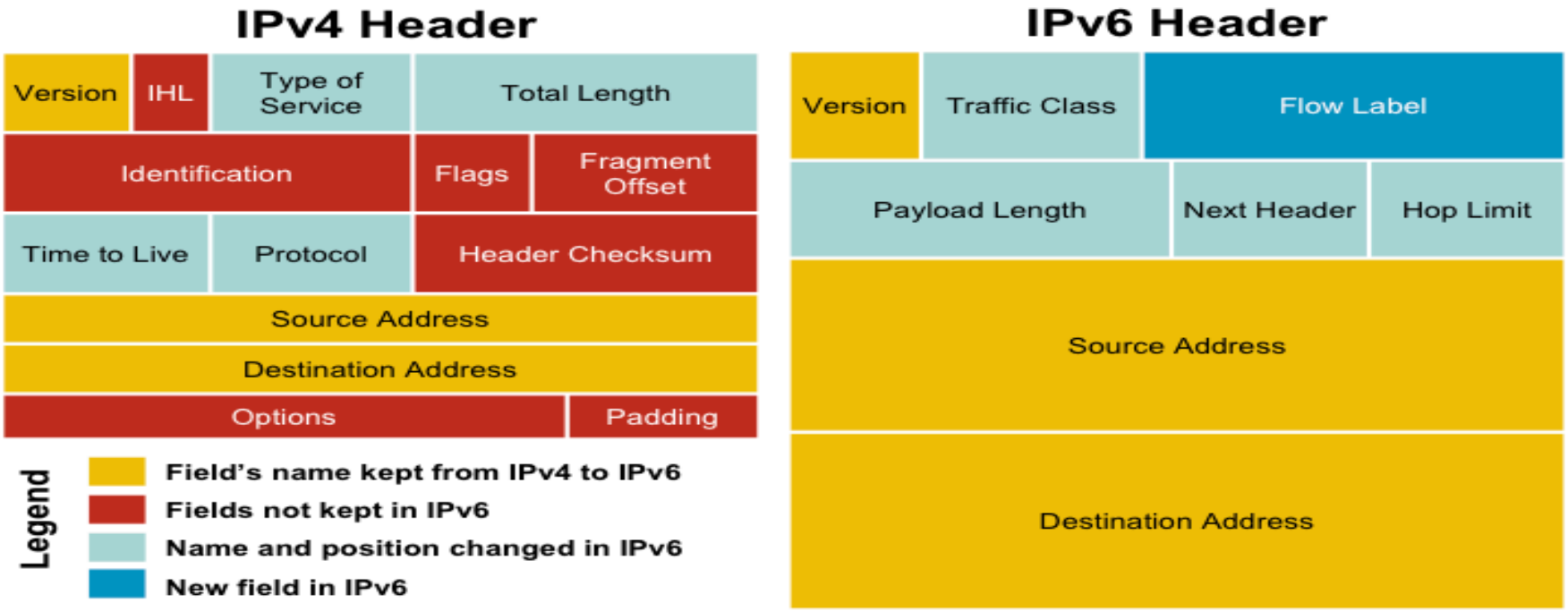
Route Source	AD
Connected Interface	0
Static Route	1
Enhanced IGRP Summary Route	5
External BGP	20
Internal Enhanced IGRP	90
IGRP	100
OSPF	110
IS-IS	115
RIP	120
EGP	140
External Enhanced IGRP	170
Internal BGP	200
Unknown	255

Questions

Agenda

- TCP/IP communication and IP Routing
- Internet Routing and Routing Protocols
- **IPv6 Address Structure**
- Training Lab Topology
- OSPF Operation
- Lab exercise: Basic router and OSPF
- BGP – operation and attributes
- BGP Scaling Techniques
- Lab: iBGP, eBGP, Peer-group, RR

Protocol Header Comparison

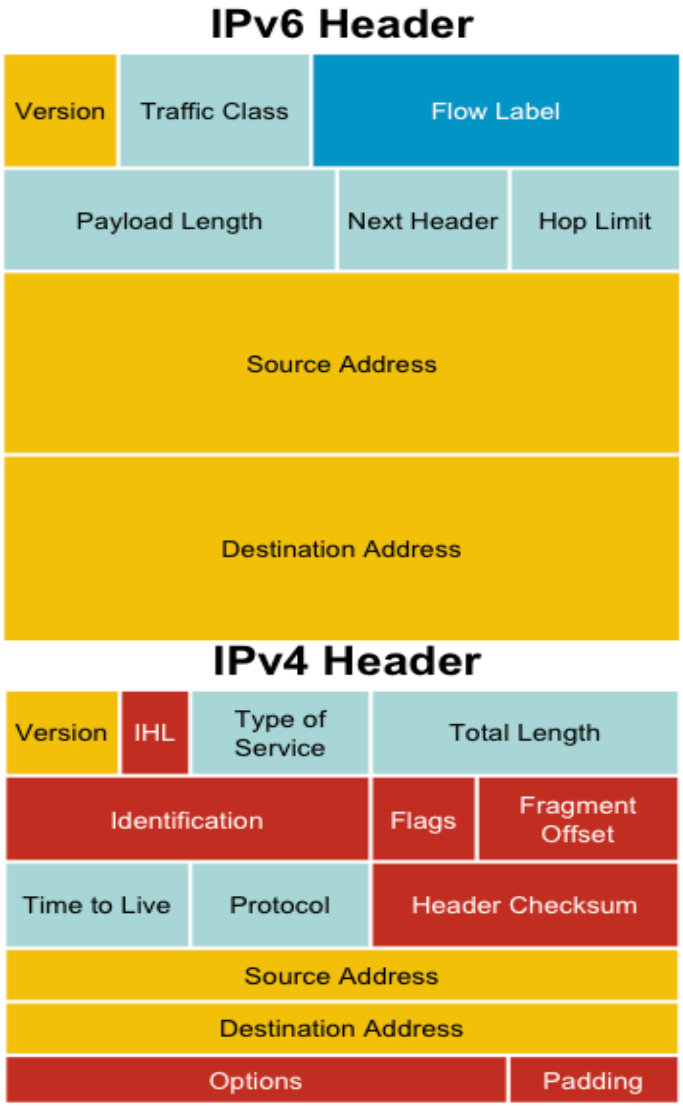


- IPv4 contains 10 basic header fields, while IPv6 has 6 basic header fields
- IPv6 header size is 40 octets compared to 20 octets (but variable) for IPv4
- So a smaller number of header fields and 64-bit aligned to enable fast processing by current processors (*fixed vs variable logic*)
- No header checksum!

Diagram Source: www.cisco.com

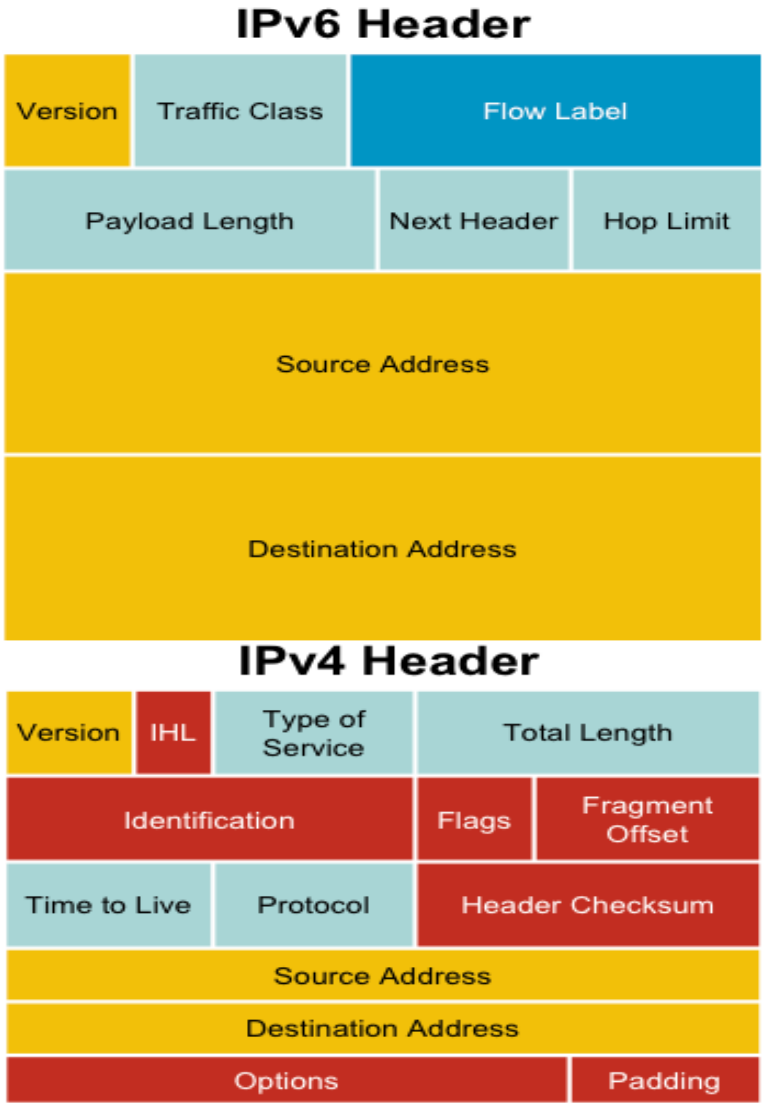
IPv6 Protocol Header Format

- Version (4-bit):
  - 4-bit IP version number (6)
- Traffic class (8-bit):
  - Similar to DiffServ in IPv4; define different classes or priorities.
- Flow label (20-bit):
  - allows IPv6 packets to be identified based on flows (multilayer switching techniques and faster packet-switching performance)



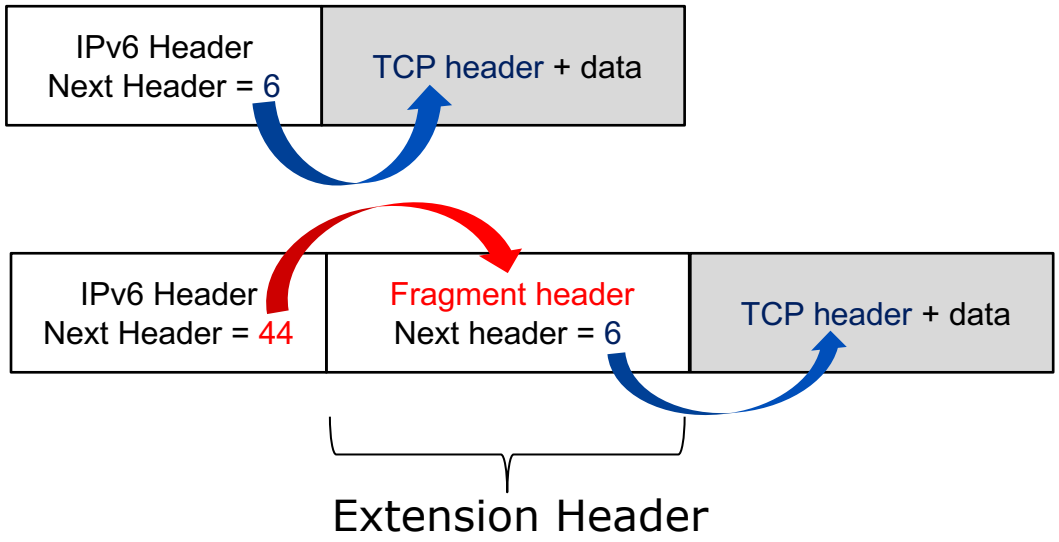
IPv6 Protocol Header Format

- Payload length (16-bit):
  - Defines the length of the IPv6 payload (including extension headers); Total Length in IPv4 includes the header.
- Next header (8-bit):
  - Identifies the type of information following IPv6 header. Could be upper layer (TCP/UDP), or an extension header (similar to Protocol field in IPv4).
- Hop limit (8-bit):
  - Similar to TTL in IPv4



IPv6 Extension Header

- IPv6 allows an optional *Extension Header* in between the IPv6 header and upper layer header
  - to carry additional Internet layer information, identified by the unique Next Header values



Next Header values:

0 Hop-by-hop option
2 ICMP
6 TCP
17 UDP
43 Source routing
44 Fragmentation
50 Encrypted security payload
51 Authentication
59 Null (No next header)
60 Destination option

IPv6 Extension Header (contd)

- An IPv6 packet may carry none or many extension headers
  - A next header value of 6 or 17 (TCP/UDP) indicates there is no extension header
    - the next header field points to TCP/UDP header, which is the payload

- Unless the next header value is 0 (*Hop-by-Hop option*), extension headers are processed only by the destination node, specified by the destination address.

Fragmentation Handling In IPv6

- In IPv6, fragmentation is only performed by the host/source nodes, and not the routers along the path (unlike IPv4)
- Each source device tracks the MTU size for each session
- When a IPv6 host has large amount of data to be sent, it will be send in a series of IPv6 packets (fragmented)
  - IPv6 hosts use Path MTU Discovery (PMTUD) to determine the most optimum MTU size along the path

Source: www.cisco.com

Path MTU Discovery

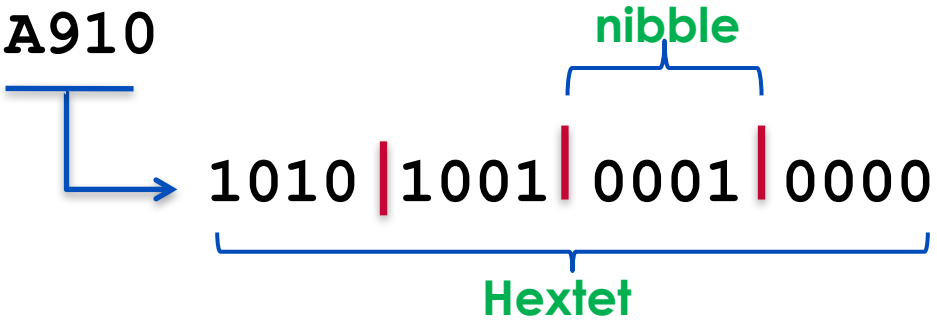
- With PMTUD, the source IPv6 device assumes the initial PMTU is the MTU of the first hop in the path
  - upper layers (Transport/Application) send packets based on the first hop MTU
  - If the device receives an "*ICMP packet too big*" message, it informs the upper layer to reduce its packet size, based on the actual MTU size (contained in the message) of the node that dropped the packet

Path MTU =1100

IPv6 Address Representation

- IPv6 address is 128 bits
- Number of IPv6 addresses : 2^128 ~ 3.4 x 10^{38}
- IPv6 address is represented in hexadecimal
  - 4-bits (nibble) represent a hexadecimal digit
  - 4 nibbles (16-bits) make a hextet
  - represented as eight hextets (4 nibbles or 16 bits), each separated by a colon (:)

2001:ABCD:1234::DC0:A910



IPv6 Address Representation (2)

2001:0DB8:0000:0000:0000:036E:1250:2B00

- Abbreviated form

2001:**0**DB8:**0000:0000:0000:0**36E:1250:2B00 ⇦ **Leading 0s**

  - Leading zeroes (**0**) in any hextet can be omitted
    2001:DB8:**0:0:0**:36E:1250:2B00 ⇦ **Sequence of 0s**
  - A double colon (**::**) can replace contiguous hextet segments of zeroes
    2001:DB8**::**36E:1250:2B00 ⇦ **Double colons**
  - (::) can only be used once!

IPv6 Address Representation (3)

- Double colons (::) representation
  - RFC5952 recommends that the rightmost set of :0: be replaced with :: for consistency

    2001:DB8:0:0:2F:0:0:5

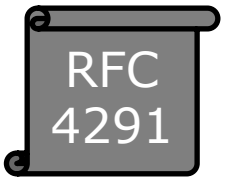
    2001:DB8:0:0:2F::5 instead of 2001:DB8::2F:0:0:5

- Prefix Representation
  - Representation of prefix is similar to IPv4 CIDR

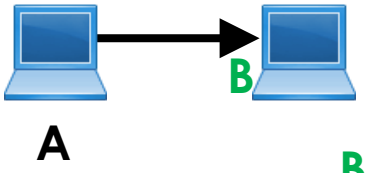
→ prefix/prefix-length

2001:DB8:12::/40

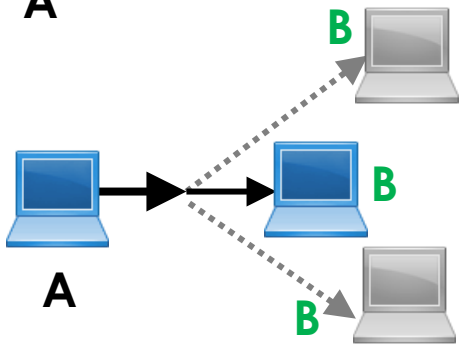
IPv6 Addressing Model



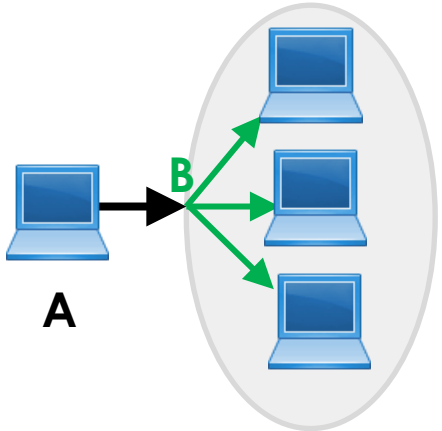
- Unicast Address
  - Assigned to a single interface
  - Packet sent only to the interface with that address



- Anycast Address
  - Same address assigned to more than one interface (on different nodes)
  - Packet for an anycast address routed to the nearest interface (routing distance)



- Multicast Address
  - group of interfaces (on different nodes) join a multicast group
  - A multicast address identifies the interface group
  - Packet sent to the multicast address is replicated to all interfaces in the group



Special Unicast Addresses

- Unspecified Address (absence of a address)

  ::/128

- Loopback (test OSI/TCP-IP stack implementation)

  ::1/128

Global Unicast Addresses

- Globally unique and routable IPv6 address
- Currently, only global unicast address with first three bits of **001** have been assigned

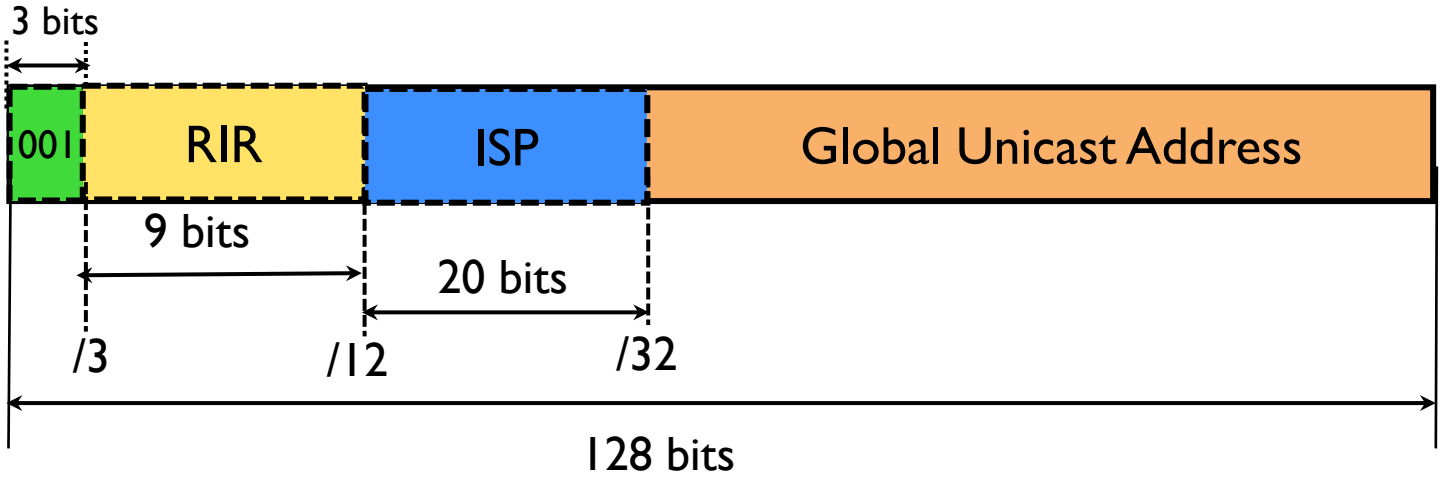
```
0010 0000 0000 0000 (2000::/3)
0011 1111 1111 1111 (3FFF::/3)
```

- IANA gives a /12 each from 2000-3FFF::/3 to each RIR

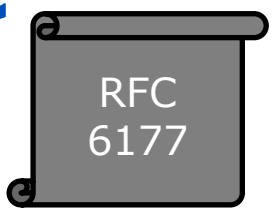
APNIC	2400::/12
ARIN	2600::/12
LACNIC	2800::/12
RIPE NCC	2A00::/12
AfriNIC	2C00::/12

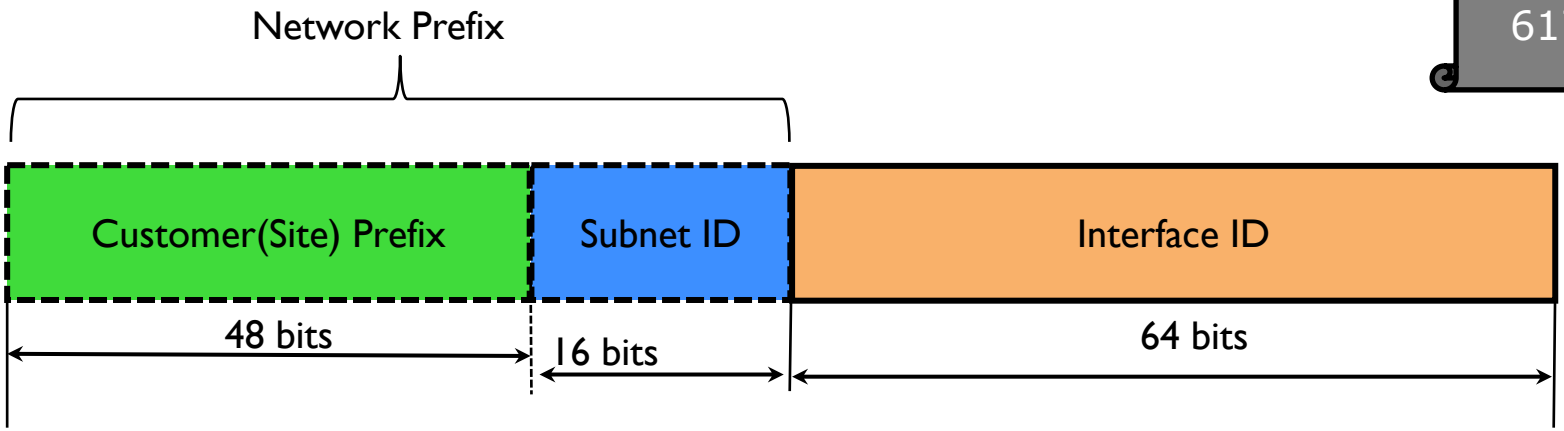
Global Unicast Addresses

- RIRs assign /32 to ISPs



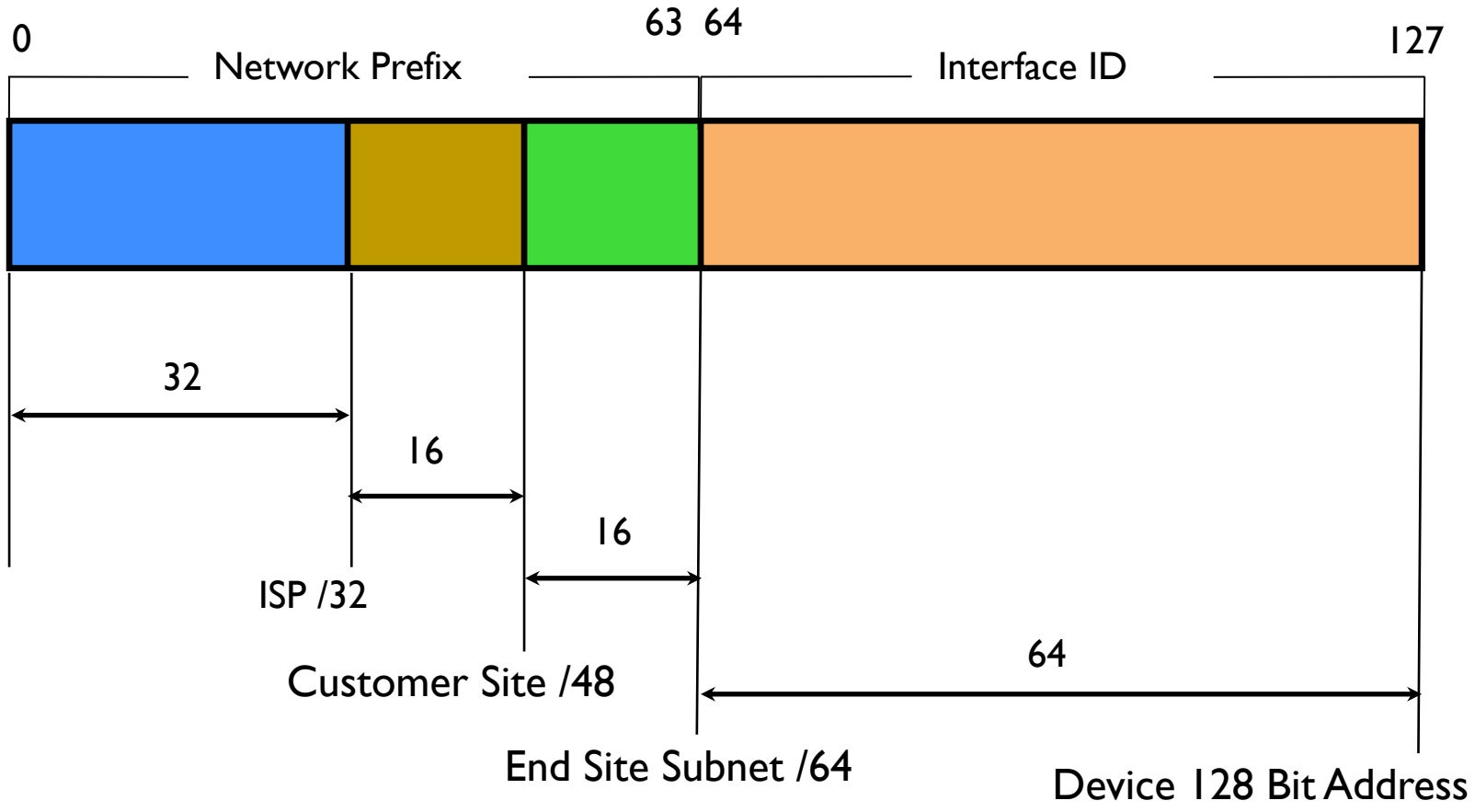
IPv6 Addressing Structure





- Customer (Site) Prefix: assigned to a customer site
  - Group of subnets
  - ISPs/RIRs 'would' assign /48 or /56 to customers
- Subnet ID: identifies the subnets (links) within a site
- Interface ID: host portion of the IPv6 address
  - how many hosts within a subnet

IPv6 Addressing Structure



Link-local Unicast Addresses

- Auto configured address (similar to APIPA)
  - Every IPv6 enabled device must have a link-local address
  - To communicate with other IPv6 devices on the same link
  - FE80::/10
- The link-local address is used by routers as the next-hop address when forwarding IPv6 packets
- All IPv6 hosts on a subnet/link, uses the router's link-local as the default gateway
  - Routers use the link-local as the source in ND-RA messages

Well-known Multicast Addresses

- Multicast addresses can only be destinations and never a source

  `FF00::/8`

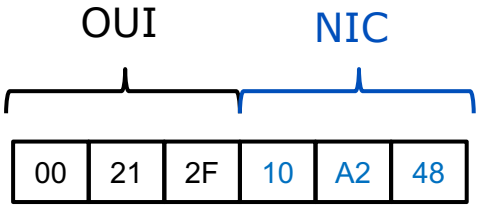
- Pre-defined multicast addresses:
  - `FF02::1` All nodes multicast
    - All IPv6 enabled devices join this multicast group
    - Packets sent to this address is received by all nodes
  - `FF02::2` All routers multicast
    - The moment IPv6 is enabled on a router (`#ipv6 unicast-routing`), the router becomes a member of this group
  - `FF02::1:FFXX:XXXX/104` Solicited Node multicast
    - NS messages (~ARP request) are sent to this address
    - Uses the least significant 24-bits of its unicast/anycast address
    - Must compute and join for every unicast (link-local & global) on a interface

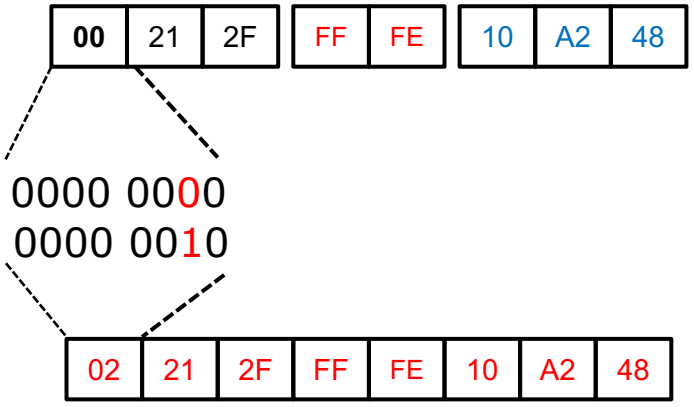
Well-known Multicast Addresses

- Pre-defined multicast addresses:
  - `FF02::1:2` All DHCP Servers/Relay Agents
    - Clients use this multicast address to discover any DHCPv6 servers/relays on the local link (link-scoped)
  - `FF05::1:3` All DHCP servers
    - Generally used by Relays to talk to servers
    - Site-scoped

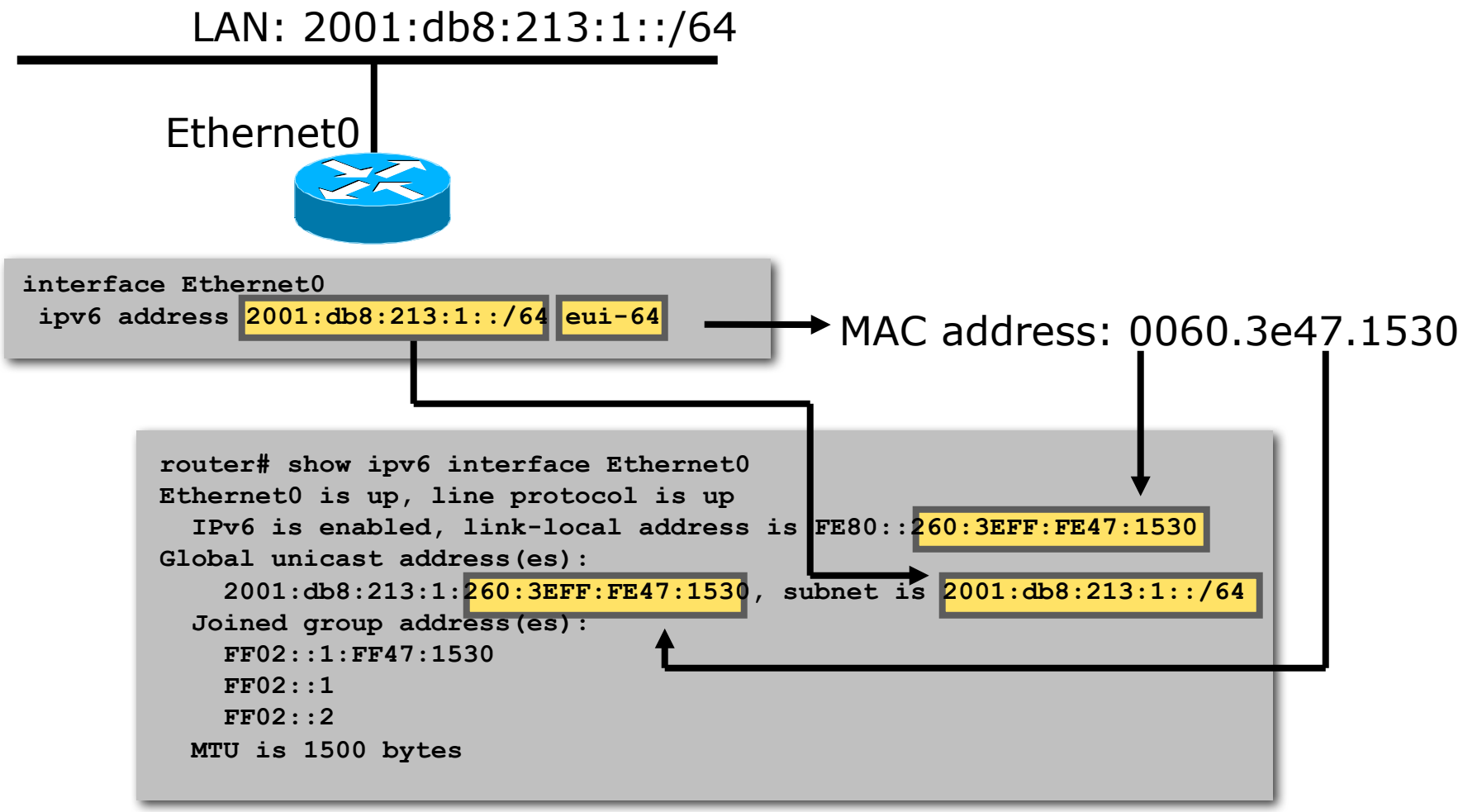
Modified EUI-64 format

- Allows IPv6 device to compute a unique 64 bit Interface ID using the interface MAC address (48 bit)
  - MAC address is split into two 24 bit halves
    - OUI and NIC
  - Then 0xFFFE is inserted between the two halves
    - 0xFFFE is reserved value, not assigned to any OEM
  - Invert 7th bit (U/L) of the OUI to get the EUI-64 address
    - addresses assigned to OEMs have this bit set to 0 to indicate global uniqueness
    - Set to 1 (invert 0) to indicate IEEE identifier is used, or 0 if otherwise.

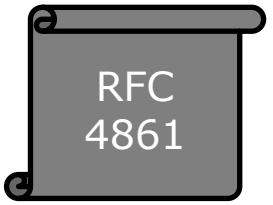




IPv6 Addressing EUI-64



ICMPv6 Neighbor Discovery



- Router Solicitation (RS):
  - sent by IPv6 host to "all routers" multicast to request RA
- Router Advertisement (RA):
  - sent by a IPv6 router to the "all nodes" multicast (200 secs)
  - IPv6 prefix/prefix length, and default gateway
- Neighbor Solicitation (NS):
  - sent by IPv6 host to the "solicited node" multicast to find the MAC address of a given IPv6 address (~ARP request).
- Neighbor Advertisement (NA):
  - sent in response to a NS and informs of its MAC address.
- ICMPv6 Redirect:
  - informs the source of a better next-hop

IPv6 Neighbor Discovery (ND)

- Host A would like to communicate with Host B
- Host A IPv6 address `2406:6400::10`
- Host A IPv6 link-local `fe80::226:bbff:fe06:ff81`
- Host A MAC address `00:26:bb:06:ff:81`
- Host B IPv6 global address `2406:6400::20`
- Host B Link-local **UNKNOWN** (if GW outside the link)
- Host B MAC address **UNKNOWN**
- How will Host A create L2 frame and send to Host B?

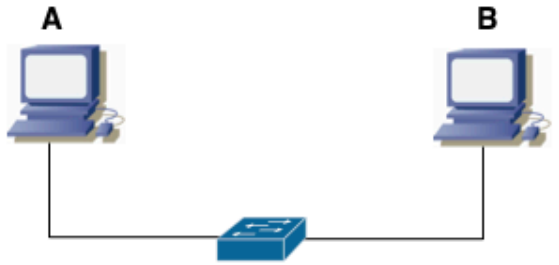
IPv6 Neighbour Discovery (ND)

Host A
IPV6 global address: 2406:6400::0010
IPv6 Link local: fe80::0226:bbff:fe06:ff81
MAC address: 00:26:bb:06:ff:81
Listen to other then above:
FF02::1 [All node multicast]
FF02:0:0:0:0:1:ff00:0010 [Solicited node m.cast unicast]
FF02:0:0:0:0:1:ff06:ff81 [Solicited node m.cast link local]

Packet S: 2406:6400::0010 D:2406:6400::0020

ICMP6 NS Type 135 S: fe80::0226:bbff:fe06:ff81 D:FF02:0:0:0:0:1:ff00:0020

Frame S: 00:26:bb:06:ff:81 D 33:33:ff:00:00:20
Ethernet reserved IPv6 m.cast: 33:33:xx:xx:xx:xx



Multicast enable switch: Unicast by IGMP snooping
Non multicast enable switch: broadcast, PC LAN card filter or discard

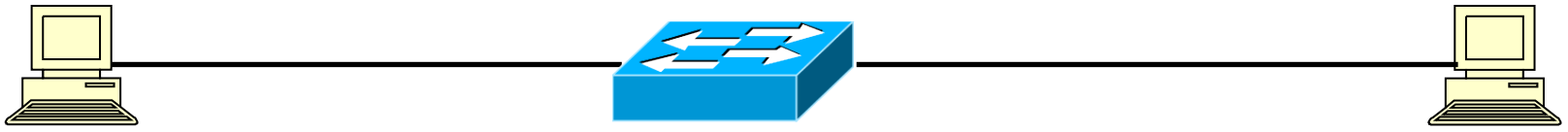
Host B
IPV6 global address: 2406:6400::0020
IPv6 Link local: fe80::0226:bbff:fe06:ff82 [Unknown to A]
MAC address: 00:26:bb:06:ff:82 [Unknown to A]
Listen to other then above:
FF02::1 [All node multicast]
FF02:0:0:0:0:1:ff00:0020 [Solicited node m.cast unicast]
FF02:0:0:0:0:1:ff06:ff82 [Solicited node m.cast link local]

Packet S: 2406:6400::0020 D:2406:6400::0010

ICMP6 NA Type 136 S: fe80::0226:bbff:fe06:ff82 D:fe80::0226:bbff:fe06:ff81

Frame S: 00:26:bb:06:ff:82 D 00:26:bb:06:ff:81

IPv6 Address Resolution



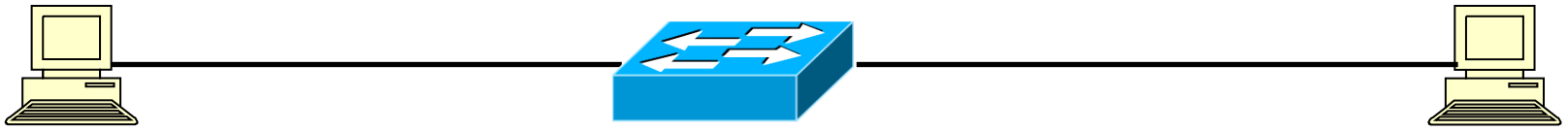
1 ICMPv6 NS Type135 — Multicast

SMAC: 00:26:BB:06:FF:81	**DMAC:** 33:33:FF:00:00:20
Source IPv6: FE80::0226:BBFF:FE06:FF81	**Destination IPv6:** FF02:0:0:0:0:1:FF00:0020
Payload	

2 ICMPv6 NA Type136 — Unicast

SMAC: 00:26:BB:06:FF:82	**DMAC:** 00:26:BB:06:FF:81
Source IPv6: FE80::0226:BBFF:FE06:FF82	**Destination IPv6:** FE80::0226:BBFF:FE06:FF81
Payload	

IPv6 Address Resolution



3

IPv6 Packet

Unicast

SMAC: 00:26:BB:06:FF:81	**DMAC:** 00:26:BB:06:FF:82
Source IPv6: 2406:6400::0010	**Dest IPv6:** 2406:6400::0020
Payload	

Unicast

4

IPv6 Packet

SMAC: 00:26:BB:06:FF:82	**DMAC:** 00:26:BB:06:FF:81
Source IPv6: 2406:6400::0020	**Dest IPv6:** 2406:6400::0010
Payload	

IPv6 Address Auto-configuration

- Stateless address auto-configuration (SLAAC)
  - No manual configuration required
  - Gets the IPv6 prefix and prefix length through RA (local router)
  - EUI-64 for interface ID

- Stateful - DHCPv6
  - To track address assignments

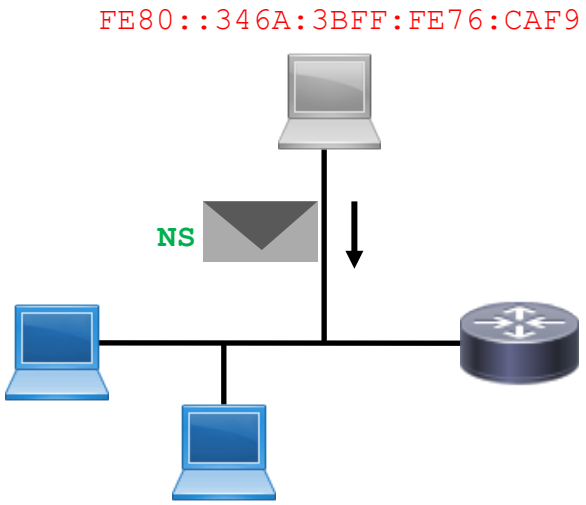
Stateless Address Autoconfig (1)

RFC 2462

When a host joins a link/subnet:

- It auto-generates a link-local using the `FE80::/10` prefix and EUI-64:
  - Ex: `FE80::346A:3BFF:FE76:CAF9`

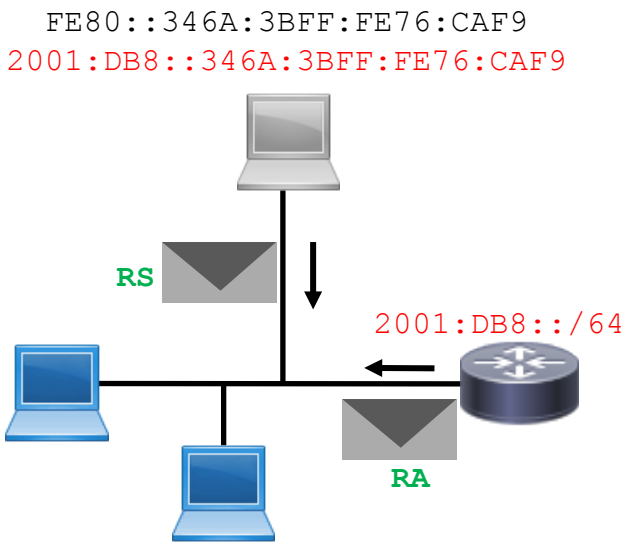
- DAD is performed on the link-local:
  - NS message is sent to the "solicited-node" multicast (`FF02::1:FF76:CAF9`), with `::/128` as the source
  - If no NA message is received back, the generated address can be used
    - would send a NA message to the "all-nodes" multicast (`FF02::1`)



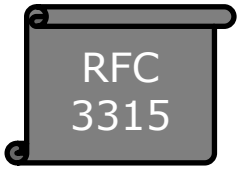
Stateless Address Autoconfig (2)

Once the node has a link-local address:

- sends a RS message to the ”all-routers” multicast (FF02::2)
  - link-local as the source address
- The router responds with a RA message
  - IPv6 prefix and prefix length
  - link-local is the source
  - Managed and Other flags are not set!
- The node generates the IPv6 address
  - uses the received prefix (2001:DB8::/64)
  - Interface ID (EUI-64)
  - 2001:DB8::346A:3BFF:FE76:CAF9
  - DAD not necessary (link-local validated for the same interface!)



Stateful Autoconfig – DHCPv6 (1)



DHCPv6 is used:

- If there are no router(s) on the subnet/link, OR
- If the RA message specifies to get addressing information via DHCPv6

If the router's RA message has the:

- O (other) flag set: *stateless DHCPv6*
  - auto-generate IPv6 address (IPv6 prefix, prefix length in the RA)
  - obtain other information (DNS server, domain) via DHCPv6
- M (managed) flag set:
  - obtain all addressing information via DHCPv6
  - 'O' flag is redundant

Stateful Autoconfig – DHCPv6 (2)

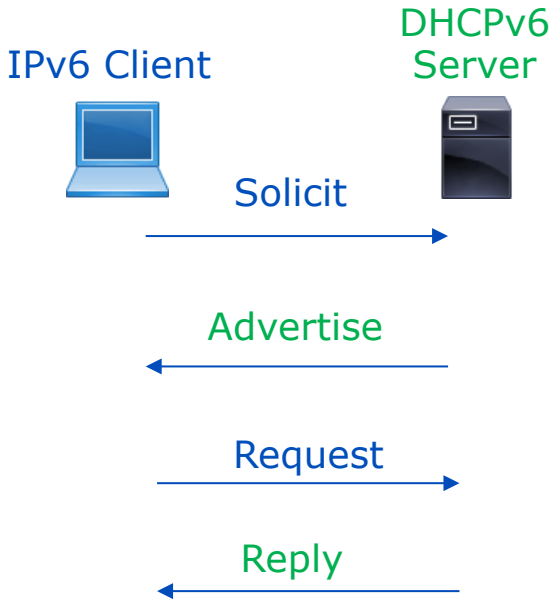
1. Client sends Solicit message to `FF02::1:2` to find any available DHCPv6 servers

2. Server responds with an Advertise message
   - the tentative IPv6 address/prefix
   - Other parameters (DNS, domain, default gateway, lease time)
   - could receive multiple Advertise messages

3. Client selects the server, and sends a Request asking to formally request the indicated IPv6 address

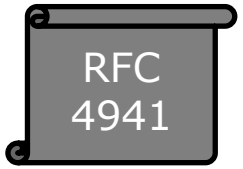
4. Server responds with a Reply to confirm the assignment

5. Performs DAD before using!



IPv6 Interface ID – Privacy

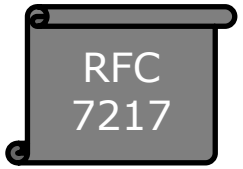
- Overcome the ability to track (interface ID based on MAC address):
  - Temporary address (changes): outgoing connections
  - Secured address: incoming connection




```
Temp > 2001:dc0:a000:4:84a3:49b6:1919:26fb
Secured> 2001:dc0:a000:4:108b:3690:9335:b7ec
Temp > 2001:dc0:a000:4:14e6:d4a3:815d:91dd
```

- Ease network management yet improve privacy:
  - Stable interface identifiers for each subnet

```
Secured> 2001:dc0:a000:4:cbb:347c:6215:1083
```



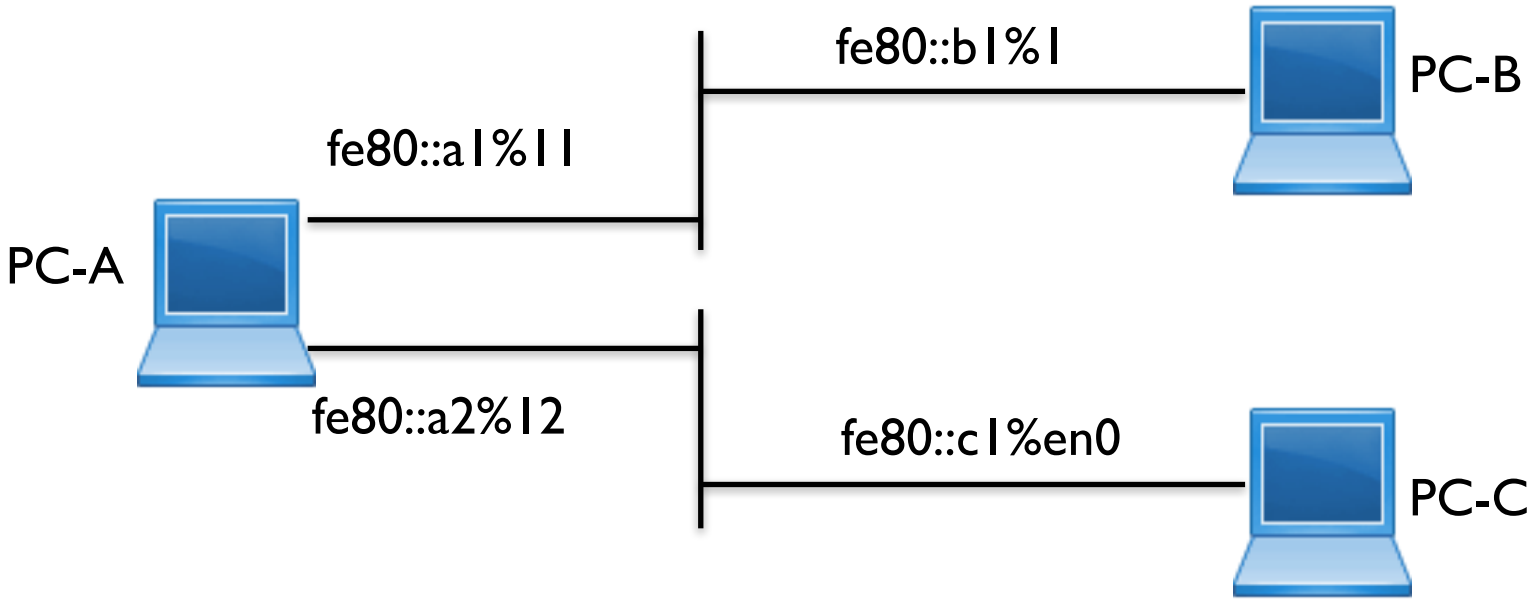
Zone IDs for Link-locals

Interface en0 - `fe80::4e0:37e4:c5d1:c845%en0`
Interface en5 - `fe80::aede:48ff:fe00:1122%en5`

- Zone IDs help uniquely distinguish which link/subnet an interface is connected to

- To ping a remote IPv6 node, use your interface zone ID (so that the response packet has a path)

Quiz of Zone ID



- Please write down the commands:
  - PC-A pings PC-B
  - PC-A telnet PC-C

Questions

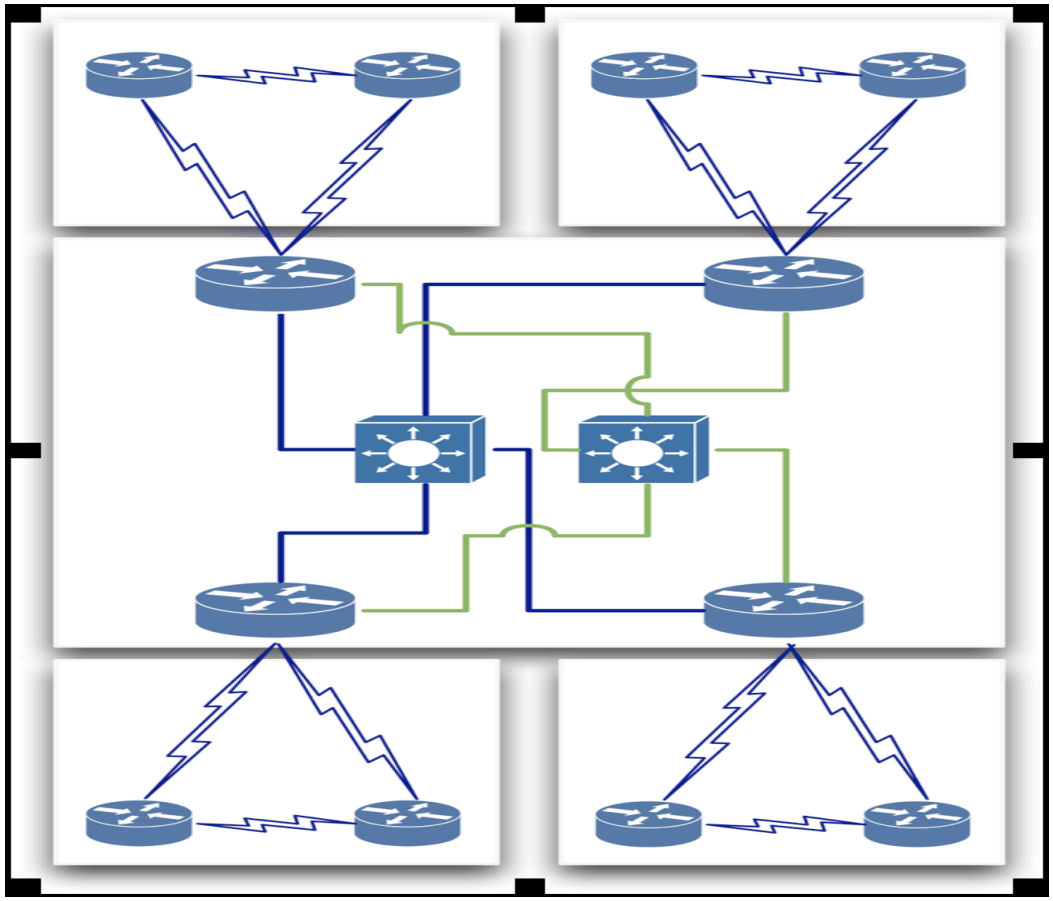
Agenda

- TCP/IP communication and IP Routing
- Internet Routing and Routing Protocols
- IPv6 Address Structure
- **Training Lab Topology**
- OSPF Operation
- Lab exercise: Basic router and OSPF
- BGP – operation and attributes
- BGP Scaling Techniques
- Lab: iBGP, eBGP, Peer-group, RR

Training ISP Network Topology

- Scenario:
  - Training ISP has 4 main operational regions
  - Each region has 2 small POPs
  - Each region has a datacenter to host content
  - Regional networks are inter-connected with redundant links

Training ISP Network Topology



Training ISP Topology Diagram

Training ISP Network Topology

- Regional Network:
  - Each regional network has 3 routers
  - 1 Core & 2 Edge
  - 2 Point-of-Presence (POPs) at each region
  - Each POP (Edge router) aggregates customers

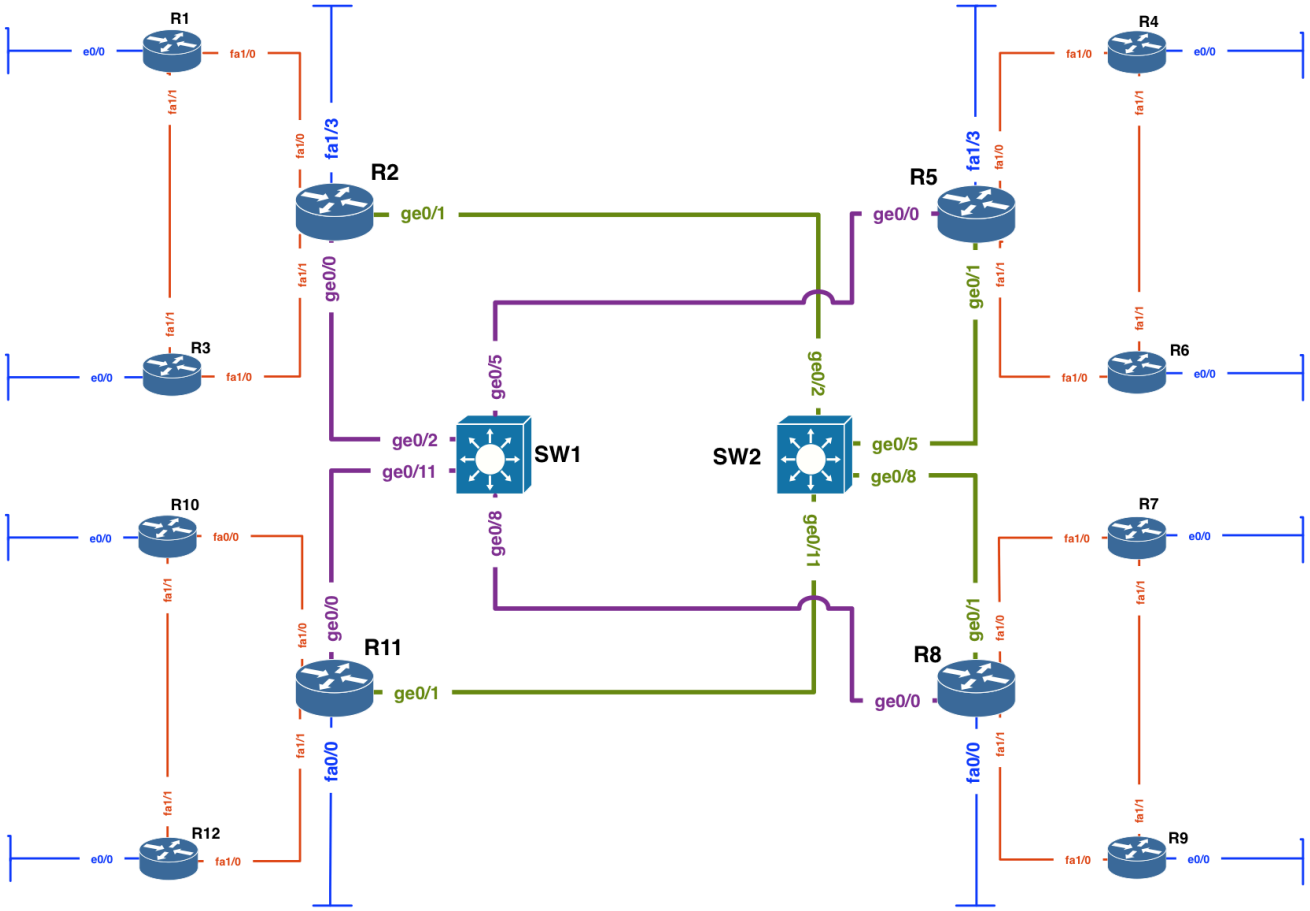
Training ISP Network Topology

- Access Network:
  - Connects customer network to Edge router
  - Usually 10 to 100 Mbps link
  - Separate routing policy from most of ISP
  - connects with dedicated (customer) IP prefix

Training ISP Network Topology

- Transport Link:
  - Interconnects regional core routers
  - High speed links (GE/10GE)
  - Training ISP has 2 transport links for redundancy
    - » Purple & Green links are connected to two carrier 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 consider scalability for next couple of years for both customer and infrastructure blocks
  - Summarization not necessary for infrastructure WAN and loopback address

Training ISP Network Topology

- Design Consideration:
  - All WAN links should be ICMP reachable for monitoring purpose (at least from trusted hosts/subnet)
  - Conservation considered for IPv4 address planning and aggregation for IPv6 address planning.

Training ISP Network Topology

- Design Consideration:
  - OSPF (v2 and v3) is the IGP in the network (carries) infrastructure IP prefix
  - Each region is a separate OSPF area
  - Transport core is in Area 0
  - Customer connects through static or eBGP (not OSPF)
  - iBGP will carry external prefixes within ISP core as well as customer prefixes

Training ISP IPV6 Addressing Plan

- IPv6 address plan consideration:
  - Big IPv6 address space can result in very very large routing table size
  - Most upstream apply IPv6 prefix filters – aggregation
    - » anything smaller than /48 might be dropped

IPv6 Address Plan: ISP Infra

- ISPs receive /32 from their RIR
- Dedicate a /40 (or /48) for the backbone infra
  - Every infrastructure assignment from this block!
  - Carried by IGP (not iBGP)
- Loopbacks
  - Generally one /48 (/60 and /64 also common) for all loopbacks
  - /128 as loopback
- Point-to-Point links
  - Dedicate a /48 for all PtP links
  - Assign /64 per link (RFC3177); RFC6164 recommends /127
    - Reserve /64 per link but use /127

IPv6 Address Plan: Customer

- Consider regional delegation
  - Aggregation in mind!
  - /40 per region?
- One /48 per customer
  - Could be transit customers or leased line customers
  - Could be given additional /48s as they grow
- Common to see ISPs give:
  - /56 to mid-sized customers
  - /64 or /60 for very small customers
  - Please share your experience

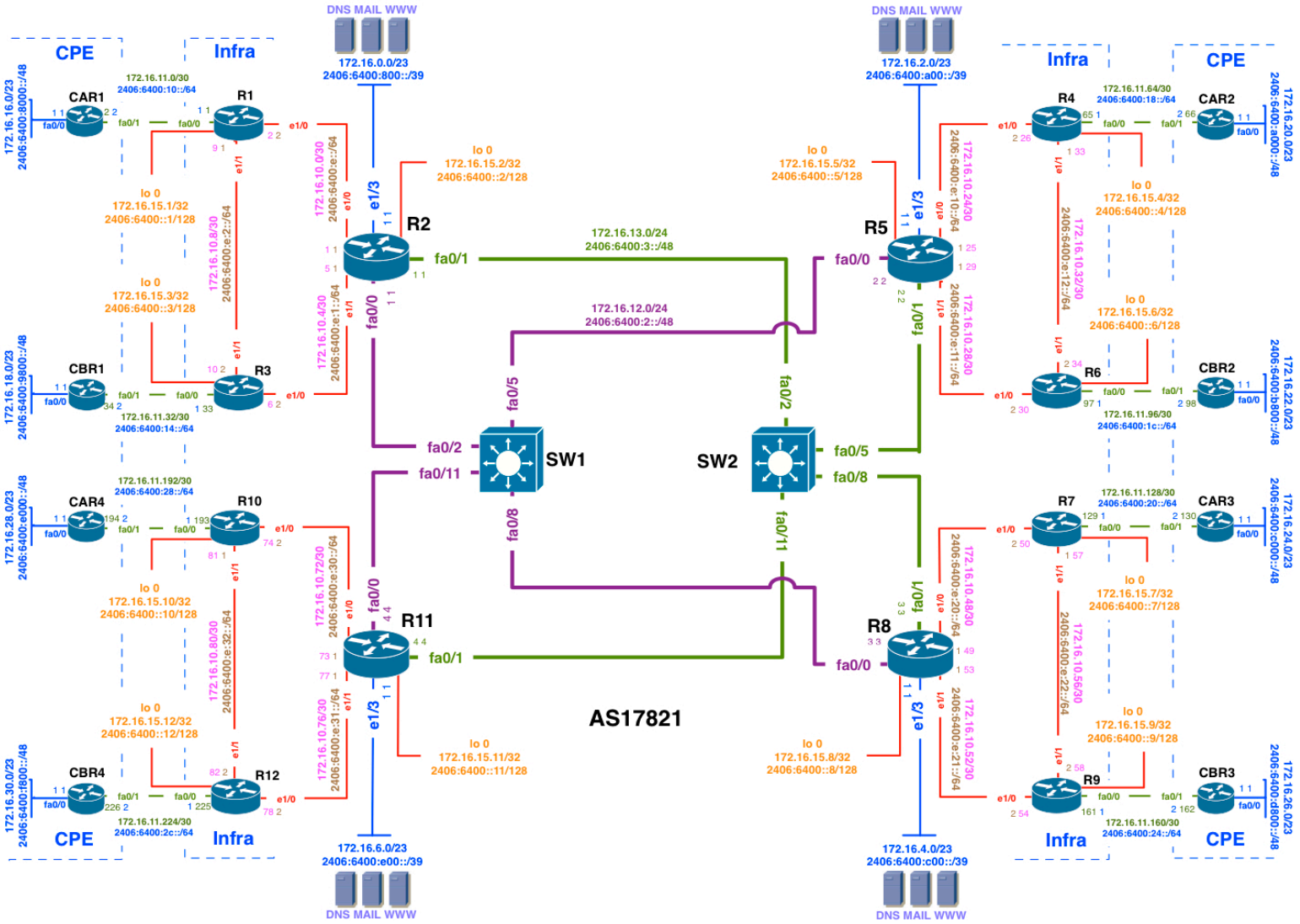
IPv6 Address Plan: Customer WAN links

- Either use from their own /48 block
  - /64 from their block
- Dedicate a /48 block for customer WAN links
  - Helps to monitor customer links
  - Not to be mistaken with the trusted infra PtP block!
  - Actual addressing still the same:
    - Reserve /64 and use /127
- Carried in iBGP (not IGP)
  - Aggregated at the GW router or POP routers

Addressing Planning

- RIRs will usually allocate the next block to be contiguous with the first allocation
  - Minimum allocation is /32
  - Very likely that subsequent allocation will make it /31
    - consider in your plan!

Training ISP IPV6 Addressing Plan



Example Address Plan

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

Training ISP IPV6 Addressing Plan

-

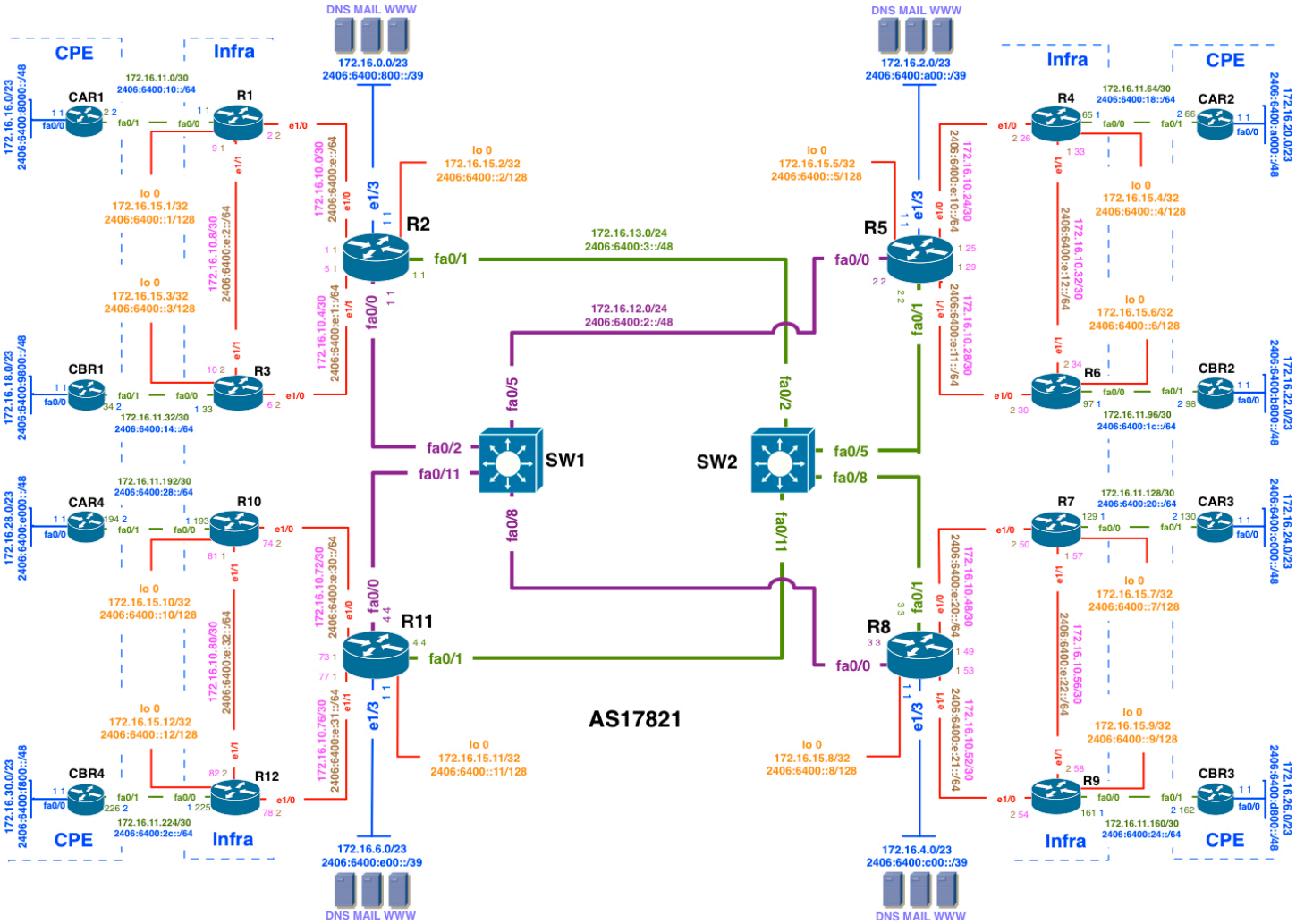
Table 1: Top level distribution infrastructure & customer					
Block#	**Prefix**	**Description**	**Reverse Domain**	**SOR**	**Registration**
1	**2406:6400::/32**	***Parent Block***	0.0.4.6.6.0.4.2.ip6.arpa.	N/A	APNIC
2	2406:6400:0000:0000::/36	Infrastructure	0.0.0.4.6.6.0.4.2.ip6.arpa.	No	Optional
	2406:6400:1000:0000::/36				
	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

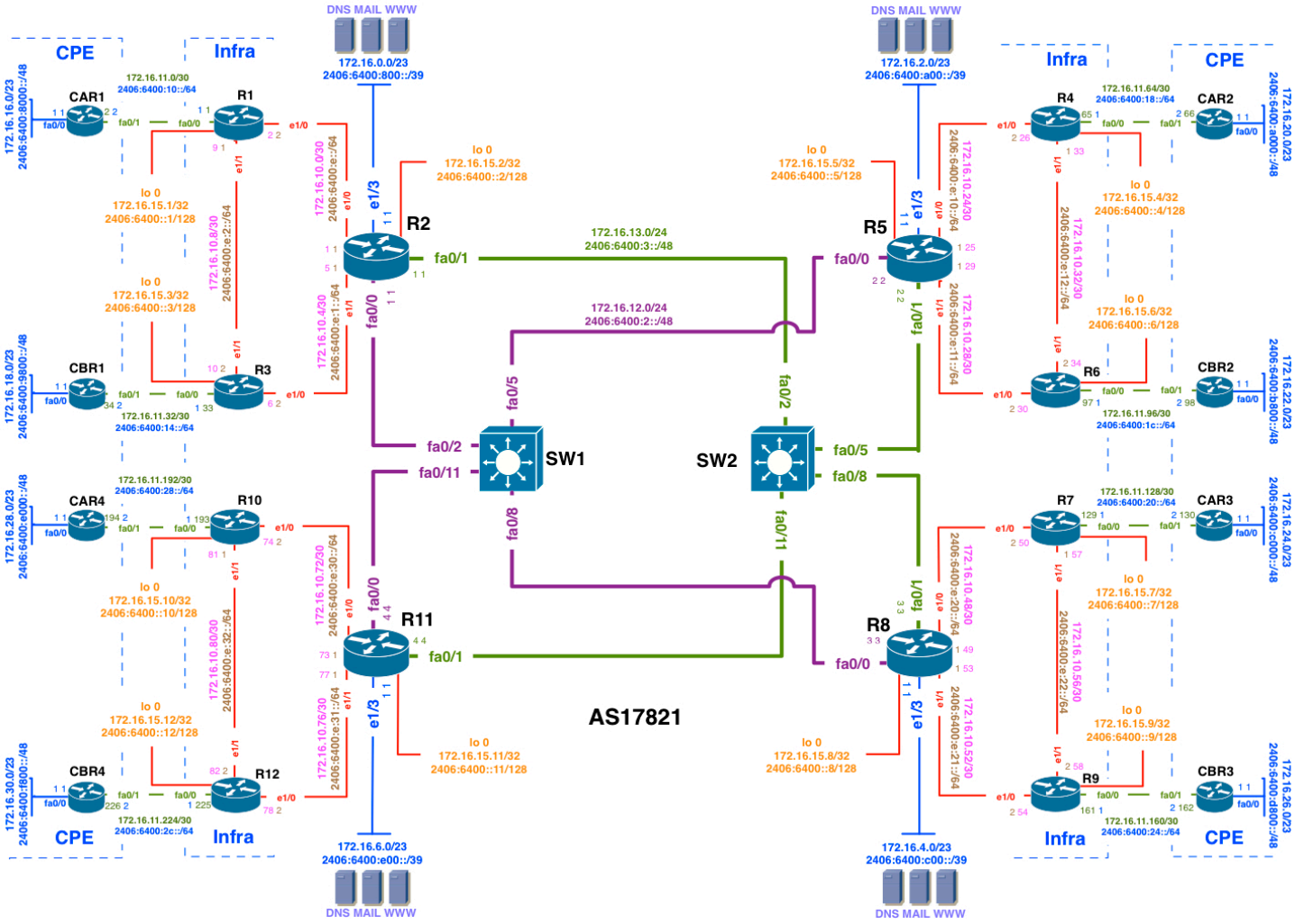
Block#	Prefix	Description	Reverse Domain	SOR	Registration
2	***2406:6400:0000:0000::/36***	***Infrastructure***	0.0.0.4.6.6.0.4.2.ip6.arpa.	No	Optional
11	2406:6400:0000:0000::/40	Loopback, Transport & WAN [Infra+CS]	0.0.0.0.4.6.6.0.4.2.ip6.arpa.	No	Optional
	2406:6400:0100:0000::/40				
	2406:6400:0200:0000::/40				
	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	Recommeded
	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	Recommeded
	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	Recommeded
	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	Recommeded
	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



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	Recommeded
	2406:6400:0001:0000::/48				
21	2406:6400:0002:0000::/48	Purple Transport		No	Recommeded
22	2406:6400:0003:0000::/48	Green Transport		No	Recommeded
	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	Recommeded
	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	Recommeded
	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	Recommeded
	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	Recommeded
	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	Recommeded
	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	Recommeded
	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	Recommeded
	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	Recommeded
	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	Recommeded
	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

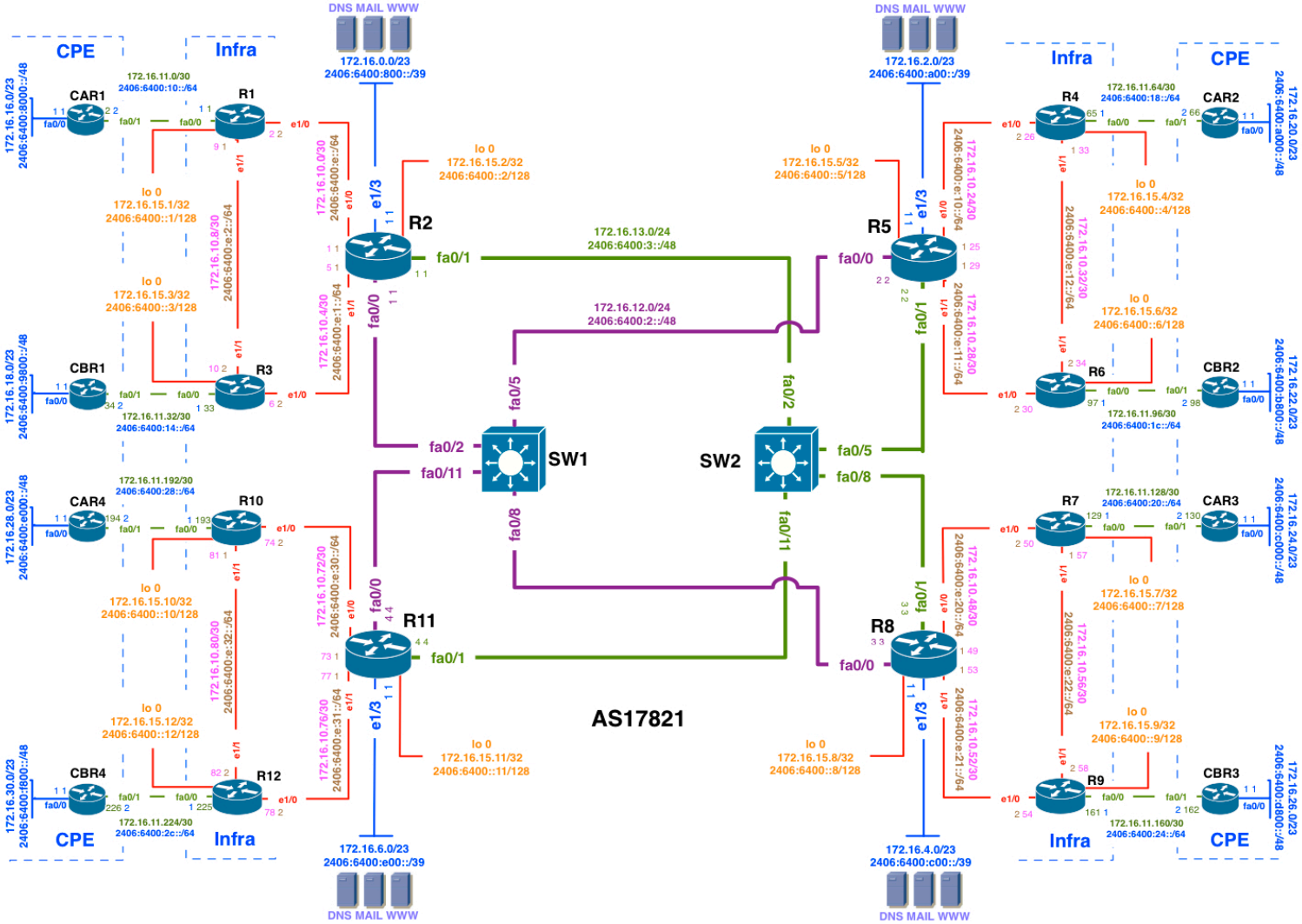
Block#	Prefix	Description	Reverse Domain
24	2406:6400:0010:0000::/45	WAN CS Link Region1 Summary [R2]	
25	2406:6400:0010:0000::/46	WAN CS Link Region1 POP1 Summary [R1]	
26	2406:6400:0014:0000::/46	WAN CS Link Region1 POP2 Summary [R3]	

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	Recommeded
			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 10: Further detail Infra WAN

Block#	Prefix	Description	PTR Record	SOR	Registration
23	2406:6400:000E:0000::/48	WAN Prefix Infra Link		No	Recommeded
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	Recommeded
	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	Recommeded
	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	Recommeded
	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	Recommeded
	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	Recommeded
	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	Recommeded
	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	Recommeded
	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	Recommeded
	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				

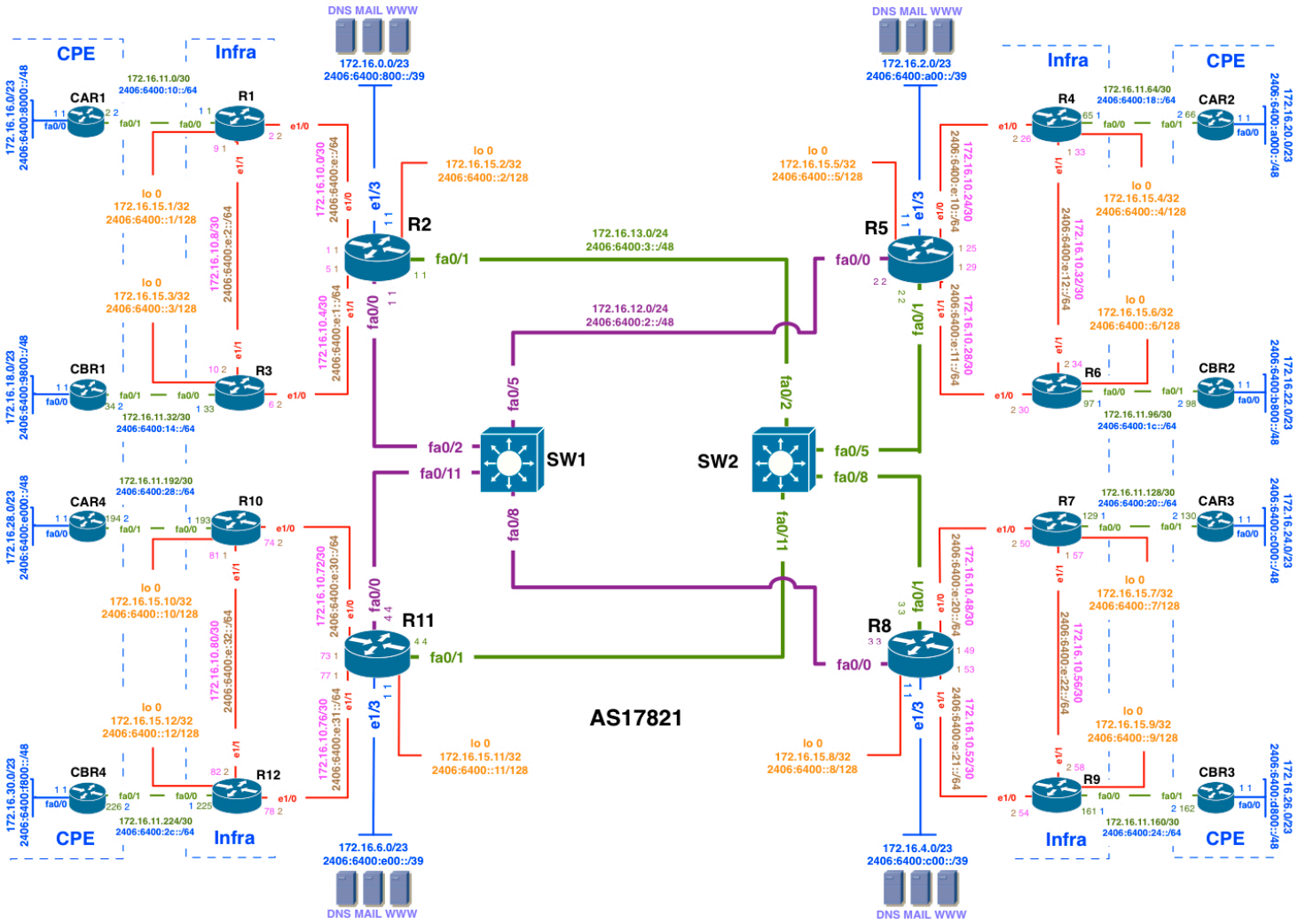
Training ISP IPV6 Addressing Plan

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				

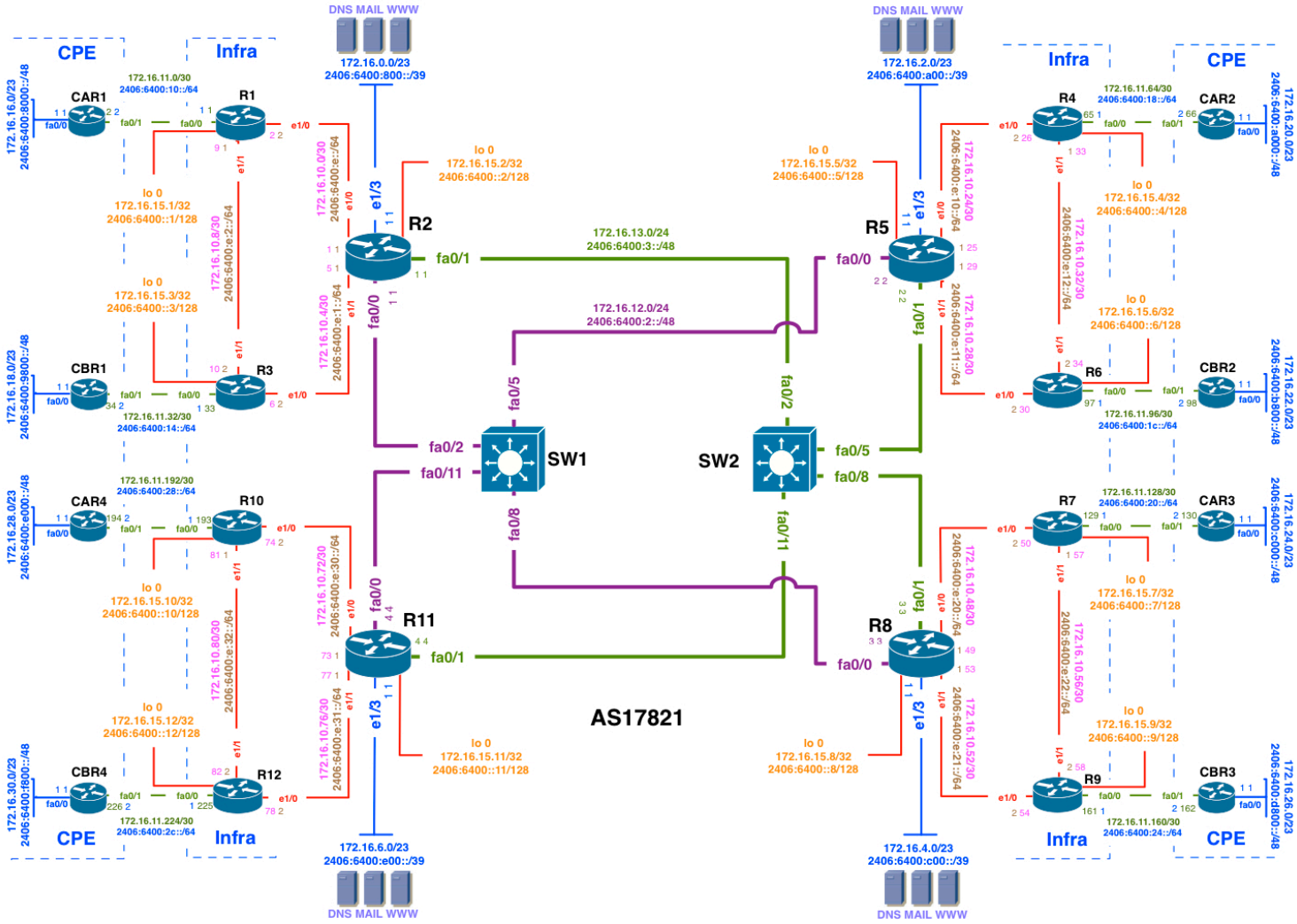
Training ISP IPV6 Addressing Plan

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				

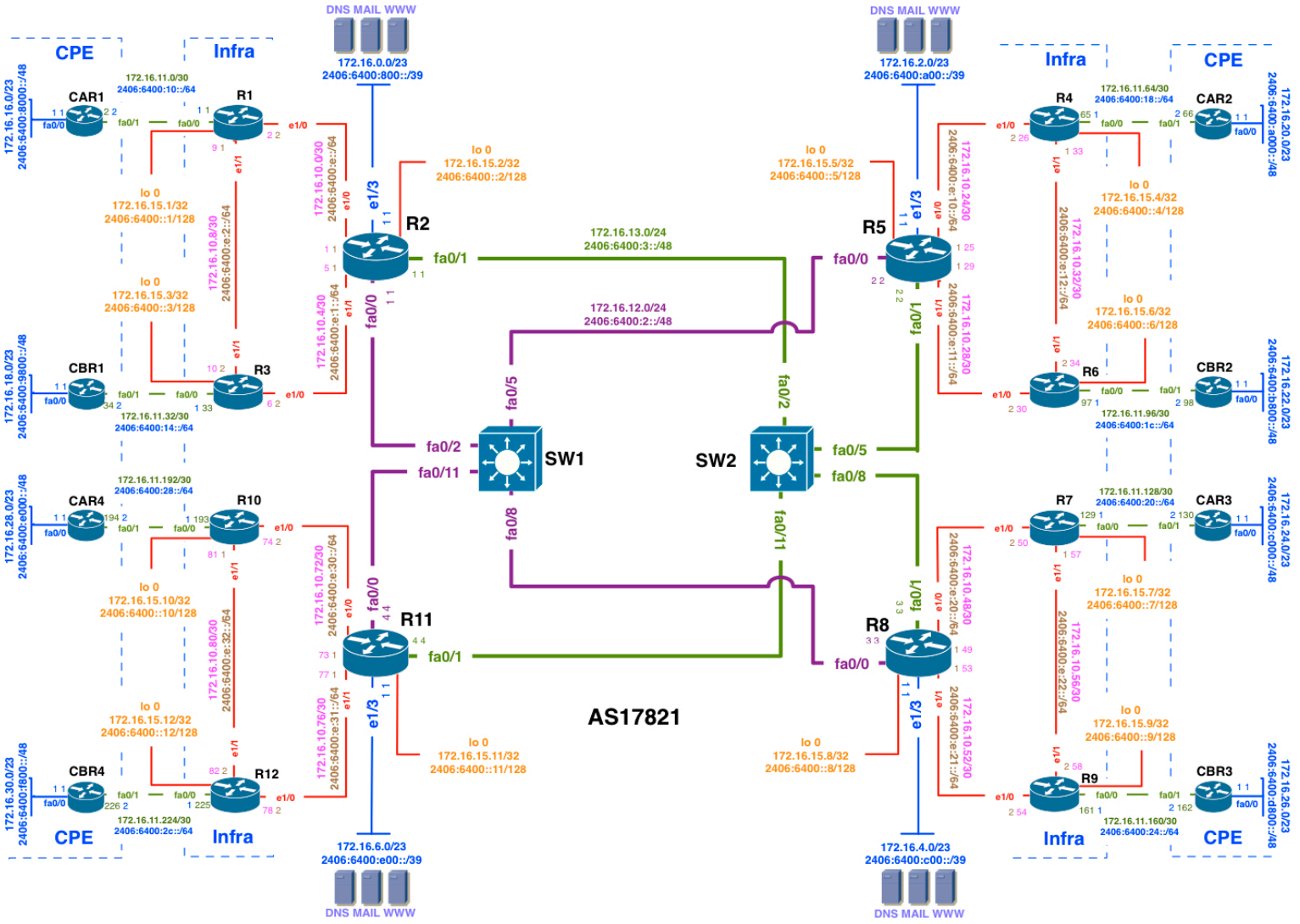
Training ISP IPV6 Addressing Plan

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

Table 23: Detail customer block Region 3					
Block#	**Prefix**	**Description**	**Reverse DNS**	**SOR**	**Registration**
	2406:6400:C000:0000::/40	1st Customer block POP1 [R7]			
	2406:6400:C000:0000::/48	1st Customer prefix POP1 [R7]		Yes	Yes
	2406:6400:C001:0000::/48				
	2406:6400:C002:0000::/48				
	2406:6400:C003:0000::/48				
	2406:6400:C004:0000::/48				
	2406:6400:C005:0000::/48				
	2406:6400:C006:0000::/48				
	2406:6400:C007:0000::/48				
	2406:6400:D800:0000::/40	1st Customer block POP2 [R9]			
	2406:6400:D800:0000::/48	1st Customer prefix POP2 [R9]		Yes	Yes
	2406:6400:D801:0000::/48				
	2406:6400:D802:0000::/48				
	2406:6400:D803:0000::/48				
	2406:6400:D804:0000::/48				
	2406:6400:D805:0000::/48				
	2406:6400:D806:0000::/48				
	2406:6400:D807:0000::/48				

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				

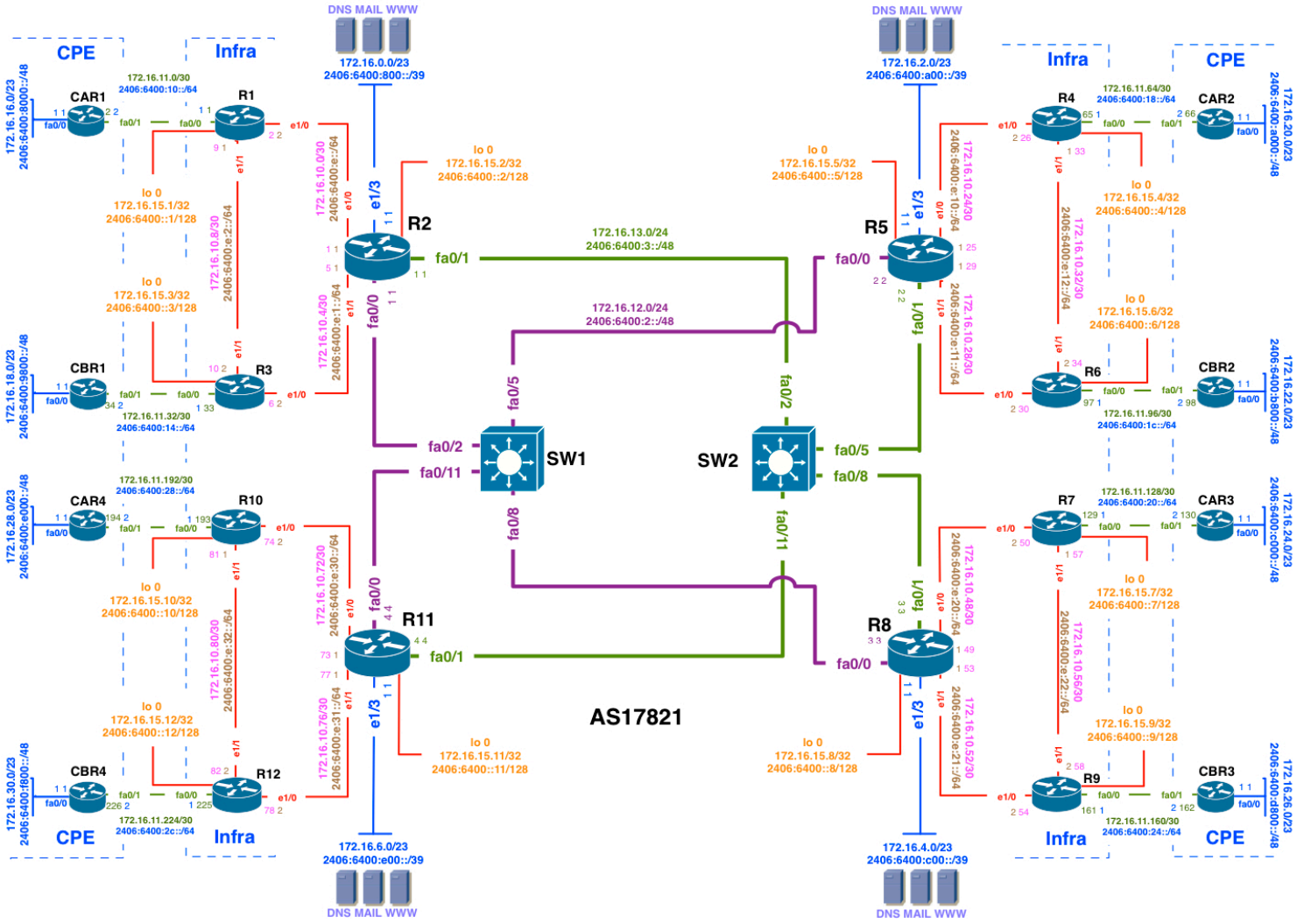
Training ISP IPV6 Addressing Plan

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

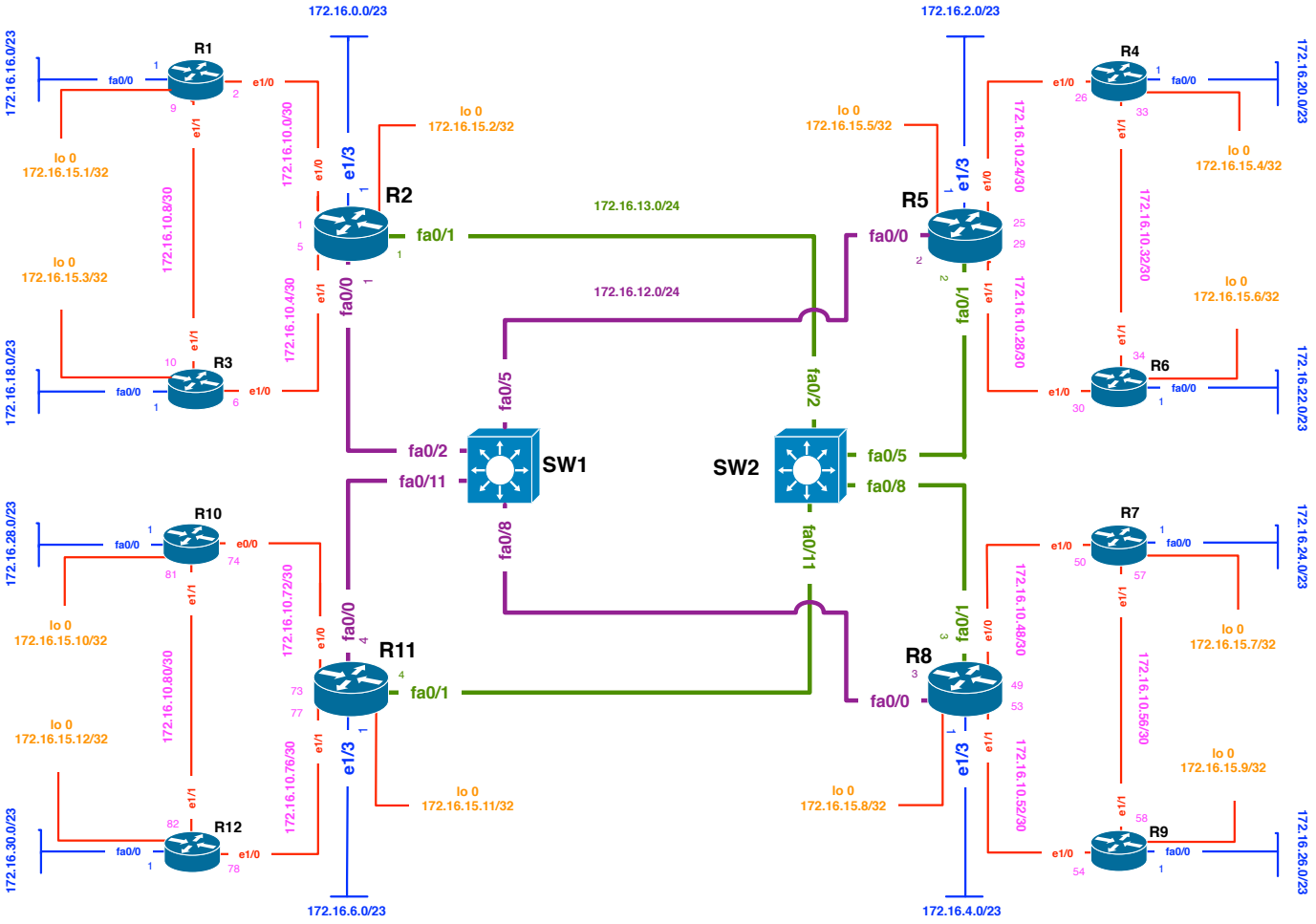
25	172.16.12.0	/24	Transport link PURPLE	No	
26	172.16.13.0	/24	Transport link GREEN	No	
27	172.16.15.0	/24	Loopback	No	

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

Agenda

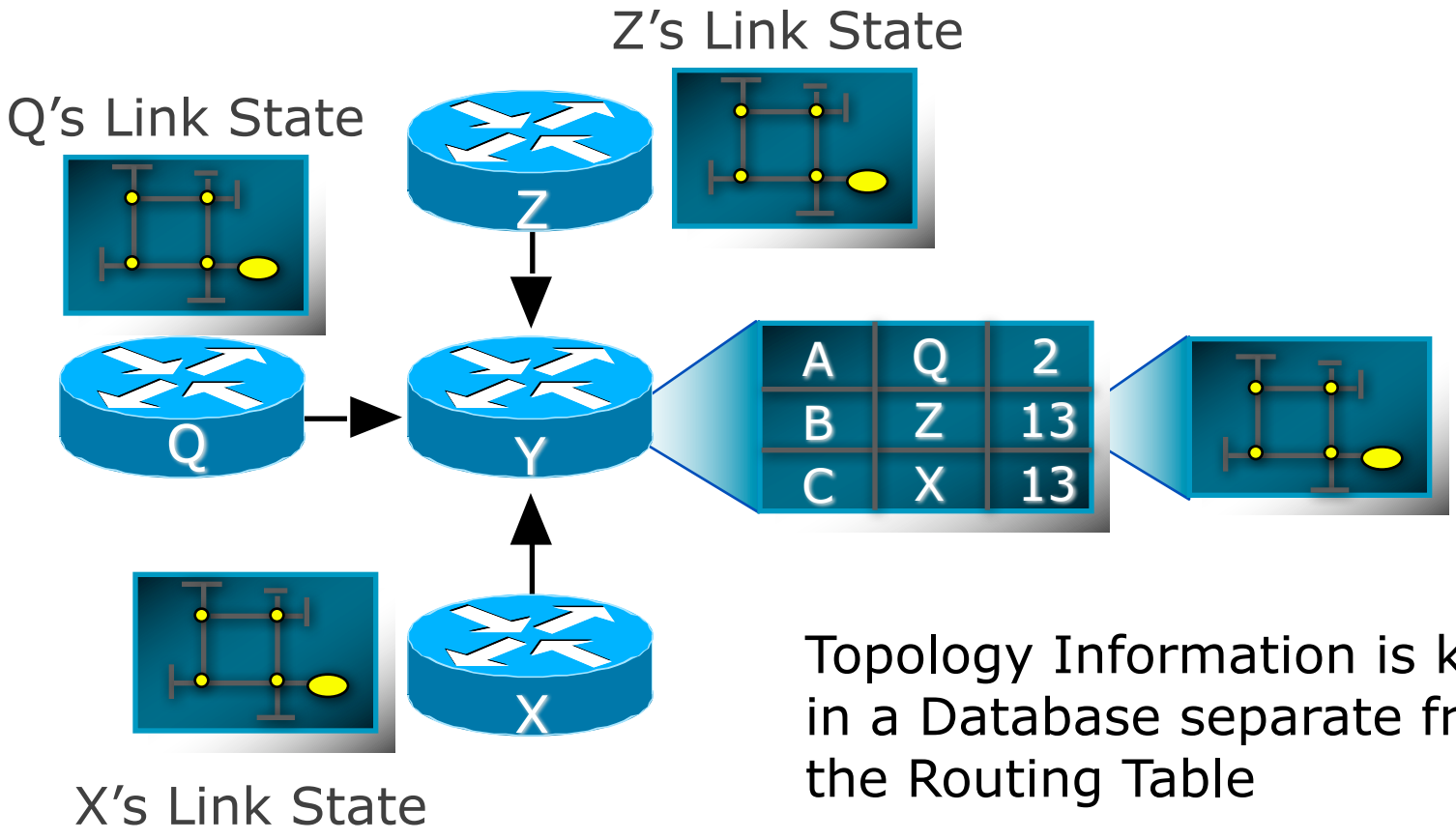
- TCP/IP communication and IP Routing
- Internet Routing and Routing Protocols
- IPv6 Address Structure
- Training Lab Topology
- **OSPF Operation**
- Lab exercise: Basic router and OSPF
- BGP – operation and attributes
- BGP Scaling Techniques
- Lab: iBGP, eBGP, Peer-group, RR

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



Topology Information is kept in a Database separate from the Routing Table

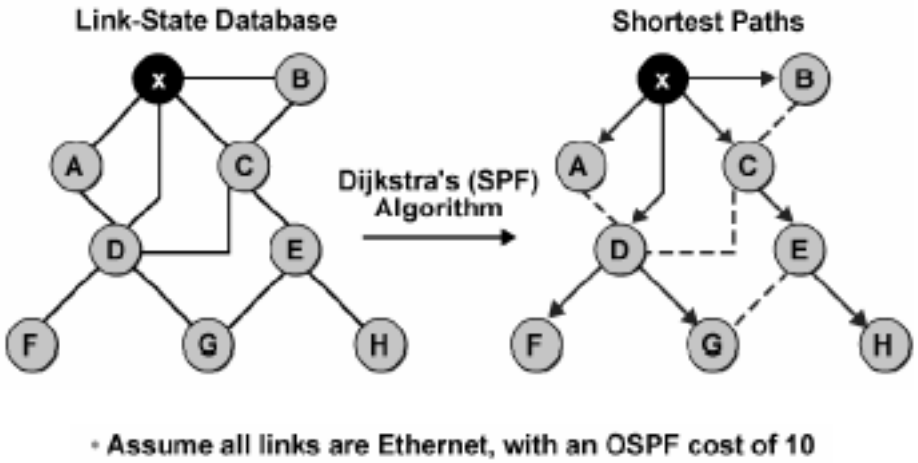
What is Link State Routing

- 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

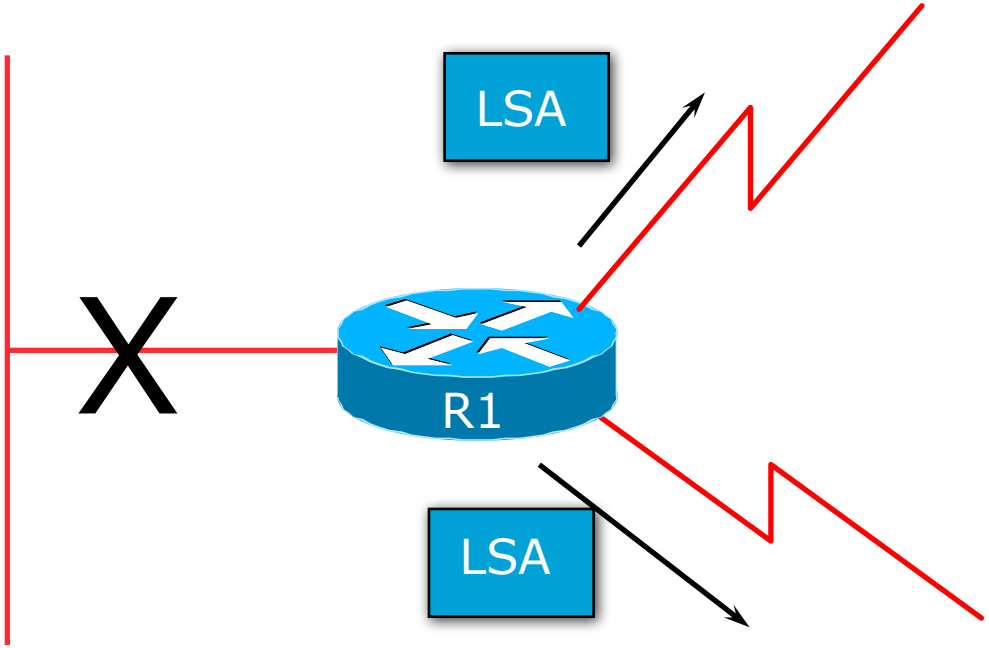
- Neighbor Table
  - List of all recognized neighboring router to whom routing information will be interchanged
- Topology Table
  - Also called LSDB which maintain list of routers and their link information i.e network destination, prefix length, link cost etc
- Routing table
  - Also called forwarding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

Shortest Path First (SPF) Tree



- 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

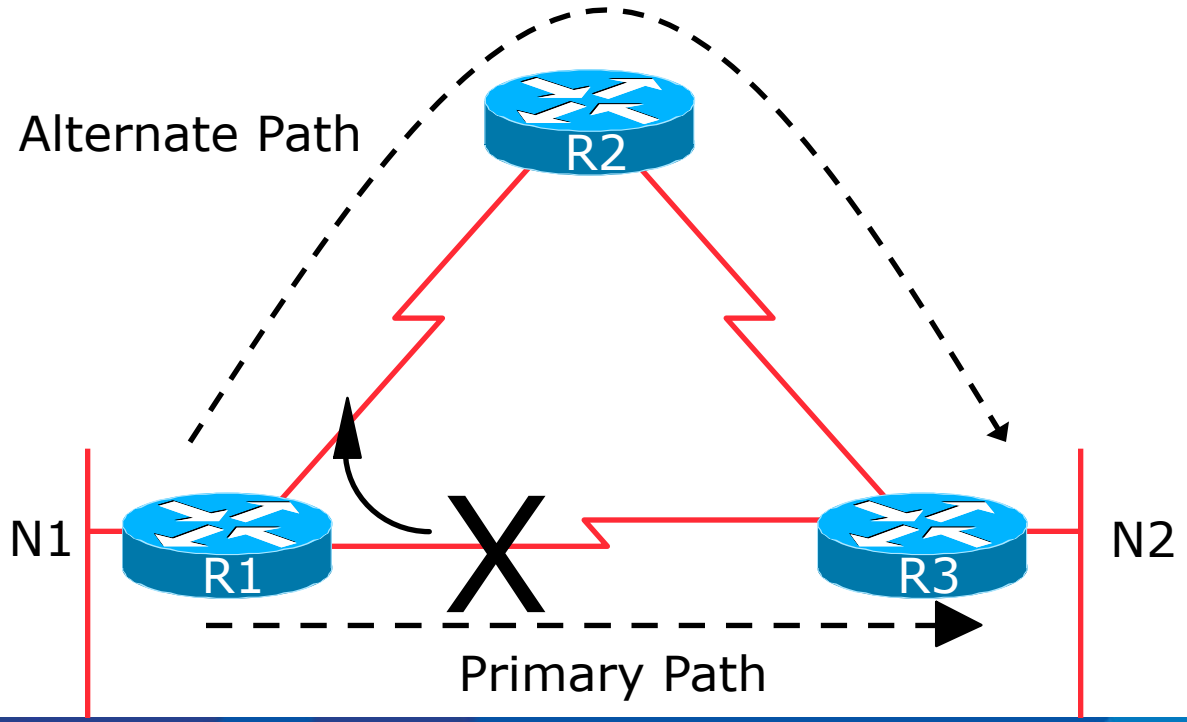
Low Bandwidth Utilisation



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

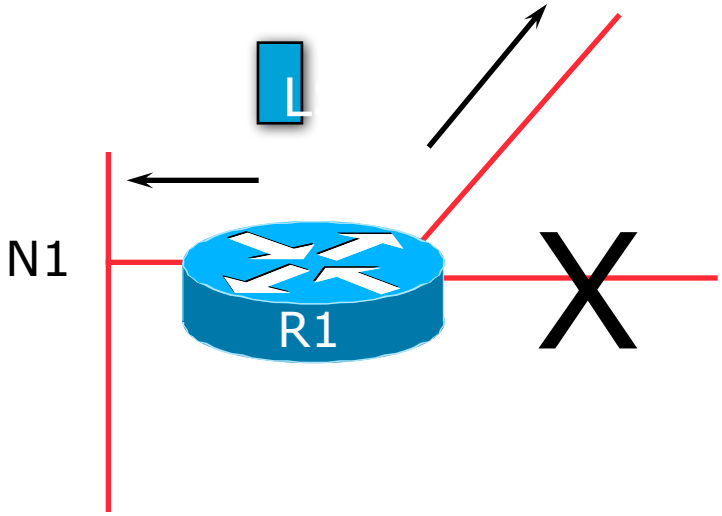
Fast Convergence

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



Fast Convergence

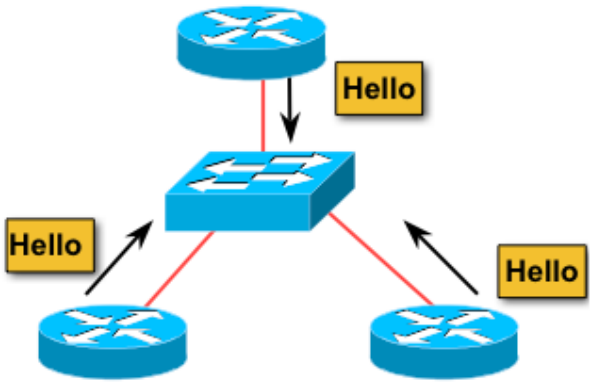
- Finding a new route
  - LSA flooded throughout area
  - Acknowledgement based
  - Topology database synchronised
  - Each router derives routing table to destination network



Basic OSPF Operation

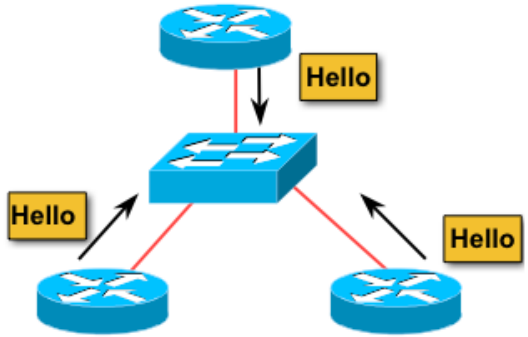
- Neighbor discovery
  - Send L3 multicast message (hello) to discover neighbors
- Exchanging topology table (LSDB)
  - Send L3 multicast message (DBD packets)
- Use SPF algorithm to select best path
  - Each router independently calculates best path from an identical topology database of an OSPF network 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**

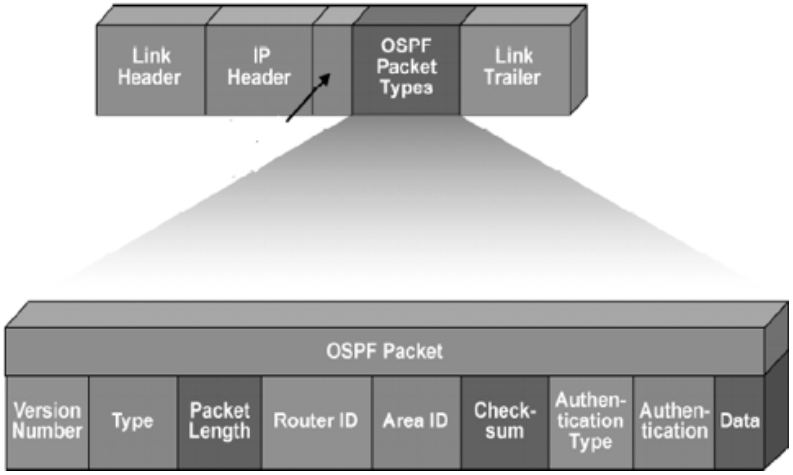
Maintaining Routing Information

- 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)

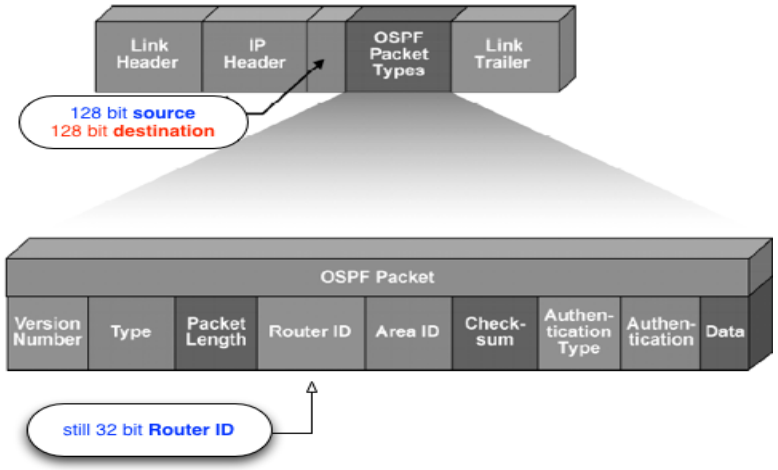
Format of OSPF Packet Header Field

- 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



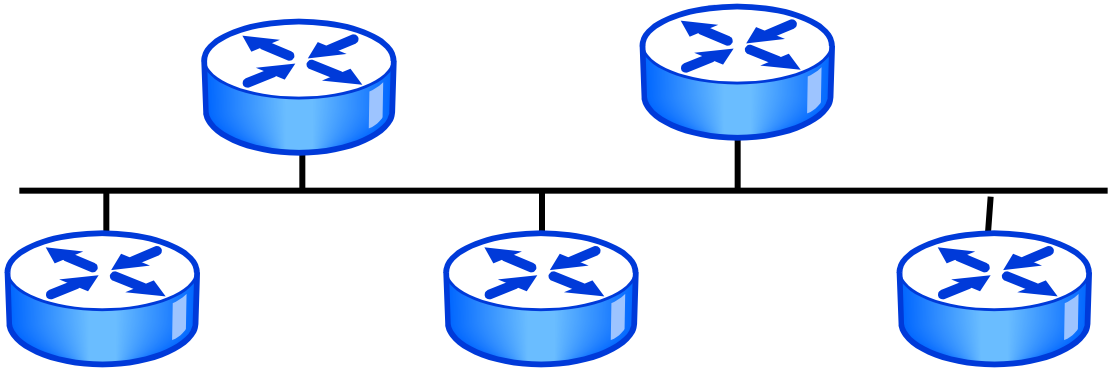
- OSPFv3 still use 32 bit number as router ID
- So OSPFv3 operation and packet types are same as OSPFv2
- Change will be in IP header where source address will be interface address and destination will be FF02::5 which is 128 bit address.
- Change will be in DBD [t2] and LSU packet [t4] to carry 128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

OSPF Network Topology

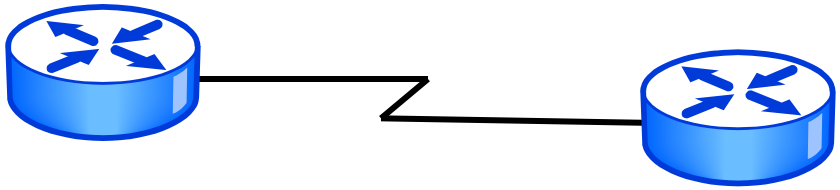
- OSPF network can made up of different types of network links
- Neighbor relationship behavior will also be different for each network type
- It is important for OSPF to be configured correctly based on its network types to be functioned properly
- Some network type create neighbor relationship automatically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

OSPF Network Topology

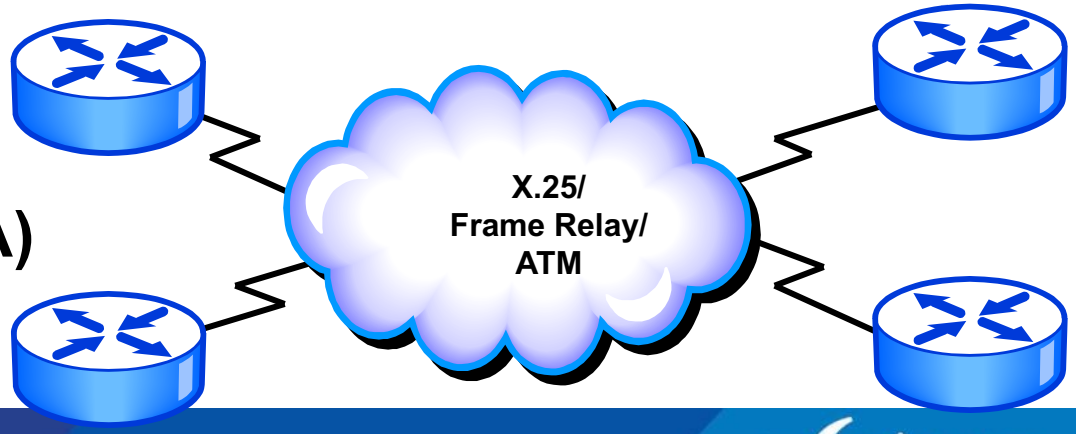
Broadcast Multi-access



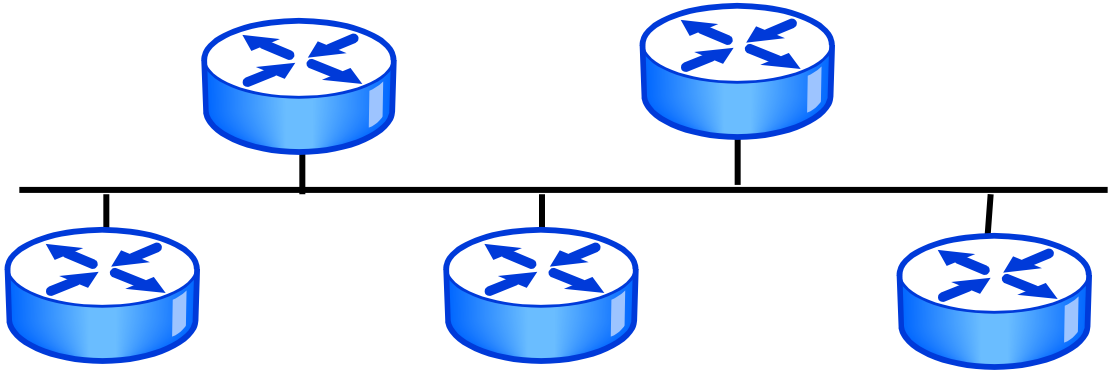
Point-to-Point



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

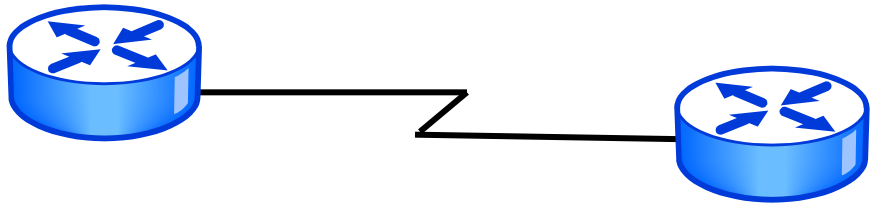
Broadcast Multi-access Network

- 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

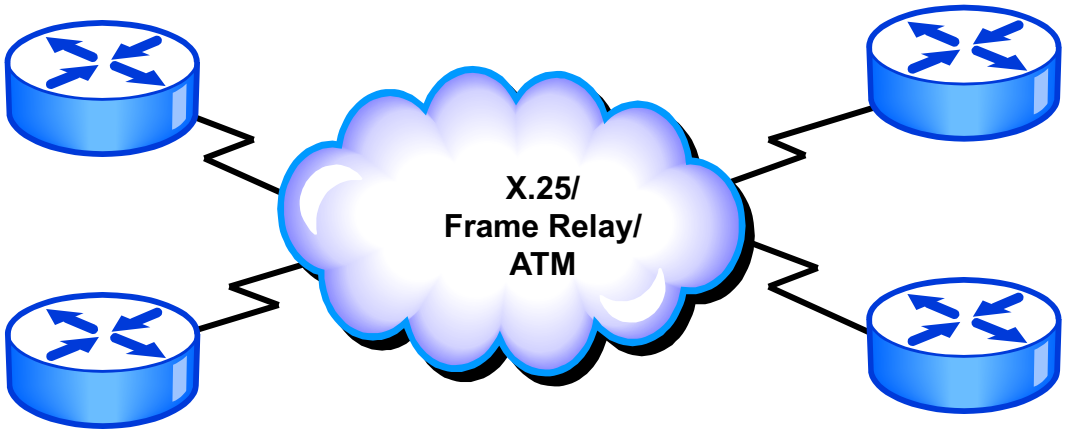
- Router with the highest priority value is the DR, Second highest is BDR
- In the event of tie router with the highest IP address on an interface become DR and second highest is BDR
- DR/BDR election can be manipulated by using router-ID command.
- In practice loopback IP address is used as router ID 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

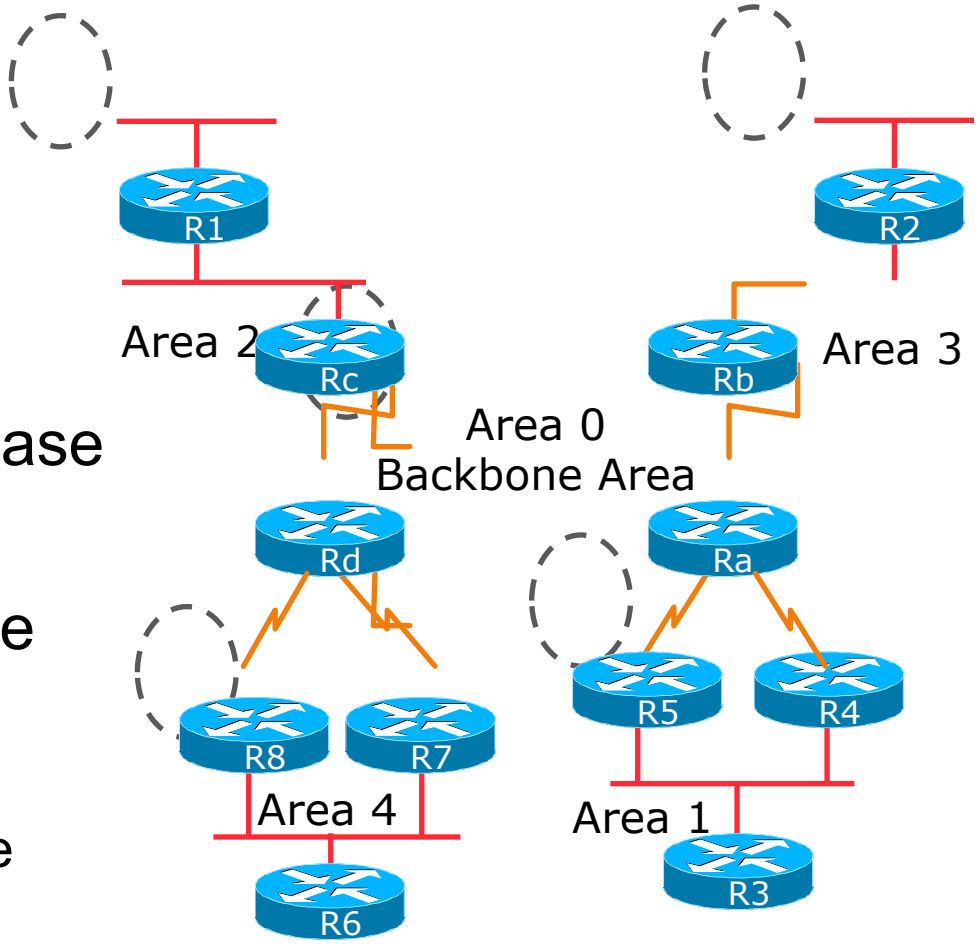
Non Broadcast Multi-access Network



- 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

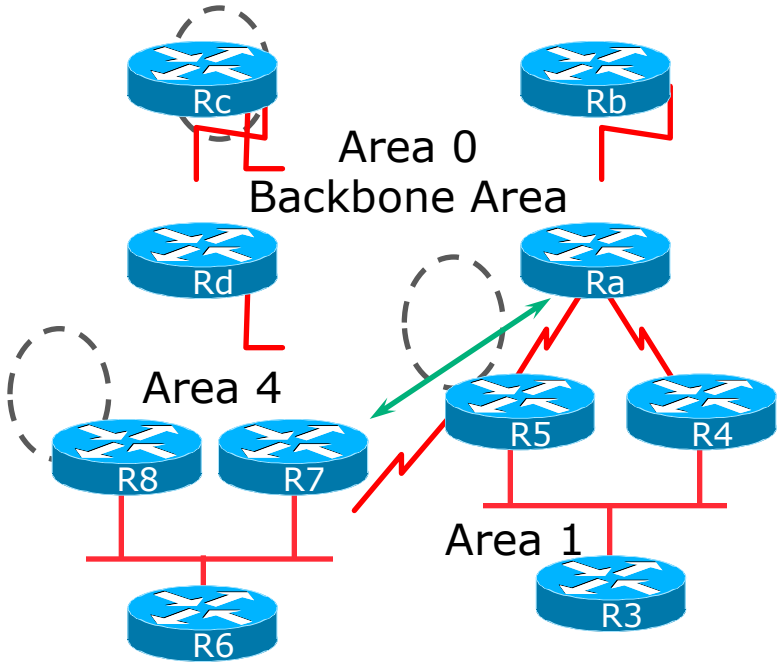
OSPF Areas

- 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

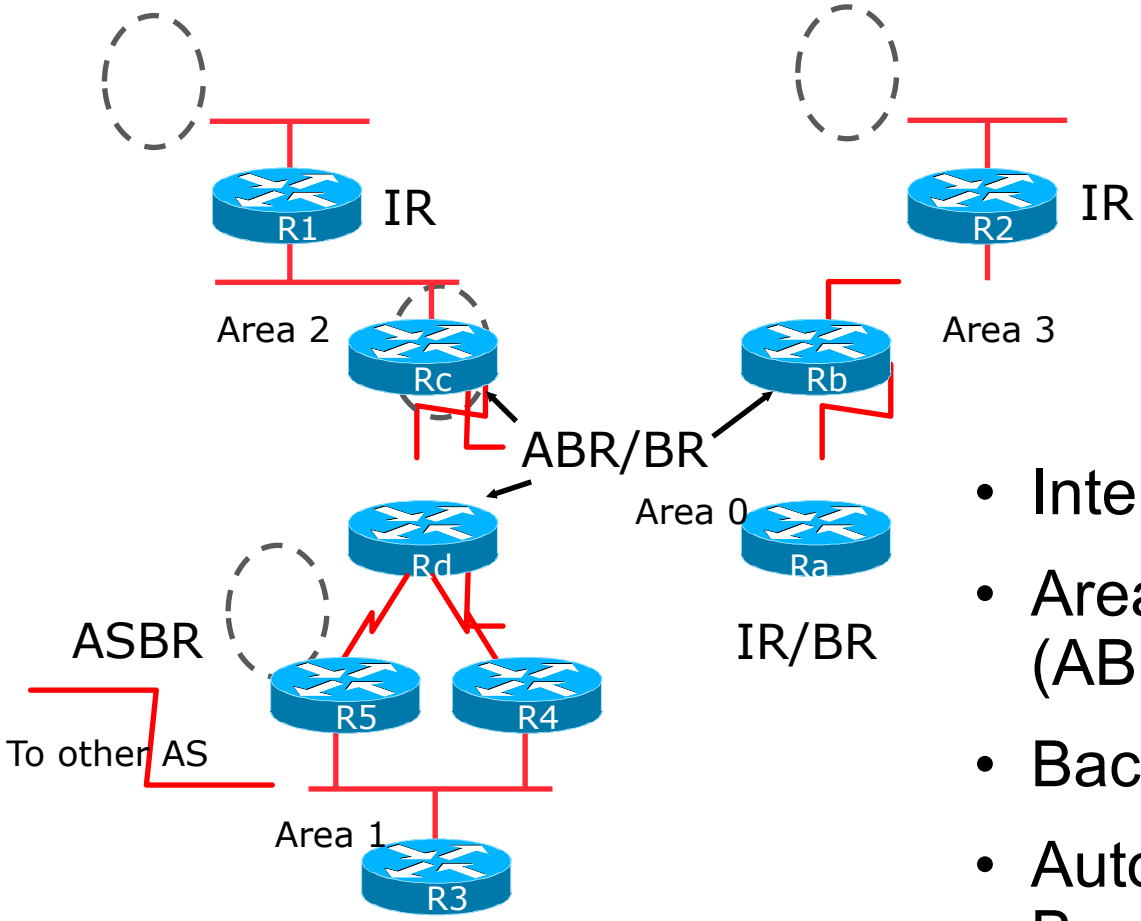


Virtual Links between OSPF Areas

- 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

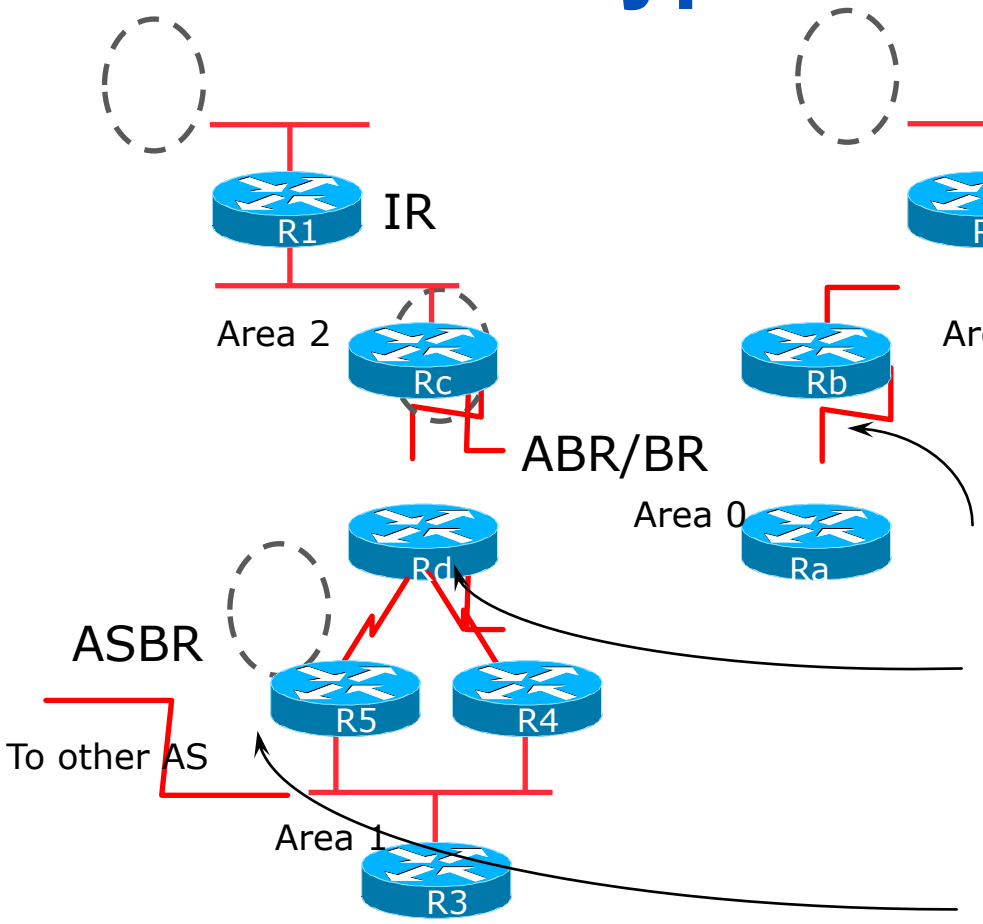


Classification of Routers



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

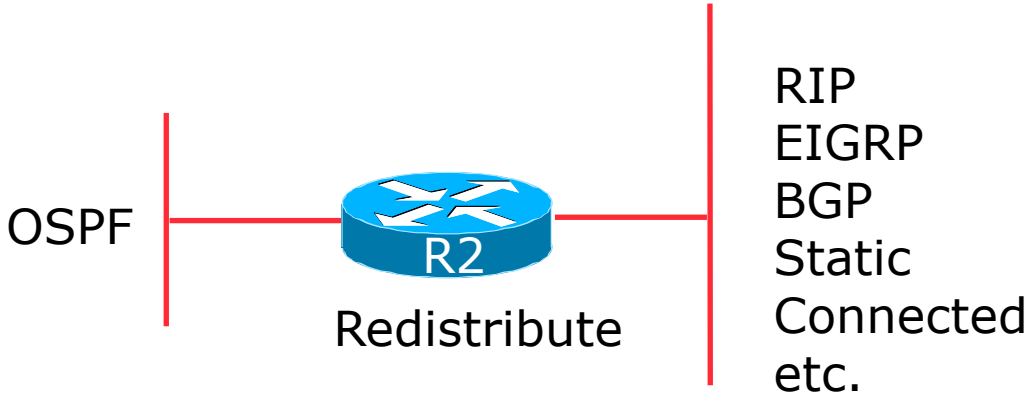
OSPF Route Types



- 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

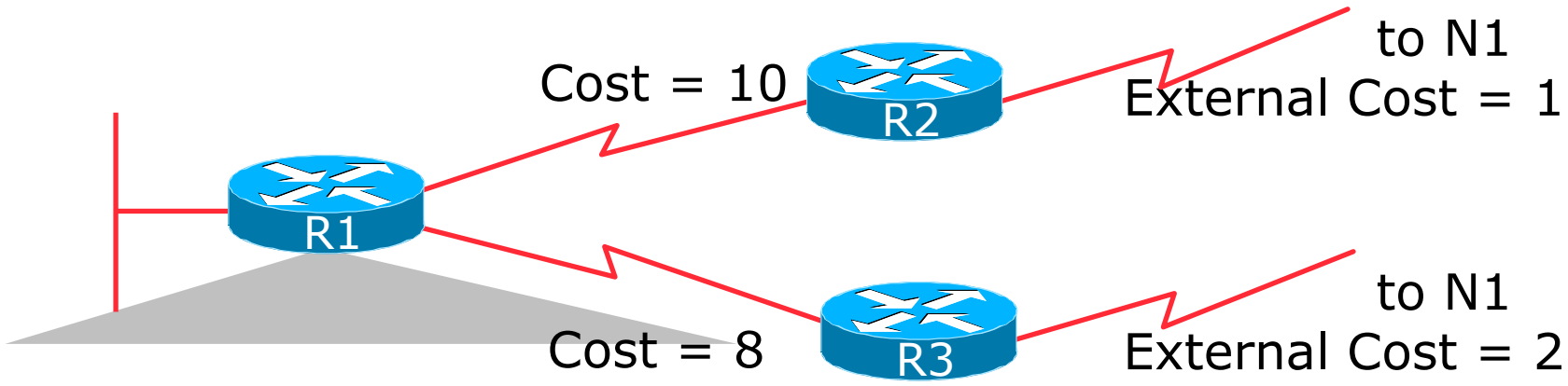
External Routes

- 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

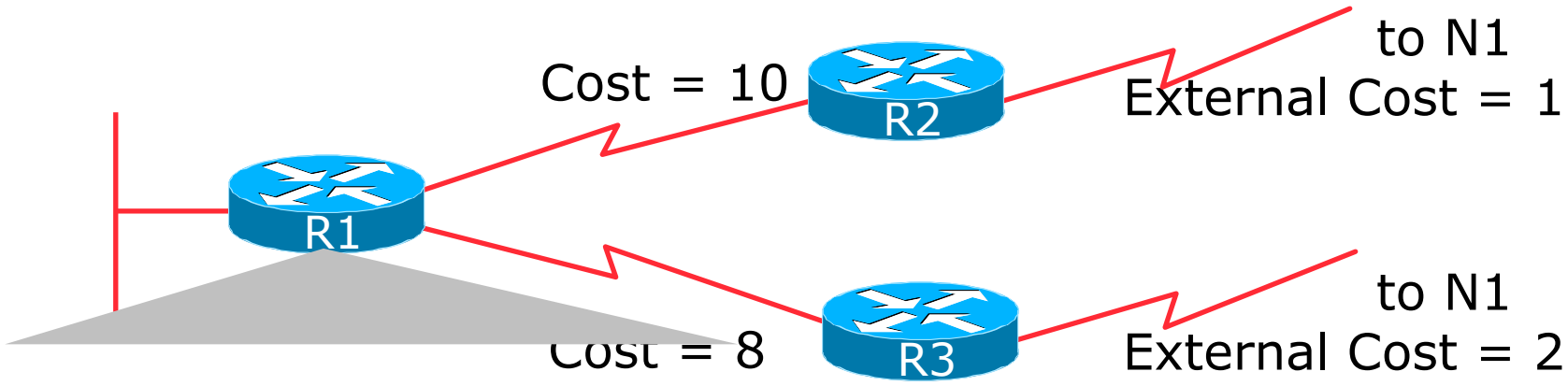


Network	Type 1	Next Hop
N1	11	R2
N1	10	R3

Selected Route

External Routes

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



Network	Type 1	Next Hop
N1	1	R2
N1	2	R3

Selected Route

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)

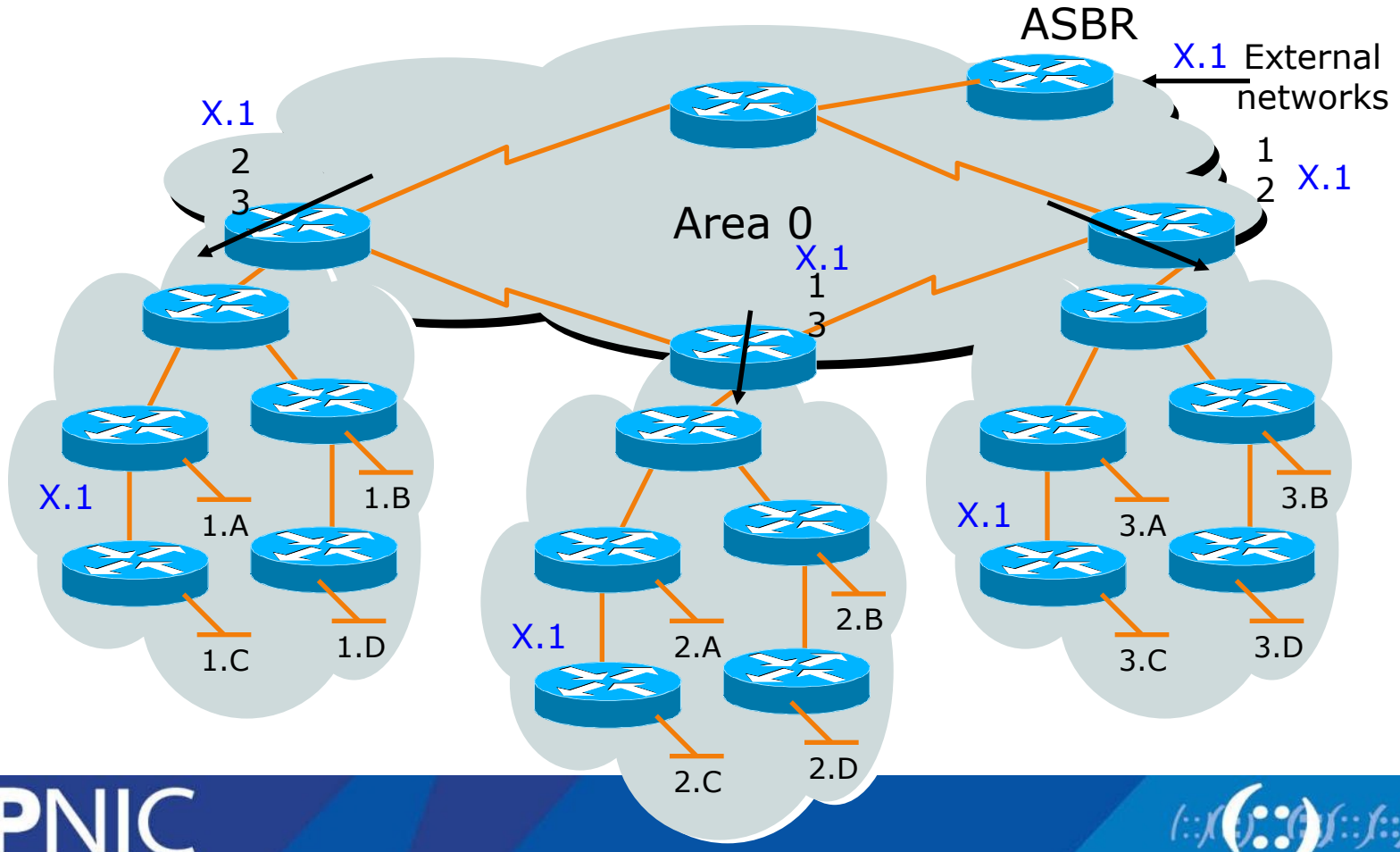
- Defines routes to destination external to the AS
- Default route is also sent as external
- Two types of external LSA:
  - E1: Consider the total cost up to the external destination
  - E2: Considers only the cost of the outgoing interface to the external destination
- (Type 7 LSAs used to describe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

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

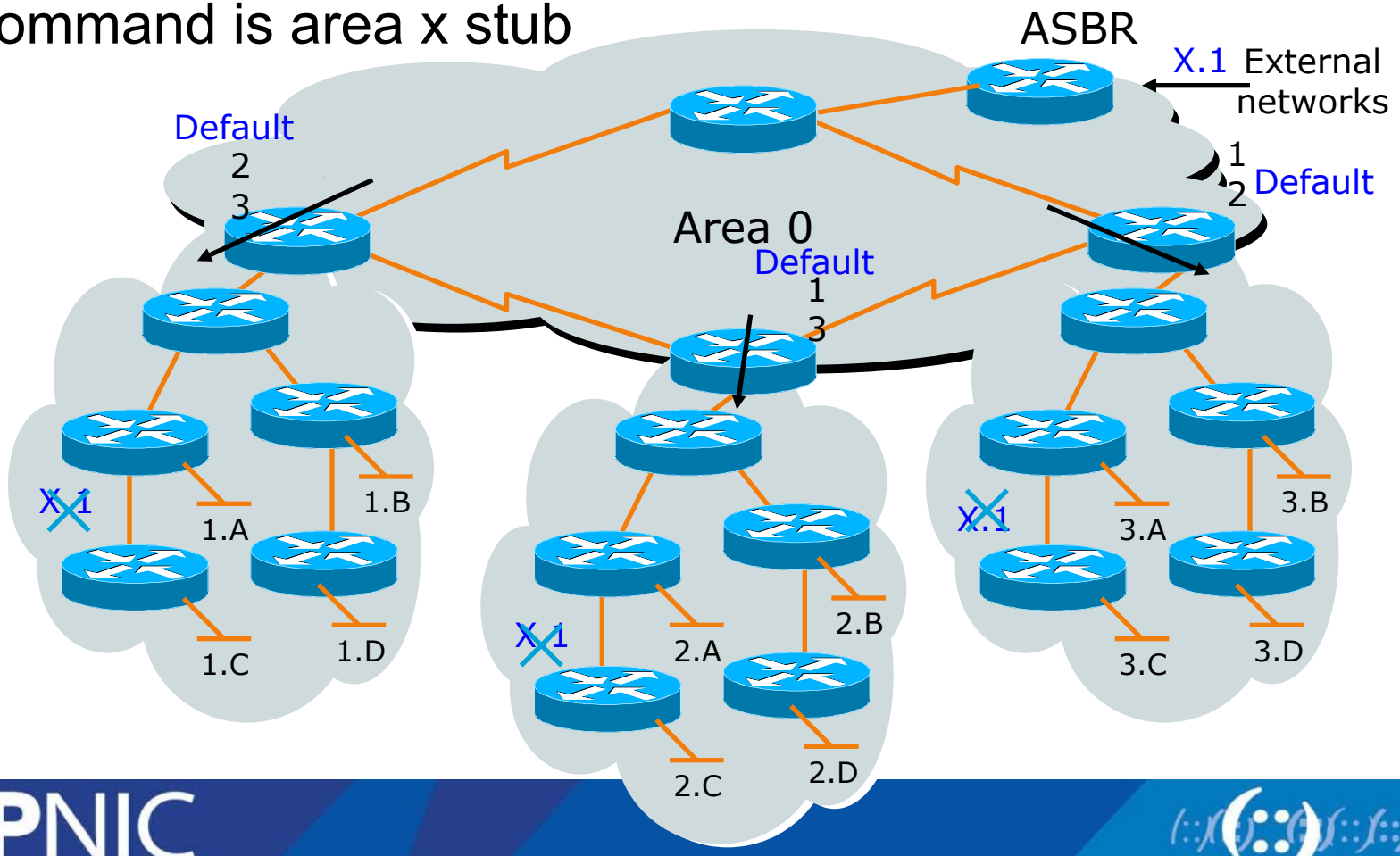
Regular Area (Not a Stub)

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



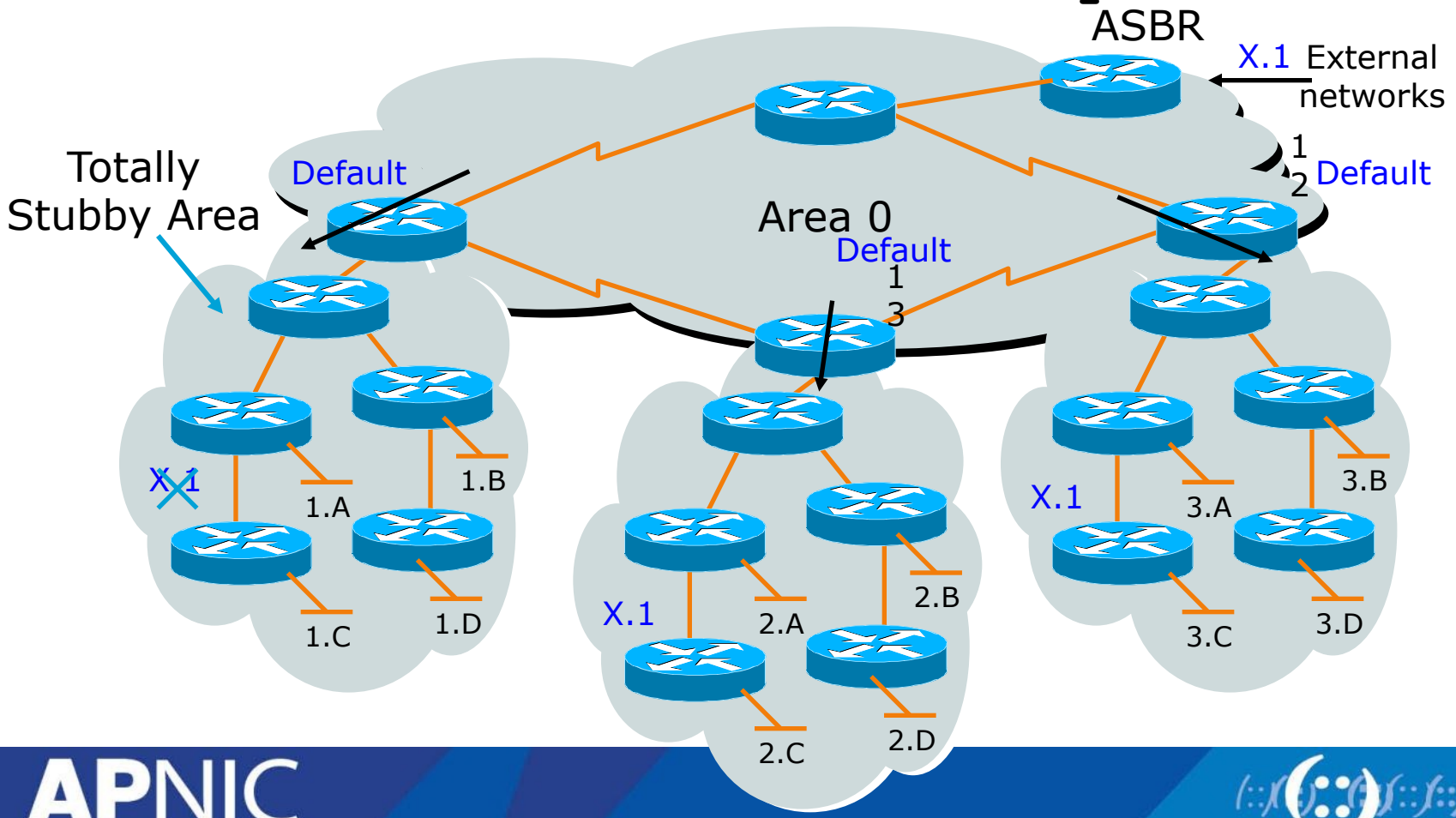
Normal Stub Area

- 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

Agenda

- TCP/IP communication and IP Routing
- Internet Routing and Routing Protocols
- IPv6 Address Structure
- Training Lab Topology
- OSPF Operation
- **Lab exercise: Basic router and OSPF**
- BGP – operation and attributes
- BGP Scaling Techniques
- Lab: iBGP, eBGP, Peer-group, RR

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 OSPFv3:

- 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 OSPF3 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 unreachables
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 unreachables
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 unreachables
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 unreachables
 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

Agenda

- TCP/IP communication and IP Routing
- Internet Routing and Routing Protocols
- IPv6 Address Structure
- Training Lab Topology
- OSPF Operation
- Lab exercise: Basic router and OSPF
- **BGP – operation and attributes**
- BGP Scaling Techniques
- Lab: iBGP, eBGP, Peer-group, RR

Border Gateway Protocol - BGP

- Runs over TCP (port 179)
  - TCP connection required before BGP session
  - Need to be reachable!
- Path vector routing protocol
  - Best path selection based on path attributes
  - Route: destination and the attributes of the path to reach the destination
- Incremental BGP updates

Internal & External BGP

- eBGP used to:
  - Exchange networks/routes between ASes
    - Aggregates and sub-aggregates
  - Implement routing policies
    - To manipulate inbound and outbound traffic
- iBGP is used to:
  - Carry customer networks/prefixes
  - Internet routes (some or all) across the AS backbone

BGP Message Types

- Open:
  - After a TCP connection has been established between two BGP routers, an Open message is sent
    - Once the open message is confirmed (keepalive), the BGP session is established – become BGP peers/neigbors!
  - Contains:
    - Sender's ASN
    - BGP version
    - BGP router ID
    - Hold-time (3 x keepalive interval)

BGP Message Types

- Keepalive:
  - Exchanged initially to acknowledge Open messages
  - Exchanged periodically (60 secs) to maintain BGP session
    - Dataless packet

- Update:
  - BGP peers exchange network information through Update messages
    - One update for each path!
  - Contains:
    - Withdrawn routes – no more reachable
    - Path attributes – attributes for this path to reach the destinations specified by the NLRI
    - NLRI – list of networks reachable through this path <prefix, length>

BGP Message Types

- Notification:
  - Sent when an error condition is detected
  - The BGP session is torn down immediately!
  - Contains:
    - Error code
    - Error sub-code
    - Data related to error

BGP Neighbor States

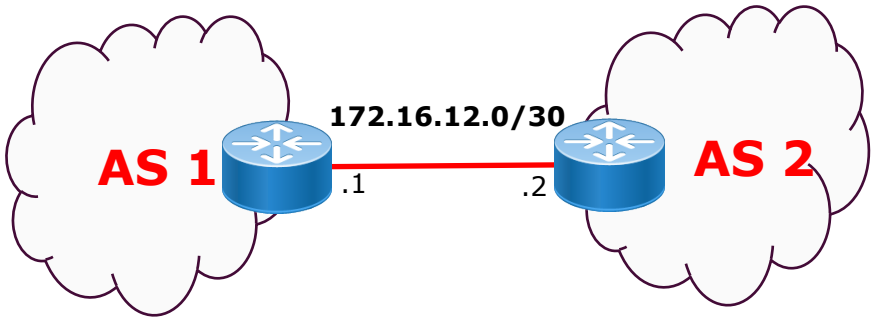
- A BGP router goes through six different states
  - Idle
    - The router is looking for a route to its neighbor
  - Connect
    - BGP router moves from Idle to Connect state if it has found a route to its neighbor, and has started the 3-way TCP handshake
    - If the 3-way handshake is complete, sends an Open message
  - Active
    - A router transitions to Active state if the initial 3-way handshake was not successful
    - Initiates a new 3-way handshake
    - If the 3-way handshake is complete, sends an Open message
    - Else, falls back to Idle

BGP Neighbor States

- A BGP router goes through six different states
  - Open Sent
    - An Open message has been sent to the neighbor
    - If it receives a keepalive, moves to Open Confirm, else back to Active
  - Open Confirm
    - Has received an acknowledgment for its Open message, and is waiting for the initial keepalive
    - If it receives the intial keepalive, transitions to Established
  - Established
    - The BGP neighbor relationship (session) is established!
    - Routing information can now be exchanged

BGP Neighbor Relationship

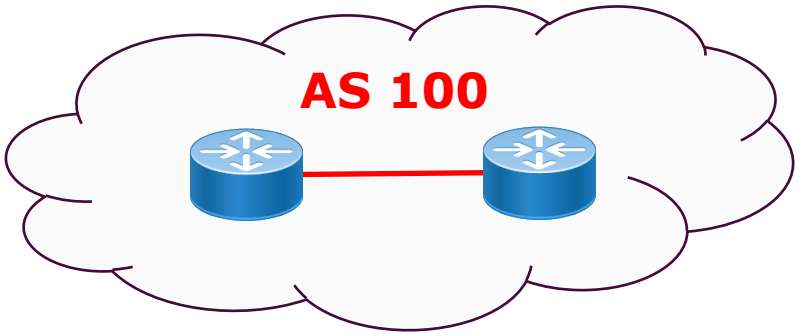
- eBGP neighbors/peers
  - BGP session established between routers in different ASes
  - Generally directly connected!
    - Session established using directly connected intf IP
    - Peering address must match the TCP session!
  - Else, we need a static route to reach the neighbor and change the eBGP TTL value (default 1)




```
router bgp 1
 neighbor 172.16.12.2 remote-as 2
!
 address-family ipv4
  neighbor 172.16.12.2 activate
 !
```

BGP Neighbor Relationship

- iBGP neighbors/peers
  - BGP session established between routers within the same AS
  - Does not need to be directly connected
    - IGP ensure reachability (TCP connection)
  - Generally using loopback addresses

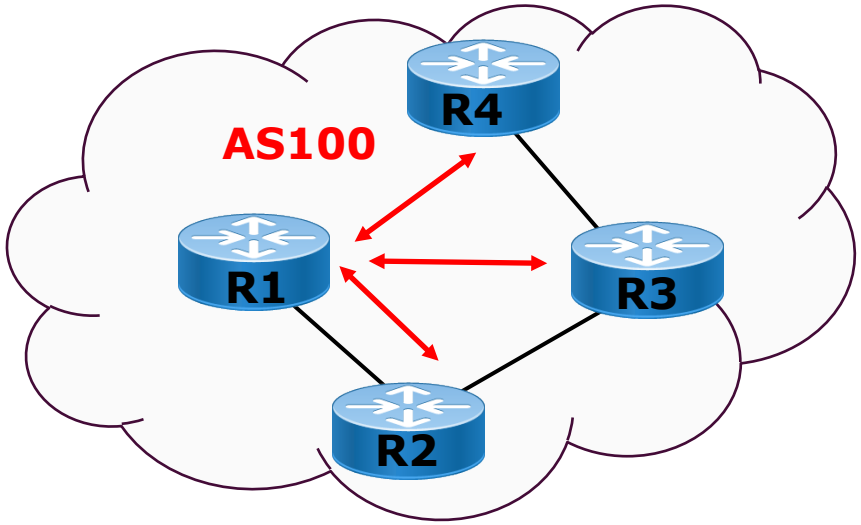



```
router bgp 100
 neighbor 10.10.10.2 remote-as 100
!
```

iBGP Operation

- iBGP routers must:
  - Originate directly connected routes
  - Carry routes learned from outside the AS to all routers within the AS
    - Fully-meshed instead of redistributing!
    - Advertise routes learned from eBGP peers to all iBGP peers!
  - To prevent routing loops (in a fully-meshed network)
    - iBGP routers are not allowed to advertise iBGP learned routes to other iBGP peers!

iBGP full-mesh




```
router bgp 100
 neighbor 10.10.10.2 remote-as 100
 neighbor 10.10.10.3 remote-as 100
 neighbor 10.10.10.4 remote-as 100
 !
```

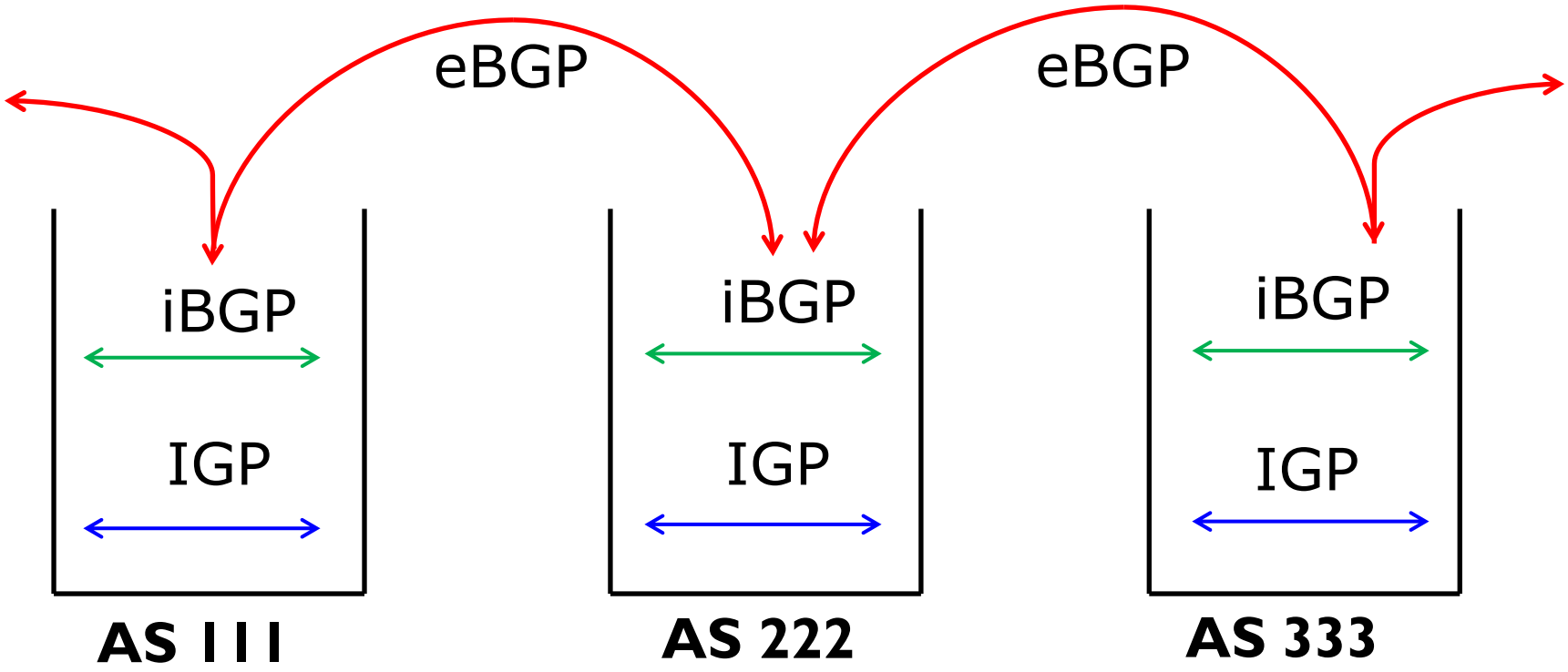
Sourcing iBGP from Loopback

- By default, routers use the exit-interface address as the source address for locally originated packets (updates)
  - If the BGP TCP session was established using any other interface (loopbacks) addresses, the source address for BGP updates must match!

- The `update-source loopback` command achieves this

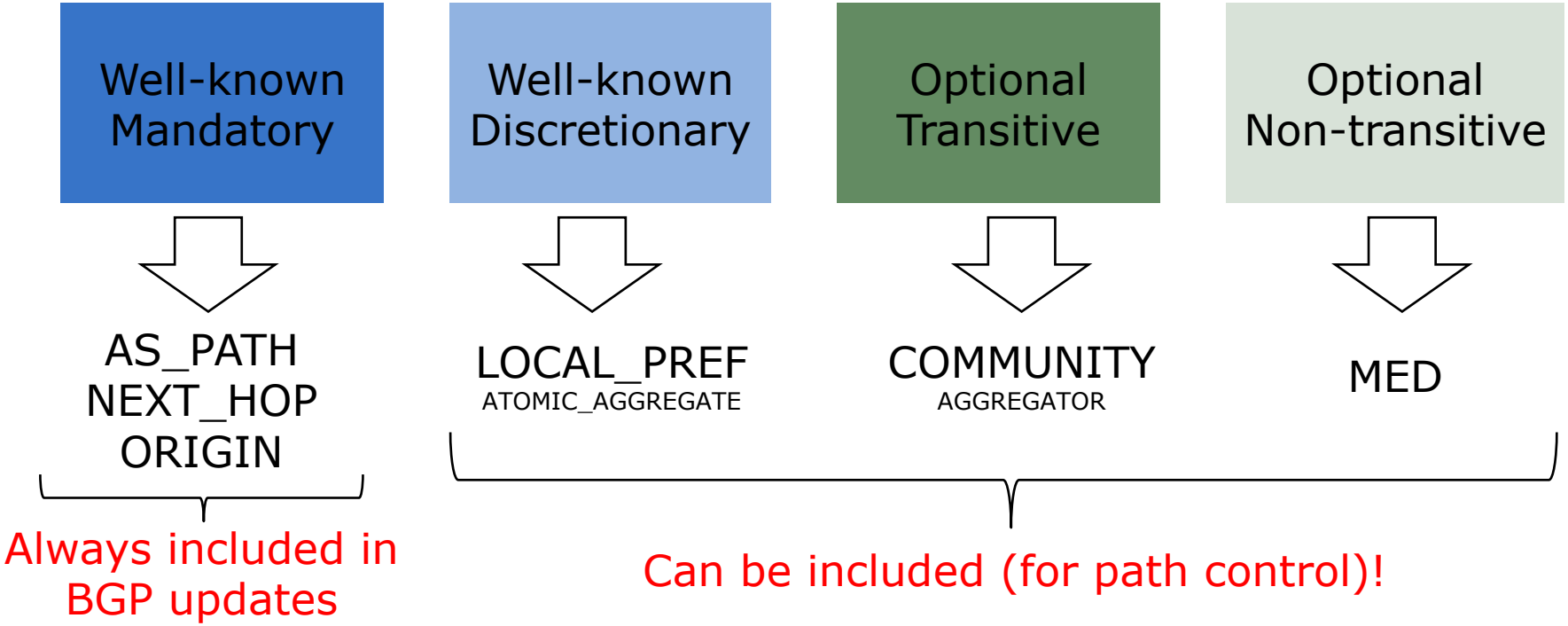
```
router bgp 100
 neighbor 10.10.10.1 remote-as 100
 neighbor 10.10.10.1 update-source loopback 0
!
```

How it all works?



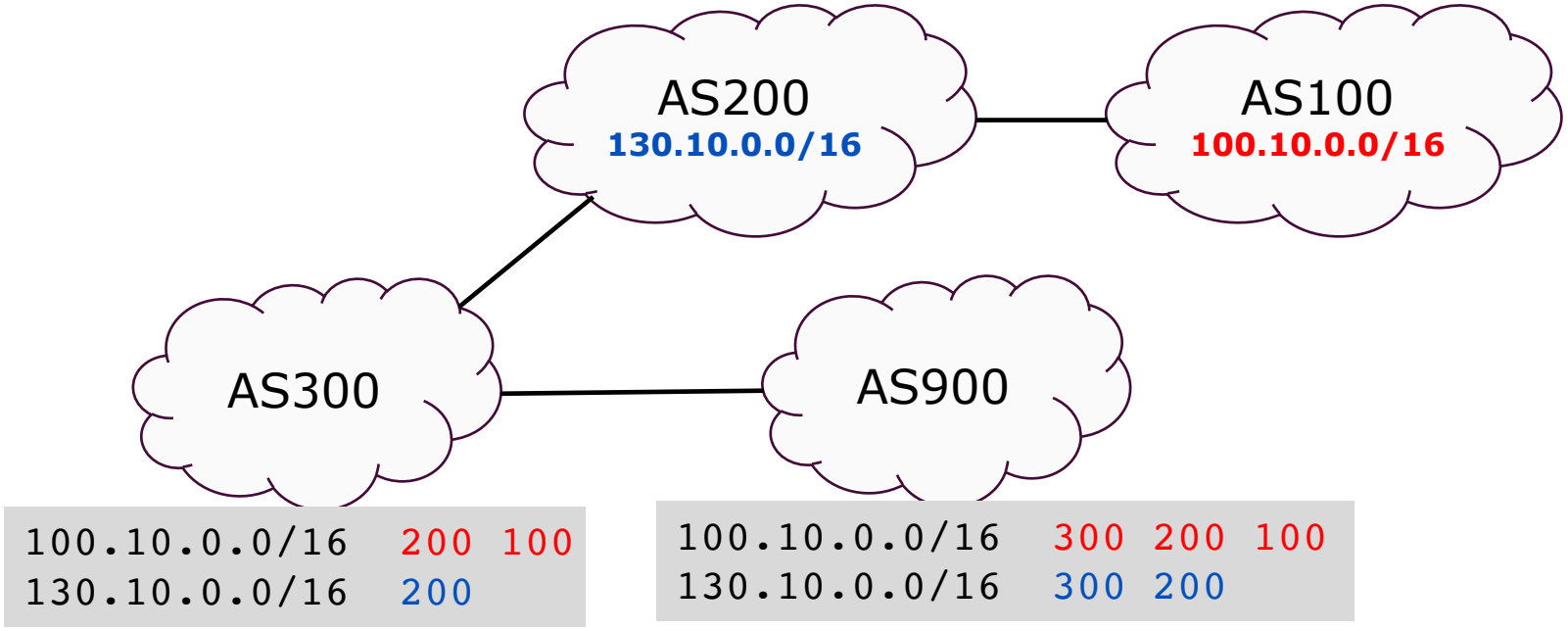
BGP Path Attributes

- Attributes describe the path to a network(s)/NLRI
  - Used to enforce routing policies for path control!



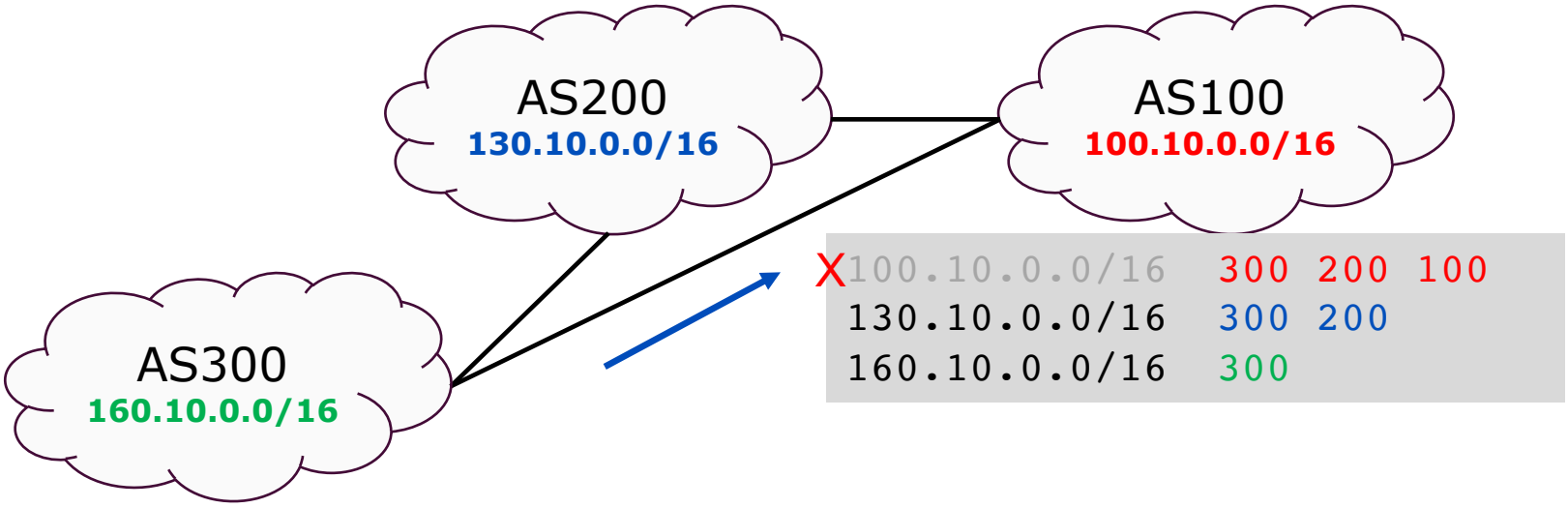
AS_PATH

- Indicates the list of ASes a route has passed through to reach the local AS
  - the list of ASes to reach a destination
  - can influence path selection!



AS_PATH

- Used to ensure a loop-free exchange of routing info between ASes
  - If own AS is seen in an update from an eBGP peer, loop is detected (Update is dropped)!

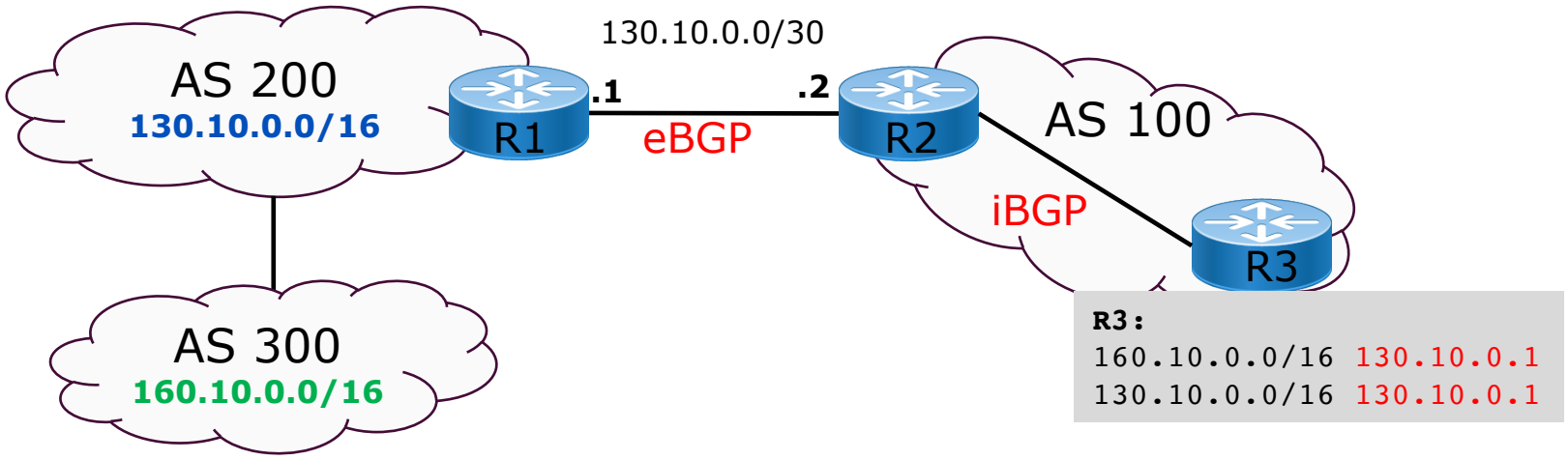


NEXT_HOP

- Indicates the next hop address to reach the destination
  - Source of the update packet!
- For eBGP
  - eBGP neighbor address (to reach the next AS)
- For iBGP
  - Generally the loopback address

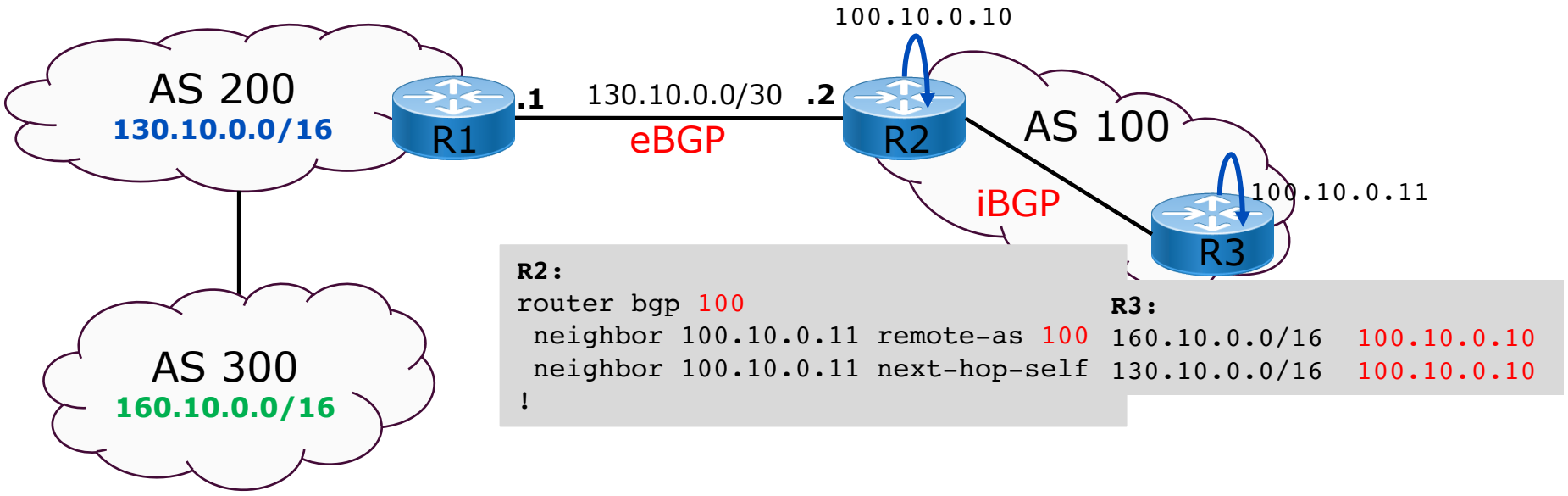
NEXT_HOP

- eBGP learned routes are advertised to iBGP peers without changing the next hop
  - Routers within the AS need to be able to reach the next hop (IGP or static)
  - Else, external routes not installed in the routing table!



NEXT-HOP-SELF

- Override the eBGP next hop default behavior with `next-hop-self` command
  - Advertises itself as the next hop for external routes
    - Reachable through IGP



R2:

```
router bgp 100
 neighbor 100.10.0.11 remote-as 100
 neighbor 100.10.0.11 next-hop-self
!
```

R3:

```
160.10.0.0/16   100.10.0.10
130.10.0.0/16   100.10.0.10
```

ORIGIN

- Indicates the origin of the route
  - IGP (i)
    - Interior to the originating AS (advertised with the **`network`** command)
  - EGP (e)
    - Generated by EGP (obsolete!)
  - Incomplete (?)
    - Route's origin is unknown (usually redistributed)

ORIGIN

```
R1# show ip bgp
BGP table version is 24, local router ID is 172.16.1.2
Status codes: s suppressed, d damped, h history, * valid, > b
internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric  LocPrf  Weight  Path
*> 192.208.10.0     192.208.10.5             0               0  300 i
*> 172.16.1.0       0.0.0.0                  0           32768  i
<output omitted>
```

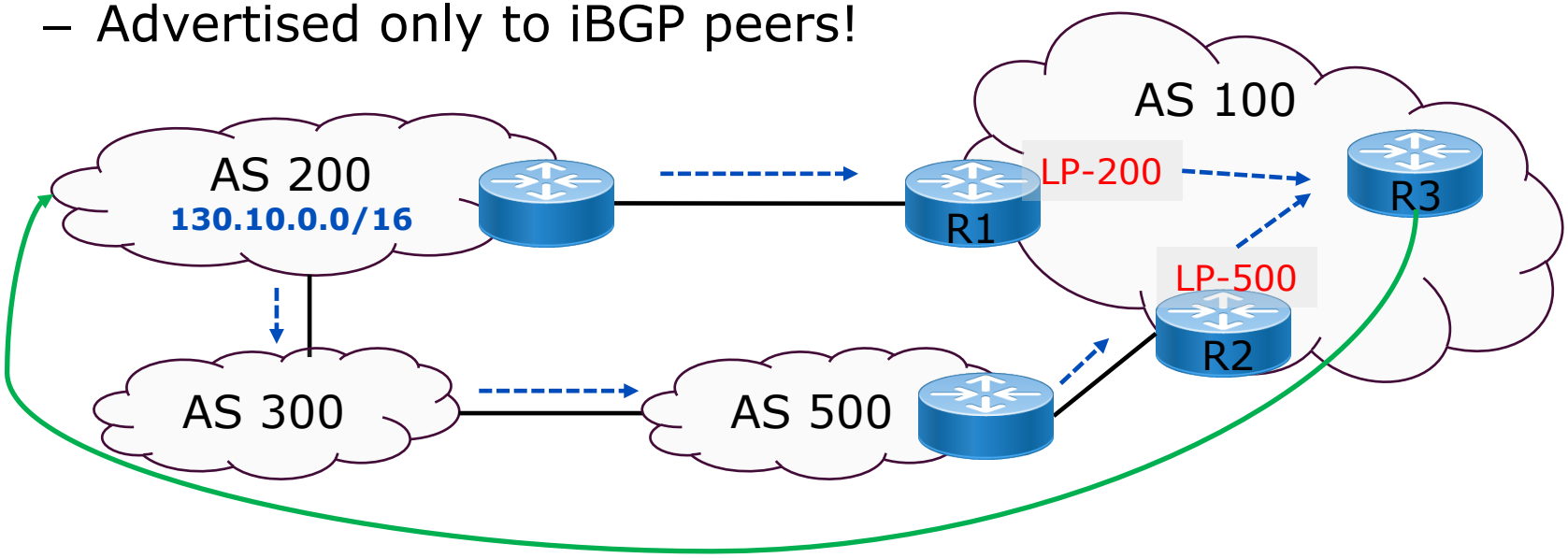
i = Route generated by the **network** command.

```
R1# show ip bgp
<output omitted>
   Network          Next Hop            Metric  LocPrf  Weight  Path
*> 10.1.1.0/24       0.0.0.0                  0           32768  ?
*> 192.168.1.0/24    10.1.1.2                84           32768  ?
*> 192.168.2.0/24    10.1.1.2                74           32768  ?
<output omitted>
```

? = Route generated by unknown method (usually redistributed).

LOCAL_PREF

- Local preference tells routers within the AS (local) the preferred path to exit the AS
  - Path with highest local_pref wins
    - Outbound traffic!
- Local to the AS
  - Advertised only to iBGP peers!



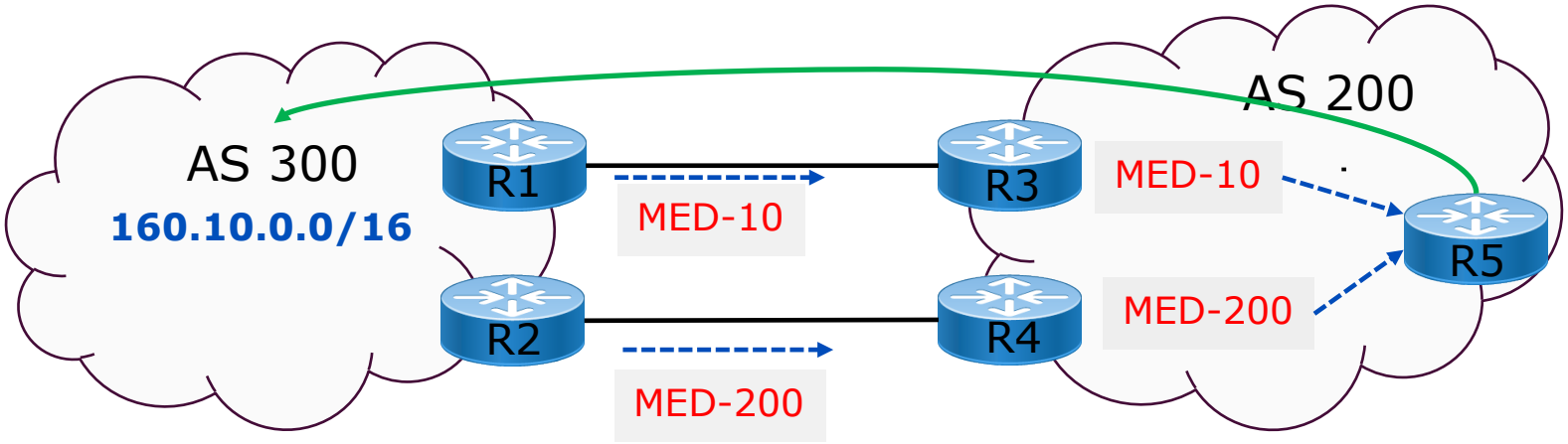
COMMUNITY

- Used to group prefixes (incoming/outgoing) and apply policies to the communities
  - A prefix can belong to more than one community
- Is (was?) a 32-bit integer
  - Represented as two 16-bit integers [`ASN:number`]
    - Works well for 2-byte ASN
- With 4-byte ASNs
  - Common to see [`private-ASN:number`]
  - RFC 8092 (BGP Large Communities): 96-bit integer
    - [32-bit ASN:32-bit:32-bit]

MED

- Multi-exit discriminator is inter-AS non-transitive
  - Indicates to neighbor AS about the preferred entry points into the local AS (incoming traffic)

- The path with lowest MED wins!



BGP Operation

- BGP learns routes from iBGP and eBGP peers
  - Selects best path based on the attributes
  - Installs best path in the routing table
  - Advertises the best paths to its other BGP peers
    - eBGP learned routes to iBGP peers
    - iBGP learned routes to eBGP peers

Advertising Networks in BGP

- The `network` statement
  - allows BGP to inject routes into BGP table and advertise to neighbors only if it already exists in the routing table!

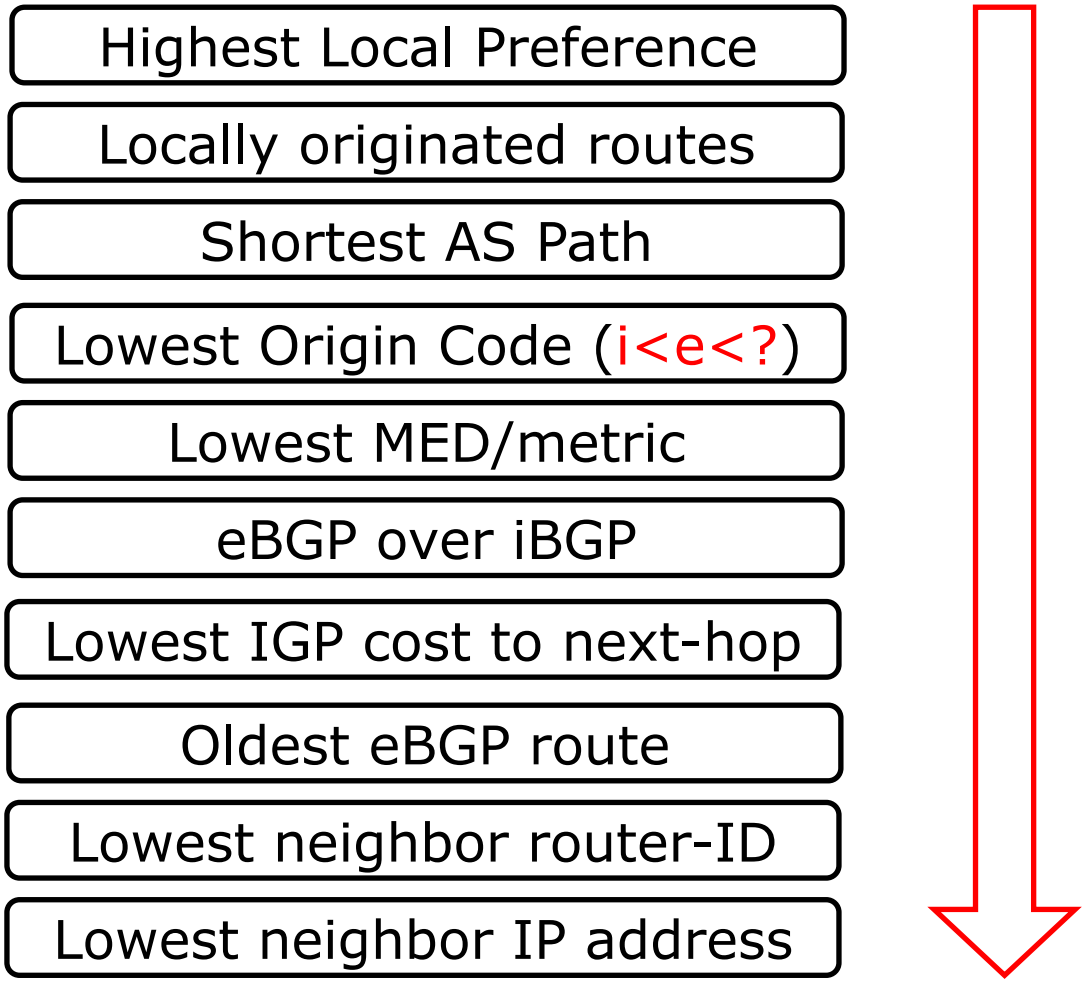
```
router bgp 100
 address-family ipv4 unicast
  network <prefix> mask <subnet-mask>
 address-family ipv6 unicast
  network <prefix/length>
```

- BGP "Synchronization Rule":
  - iBGP learned routes should not be installed in the routing table nor advertised to eBGP peers unless the route was learned through an IGP first!
    - Prevents black-hole routes!

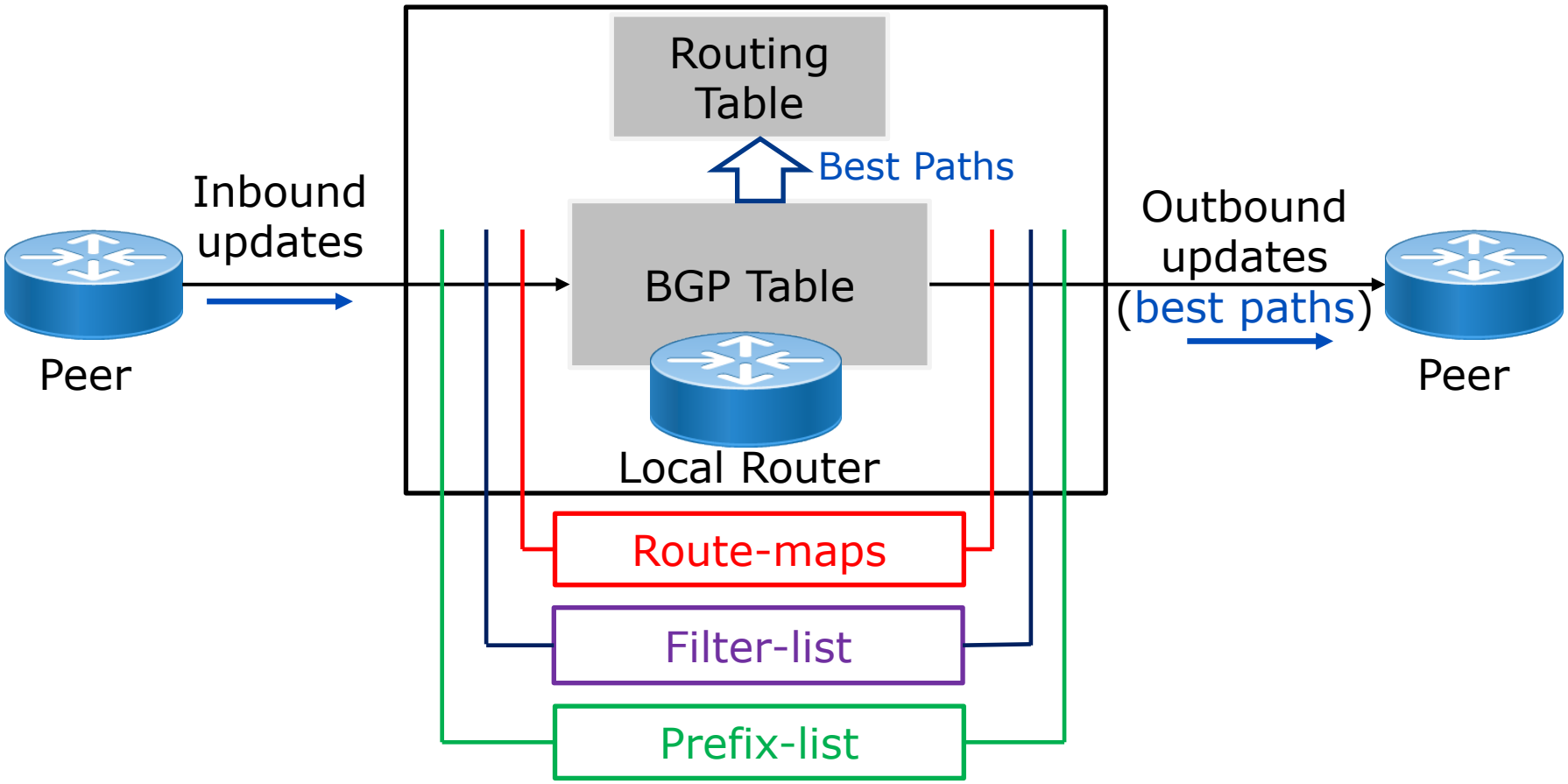
BGP Tables

- Neighbor Table
  - List of all BGP neighbors
- BGP Table
  - List of routes learned from all BGP neighbors
  - (And locally originated routes!)
- Routing (Forwarding) Table
  - All best paths
    - selected based on attributes and whose next-hops are reachable!

BGP Best Path Selection



Influence Path Selection – Policy



Policy Tools

- Prefix-list
  - To filter routes/prefixes
- Filter-list
  - To filter based on AS-path
  - To apply AS-path ACLs
- Route-map
  - modify attributes based on condition matches

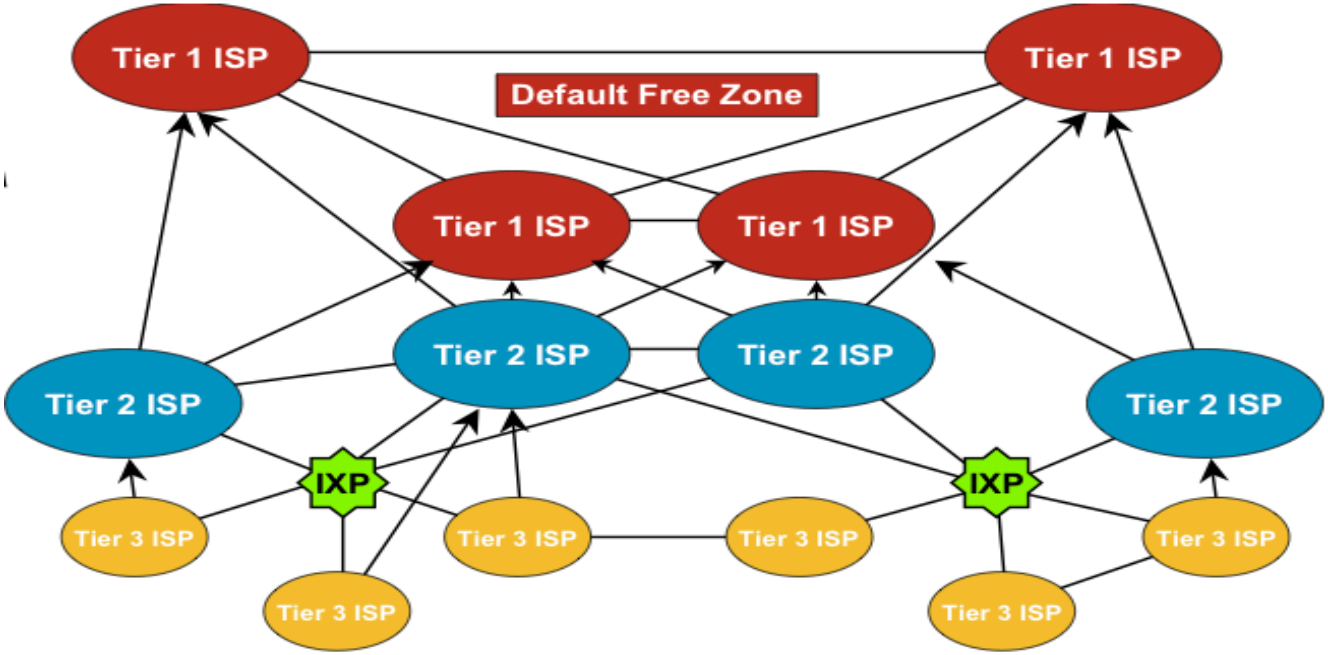
Exchanging Routes

- Pay someone to advertise your networks
  - TRANSIT
  - Make sure they have good onward peering/transit!

- Interconnect with as other ASes to exchange locally originated routes and traffic
  - PEERING
  - Private Peering
    - Between two ASes
  - Public Peering
    - at an IXP (domestic/global)

ISP Hierarchy

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



Questions

Agenda

- TCP/IP communication and IP Routing
- Internet Routing and Routing Protocols
- IPv6 Address Structure
- Training Lab Topology
- OSPF Operation
- Lab exercise: Basic router and OSPF
- BGP – operation and attributes
- **BGP Scaling Techniques**
- Lab: iBGP, eBGP, Peer-group, RR

BGP Peer Group

- Problem: number of BGP updates in a iBGP mesh
  - BGP updates generated for each neighbor individually
    - CPU wasted on repeat calculations
  - iBGP neighbors receive the same update
    - Contain same info
- Solution: Peer Groups
  - Group neighbors with the same outbound update policy
  - Updates are generated once per group

BGP Peer Group

- Still need to establish TCP sessions individually
- Useful when many neighbors have the same outbound policies
  - Runs through the outbound filters only once for the group (applied to all members)
- Members can have a different inbound policy!
- Simplifies configuration
  - Define the peer group
  - Add neighbors to the peer group
    - Still need to configure peering individually
  - Apply filters (outbound) to the group

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 (RFC 2270)
  - Also at IXPs where ISP policy is generally the same for each IX peer

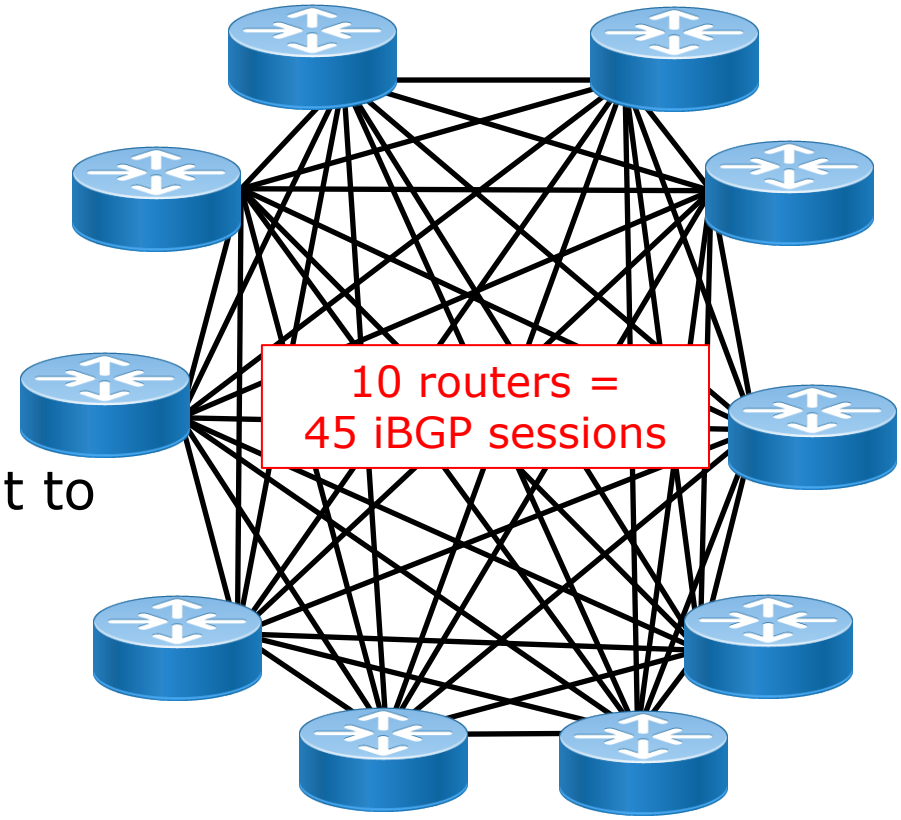
BGP Loop Prevention

- eBGP
  - AS-PATH attribute
  - If the local ASN is seen in a route received from a eBGP peer, a routing loop has occurred
    - Drop the route!

- iBGP
  - BGP router is not allowed to advertise iBGP learned routes to other iBGP peers within the AS
  - How do all iBGP routers learn about each other's networks?
    - iBGP full-mesh!

Scaling iBGP mesh

- Number of iBGP sessions
  - n(n-1)
    - 10 routers = 45 sessions
    - 100 routers = 4950 sessions
- Number of BGP updates
  - Every update needs to be sent to all iBGP peer



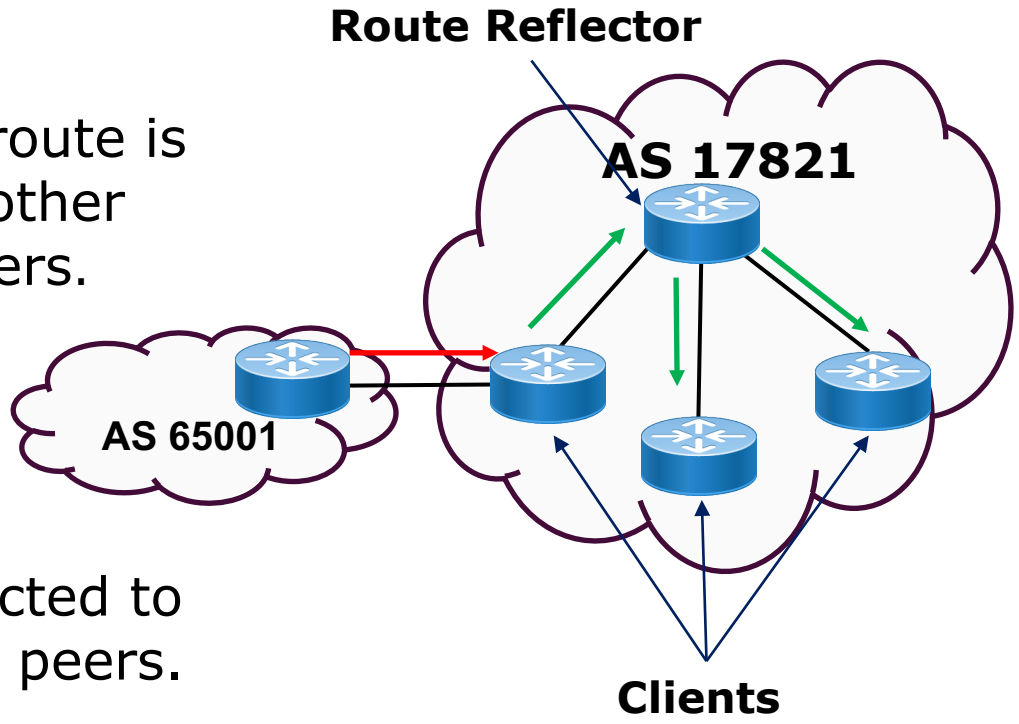
Solution

- **Route Reflection** (RFC4456)
  - RR client peers only with the RR
  - RR and its clients form a CLUSTER
  - Non-clients
  - An AS can have multiple clusters

RRs do not affect the actual traffic path;
only affects the path BGP messages take!

RR Operation

- When a RR receives an Update:
  - If from a client peer, the route is reflected (advertised) to other clients, and non-client peers.
  - If from a non-client, only reflected to client peers
  - If from a eBGP peer, reflected to both client and non-client peers.



Routing loops can happen with RRs!

Avoiding loops in RR

- **Originator_ID** attribute
  - The BGP router id of the originator, created by the RR
    - If you see your router id in the Originator_Id attribute, loop has occurred.

- **Cluster_List** attribute
  - In a cluster with a single RR, the Cluster ID is the router ID of the RR
  - If more than one RR in a cluster, a 4-byte Cluster ID configured
  - Cluster_List reflects the sequence of clusters a route has passed through
  - When a RR reflects routes (from clients) to non-clients, it appends the local Cluster ID to the Cluster_List
    - When a RR receives an update, if the local Cluster ID is seen in the list, a loop has occurred! (drops the update)

RR example

- RR configuration

```
router bgp 17821
 <output-omitted>
 neighbor 2406:6400:2 remote-as 17821
 neighbor 2406:6400:2 route-reflector-client
 neighbor 2406:6400:3 remote-as 17821
 neighbor 2406:6400:3 route-reflector-client
 neighbor 2406:6400:4 remote-as 17821
 neighbor 2406:6400:4 route-reflector-client
<output-omitted>
```

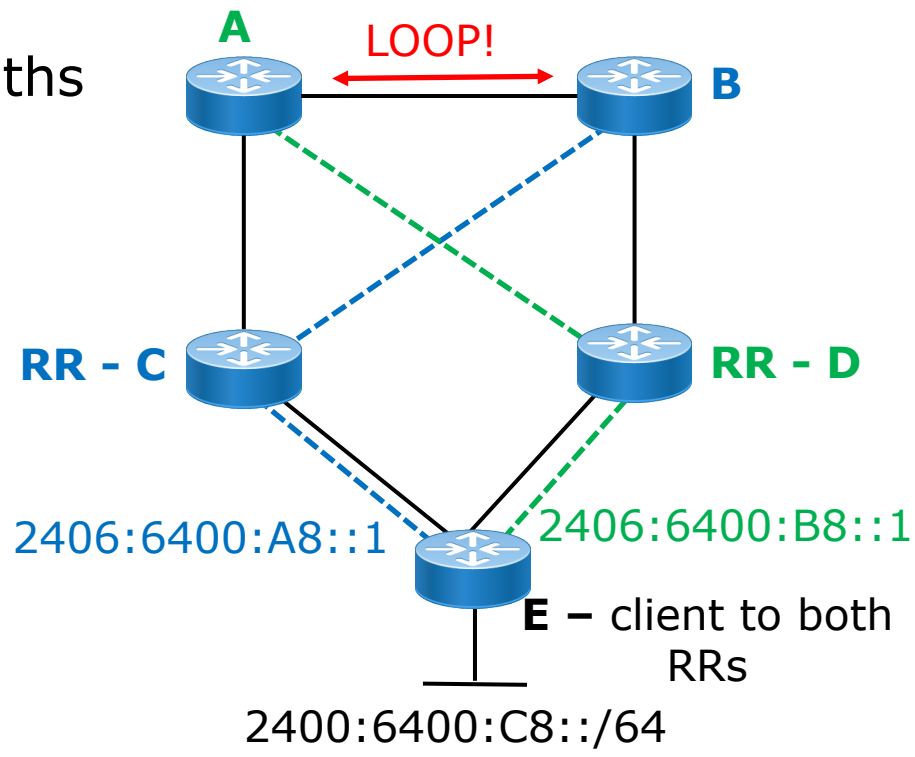
- RR Client config

```
router bgp 17821
 <output-omitted>
 neighbor 2406:6400:1 remote-as 17821
 <output-omitted>
```

Only peers with the RR

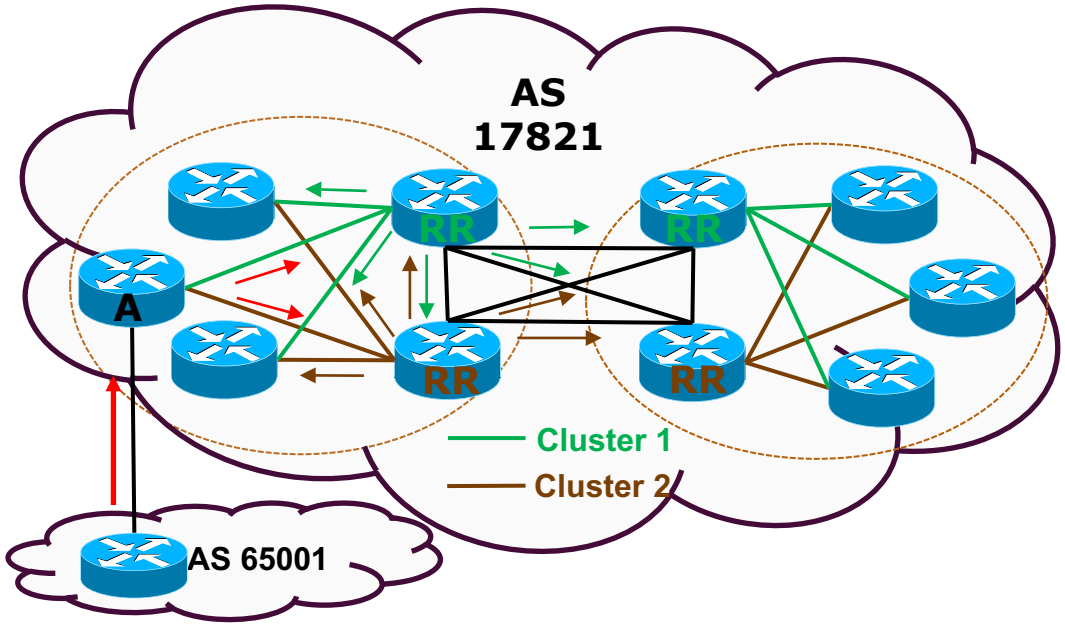
RR Selection

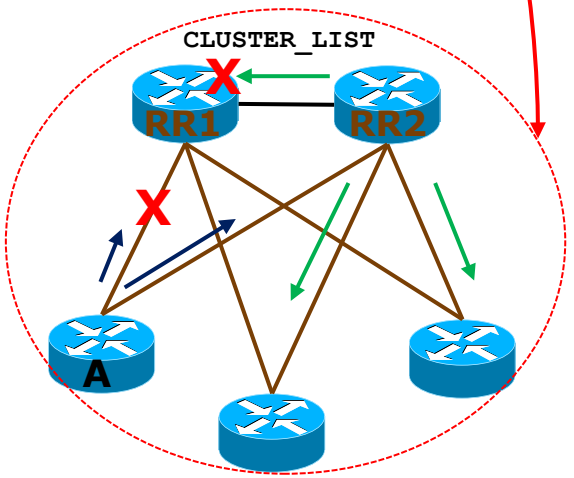
- Best practice is to follow the physical topology
  - Ensures traffic forwarding paths wont be affected
  - Prevents routing loops



RR Redundancy

- Most ISP networks would overlay two clusters
  - Each client peers with RRs in different clusters (same POP) for redundancy (NEVER two RRs in the same cluster!)
  - All RRs fully-meshed!
  - Can have full-mesh between clients in the same cluster

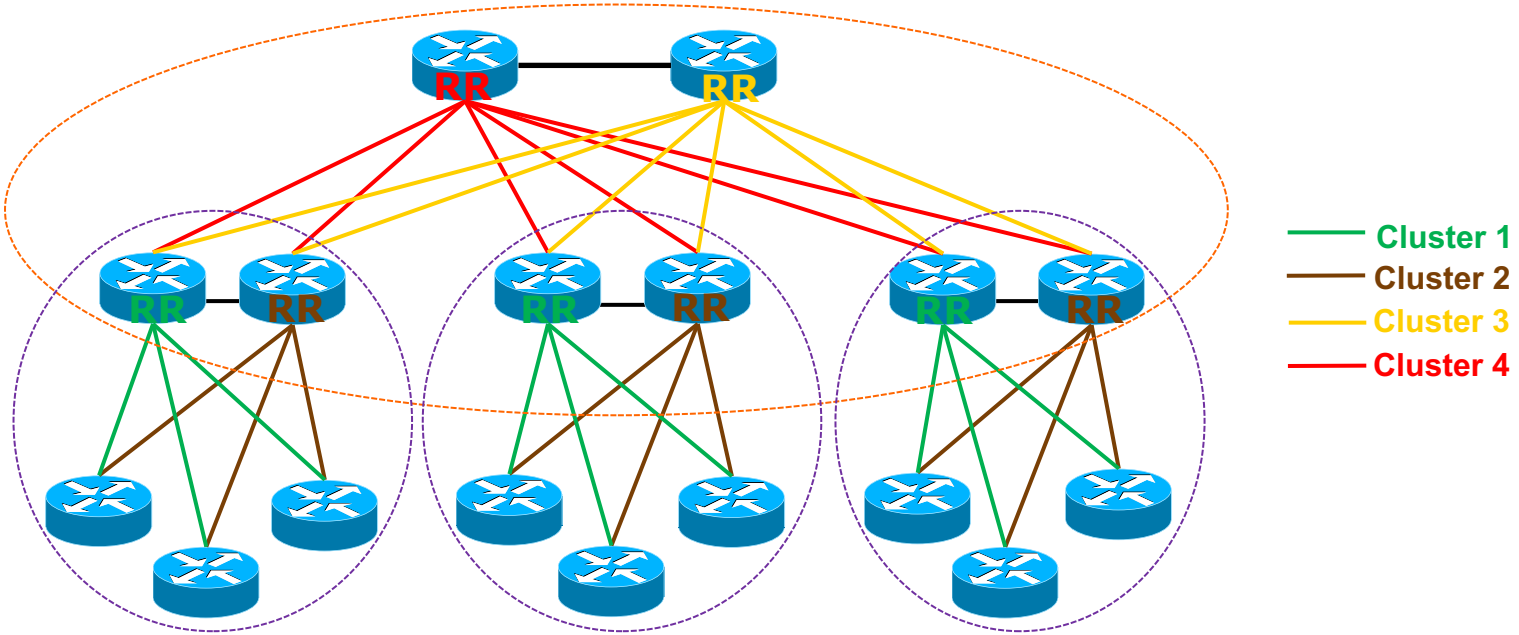





```
router bgp 17821
  bgp cluster-id 1.1.1.1
```

RR Redundancy – Best Practice

- A hierarchical RR design
  - RRs of some clusters are clients of other RRs in different clusters
  - Less iBGP sessions
  - Easier to manage and scale



Questions

Agenda

- TCP/IP communication and IP Routing
- Internet Routing and Routing Protocols
- IPv6 Address Structure
- Training Lab Topology
- OSPF Operation
- Lab exercise: Basic router and OSPF
- BGP – operation and attributes
- BGP Scaling Techniques
- **Lab: iBGP, eBGP, Peer-group, RR**

Case study- Deployment IPv6 in EGP

- Scenario:
  - MP-BGP 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

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:

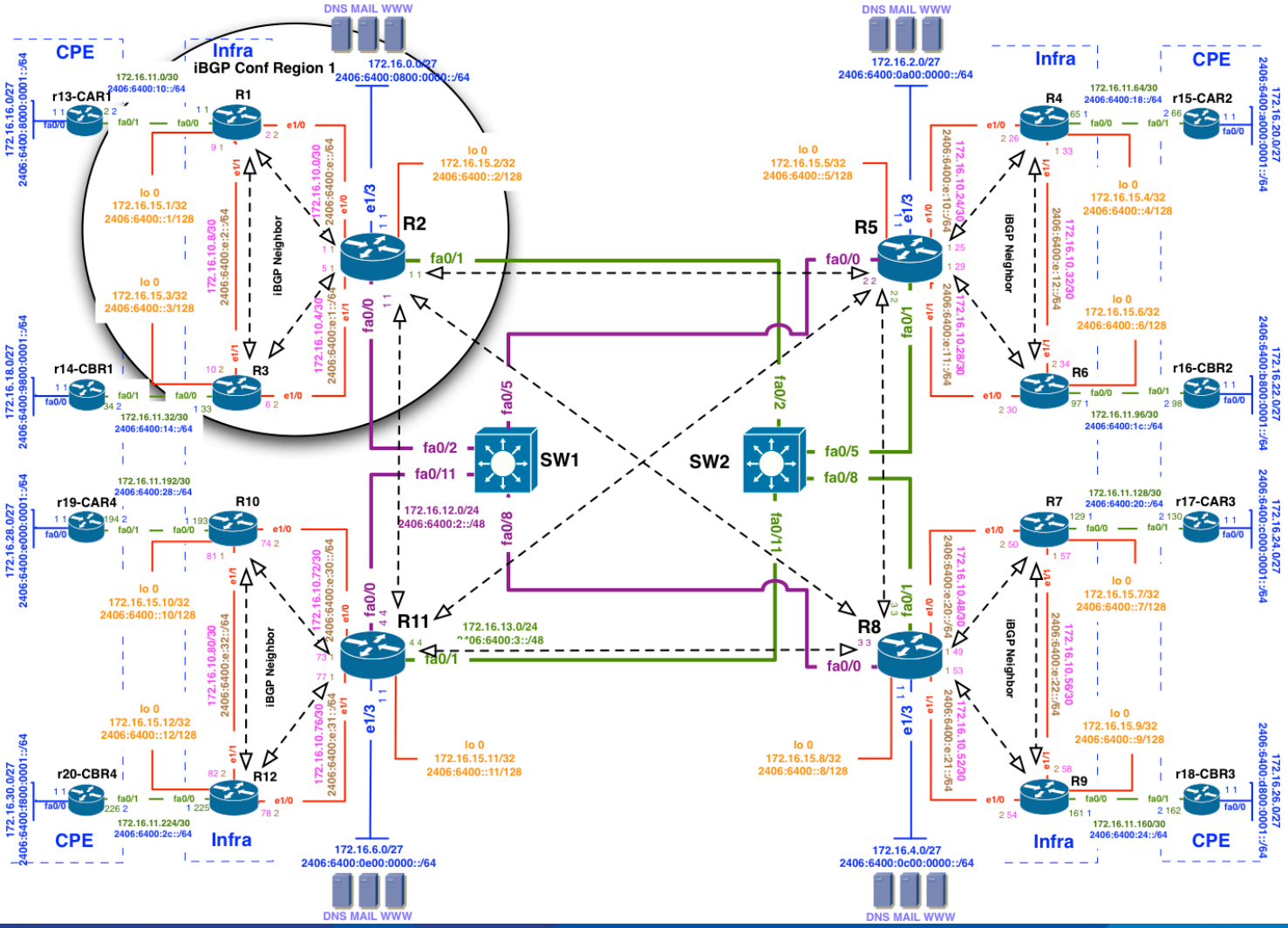
```
router bgp 17821
address-family ipv6
!
neighbor 2406:6400:0000:0000::2 remote-as 17821
neighbor 2406:6400:0000:0000::2 update-source loopback 0
neighbor 2406:6400:0000:0000::2 activate
!
network 2406:6400:0100:0000::/48
```

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

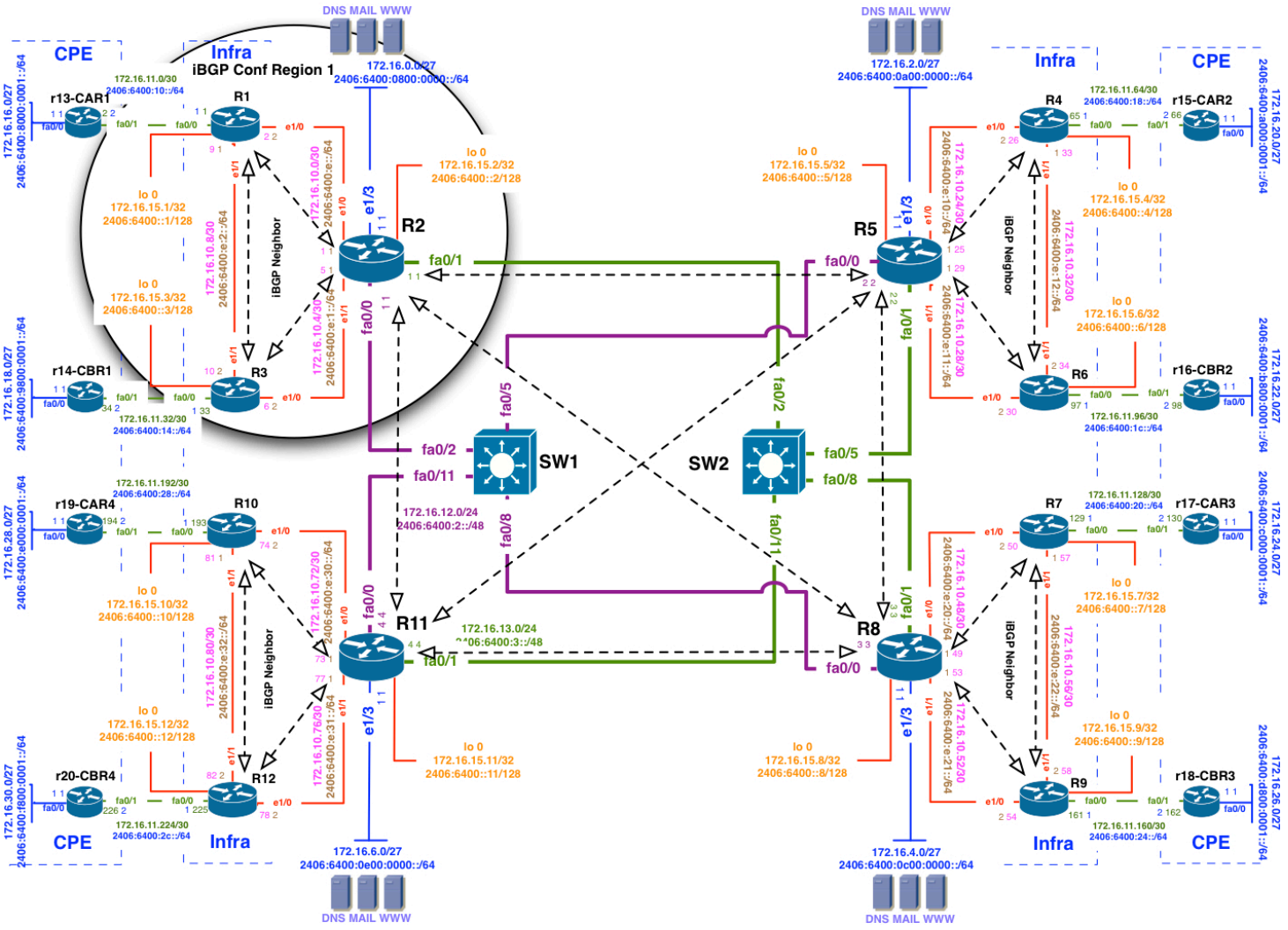
```
config t
router bgp 17821
address-family ipv6
no synchronization
neighbor 2406:6400:0000:0000::1 remote-as 17821
neighbor 2406:6400:0000:0000::1 update-source loopback 0
neighbor 2406:6400:0000:0000::1 activate
neighbor 2406:6400:0000:0000::3 remote-as 17821
neighbor 2406:6400:0000:0000::3 update-source 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

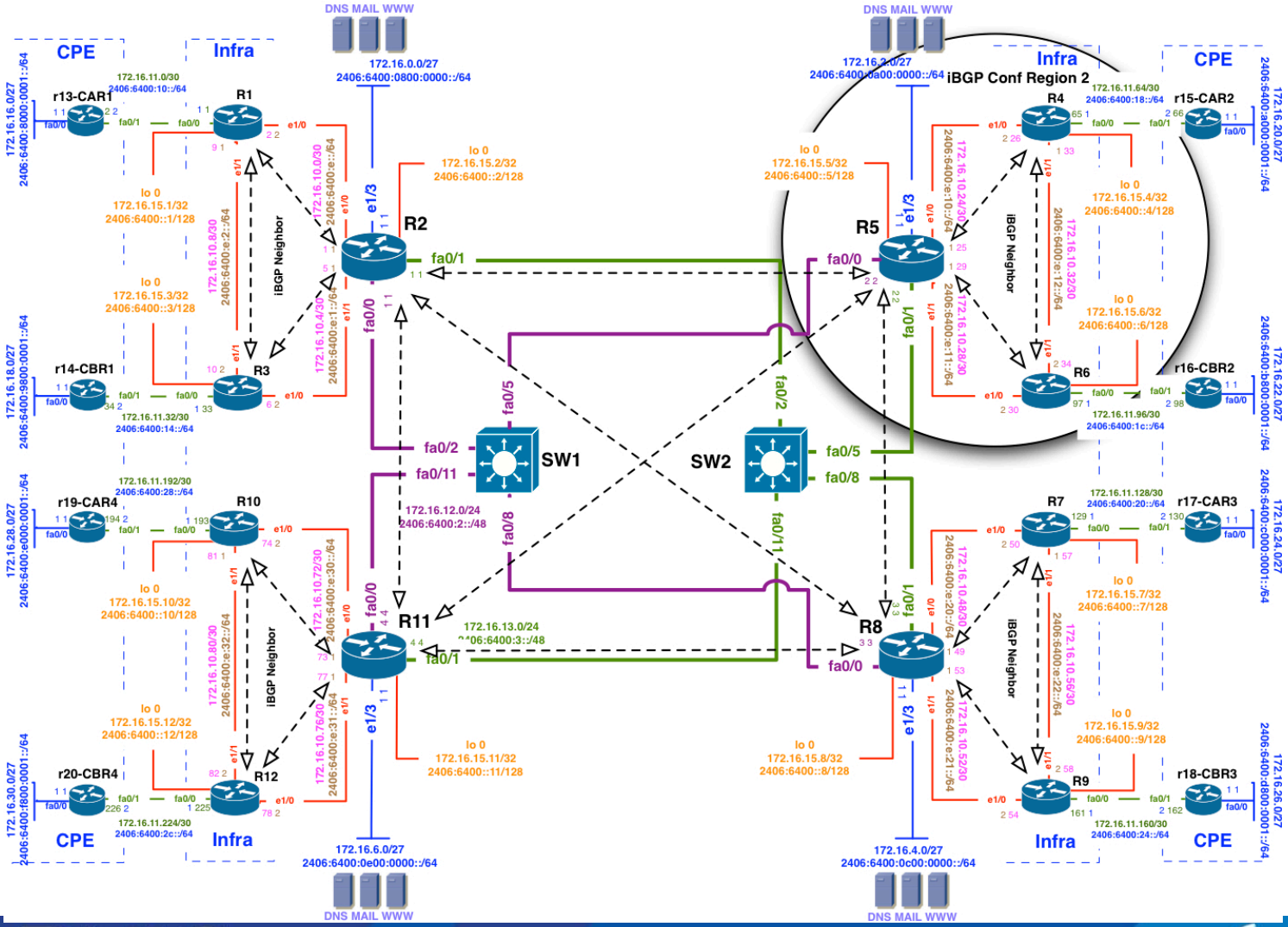


iBGP Full Mesh Issue

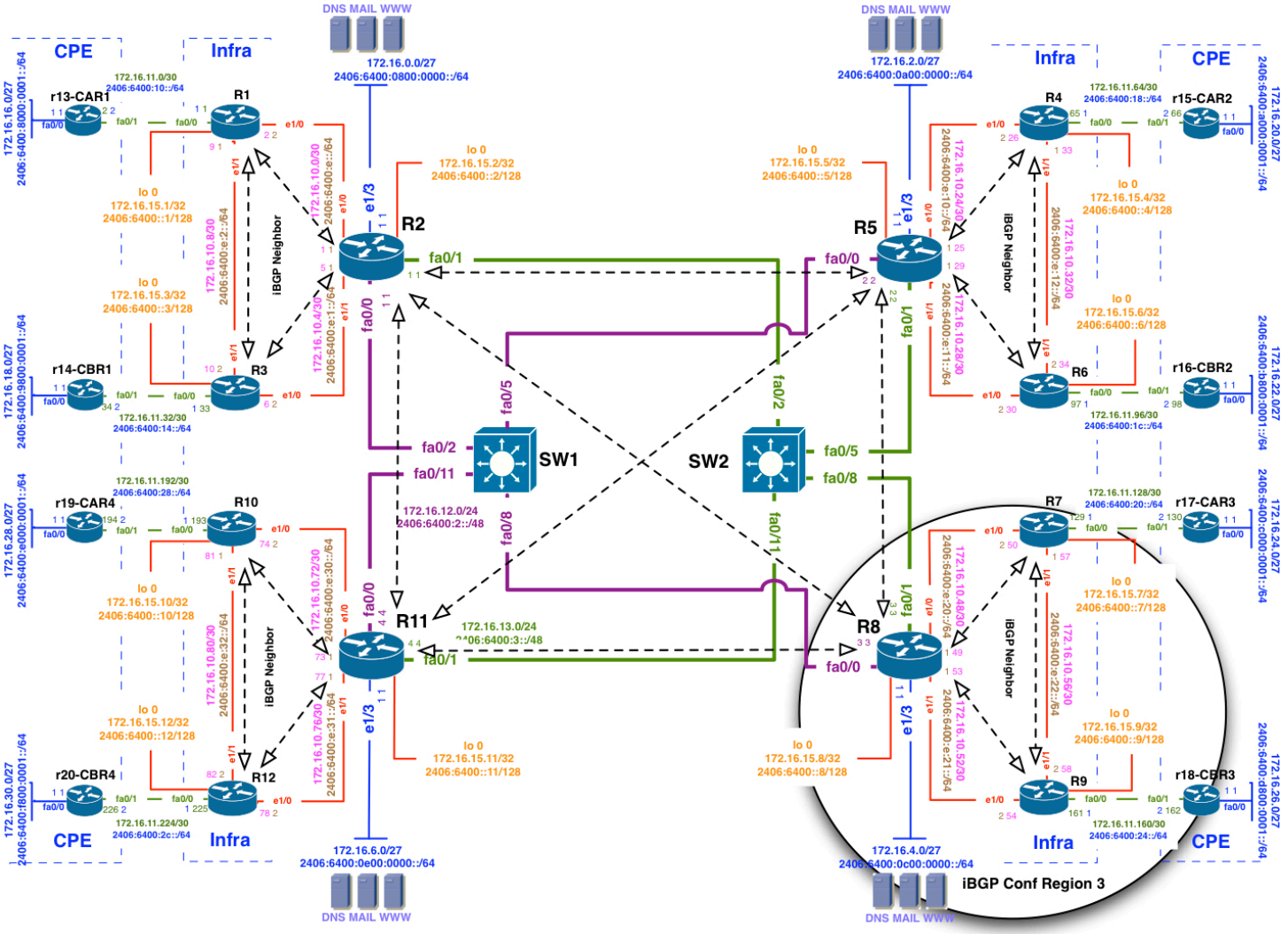
Route reflector configuration:

```
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
```

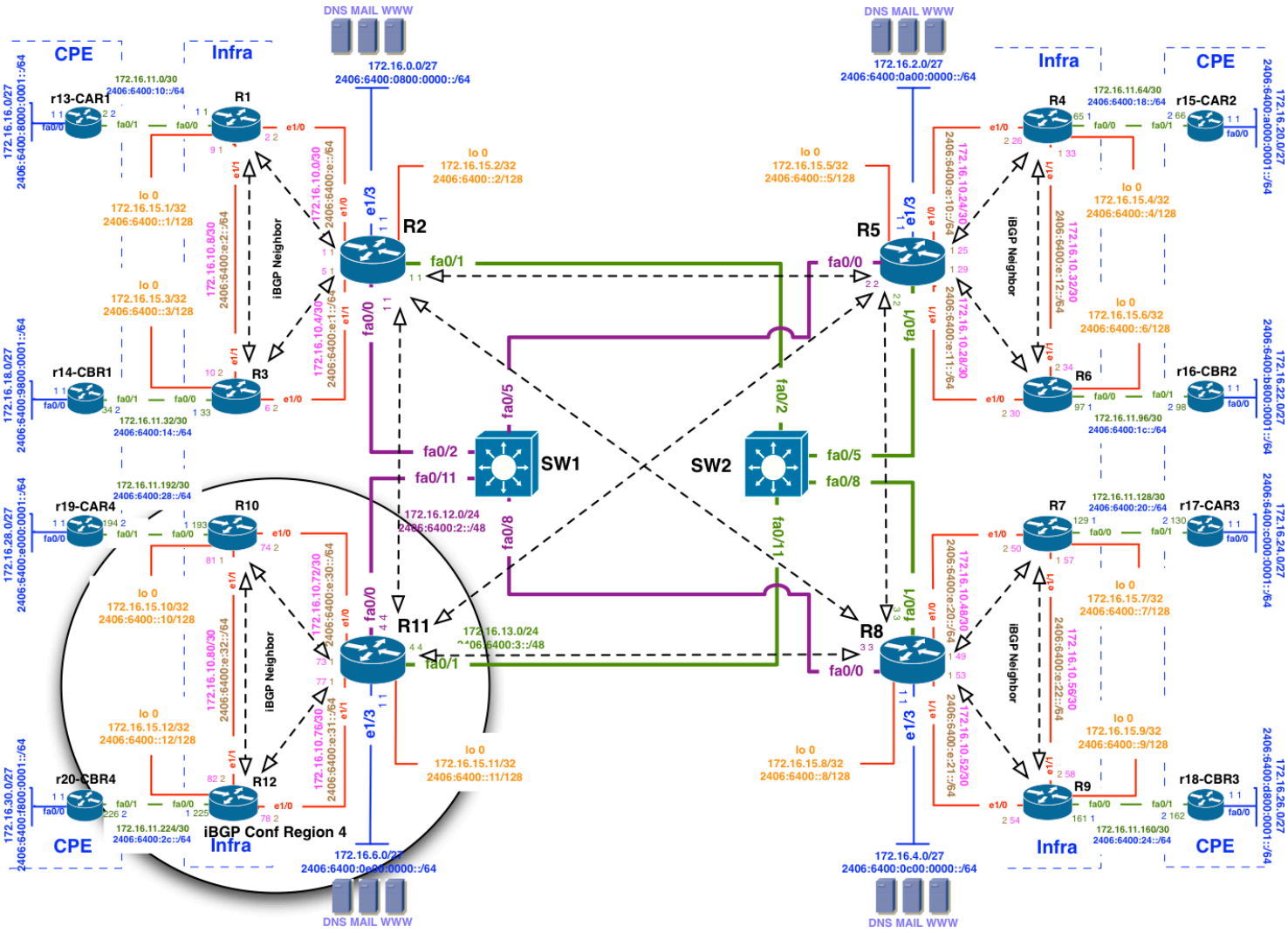
iBGP Peering For Region 2



iBGP Peering For Region 3



iBGP Peering For Region 4



IPv6 IOS Command For eBGP

Adding eBGP Neighbor:

```
router bgp 17821
address-family ipv6
!
neighbor 2406:6400:000D:0000::5 remote-as 45192
neighbor 2406:6400:000D:0000::5 activate
```

eBGP neighbor is always recommended with directly connected interface

References

- IPv6 Addressing – ISP/IXP
  - Philip Smith
  - http://www.bgp4all.com.au/dokuwiki/_media/workshops/02-ipv6-addressing.pdf
- Cisco ISP Essentials: Barry Greene & Philip Smith

Thank You!

END OF SESSION