Intro to Routing Workshop

July 2-5, 2013 – Nuku'alofa, Tonga

In conjunction with PacNOG

APNIC

Presenter

Champika Wijayatunga

Training Unit Manager, APNIC

Champika is responsible for managing its training activities in the Asia Pacific region and brings several years of experience, having worked in a number of countries in the IT industry, academia, research, and training environments.

Areas of interests:

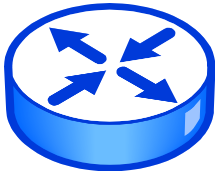
Internet Resource Management, IPv6, DNS/DNSSEC, Network Security, Routing

Contact: champika@apnic.net

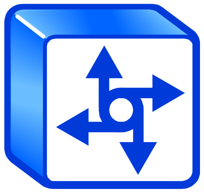
Routing Fundamentals

APNIC

Graphics / Symbols Used



Router
(layer 3, IP datagram forwarding)



Network Access Server
(layer 3, IP datagram forwarding)

Ethernet switch
(layer 2, packet forwarding)

What is a Routing Protocol?

- A set of rules defined to facilitate the exchanges of routing information between routers (Layer 3 device) inside networks

- Builds routing tables dynamically based on updates from its neighbours
  - Allows the router to find the best path in a network that has more than one path to a remote network.

- Maintains connectivity between devices within the network.

Routing Protocol Behaviour

- Updates Layer 3 routing devices, to route the data across the best path
- Learns participating routers advertised routes to discover their neighbors
- Learned routes are stored inside the routing table

What is Routing?

- Routing is the method of delivering an item from one location to another
  - Example Post Mail = delivery is being done via Post Office

- In a router network environment, it forwards traffic to a logical device destination interface.

- Routers perform two functions to deliver the packets to their destination:

  1. **Routing:** Learning the logical topology of the network to store the path inside the routing table to where the traffic should flow

  2. **Switching**: Forwarding the packets from an inbound interface to the outbound interface within the router

Distinction between *Routed* and *Routing* Protocols

- Routed protocols
  - Layer3 datagram that carry the information required in transporting the data across the network

- Routing protocols
  - Handles the updating requirement of the routers within the network for determining the path of the datagram across the network

Routing and *Routed* Protocols

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

Routing Requirements

- Activation of the protocol suite from such devices participating in the network
- Knowledge of the network destination
  - Must have an available entry in the routing table
  - Must have a valid and current route entry
- Interface presenting the best route path
  - Outbound interface with the lowest metric path

Routing Information

A routing table entry must contain the following information:

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

Network Field

- Contains information of entries
  - Networks learned (destination logical network or subnets)
    - Manually (static or default routes)
    - Dynamically (learned from routing protocol as dynamic routes)

- Information recorded is the entry on where to forward traffic to its destination when the datagram is received.

Outgoing Interface Field

- Interface to where the router sends the datagram

- Informs the administrator of interface where the update came through

Metric Field

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

Routing Protocol Metrics

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

Administrative Distance

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

Administrative Distance Chart (Cisco)

Route sources	Default distance
Connected interface	**0**
Static route out an interface	**0**
Static route to a next hop	**1**
External BGP	**20**
IGRP	**100**
OSPF	**110**
IS-IS	**115**
RIP v1, v2	**120**
EGP	**140**
Internal BGP	**200**
Unknown	**255**

Next Hop Field

- Contains the destination address of the next forwarding router
  - Address of the next hop (outgoing interface) usually within the same subnet
  - iBGP (exemption to the rule)

- Identifies the next hop so that the router can create the Layer2 frame with the destination address

Routing Table Sample (Cisco)

Cisco-router#sh ip route

Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B - BGP
D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
E1 - OSPF external type 1, E2 - OSPF external type 2, E - EGP
i - IS-IS, L1 - IS-IS level-1, L2 - IS-IS level-2, ia - IS-IS inter area
* - candidate default, U - per-user static route, o - ODR
P - periodic downloaded static route

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

Sample only

202.41.143.0/24 is variably subnetted, 2 subnets, 2 masks
S 202.41.143.0/24 is directly connected, Null0
C 202.41.143.17/32 is directly connected, Loopback0
O E2 10.110.0.0 [110/5] via 10.119.254.6, 0:01:00, Ethernet2
O 10.67.10.0 [110/128] via 10.119.254.244, 0:02:22, Ethernet2
B 217.170.115.0/24 [20/0] via 12.123.29.249, 5d16h

Routing Table Updates

- Routing table entry accuracy is required to make sure of the following:
  - Table entries are current and correct
  - New networks are inserted into the table
  - Best path is available to reach the destination network
  - Alternative routes are available to reach the destination network
  - Networks that is no longer available should not be seen in the routing table
- Depends on the routing protocol

Routing Decisions

- The main goal of the routing decision is to maintain a path that is valid and free from routing loops to the destination network, regardless of whether it is single hop or a multiple hop path

- The decision is made based on the metric value in the routing table
  – Using the sum of the metrics associated with the default routing protocol value and the intermediate connections

Router Traffic Forwarding

Forwards traffic to the out-bound interface

•Routing table entry ensures that the network topology is learnt

•Routing table entries that contain the information of the routes learned from other routers

Types of Routing Protocols

- Routing protocols are essentially applications inside the router designed to ensure correct and timely exchange of information within the network

- The IP routing protocol has several distinctions which can be divided into different groups
  - The first is the group is how the routing protocol handles the subnet mask sent during the routing update
  - The early routing protocols don't support this but the newer one's like RIP2, OSPF, ISIS, BGP4 support it. The distinctions are called "**classful** and **classless**"

Classful routing protocol (obsolete)

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

Classless Routing Protocol

- Classless routing protocol was designed to overcome the constraints from classful routing
  - OSPF
  - IS-IS
  - EIGRP (Cisco)
  - BGP
  - lead also to the development of RIP2

Characteristics of Classless Routing

- Support for different subnet mask values
  - Routers can be configured to have different subnet masks (VLSM)

- Supports route summarisation
  - Manual configuration

- Supports Classless Interdomain Routing (CIDR) architecture

Distance Vector Routing Protocol

- Another distinction based on the behavior of learning the path to the destination network
  - Distance vector and
  - Link-State protocols

- Distance vector routing protocol
  - Early technology of IP routing protocols (RIPv2)
  - Concept design was for small networks

Distance Vector Operation

- Maintains its own table by sending its own modified table for updates
- Sends updates to directly connected neighbors
  - is done in a periodic manner. This is commonly known as the (routing by rumor)
  - the timer needs to reach the expiration stage before the next update will be sent

NOTE: Updating of the tables affects the entire routing table, except those networks learned from interfaces where update is sent

Link-State Routing Protocol

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

- Routing protocols that support link-state are:
  - OSPF
  - IS-IS

Link-state Operation

- Link state operates through its main concern focusing on the link connected to the router (not the routes)
- Changes in the link state is propagated to its neighboring routers
  - to maintain the same image of the network topology among each neighbors
- Uses the router's CPU resources instead of bandwidth like Distance Vector
- When there is a state change the *incremental update* is sent to neighboring routers immediately,
  - remains silent if there's no change in the link state.

Link-state Operation

- Maintains the topology map of the network for easy local network table updates via incremental updates
  - OSPF = Link State Advertisement (LSA)
  - Dijkstra algorithm used to compute the new topology map of the network

- Metric used in Link-state is stated as "*cost*"
  - Equipment vendors default value setting can be overridden in manual configuration

Link-State Comparison Chart

Characteristic	**OSPF**	**IS-IS**
Hierarchical topology	X	X
Retains knowledge of all possible routes	X	X
Routes summarisation - manual	X	X
Event-triggered announcement	X	X
Load balancing – equal path	X	X
VLSM support	X	X
Routing algorithm	**Dijkstra**	**IS-IS**
Metric	**Cost**	**Cost**
Hop Count limit	**Unlimited**	**1024**
Scalability	**Large**	**Vry-Lrg**

Interior and Exterior Routing Protocols

- Other group distinctions with routing protocols are simplified as the protocols used for internal or external networks
- Interior
  - Routing protocol used to maintain routes within the organisation
  - Routing protocols that support it are:
    - *RIPv2, OSPF, IS-IS, EIGRP (Cisco)*
- Exterior
  - Routing protocol used to maintain routes connecting to different organisations
  - Exchange routing information between organisations
    - Using Exterior Gateway Protocol (EGP)
  - Border Gateway Protocol version 4 (BGP-4)

Questions?

APNIC

Addressing Issues and Solutions

APNIC

IP Addressing Issues

- Exhaustion of IPv4 addresses
  - Wasted address space in traditional subnetting
  - Limited availability of /8 subnet addresses
- Internet routing table growth
  - Size of the routing table due to higher number prefix announcement
- Tremendous growth of the Internet

IP Addressing Solutions

- Subnet masking and summarisation
  - Variable-length subnet mask definition
  - Hierarchical addressing
  - Classless InterDomain Routing (CIDR)
  - Routes summarisation (RFC 1518)
- Private address usage (RFC 1918)
  - Network address translation (NAT)
- Development of IPv6 addresses

Subnetting Overview

- Allows the creation of additional sub-networks by simply moving the network boundary to the *right*
- When the contiguous *1s* are added, it indicates by how many bits the network portion will be extended
- The sub-network is calculated by the 2^n where "*n*" is the number of extended bits.

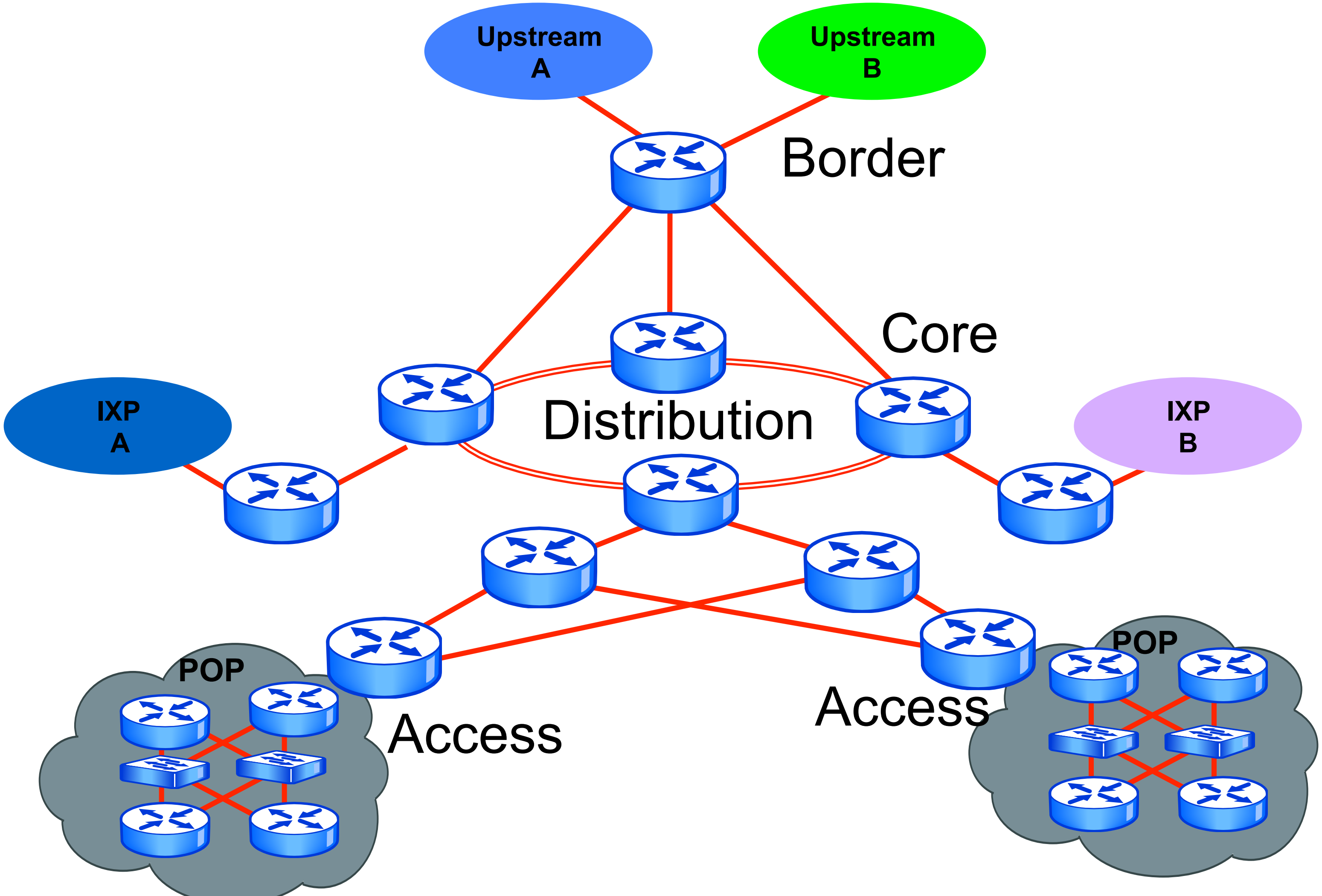
Addressing Hierarchy

- Support for easy troubleshooting, upgrades and manageability of networks

- Performance optimisation
  - Scalable and more stable
  - Less network resources overhead (CPU, memory, buffers, bandwidth)

- Faster routing convergence

Addressing Hierarchy Example



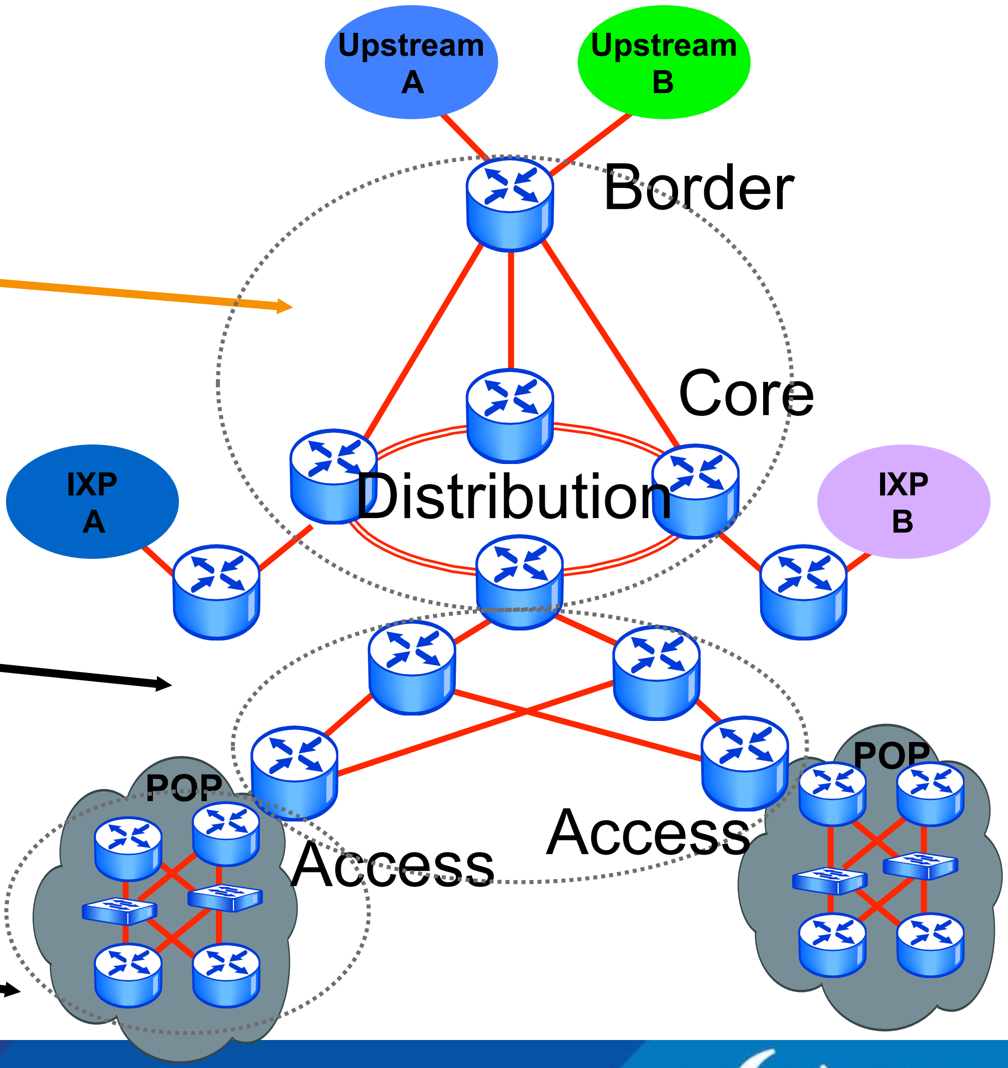
Addressing Hierarchical (cont.)

Network Number
192.168.0.0/16

Core
192.168.32.0/19

Distribution/Core
192.168.32.0/21

Access/Distribution
192.168.48.0/21

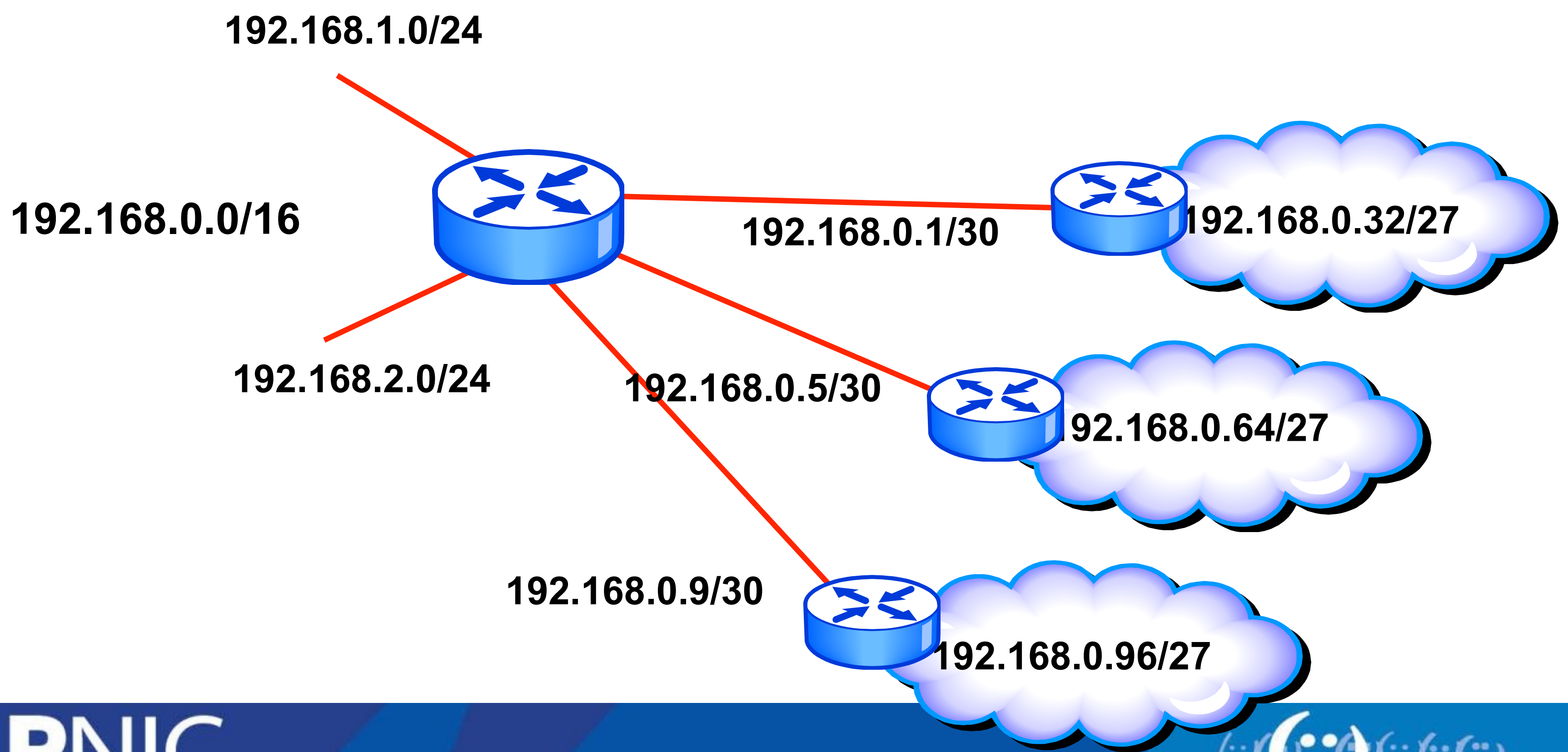


Variable Length Subnet Mask

- Allows the ability to have more than one subnet mask within a network
- Allows re-subnetting
  - create sub-subnet network addresses
- Increase the route capabilities
  - Addressing hierarchy
  - Summarisation

Calculating VLSM Example

- Subnet 192.168.0.0/24 into smaller subnet
  - Subnet mask with /27 and /30 (point-to-point)



APNIC

Calculating VLSM Example (cont.)

- Subnet 192.168.0.0/24 into smaller subnets
  - Subnet mask with /30 (point-to-point)

Description	Decimal	Binary
Network Address	192.168.0.0/30	x.x.x.000000**00**
1st valid IP	192.168.0.1/30	x.x.x.000000**01**
2nd valid IP	192.168.0.2/30	x.x.x.000000**10**
Broadcast address	192.168.0.3/30	x.x.x.000000**11**

Calculating VLSM Example (cont.)

- Subnet 192.168.0.0/24 into smaller subnets
  - Subnet mask with /27

Description	Decimal	Binary
Network Address	192.168.0.32/27	x.x.x.000**00000**
Valid IP range 192.168.0.33 - 192.168.0.62		x.x.x.000**00001**
		x.x.x.000**00010**
Broadcast address	192.168.0.63/30	x.x.x.000**11111**

Calculating VLSM Example (cont.)

- Subnet 192.168.0.0/24 into smaller subnets
  - Subnet mask with /27

Description	Decimal	VSLM	Host	Host range
1st subnet	192.168.0.0/27	**x.x.x.000**	**00000**	0-31
2nd subnet	192.168.0.32/27	**x.x.x.001**		31-63
3rd subnet	192.168.0.64/27	**x.x.x.010**		64-95
4th subnet	192.168.0.96/27	**x.x.x.011**		96-127

n = 5 (n is the remaining subnet bits)
2n – 5 = 30 host per subnet

Questions?

Summarisation of Routes

APNIC

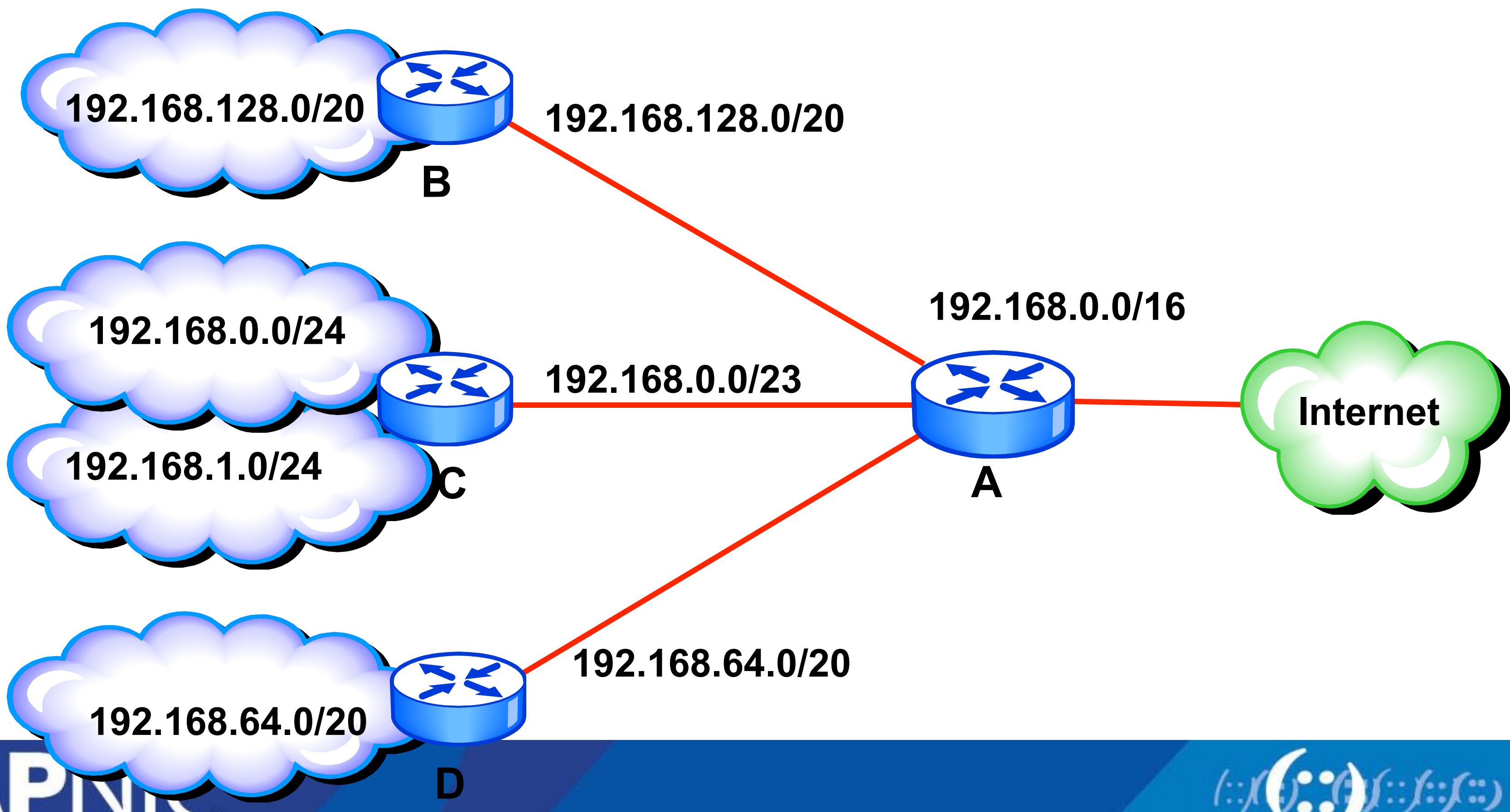
Route Summarisation

- Allows the presentation of a series of networks in a single summary address.

- Advantages of summarisation
  - Faster convergence
  - Reducing the size of the routing table
  - Simplification
  - Hiding Network Changes
  - Isolate topology changes

Summarisation Example

- Router C summarises its networks (2 x/24) before announcing to its neighbors (routers B and D)
- Router A combined the networks received from B, C, D and announce it as single /16 routing to Internet



Route summarisation

- Subnet 192.168.0.0/24 and 192.168.1.0/24 combining then to become a bigger block of address "/23"

Network	Subnet Mask	Binary
192.168.0.0	255.255.255.0	x.x.0000**0000.x**
192.168.1.0	255.255.255.0	x.x.0000**0001.x**
Summary	192.168.0.0/23	x.x.0000**0010.x**
192.168.0.0	255.255.252.0	x.x.0000**0010.x**

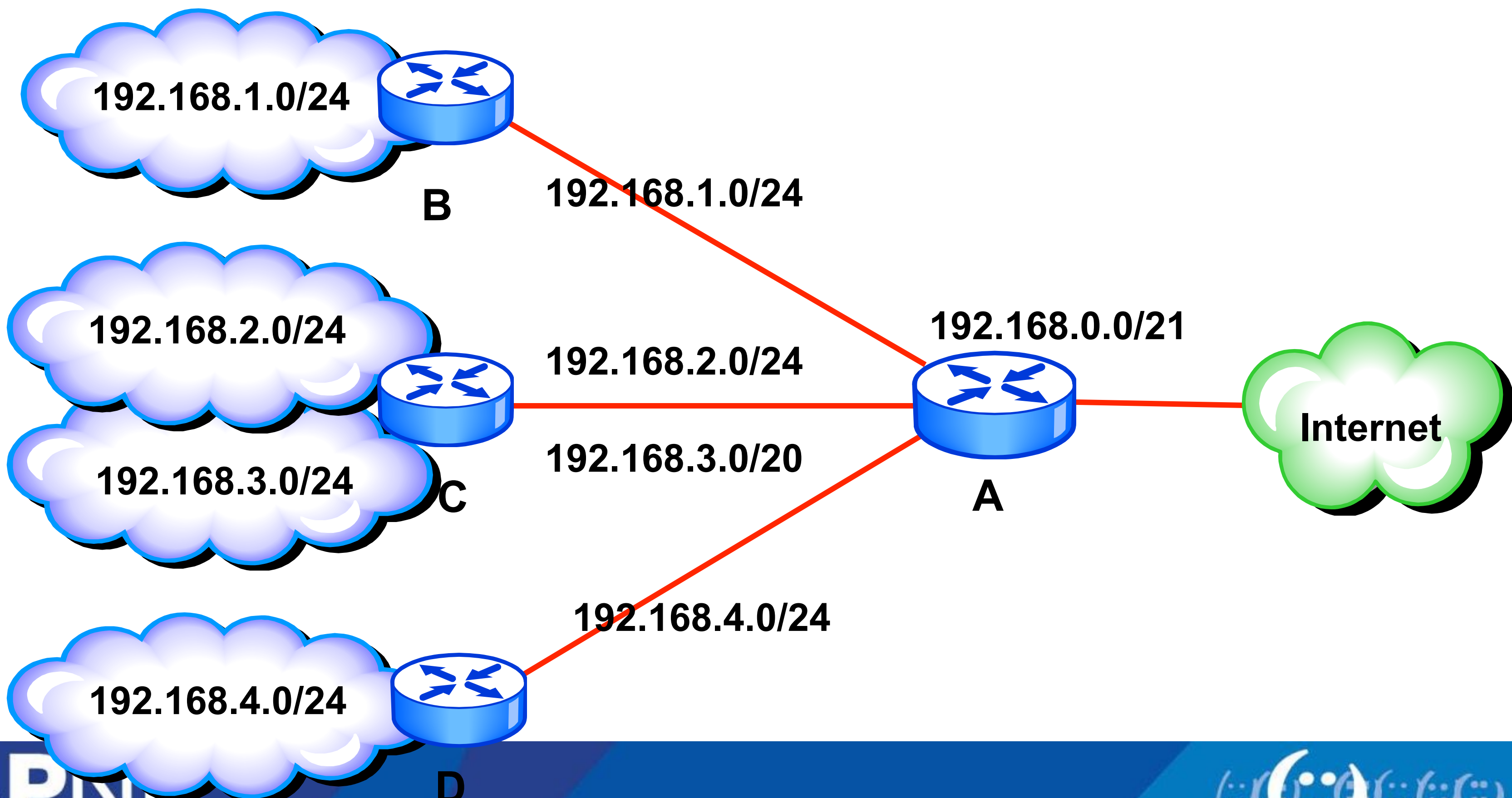
Discontiguous Networks

- A network not using routing protocols that support VLSM creates problems
  - Router will not know where to send the traffic
  - Creates routing loops or duplication

- Summarisation is not advisable for networks that are discontiguous
  - Turn off summarisation
    - Alternative solution but understand the scaling limitation
    - Find ways to re-address the network
  - Can create disastrous situation

CIDR Solution Advantage

- CIDR offers the advantage of reducing the routing table size of the network by summarising the ISP announcement into a single /21 advertisement



Questions?

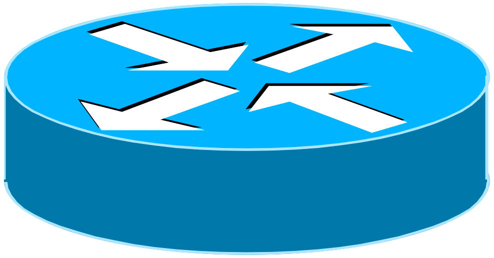
APNIC

More into Routing…

APNIC

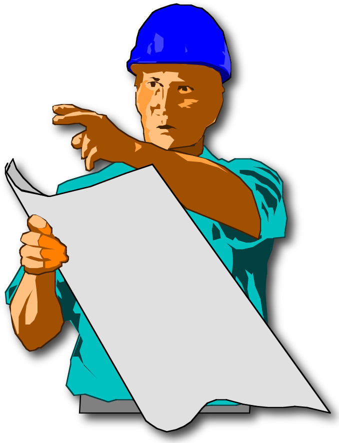
A day in a life of a router

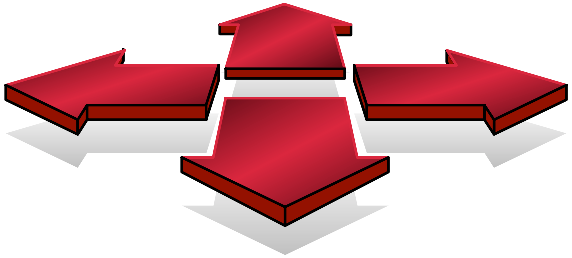
- find path
- forward packet, forward packet, forward packet, forward packet...
- find alternate path
- forward packet, forward packet, forward packet, forward packet…
- repeat until powered off



Routing versus Forwarding

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





IP Routing – finding the path

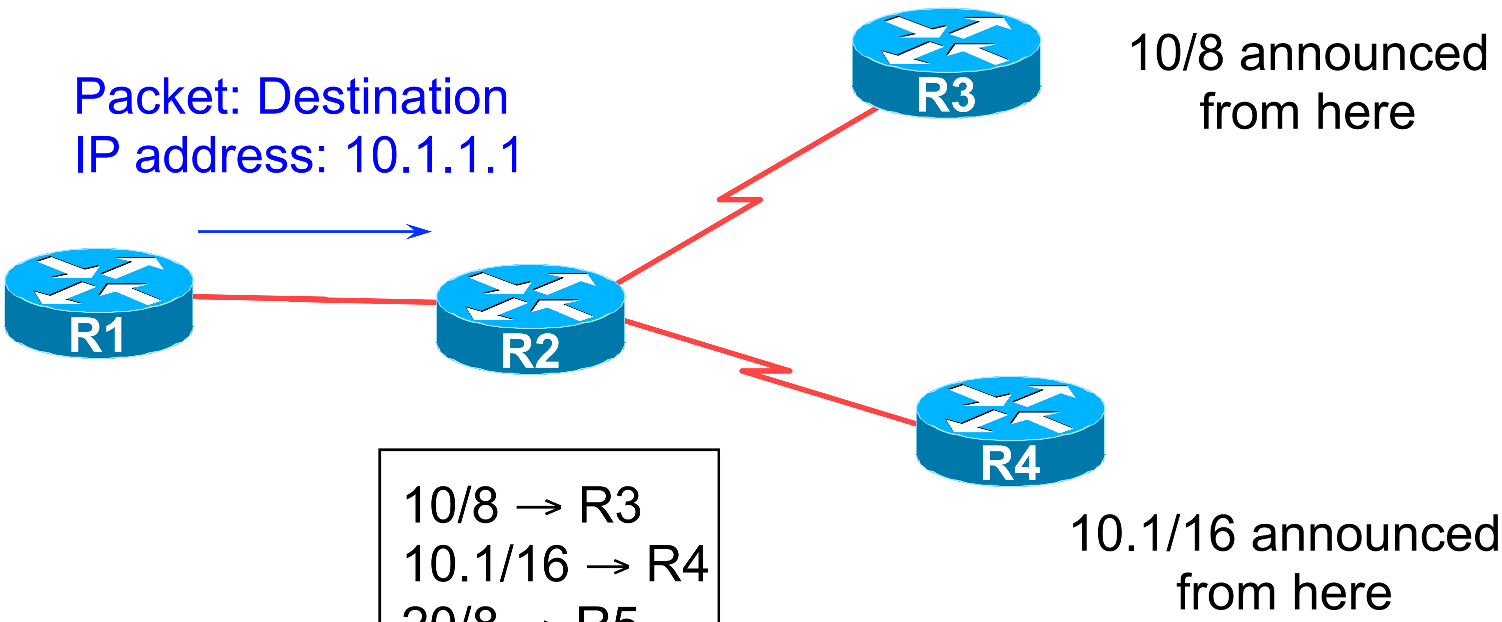
- Path derived from information received from a routing protocol
- Several alternative paths may exist
  - best path stored in forwarding table
- Decisions are updated periodically or as topology changes (event driven)
- Decisions are based on:
  - topology, policies and metrics (hop count, filtering,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

IP route lookup

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

IP route lookup

- Based on destination IP address

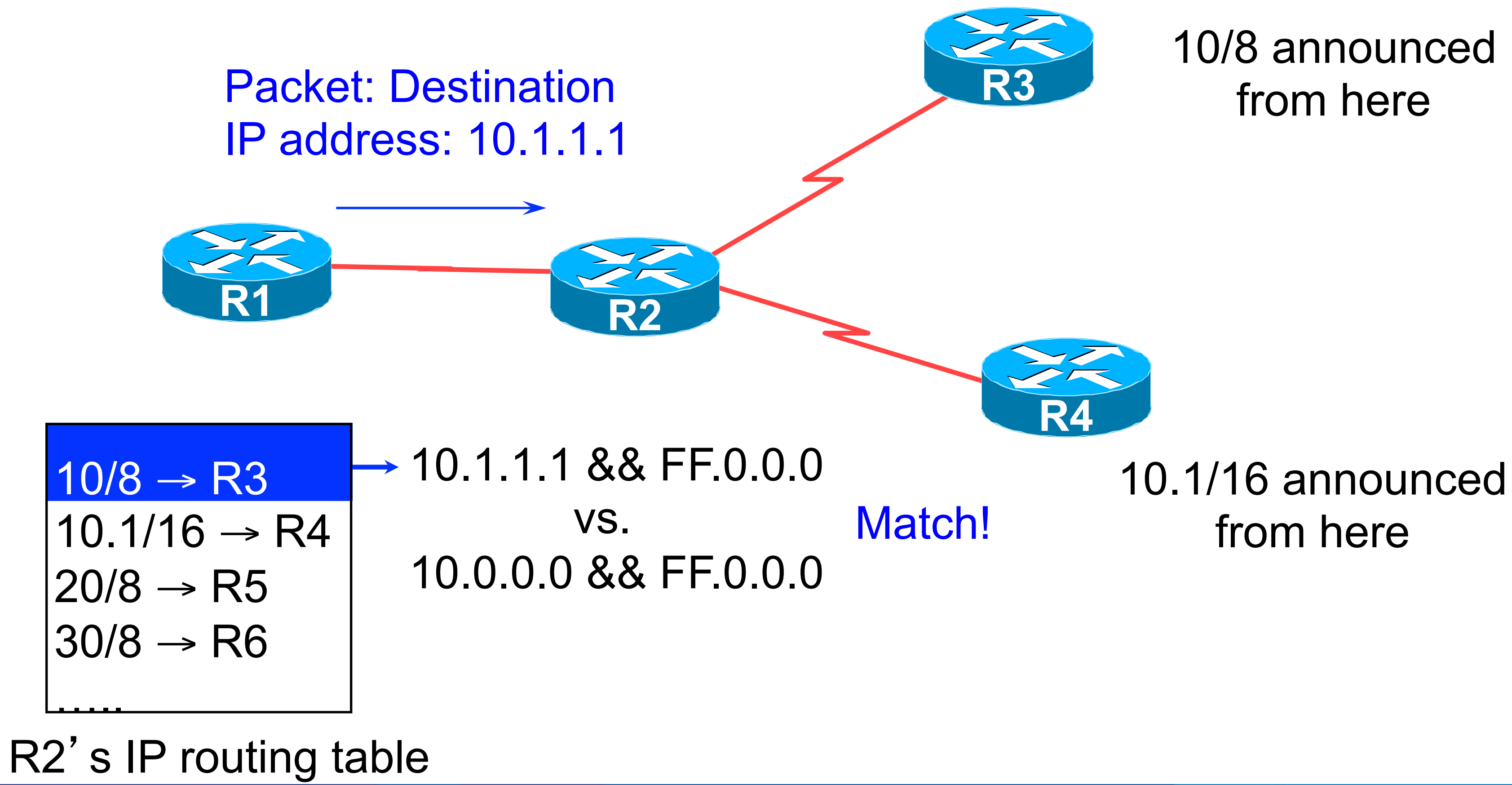


10/8 → R3
10.1/16 → R4
20/8 → R5
30/8 → R6
…..

R2's IP routing table

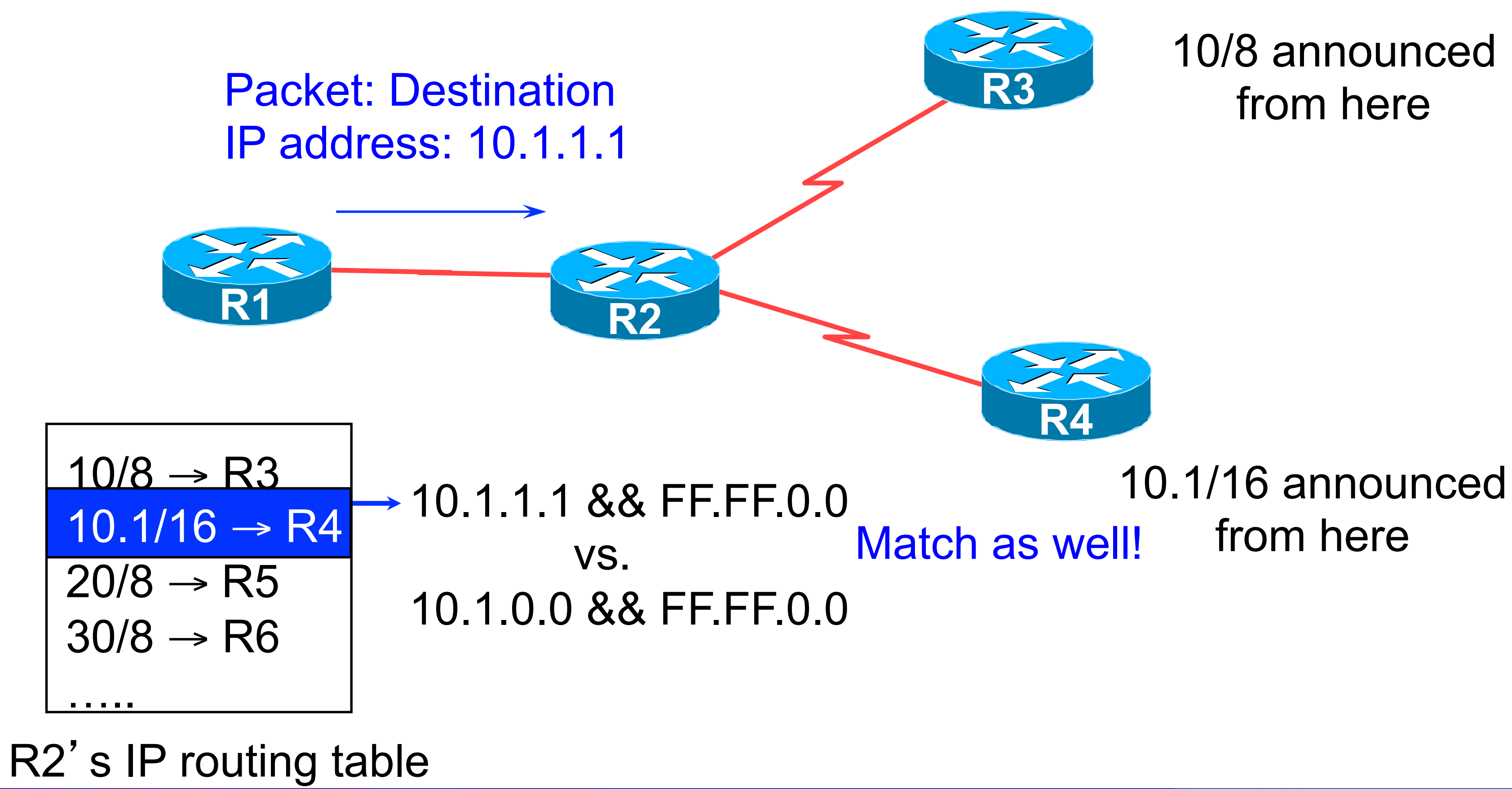
IP route lookup: Longest match routing

- Based on destination IP address



IP route lookup: Longest match routing

- Based on destination IP address

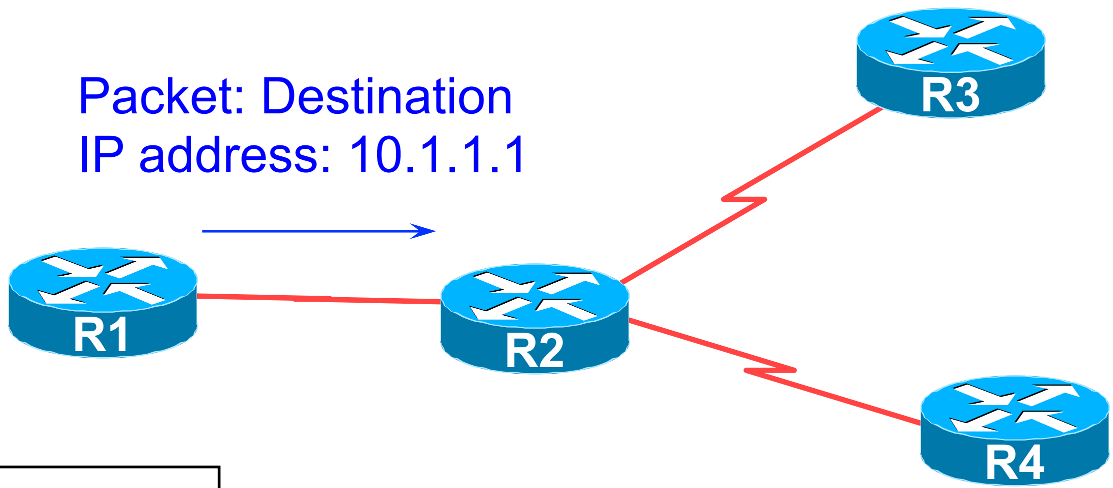


IP route lookup: Longest match routing

- Based on destination IP address

Packet: Destination IP address: 10.1.1.1

10/8 announced from here



10.1/16 announced from here

10/8 → R3
10.1/16 → R4
20/8 → R5
30/8 → R6
…..

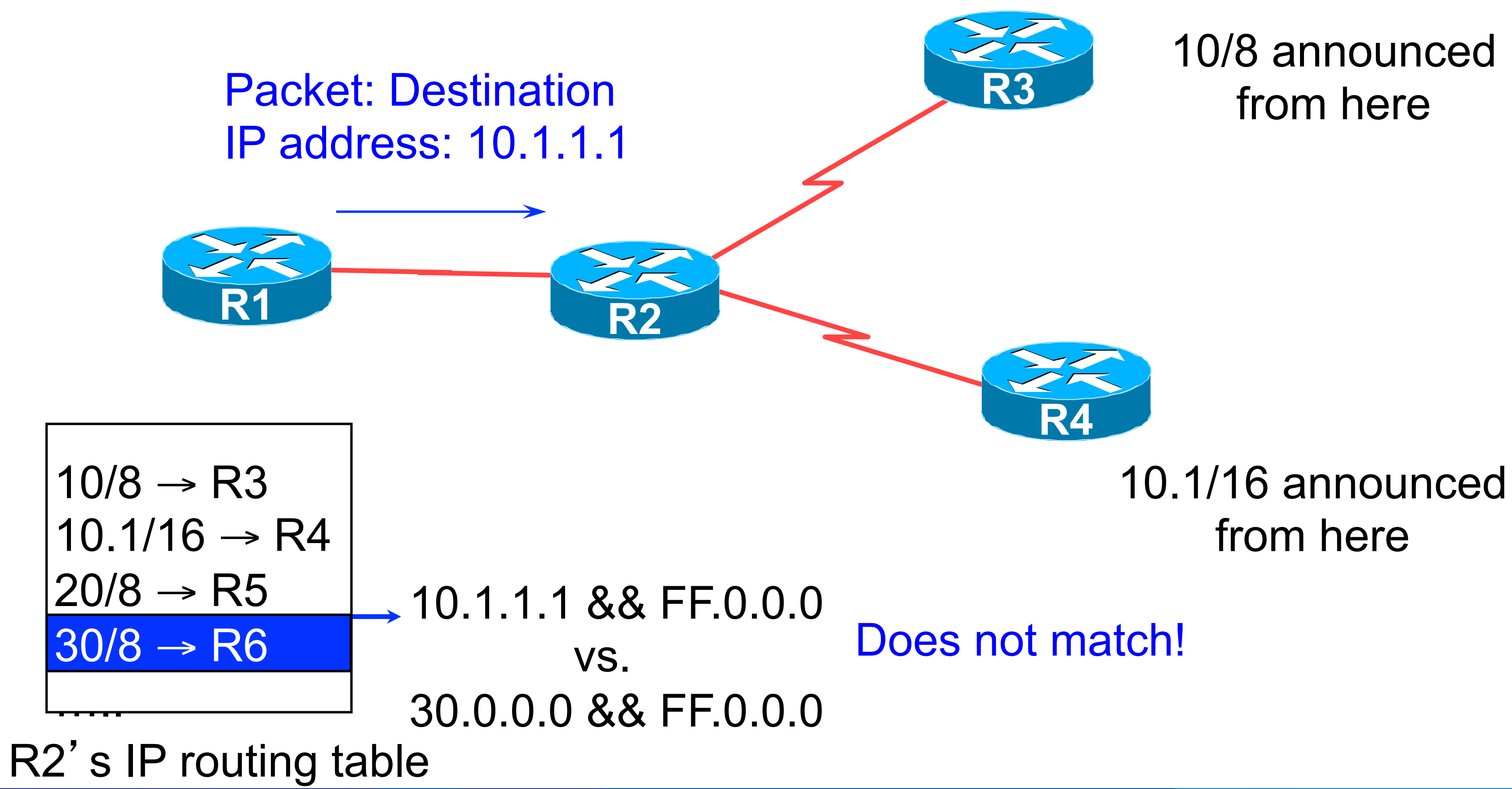
R2’s IP routing table

10.1.1.1 && FF.0.0.0
vs.
20.0.0.0 && FF.0.0.0

Does not match!

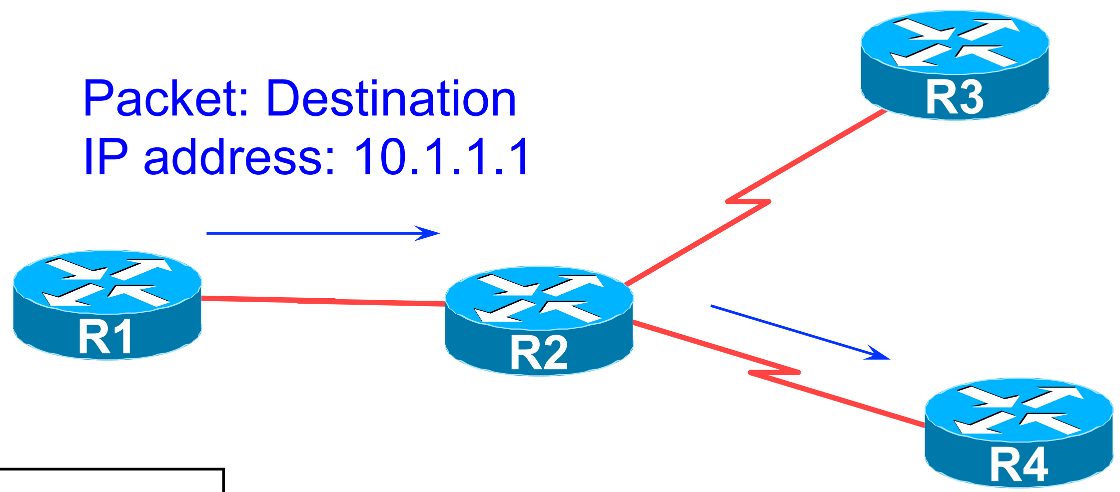
IP route lookup: Longest match routing

- Based on destination IP address



IP route lookup: Longest match routing

- Based on destination IP address



10/8 announced from here

10.1/16 announced from here

10/8 → R3
10.1/16 → R4
20/8 → R5
30/8 → R6
…..

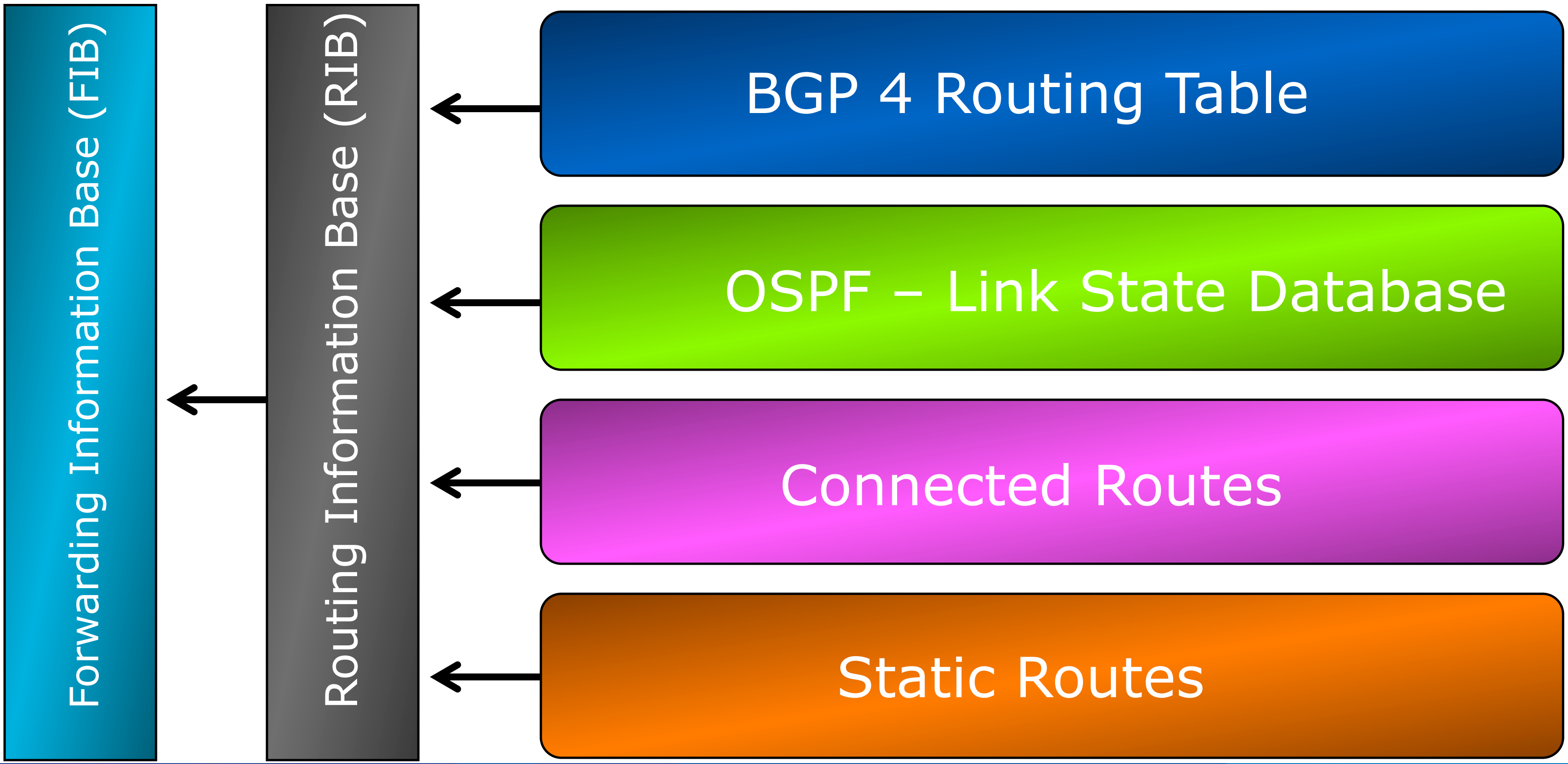
→ 10.1/16 → R4 ← Longest match, 16 bit netmask

R2's IP routing table

IP Forwarding

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

Routing Tables Feed the Forwarding Table



RIBs and FIBs

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

Explicit versus Default Routing

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

Egress Traffic

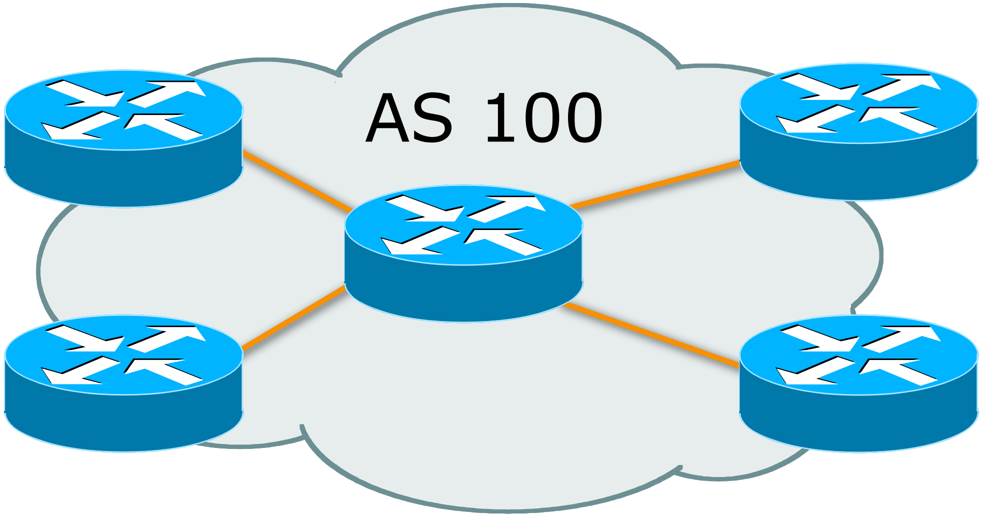
- How packets leave your network
- Egress traffic depends on:
  - route availability (what others send you)
  - route acceptance (what you accept from others)
  - policy and tuning (what you do with routes from others)
  - Peering and transit agreements

Ingress Traffic

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

Autonomous System (AS)

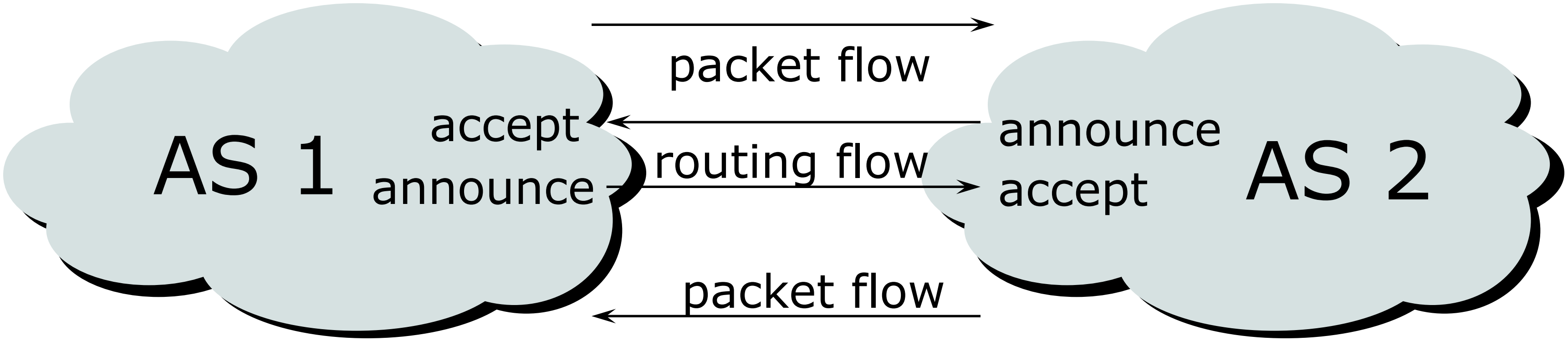
- Collection of networks with same routing policy
- Single routing protocol
- Usually under single ownership, trust and administrative control



Definition of terms

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

Routing flow and packet flow



For networks in AS1 and AS2 to communicate:

AS1 must announce to AS2

AS2 must accept from AS1

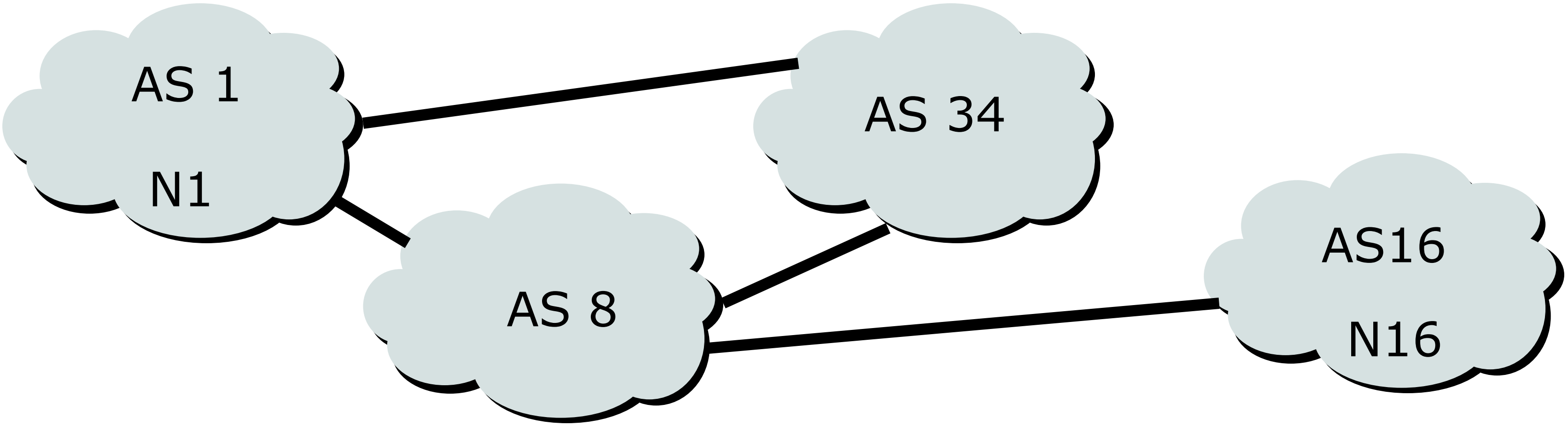
AS2 must announce to AS1

AS1 must accept from AS2

Routing flow and Traffic flow

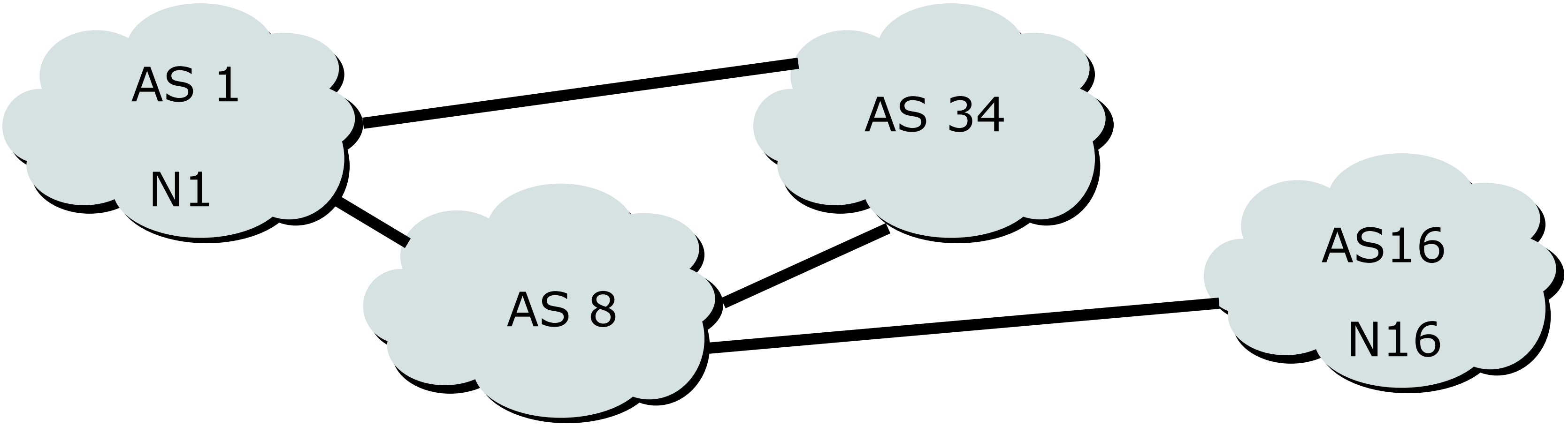
- Traffic flow is always in the opposite direction of the flow of Routing information
  - Filtering outgoing routing information inhibits traffic flow inbound
  - Filtering inbound routing information inhibits traffic flow outbound

Routing Flow/Packet Flow: With multiple ASes



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

Routing Flow/Packet Flow: With multiple ASes

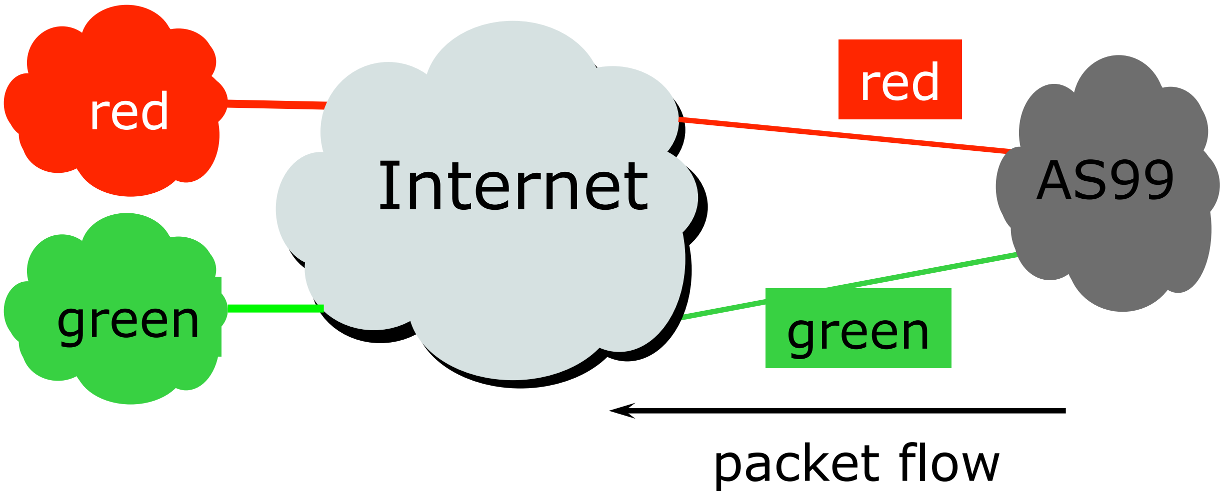


- As multiple paths between sites are implemented it is easy to see how policies can become quite complex.

Routing Policy

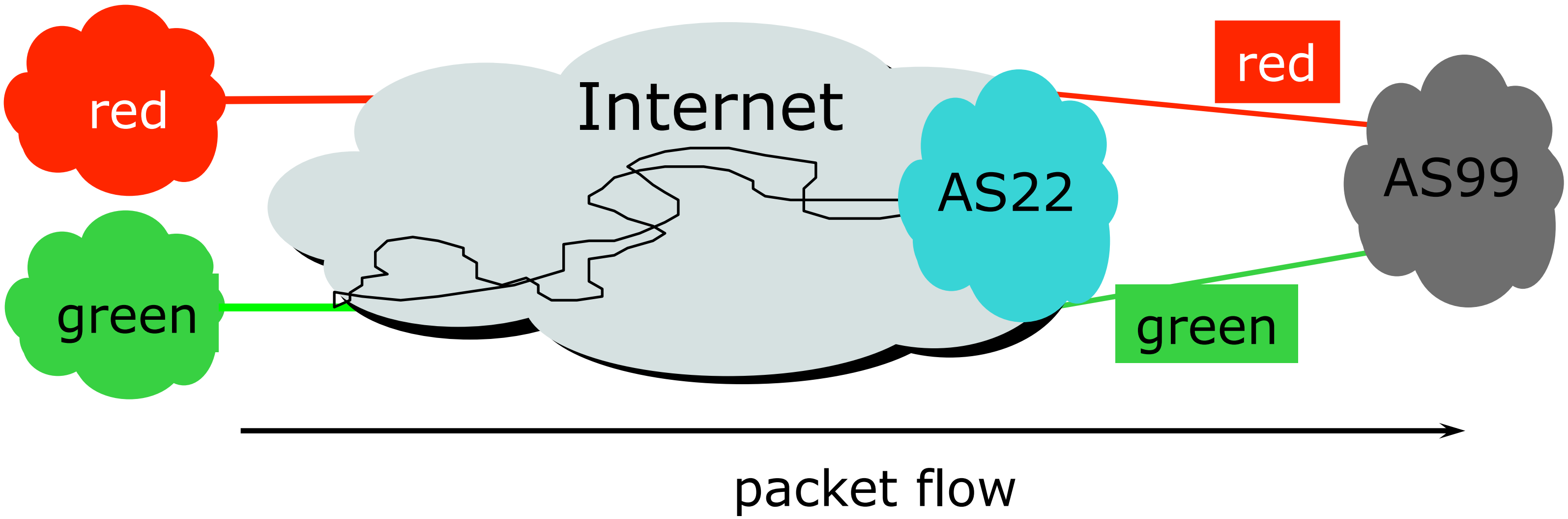
- Used to control traffic flow in and out of an ISP network
- ISP makes decisions on what routing information to accept and discard from its neighbours
  - Individual routes
  - Routes originated by specific ASes
  - Routes traversing specific ASes
  - Routes belonging to other groupings
    - Groupings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

Routing Policy Limitations



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

Routing Policy Limitations



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

Routing Policy Issues

- End June 2011:
  - 362000 prefixes
    - Not realistic to set policy on all of them individually
  - 38000 origin AS’s
    - Too many to try and create individual policies for
- Routes tied to a specific AS or path may be unstable regardless of connectivity
- Solution: Groups of AS’s are a natural abstraction for filtering purposes

1: How Does Routing Work?

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

2: How Does Routing Work?

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

3: How Does Routing Work?

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

4: How Does Routing Work?

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

Routing Protocols

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

What Is an IGP?

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

Why Do We Need an IGP?

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

What Is an EGP?

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

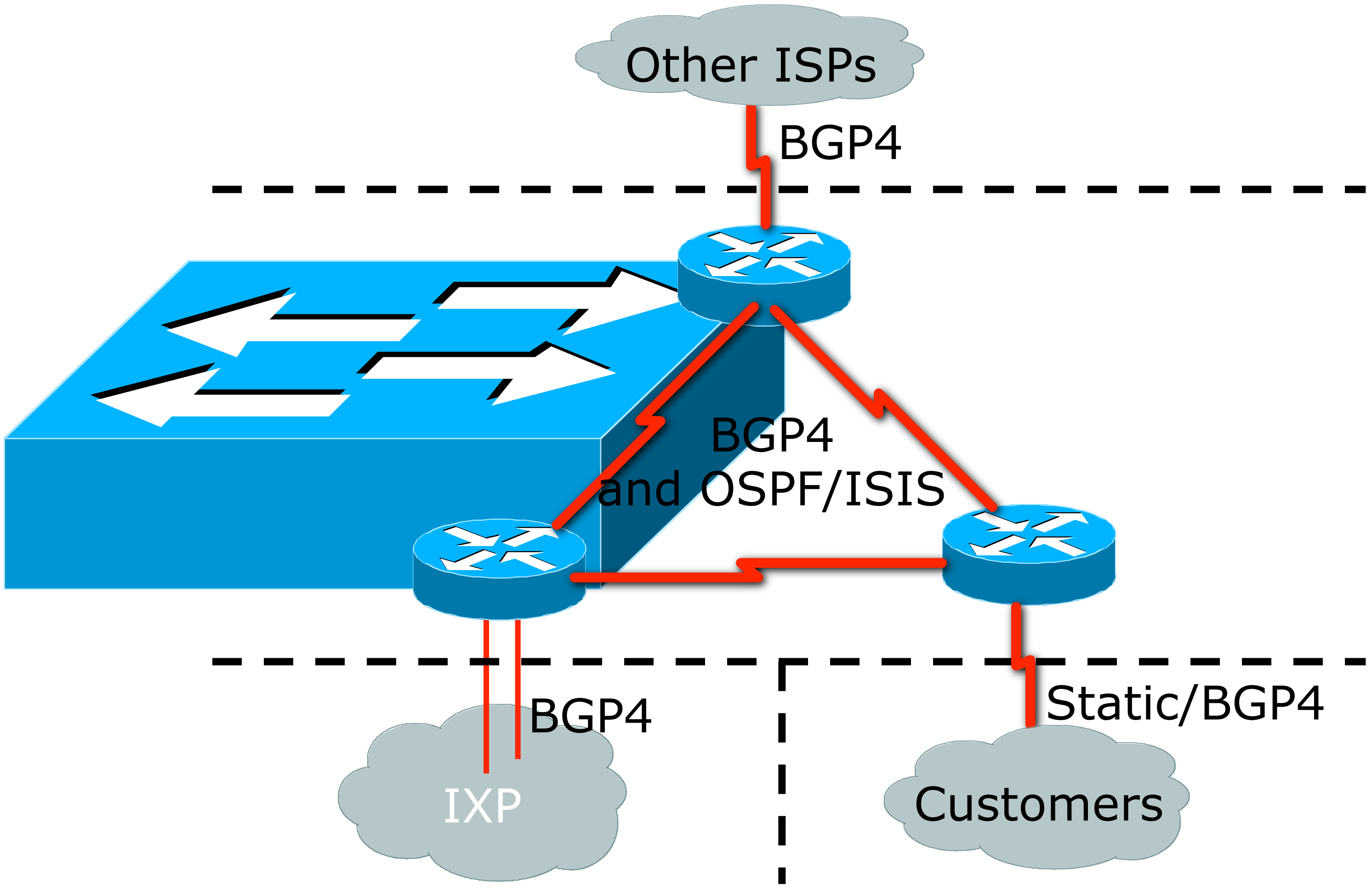
Why Do We Need an EGP?

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

Interior versus Exterior Routing Protocols

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

Hierarchy of Routing Protocols



FYI: Cisco IOS Default Administrative Distances

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

Questions?

APNIC

Network Planning Essentials

Hierarchical Network Design

- A network with different layers
  - Each level of the network has it own function

- Minimise costs
  - Avoid spending money to buy unnecessary features on equipment for each layer’s requirements
  - Save bandwidth due to modularity of design

- Scalability is the major goal
  - Fast convergence
  - Route summarisation

Disadvantages of a Flat Network

- Designed for small networks
  - Easy to design and maintain as long the network stays small
- No hierarchy
- All networking devices have the same jobs
- No layer divisions

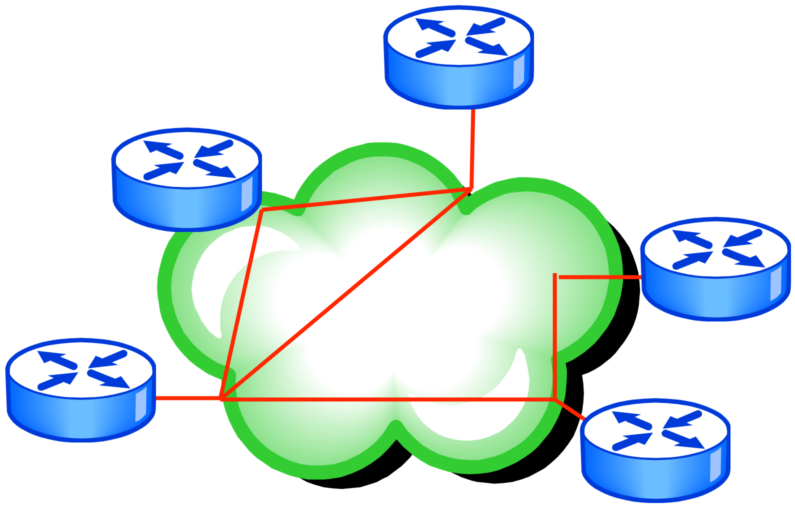
Partial Mesh vs Full-mesh Topology

- Network designers recommend mesh topology

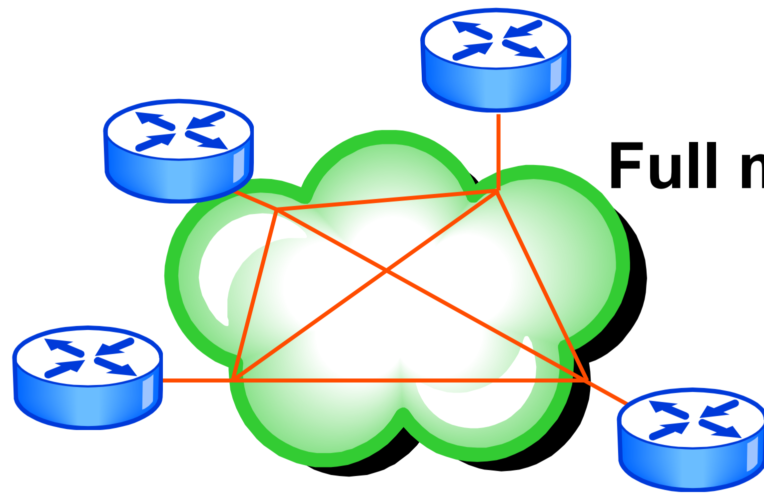
- Good performance and provides redundancy
  - Partial mesh topology
    - Has fewer connections
    - Each router may require direct connection from an intermediate link to get to another device
  - Full-mesh topology
    - All routers are connected to each other to offer good performance

Formula for full-mesh = (N x (N – 1))/2

Mesh versus Full-mesh Topology

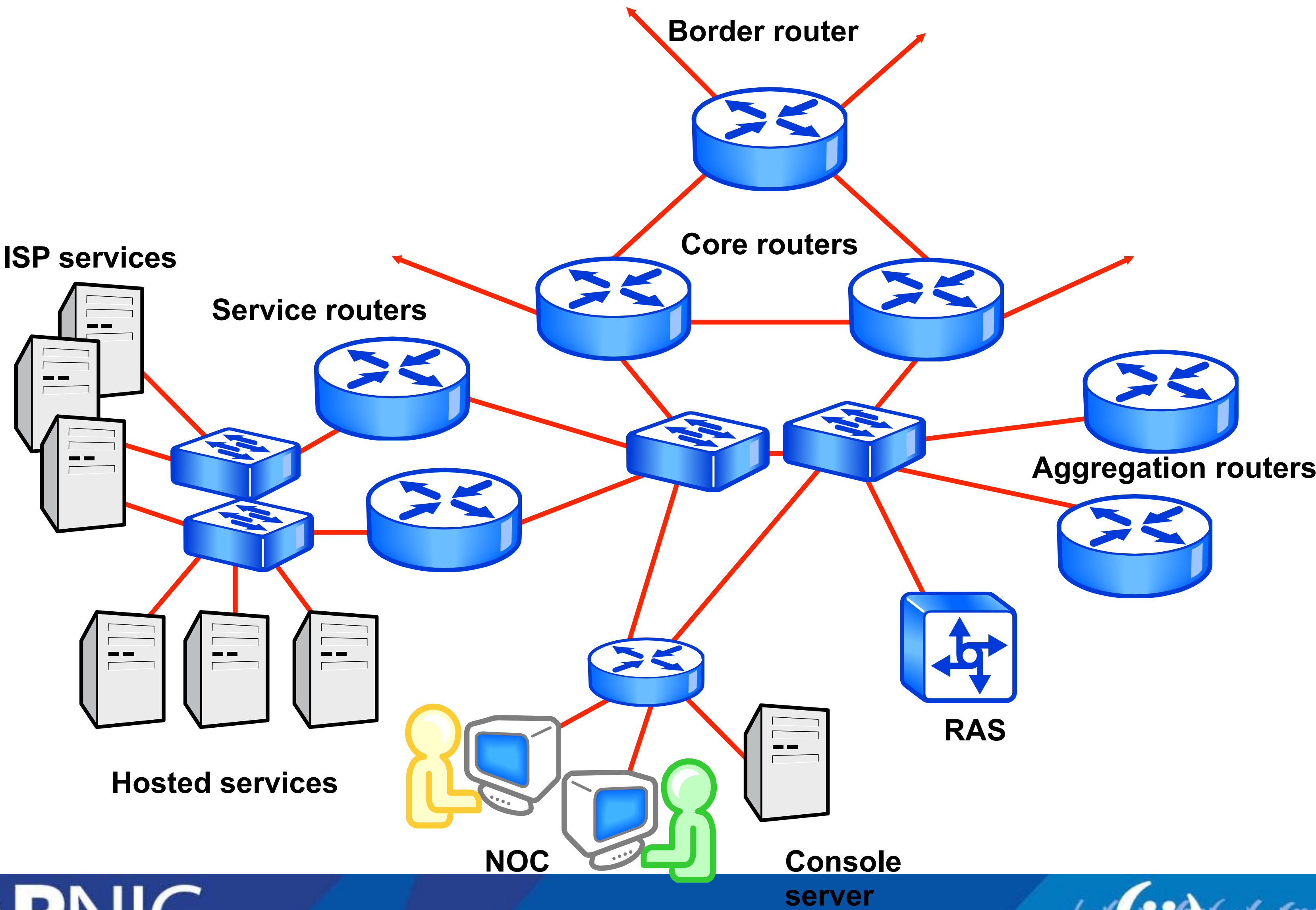


Partial mesh



Full mesh

Simple Network Plan



Border Router

- Provides connectivity to the rest of the Internet service providers in the world
- Protects the ISP network and the customers networks from the Internet
- Critical and should be correct because this is the main reason for the business connection

Core Router

- Critical for connectivity; it should be designed to have a redundant component

- When configurating this router:
  - Enable routing feature for optimised packet throughput
  - Avoid using filtering which will slow down manipulation of packets
  - Avoid usage of routing policy for filtering purposes

- Should be high-speed to switch packets easily and faster

Aggregation Router

- Aggregation or gateway router for connecting fixed line customers
- Improves routing protocol performance
- Allows summarisation of routes from an aggregated address
- Allows configuration of routing policy for customers network announcements

Services Router

- Used for services provided to customers
  - DNS, email, news
  - Hosted services (content provided)
    - Web, email, DNS
- Configured by default to have filters to allow only authorised users
- Routers with firewall features are often used as a firewall itself
- Protect the core services provided by the ISP

NOC Router

- Connects ISP essential services
  - Syslog, TACACS+, RADIUS, primary DNS

- Operations engineer network
  - Trouble isolations
  - Network monitoring
  - Research network testing
  - Staging area (option)

Access Router

- Routers designed to provide access services
  - Cable services (on demand)
  - DSL on demand service
  - Wireless services (Wifi) etc.

- Connections to this network requires proper authentication credentials

Out-of-band Console Server

- Can be typical router that has Async port configured for out-band access
- Allows remote access of routers without using the in-band network
- Access to routers through its console port
- Utilises a different network
  – Not affected if the in-band network is down
  – Only small bandwidth usage is required

Principles of Addressing

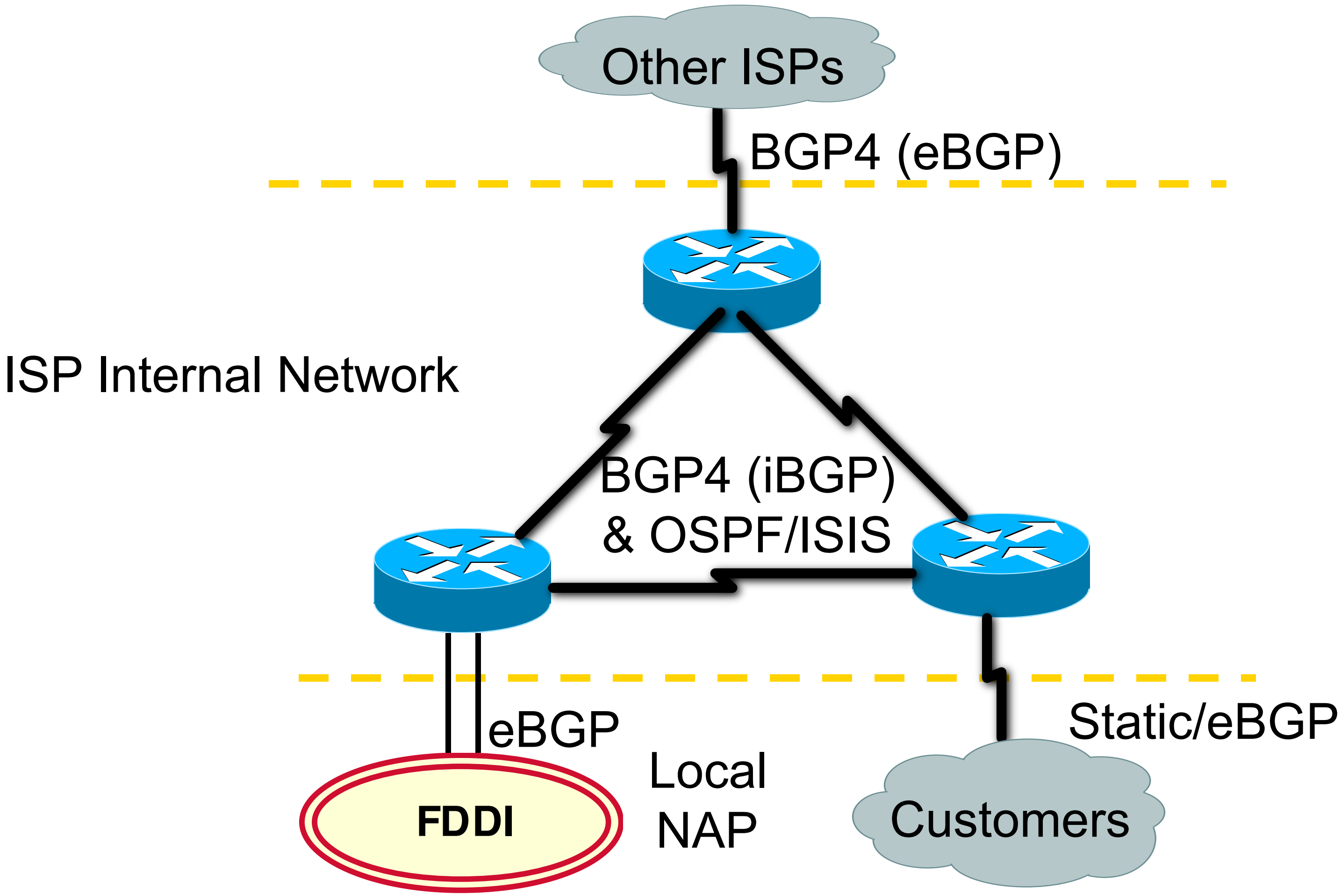
Separate customer & infrastructure address pools:

- Manageability
  - Different personnel manage infrastructure and assignments to customers
- Scalability
  - Easier renumbering
    - customers are difficult
    - infrastructure is relatively easy

Principles of Addressing

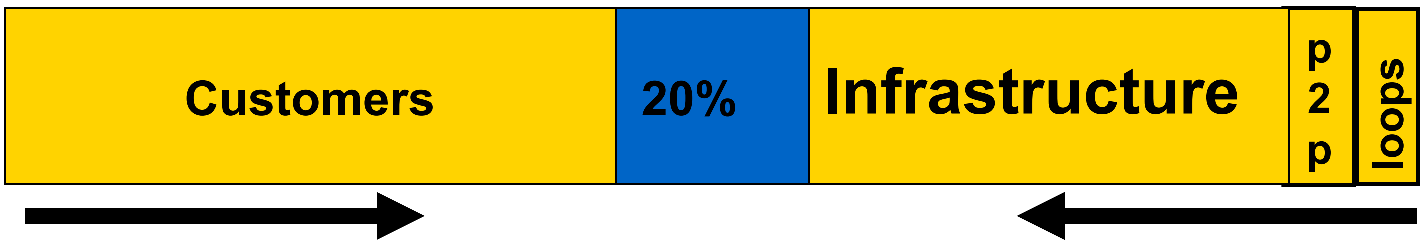
- Further separate infrastructure
  - ‘Static’ infrastructure examples
    - RAS server address pools, CMTS
    - Virtual web and content hosting LANs
    - Anything where there is no dynamic route calculation
- Customer networks
  - Carry in iBGP , do not put in IGP
    - No need to aggregate address space carried in iBGP
    - Can carry in excess of 100K prefixes

Hierarchy of Routing Protocols



Management – Simple Networks

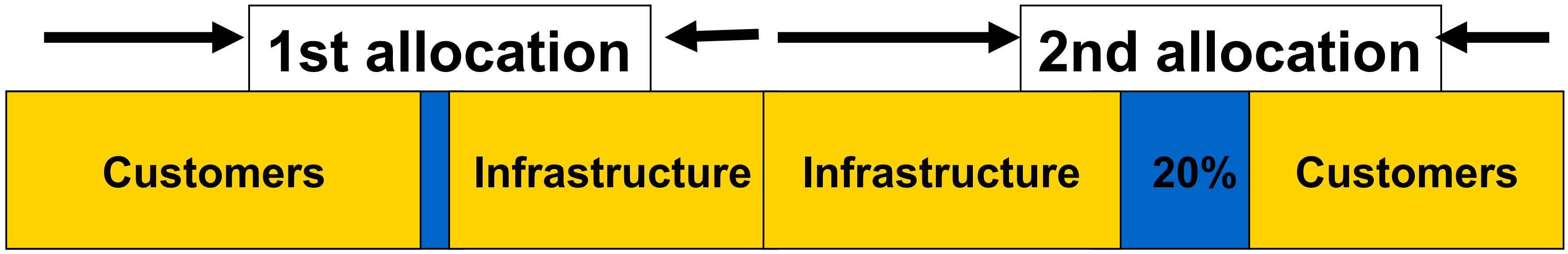
- First allocation from APNIC
  - Infrastructure is known, customers are not
  - 20% free is trigger for next request



  - Grow usage of blocks from edges
  - Assign customers sequentially

Management - Simple Network

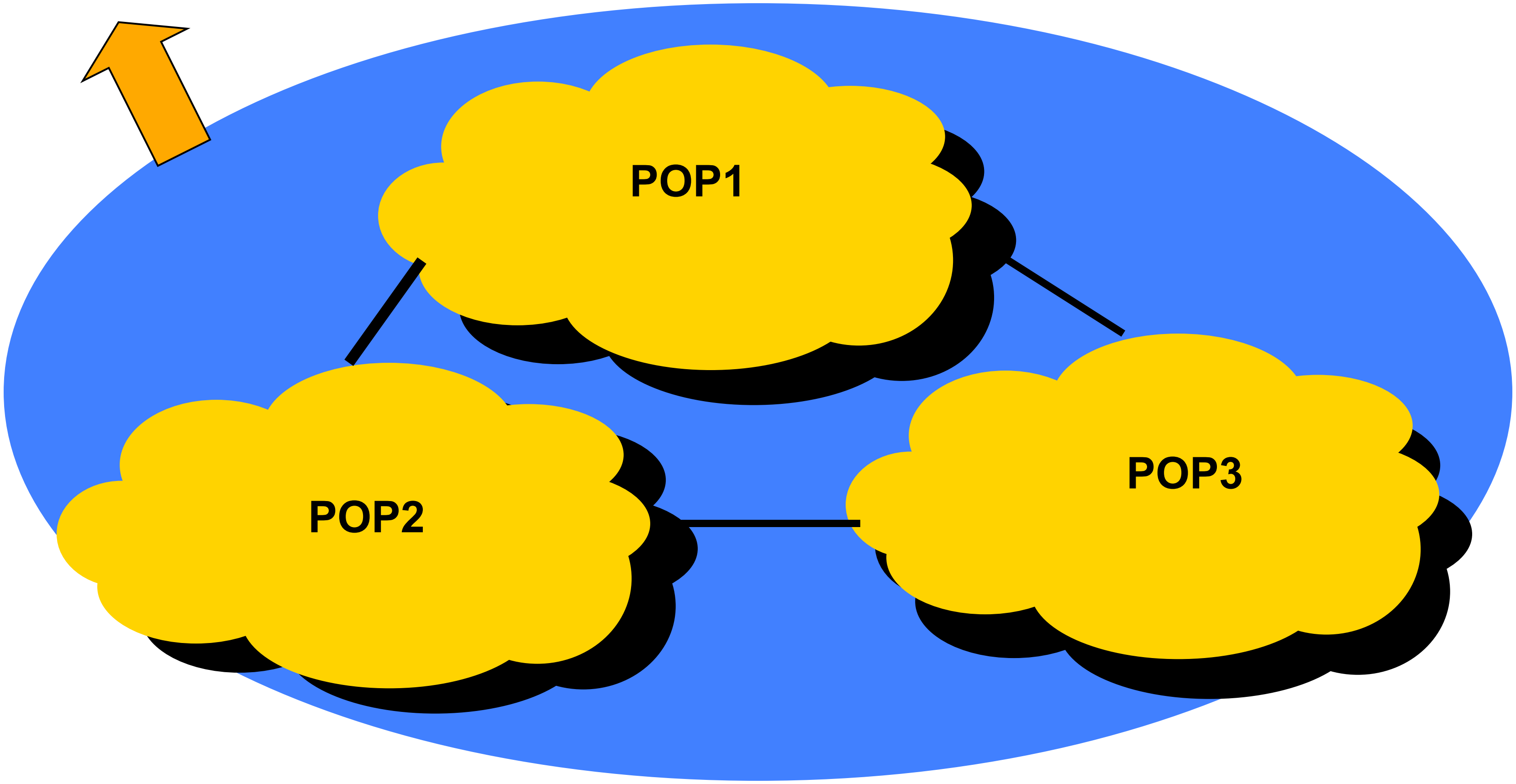
- If second allocation is contiguous



  - Reverse order of division of first block
  - Maximise contiguous space for infrastructure
    - Easier for debugging
  - Customer networks can be discontiguous

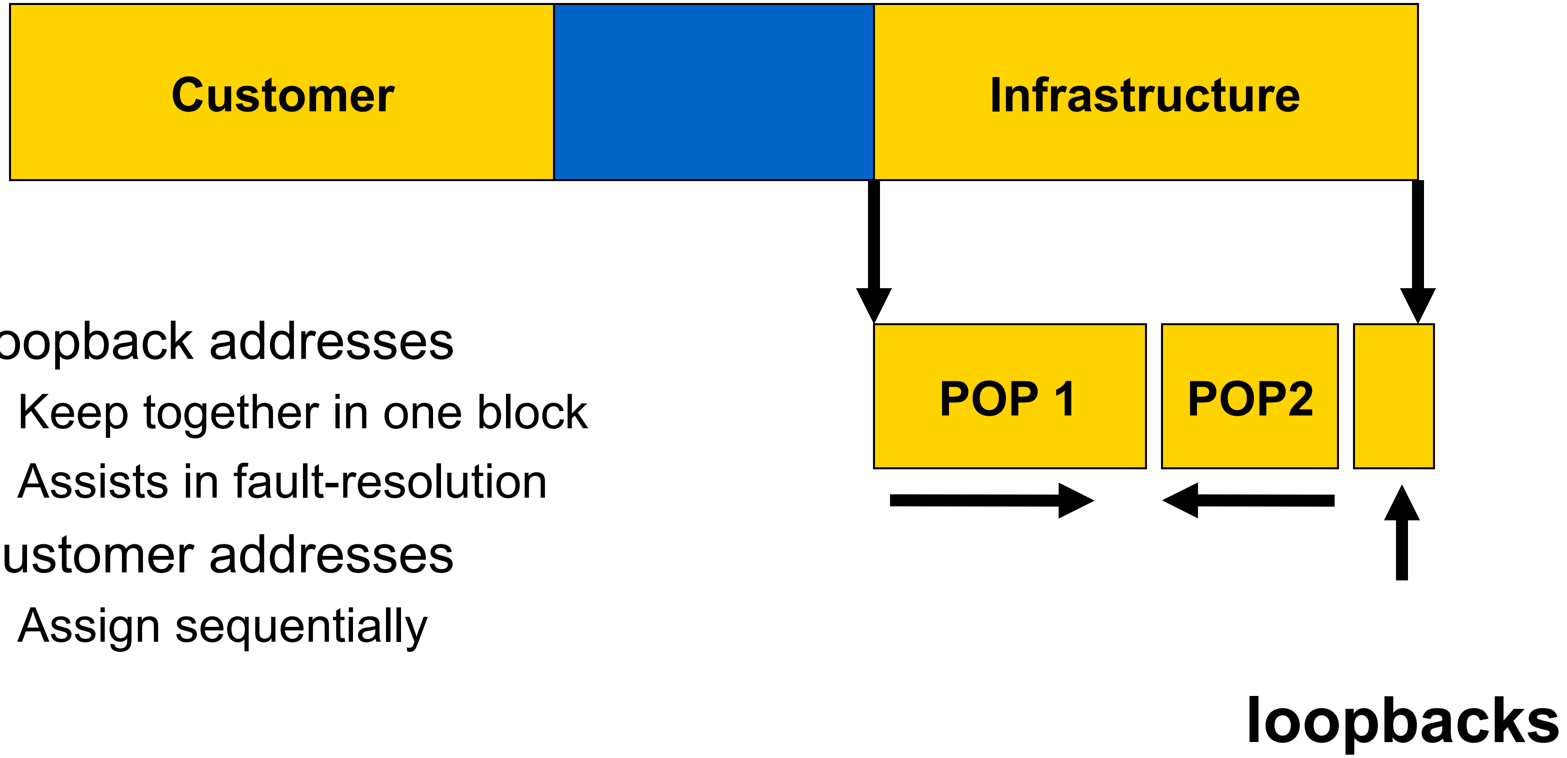
Management - Many POPs

- WAN link to single transit ISP



Management - Many POPs

- POP sizes
  - Choose address pools for each POP according to need



  - Loopback addresses
    - Keep together in one block
    - Assists in fault-resolution
  - Customer addresses
    - Assign sequentially

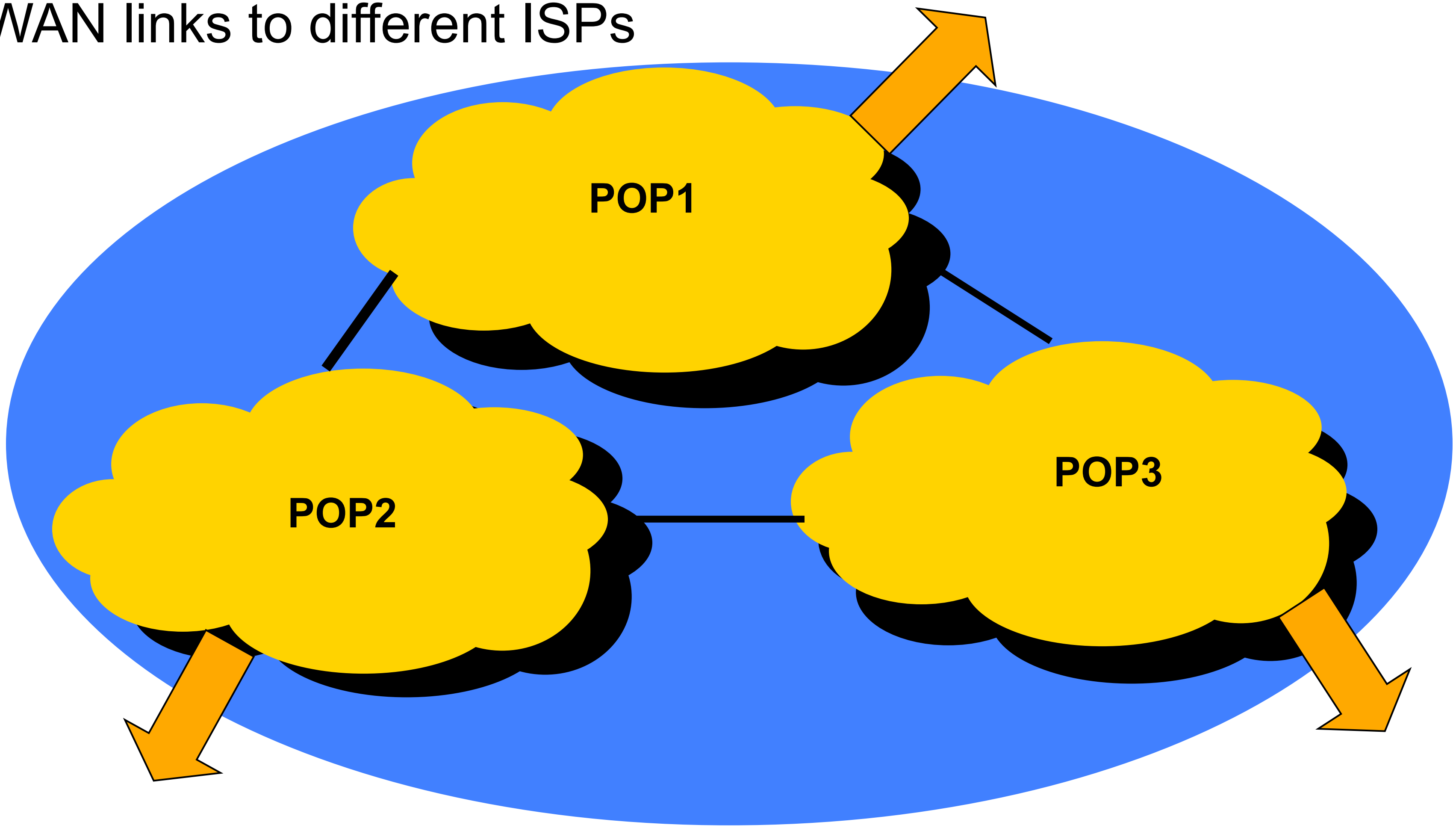
Management - Many POPs

- Minimum allocation not enough for all your POPs?
  - Deploy addresses on infrastructure first

- Common mistake:
  - Reserving customer addresses on a per POP basis

Management – Multiple Exits

- WAN links to different ISPs



Management – Multiple Exits

- Create a ‘national’ infrastructure pool

National Infrastructure	**20% free**	**POP1**	**POP2**	**POP3**

  – Carry in IGP
    - Eg. loopbacks, p2p links, infrastructure connecting routers and hosts which are multiply connected

  – On a per POP basis
    - Consider separate memberships if requirement for each POP is very large from day one.

Questions?

APNIC

Router Overview

What is a Router?

- A device in the network that processes and routes data between two points

- A device that routes data between networks using IP addressing

- A layer 3 device

- Hardware or software used to connect two or more networks

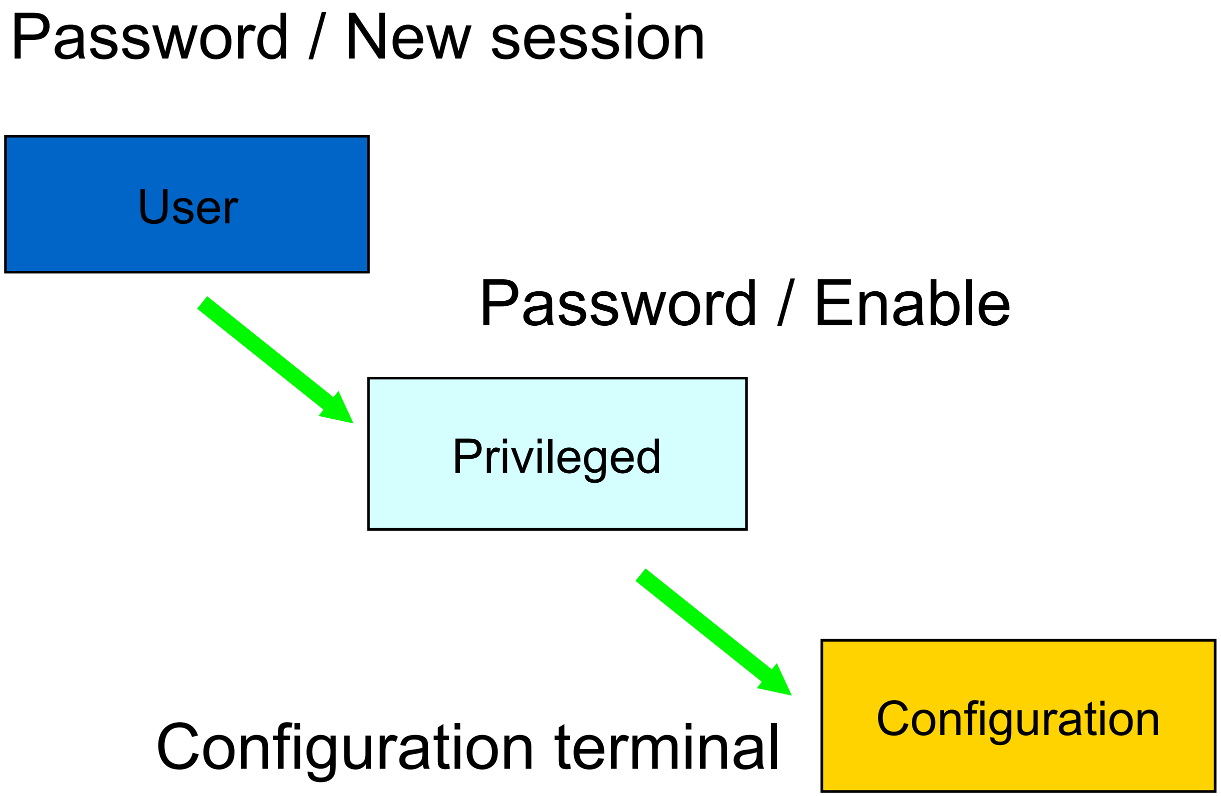
Router Basics

- Operating systems
  - IOS (Cisco)
  - Free BSD base (Juniper)
  - Quagga / Quagga (UNIX or LINUX)

- Several interfaces
  - Ethernet/Fast Ethernet, Serial, Gigabit port, Management port

- Management Interfaces
  - GUI based (web)
  - Command line interface (CLI)

Router Modes



Router Modes

- User mode
  - Check the router status and operation
  - Configuration is not visible
  - Prompt = ***router>***

- Privileged mode
  - Allows complete control to the router
  - Does not allow alteration of configuration
  - Prompt = **router#**

- Configuration mode
  - Mode to change configuration settings
  - Full control of the router configuration
  - Prompt = **router(config)#**

Router Configuration Mode

- Configuration
  - Active configuration
    - ***show running-config***
  - Startup configuration
    - ***show startup-config***

Router Components

Read Only Memory (ROM) chips:

- ROM Monitor (bootstrap program)
  - Firmware that runs when the router is boot up or reset

- Certain tasks can be done using the ROM monitor
  - Password recovery option
  - Downloading the software image using the management port

- Runs if there are no software images available on the router (with early model routers)

Router Components

Flash Memory

- Stores the software image of the router
- Usually built into the router
- Some vendors also provide external flash memory card or disk

- Allows update of router software image with less interruption of service
  - Image can be upgraded without affecting the existing image running in the router
  - Install the software then instruct the router to boot the new image after the next boot

- Allows the router to load other information
  - Router logs
    - Crash information of the router
    - Debug information

Router Component

- Non-Volatile RAM (NVRAM)
  - Stores the existing running configuration
  - Router start-up boot configuration

- Tiny memory size

- Stored configuration is very important
  - Upon router reboot / shutdown
  - Because RAM information is lost during reboot and shutdown

Router Component

- Random Access Memory (RAM)
  - Stores the current working configuration
  - Handles the tables and buffers
  - Non-permanent memory

- Broken down into two main areas
  - Main processor memory
    - Stores entry for the routing table, ARP table, and current running configuration
  - Shared processor memory
    - Buffer location for temporary stored packets for process

Router Configuration Requirements

- In configuring a router we need to address the following requirements.
  - Security
  - Manageability
  - Accessibility

Security Requirements

- To secure the router, the setup should enable the following:
  - Provide names to your router
  - Banner information
  - Configure password for the router
  - Access with privilege per user
  - Authentication and Authorisation
    - Locally configured
    - Remote server access (TACACS/RADIUS)
  - Access filters policy
  - Enable logging for auditing
  - Disable unnecessary services running

Disable Unused Access and Services

- Disable http servers running if not in use
  - http and secure http server
- Disable discovery protocol
  - CDP (Cisco)
- Disable services which can be used for reconnaissance attempts
  - Ip source-route, finger, boot server, domain-lookup, service pad

Accessibility Requirements

- Be able to manage the routers properly
  - Enable console and VTY line to
    - Allow access to the router
      - With specific host only (using filters)
    - Enable the use of privileges access
  - Provide the use of out-of-band management (console access)
  - Setup a centralised management console to control all devices

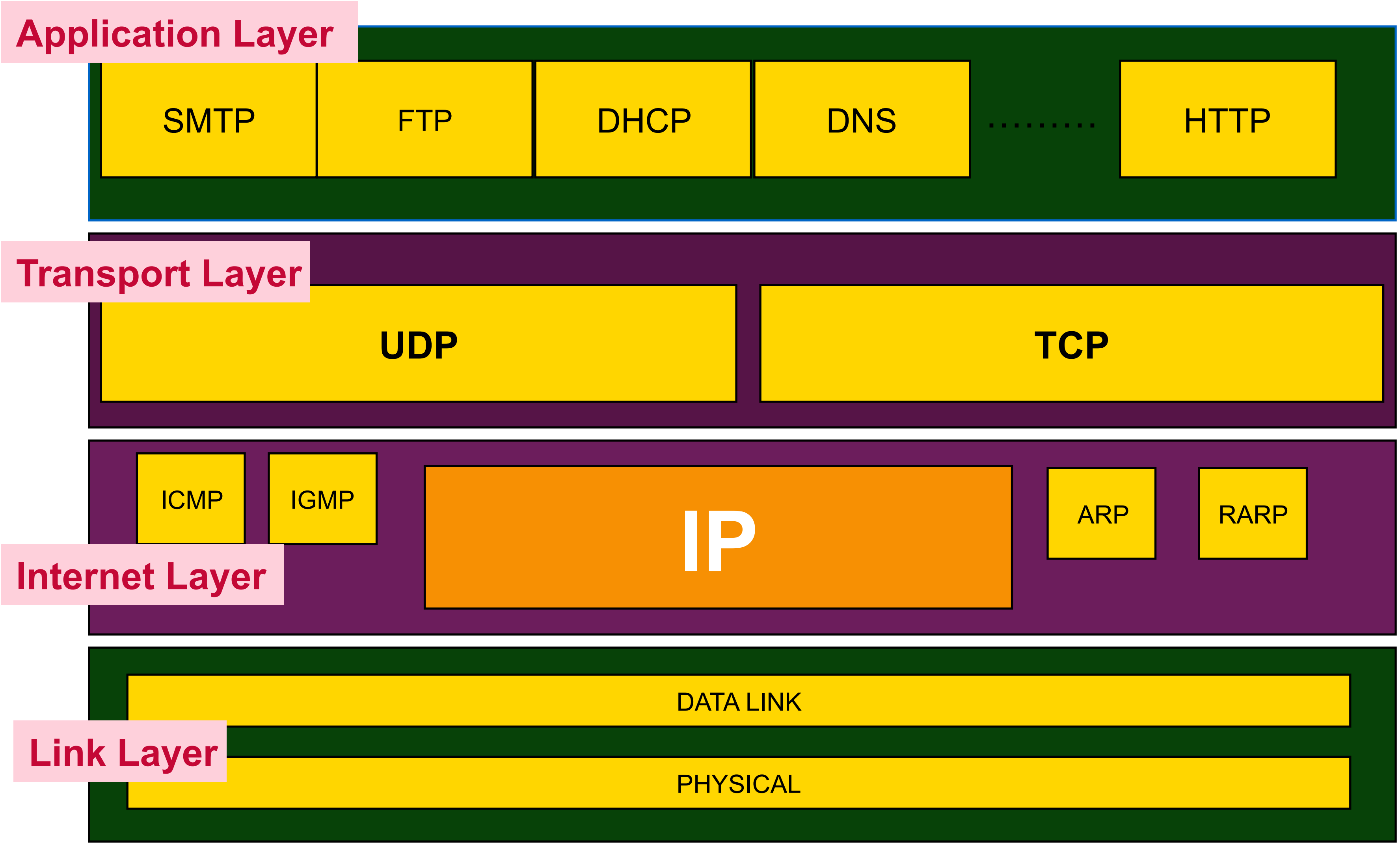
Questions?

APNIC

IPv6 Overview

APNIC

TCP/IP Protocol Structure



New Functional Improvement

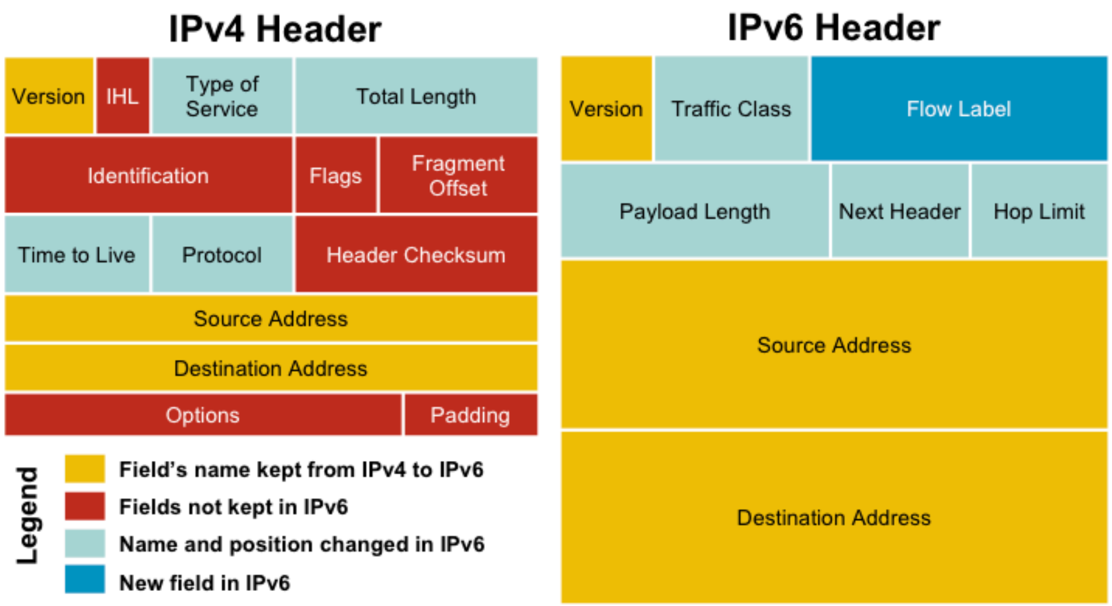
- Address Space
  - Increase from 32-bit to 128-bit address space
- Management
  - Stateless autoconfiguration means no more need to configure IP addresses for end systems, even via DHCP
- Performance
  - Fixed header size (40 bytes) and 64-bit header alignment mean better performance from routers and bridges/switches

Source: http://www.opus1.com/ipv6/whatisipv6.html

New Functional Improvement

- Multicast/Multimedia
  - Built-in features for multicast groups,management, and new "anycast" groups
- Mobile IP
  - Eliminate triangular routing and simplify deployment of mobile IP-based systems
- Virtual Private Networks
  - Built-in support for ESP/AH encrypted/ authenticated virtual private network protocols;
- No more broadcast

Protocol Header Comparison



- IPv4 contains 10 basic header field
- IPv6 contains 6 basic header field
- IPv6 header has 40 octets in contrast to the 20 octets in IPv4
- So a smaller number of header fields and the header is 64-bit aligned to enable fast processing by current processors

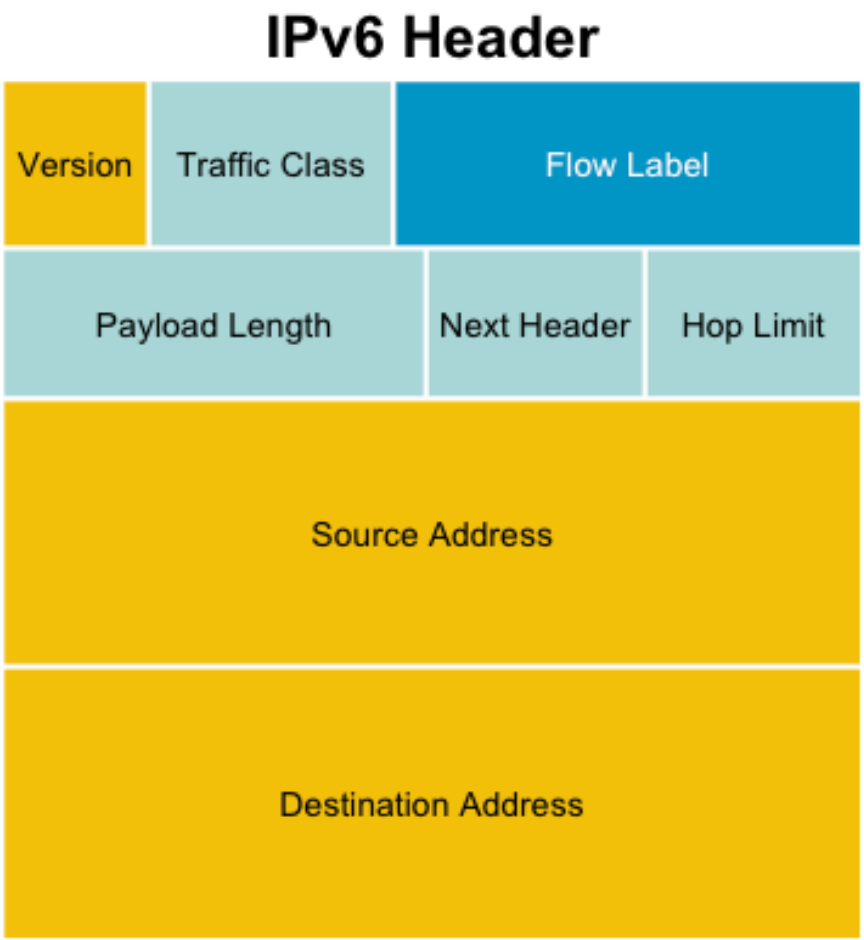
Diagram Source: www.cisco.com

IPv6 Protocol Header Fields

- Version
  - A 4-bit field, same as in IPv4. It contains the number 6 instead of the number 4 for IPv4
- Traffic class
  - An 8-bit field similar to the type of service (ToS) field in IPv4. It tags packet with a traffic class that it uses in differentiated services (DiffServ). These functionalities are the same for IPv6 and IPv4.
- Flow label
  - A completely new 20-bit field. It tags a flow for the IP packets. It can be used for multilayer switching techniques and faster packet-switching performance

IPv6 Protocol Header Format

- Payload length
  - 16-bit field is similar to the IPv4 Total Length Field
  - the length of the data carried after the header, (whereas with IPv4 the Total Length Field included the header). $2^{16} = 65536$ Octets.
- Next header
  - The 8-bit value of this field determines the type of information that follows the basic IPv6 header. It can be a transport-layer packet, such as TCP or UDP, or it can be an extension header. The next header field is similar to the protocol field of IPv4.
- Hop limit
  - This 8-bit field defined by a number which count the maximum hops that a packet can remain in the network before it is destroyed. With the IPv4 TTL field this was expressed in seconds and was typically a theoretical value and not very easy to estimate.



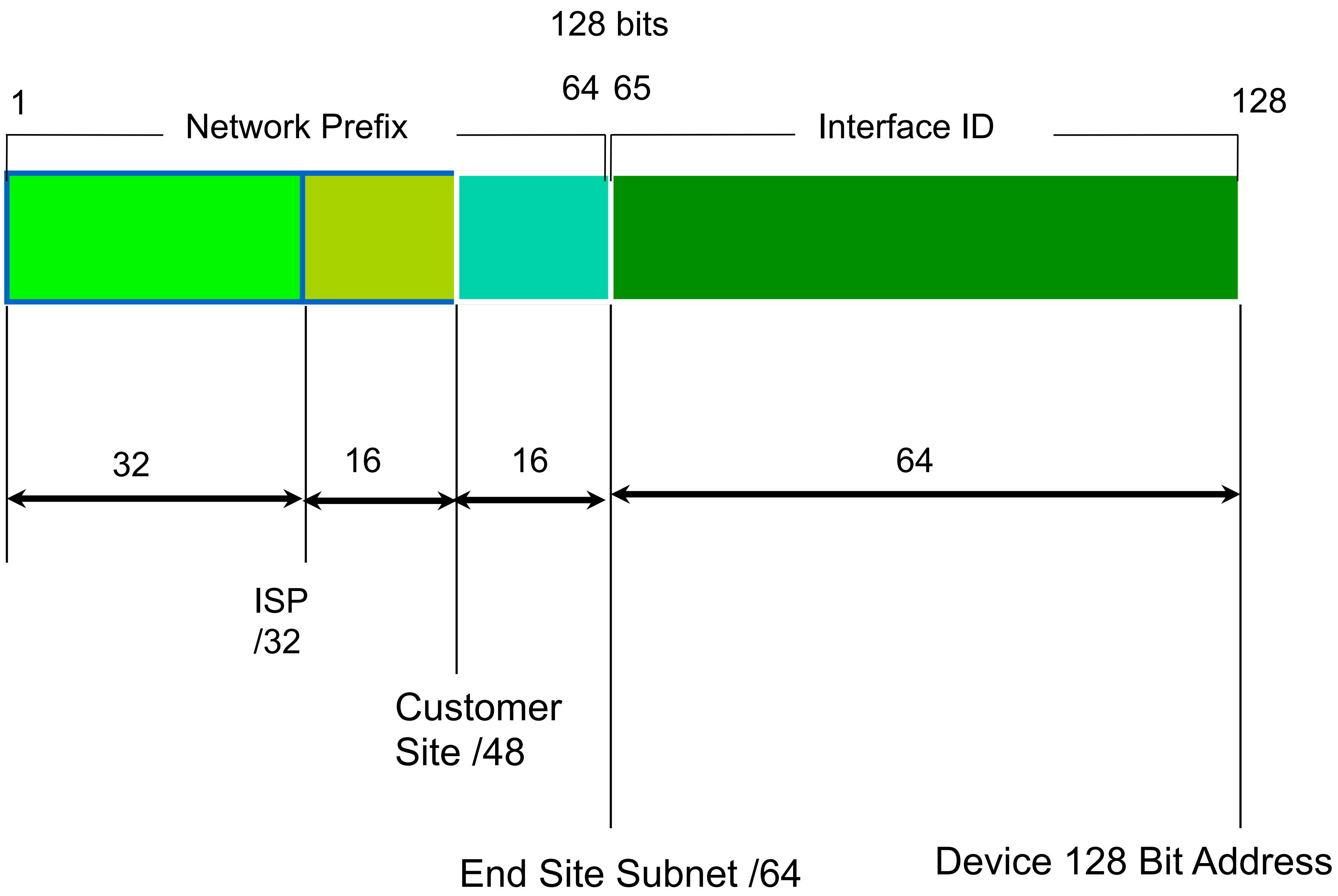
IPv6 Addressing

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

IPv6 Address Representation

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

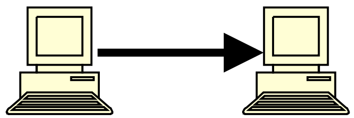
IPv6 addressing structure



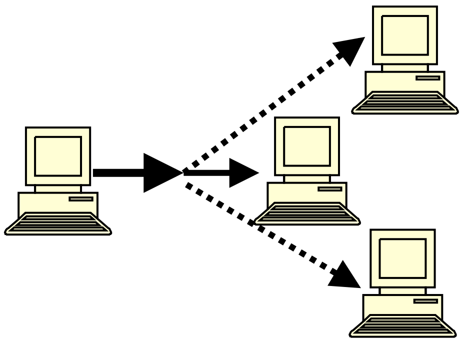
IPv6 addressing model

RFC 4291

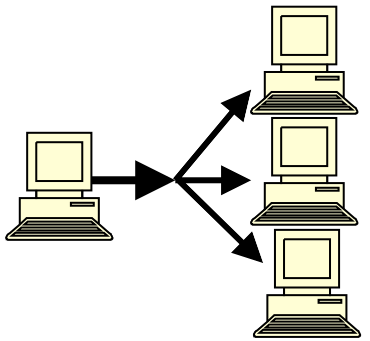
- **IPv6 Address type**
  - Unicast
    - An identifier for a single interface



  - Anycast
    - An identifier for a set of interfaces



  - Multicast
    - An identifier for a group of nodes



Addresses Without a Network Prefix

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

Local Addresses With Network Prefix

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

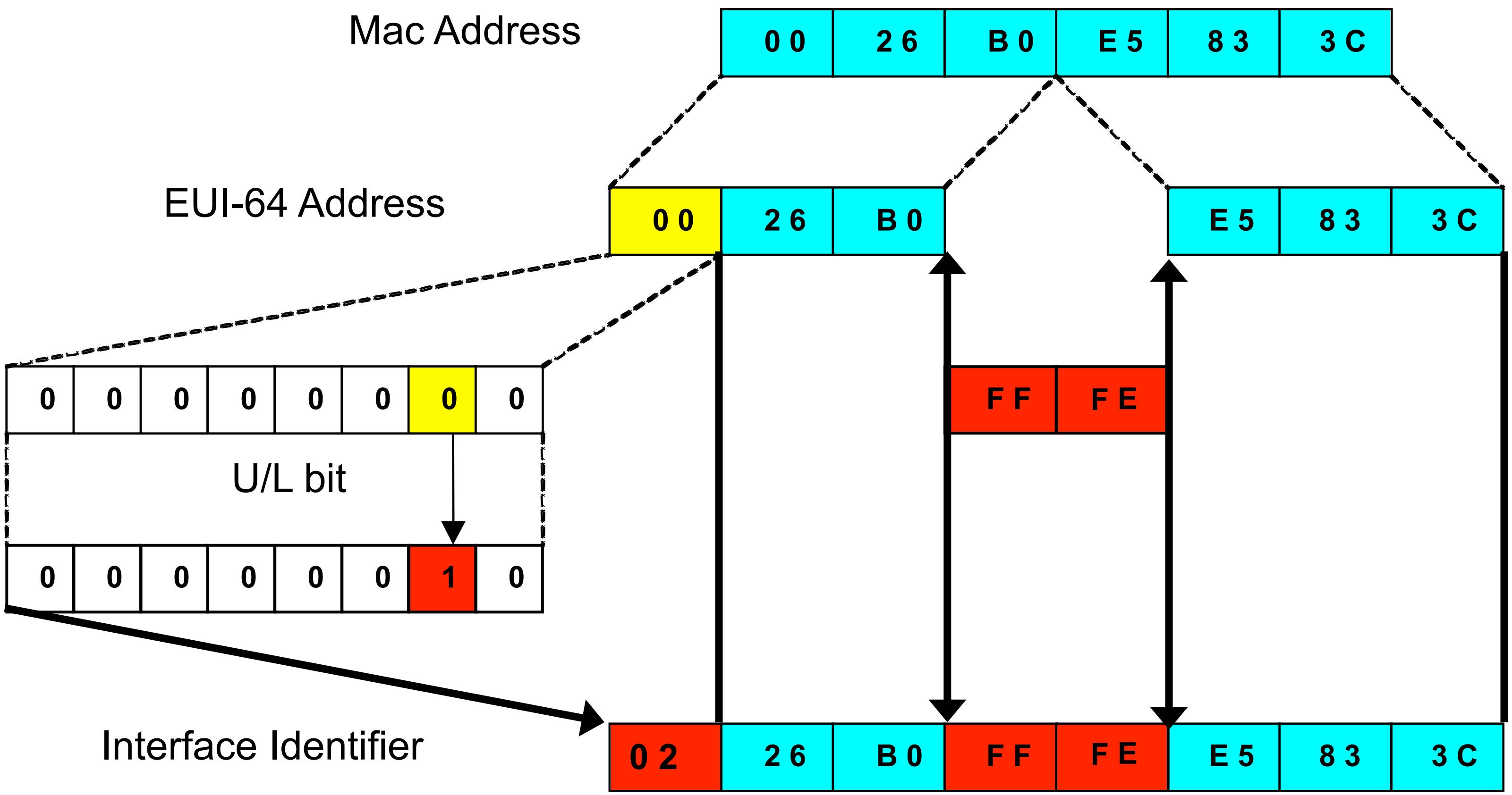
Global Addresses With Network Prefix

- IPV6 Global Unicast Address
  - Global Unicast Range: 0010 2000::/3
    0011 3000::/3
  - All five RIRs are given a /12 from the /3 to further distribute within the RIR region
    - APNIC 2400:0000::/12
    - ARIN 2600:0000::/12
    - AfriNIC 2C00:0000::/12
    - LACNIC 2800:0000::/12
    - Ripe NCC 2A00:0000::/12

Interface ID

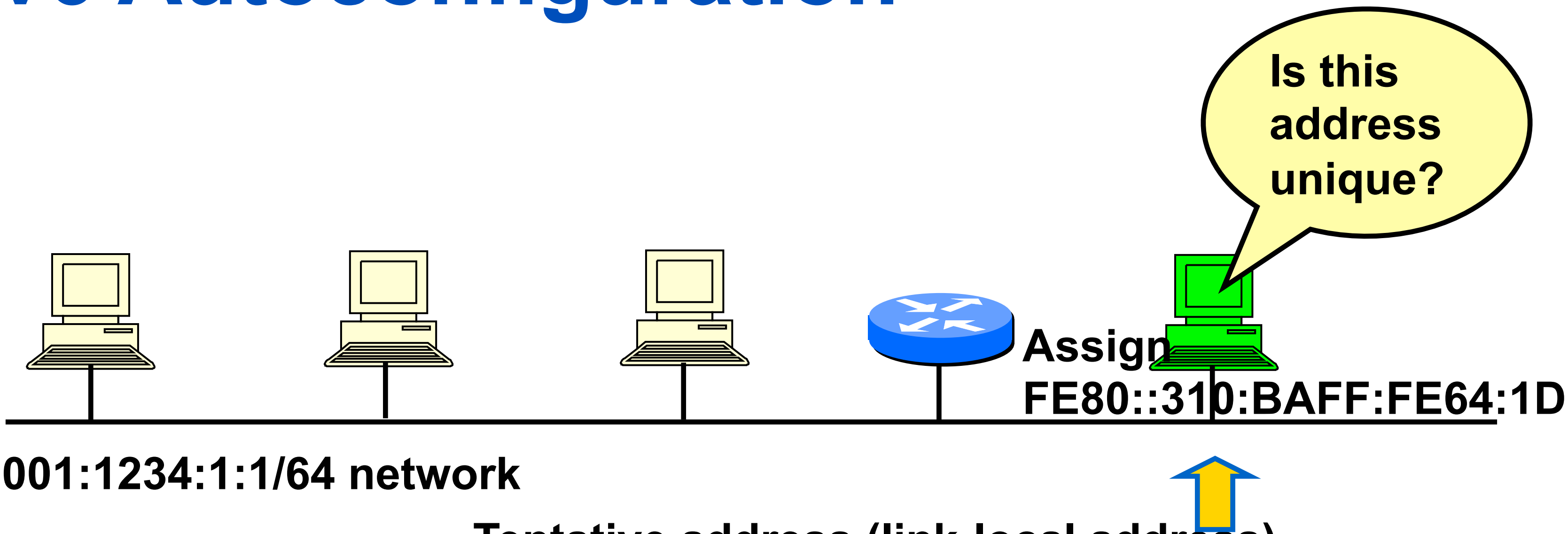
- The lowest-order 64-bit field addresses may be assigned in several different ways:
  - auto-configured from a 48-bit MAC address expanded into a 64-bit EUI-64
  - assigned via DHCP
  - manually configured
  - auto-generated pseudo-random number
  - possibly other methods in the future

EUI-64



Mac Address
0 0
2 6
B 0
E 5
8 3
3 C
EUI-64 Address
0 0
2 6
B 0
E 5
8 3
3 C
0 0 0 0 0 0 0 0
U/L bit
0 0 0 0 0 0 1 0
F F
F E
Interface Identifier
0 2
2 6
B 0
F F
F E
E 5
8 3
3 C

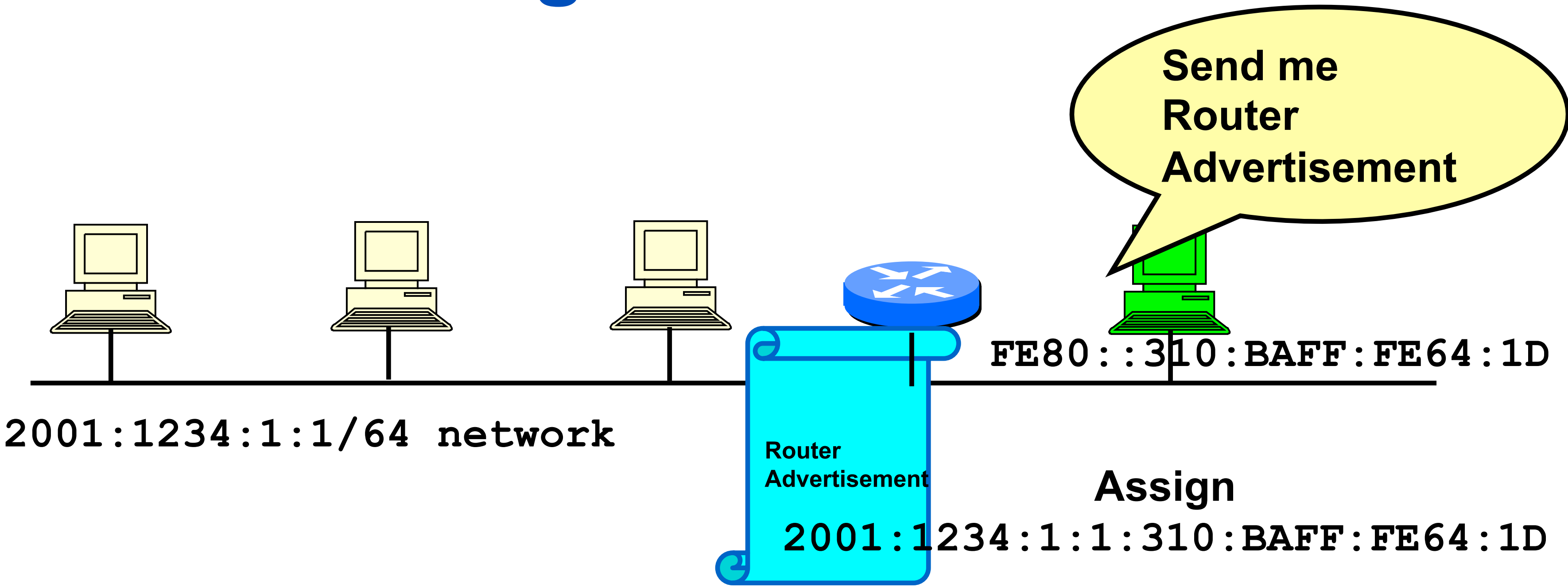
IPv6 Autoconfiguration



Tentative address (link-local address)
Well-known link local prefix +Interface ID (EUI-64)
Ex: FE80::310:BAFF:FE64:1D

1. A new host is turned on.
2. Tentative address will be assigned to the new host.
3. Duplicate Address Detection (DAD) is performed. First the host transmit
   - a Neighbor Solicitation (NS) message to all-nodes multicast address (FF02::1)
5. If no Neighbor Advertisement (NA) message comes back then the address is unique.
6. FE80::310:BAFF:FE64:1D will be assigned to the new host.

IPv6 Autoconfiguration



1. The new host will send Router Solicitation (RS) request to the all-routers multicast group (FF02::2).
2. The router will reply Routing Advertisement (RA).
3. The new host will learn the network prefix. E.g, 2001:1234:1:1/64
4. The new host will assigned a new address Network prefix+Interface ID E.g, 2001:1234:1:1:310:BAFF:FE64:1D

IPv6 Subnetting

- Network engineers must have a solid understanding of subnetting
  - Important for address planning
- IPv6 subnetting is similar (if not exactly the same) as IPv4 subnetting
- Note that you are working on hexadecimal digits rather than binary
  - 0 in hex = 0000 in binary
  - 1 in hex = 0001 in binary

IPv6 Subnetting (Example)

- Provider A has been allocated an IPv6 block **2001:0DB8::/32**
- Provider A will delegate /48 blocks to its customers
- Find the blocks provided to the first 4 customers

IPv6 Subnetting (Example)

Original block: **2001:0DB8::/32**

Rewrite as a /48 block: **2001:0DB8:0000:/48**

This is your network prefix!

How many /48 blocks are there in a /32?

$$\frac{/32}{/48} = \frac{2^{128-32}}{2^{128-48}} = \frac{2^{96}}{2^{80}} = 2^{16}$$

Find only the first 4 /48 blocks…

IPv6 Subnetting (Example)

Start by manipulating the LSB of your network prefix – write in BITS

2001:0DB8:0000::/48

In bits

2001:0DB8:	**0000 0000 0000 0000**	**::/48**	→	**2001:0DB8:0000::/48**
2001:0DB8:	**0000 0000 0000 0001**	**::/48**	→	**2001:0DB8:0001::/48**
2001:0DB8:	**0000 0000 0000 0010**	**::/48**	→	**2001:0DB8:0002::/48**
2001:0DB8:	**0000 0000 0000 0011**	**::/48**	→	**2001:0DB8:0003::/48**

Then write back into hex digits

Exercise 1.1: IPv6 subnetting

- Identify the first four /64 address blocks out of 2001:DB8:0::/48
  1. ______________________
  2. ______________________
  3. ______________________
  4. ______________________

Exercise 1.2: IPv6 subnetting

- Identify the first four /36 address blocks out of 2001:DB8::/32
  1. ____________________
  2. ____________________
  3. ____________________
  4. ____________________

Exercise 1.3: IPv6 subnetting

- Identify the first six /37 address blocks out of 2400:ABCD::/32

1. ____________________________
2. ____________________________
3. ____________________________
4. ____________________________
5. ____________________________
6. ____________________________

Questions?

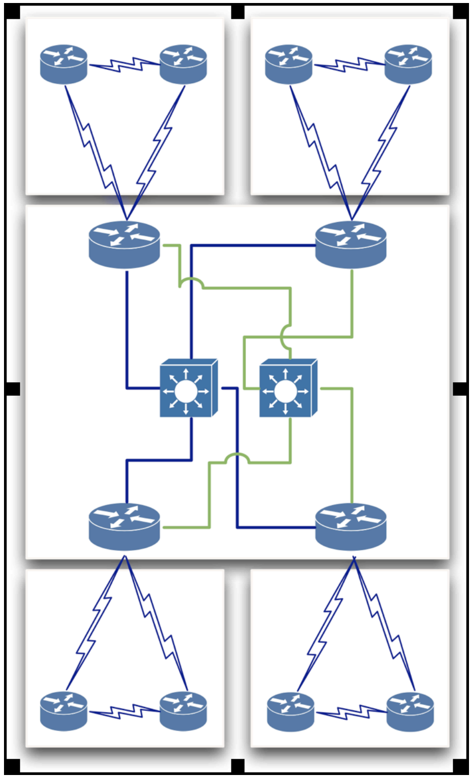
Case Study – Topology Design and Address Planning

APNIC

Training ISP Network Topology

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

Training ISP Network Topology



Training ISP Topology Diagram

Training ISP Network Topology

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

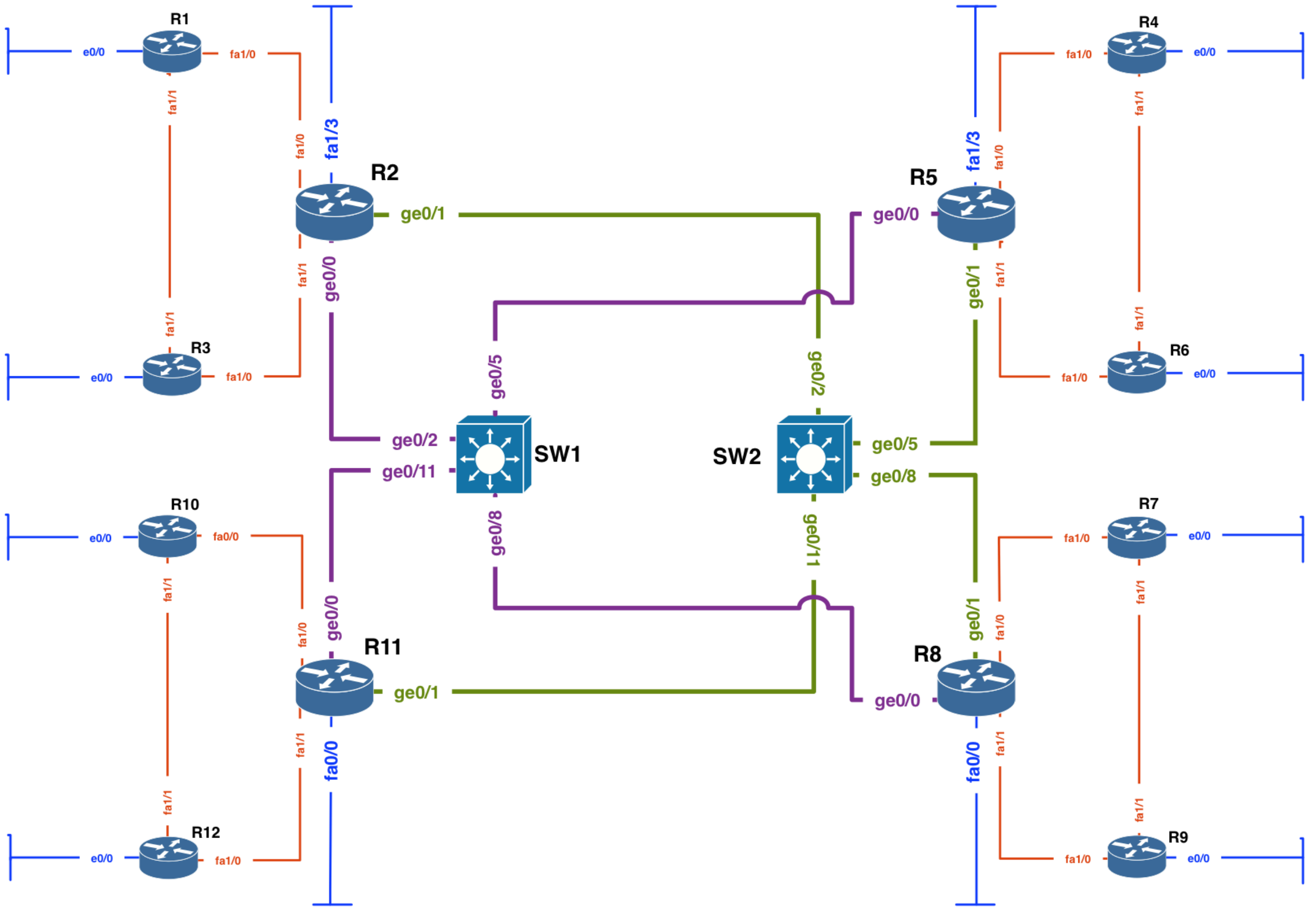
Training ISP Network Topology

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

Training ISP Network Topology

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

Training ISP Network Topology



Training ISP Core IP Backbone

Training ISP Network Topology

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

Training ISP Network Topology

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

Training ISP Network Topology

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

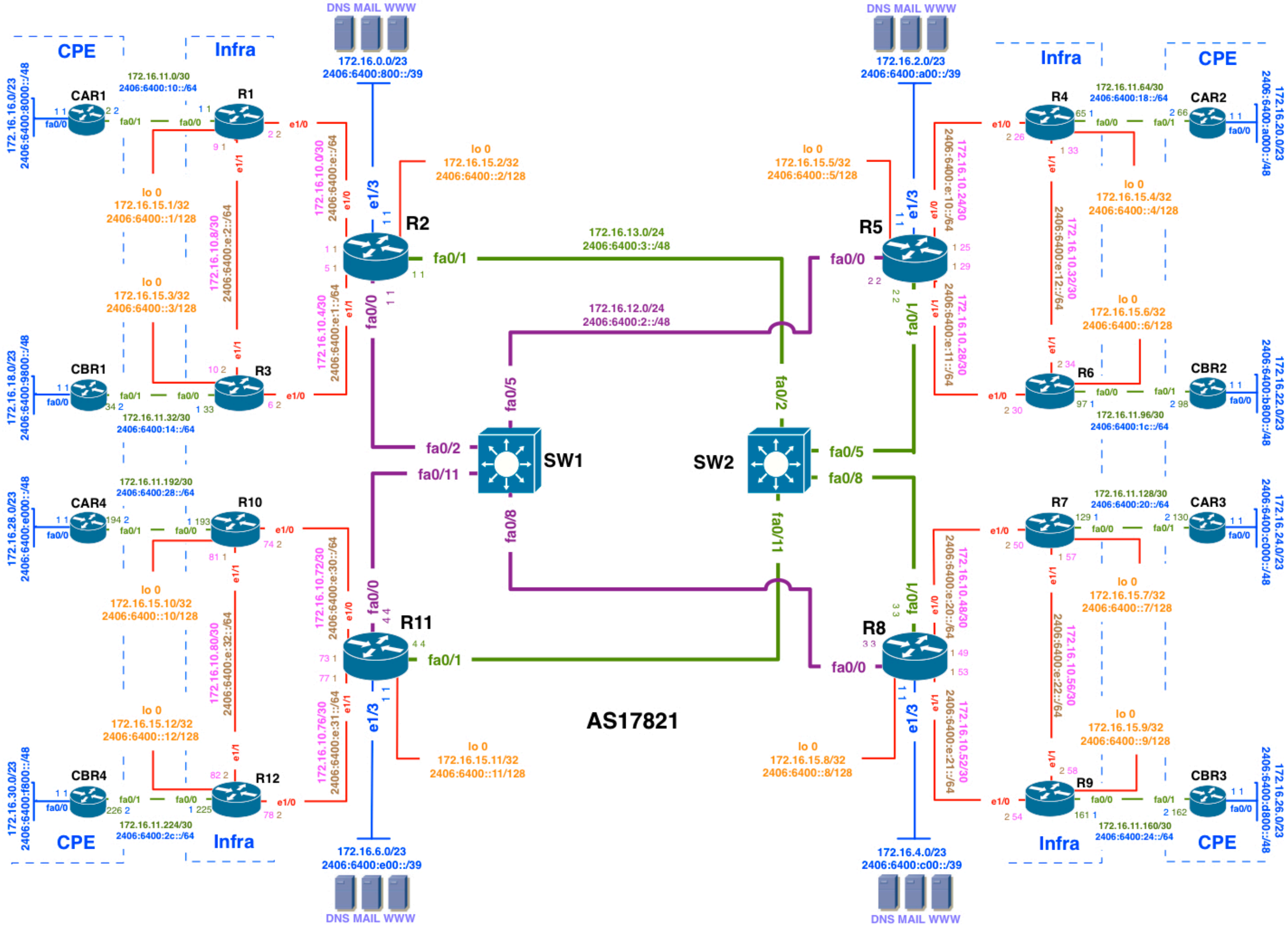
Training ISP IPv6 Addressing Plan

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

Training ISP IPv6 Addressing Plan

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

Training ISP IPv6 Addressing Plan



Training ISP IPV6 Addressing Plan

- IPv6 Address Plan:

Table 1: Top level distribution infrastructure & customer

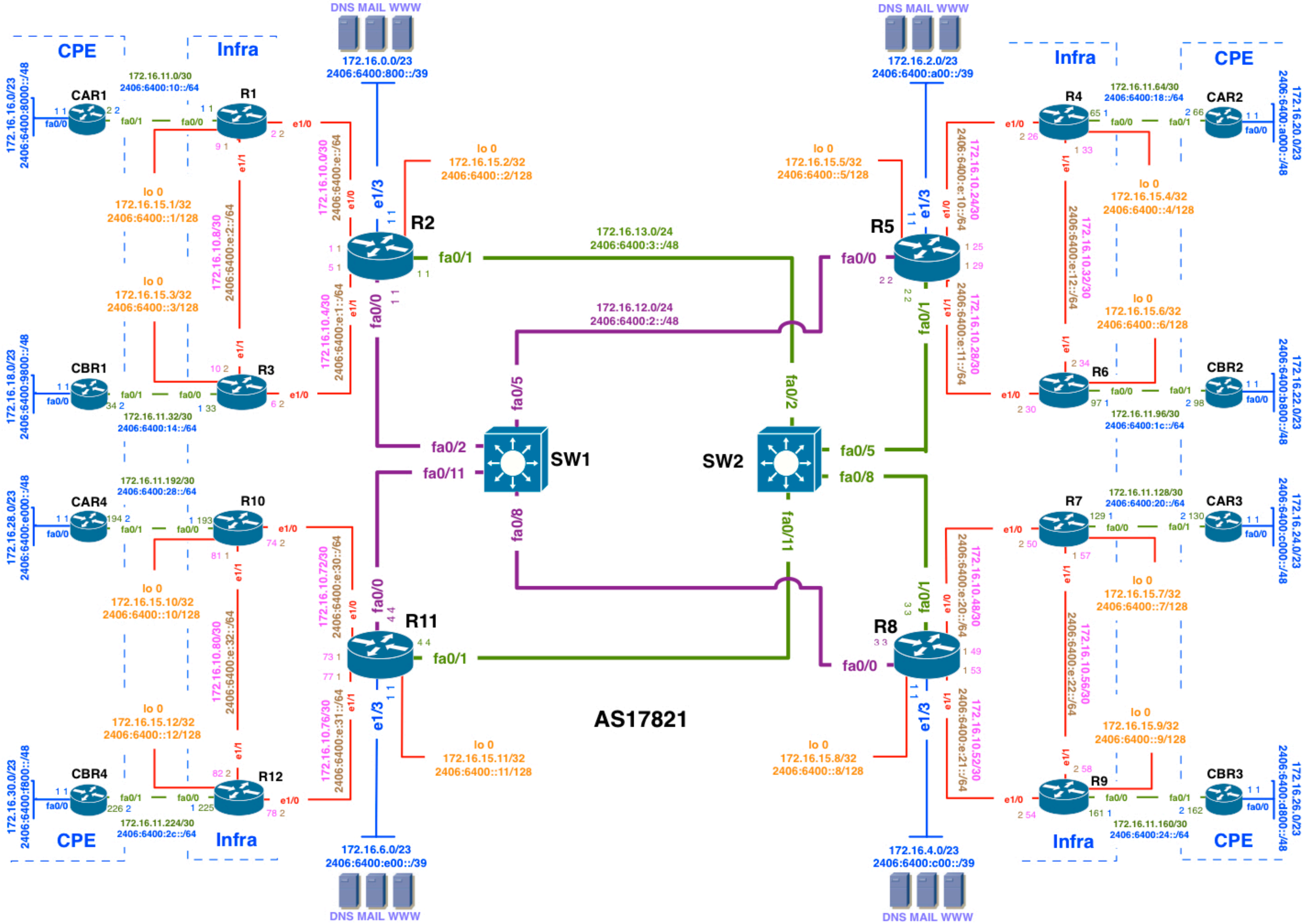
Block#	Prefix	Description	Reverse Domain	SOR	Registration
1	**2406:6400::/32**	***Parent Block***	0.0.4.6.6.0.4.2.ip6.arpa.	N/A	APNIC
2	2406:6400:0000:0000::/36	Infrastructure	0.0.0.4.6.6.0.4.2.ip6.arpa.	No	Optional
	2406:6400:1000:0000::/36				
	2406:6400:2000:0000::/36				
	2406:6400:3000:0000::/36				
	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

- IPv6 Address 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

- IPv6 Address 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				

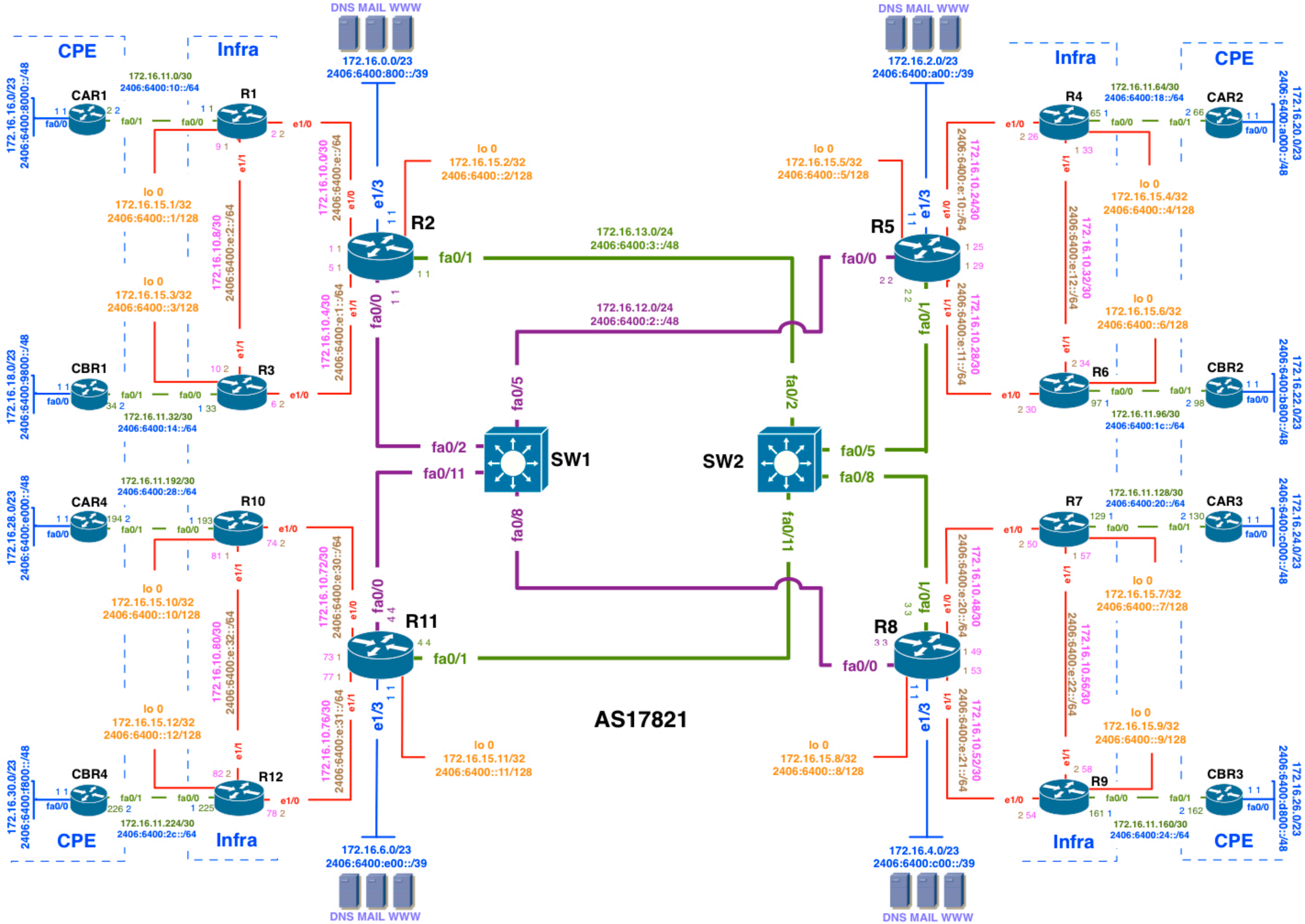
Training ISP IPV6 Addressing Plan

- IPv6 Address 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 3 DC Summary [R8]	

Training ISP IPV6 Addressing Plan



Training ISP IPV6 Addressing Plan

- IPv6 Address Plan:

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				

Table 5: Further detail loopback, transport & infrastructure WAN

Training ISP IPV6 Addressing Plan

- IPv6 Address 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

- IPv6 Address Plan:

Table 7: CS link WAN summarization options

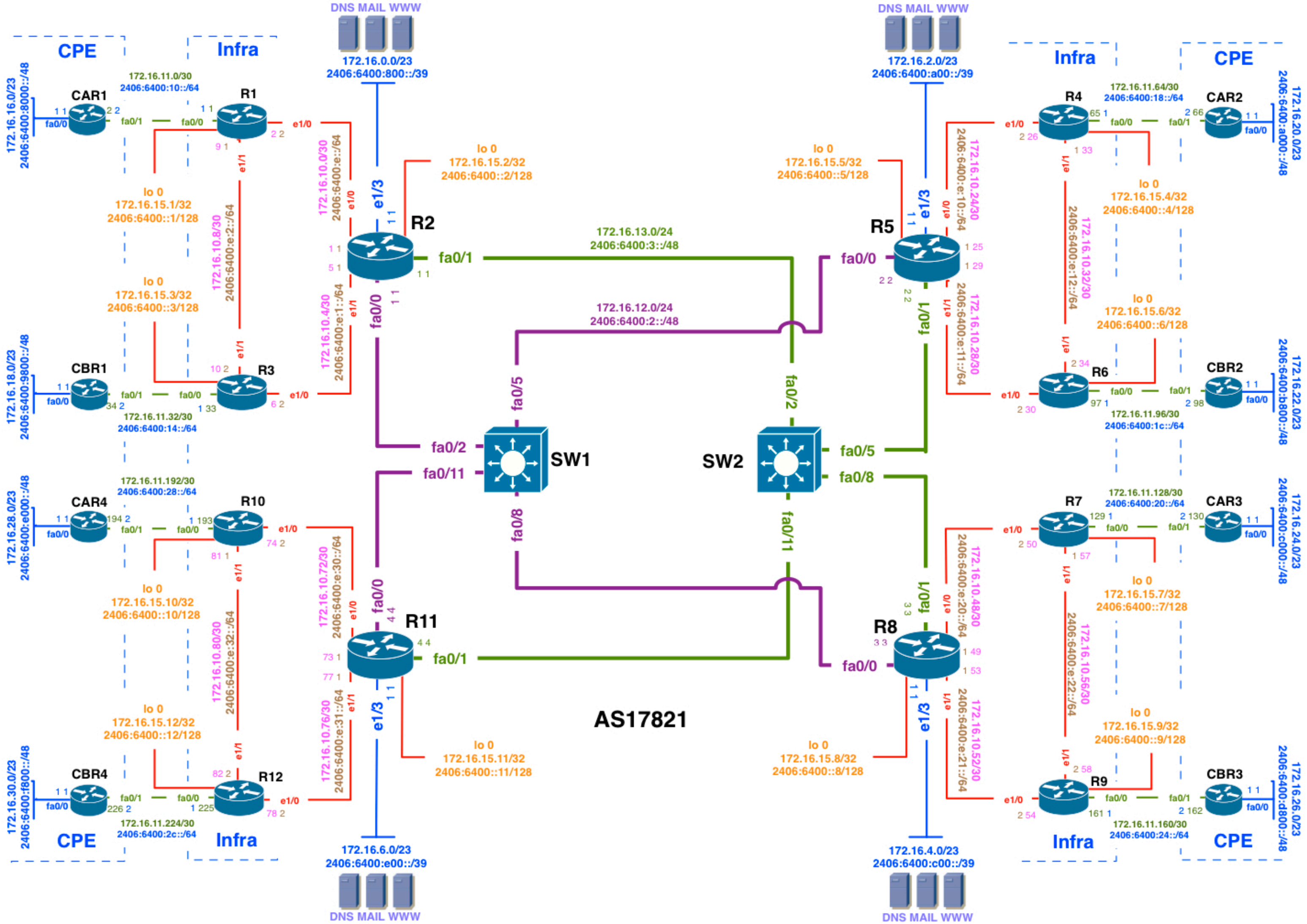
Block#	Prefix	Description	Reverse Domain
24	2406:6400:0010:0000::/45	WAN CS Link Region1 Summary [R2]	
25	2406:6400:0010:0000::/46	WAN CS Link Region1 POP1 Summary [R1]	
26	2406:6400:0014:0000::/46	WAN CS Link Region1 POP2 Summary [R3]	

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

- IPv6 Address 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

- IPv6 Address Plan:

Table 9: Further detail transport

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

Training ISP IPV6 Addressing Plan

- IPv6 Address 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

- IPv6 Address 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

- IPv6 Address 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

- IPv6 Address 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

- IPv6 Address 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

- IPv6 Address Plan:

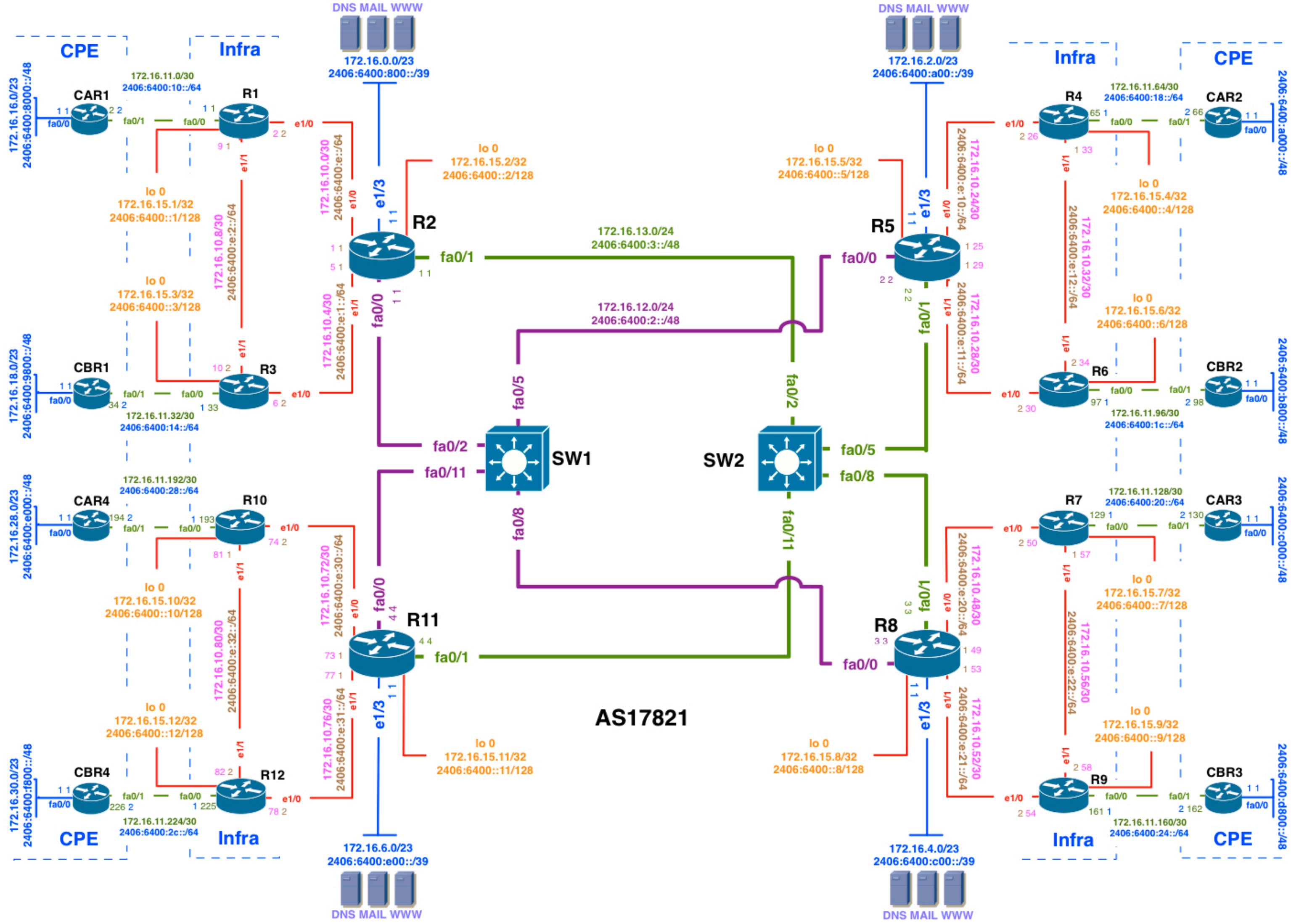
Table 16: Summarization oprions customer block Region 1			
Block#	**Prefix**	**Description**	**Reverse Domain**
	2406:6400:8000:0000::/35	Customer block Region 1 [R2]	
	2406:6400:8000:0000::/37	Customer block POP1 [R1]	
	2406:6400:8800:0000::/37	Customer block future use/POP	
	2406:6400:9000:0000::/37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

Training ISP IPV6 Addressing Plan

- IPv6 Address 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

- IPv6 Address Plan:

Table 18: Customer block Region 2

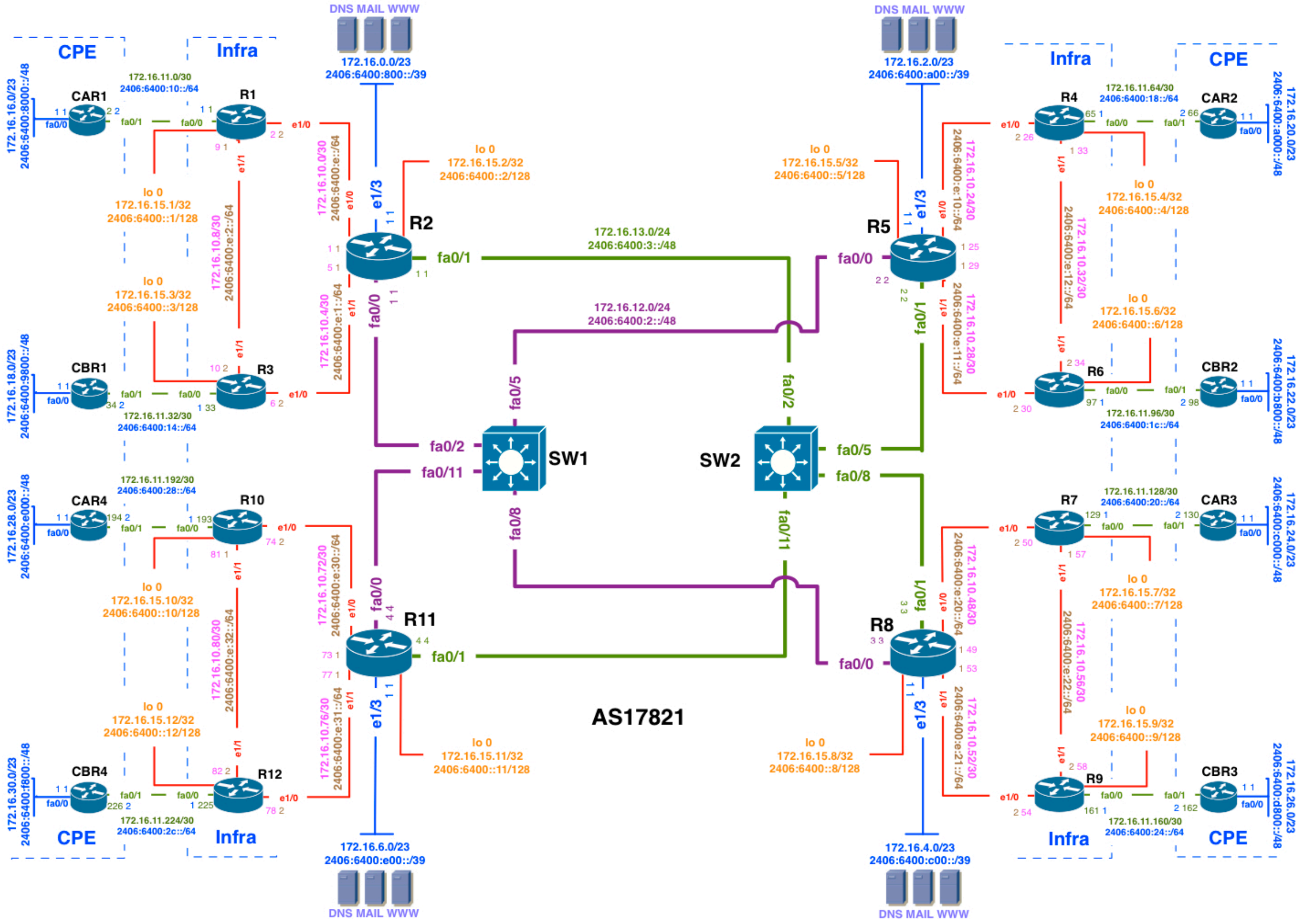
Block#	Prefix	Description	Reverse DNS	SOR	Registration
8	2406:6400:a000:0000::/35	Customer block Region 2			
	2406:6400:A000:0000::/40	Customer block POP1 [R4]		>= /48 Yes	Yes
	2406:6400:A100:0000::/40				
	2406:6400:A200:0000::/40				
	2406:6400:A300:0000::/40				
	2406:6400:A400:0000::/40				
	2406:6400:A500:0000::/40				
	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

- IPv6 Address 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



Training ISP IPV6 Addressing Plan

- IPv6 Address 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

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

- IPv6 Address Plan:

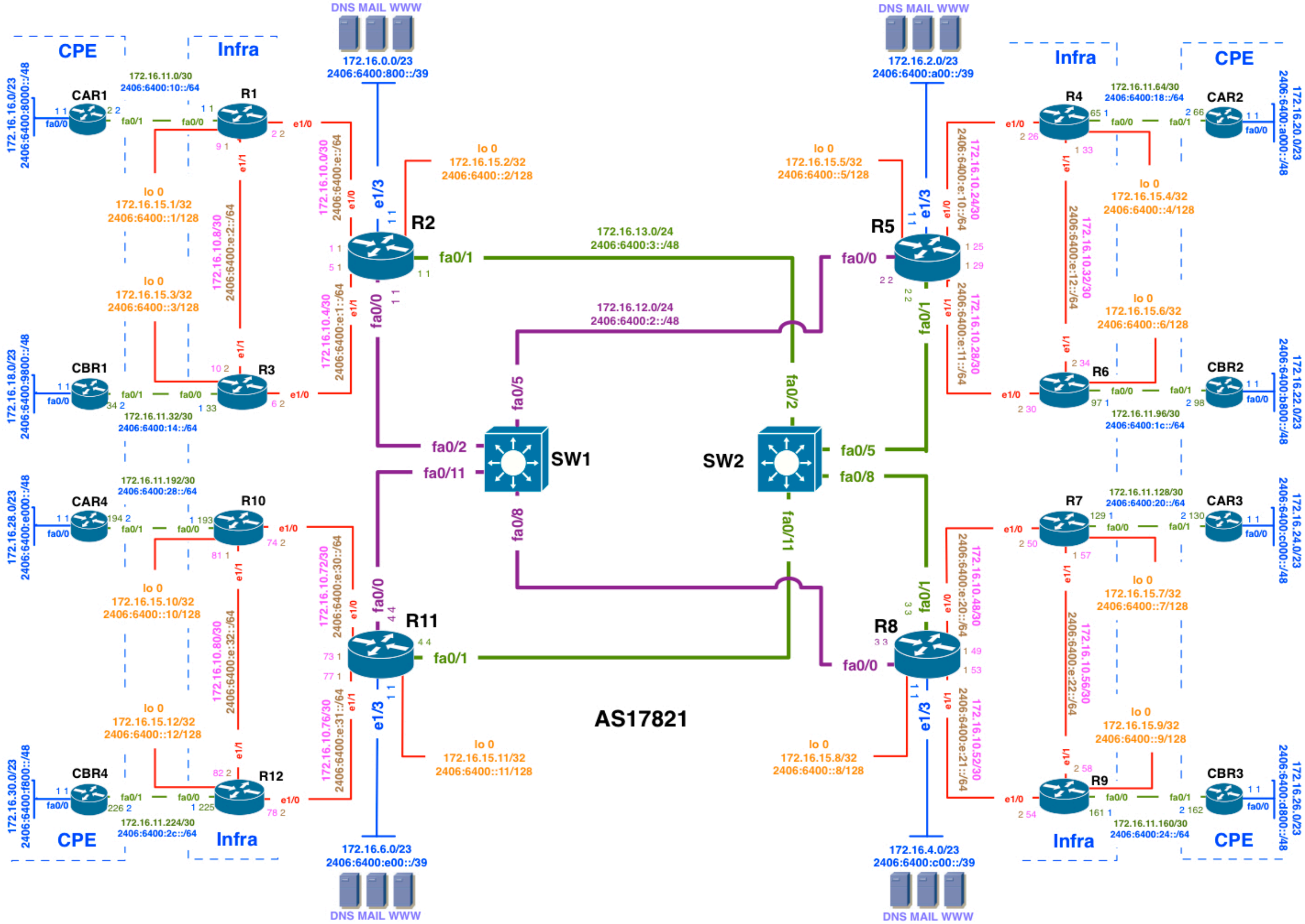
Table 22: Summarization oprions customer block Region 3			
Block#	**Prefix**	**Description**	**Reverse Domain**
	2406:6400:c000:0000::/35	Customer block Region 3 [R8]	
	2406:6400:C000:0000::/37	Customer block POP1 [R7]	
	2406:6400:C800:0000::/37	Customer block future use/POP	
	2406:6400:D000:0000::/37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

Training ISP IPV6 Addressing Plan

- IPv6 Address 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

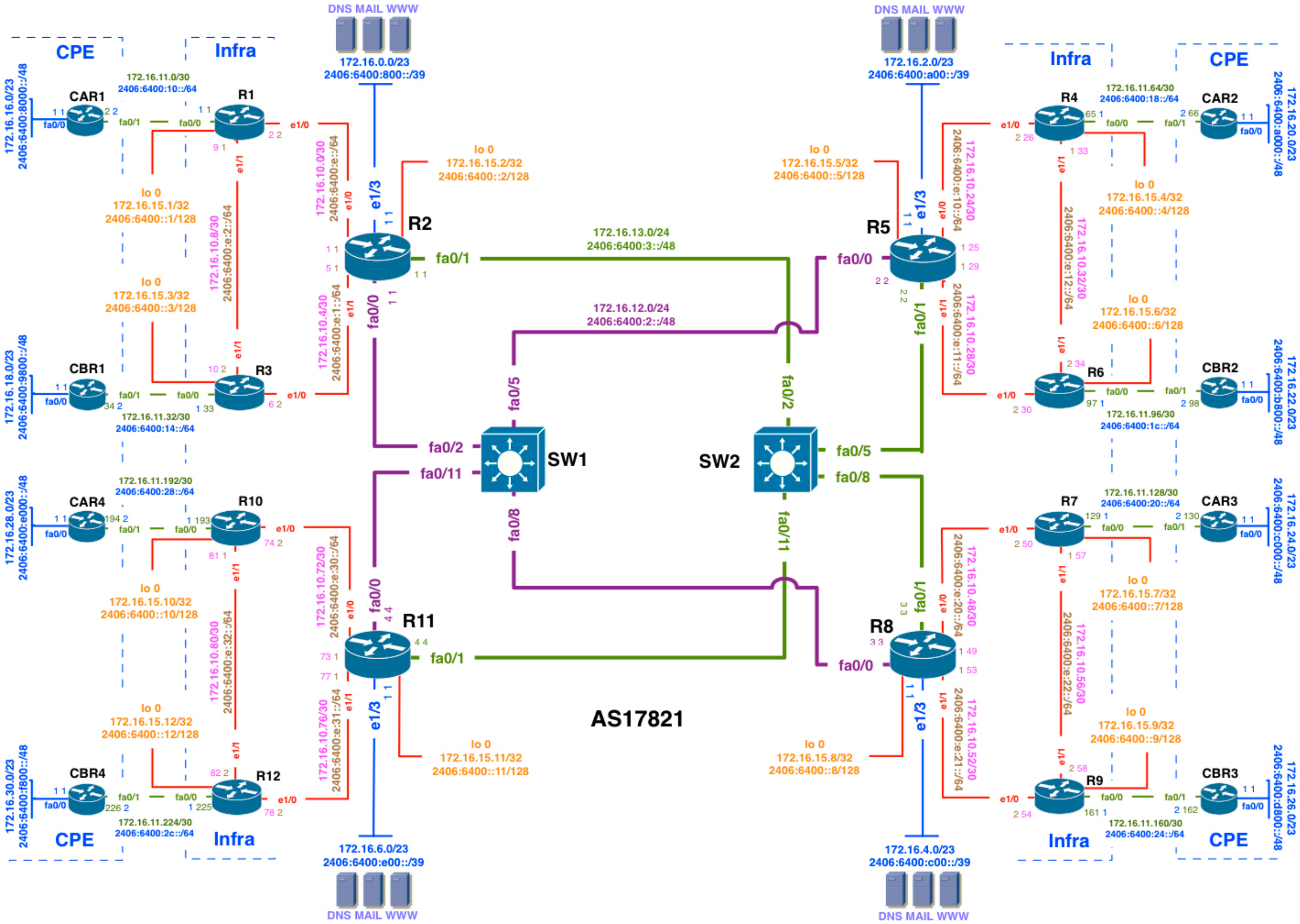
Block#	Prefix	Description	Reverse DNS	SOR	Registration
10	2406:6400:e000:0000::/35	Customer block Region 4			
	2406:6400:E000:0000::/40	Customer block POP1 [R10]		>= /48 Yes	Yes
	2406:6400:E100:0000::/40				
	2406:6400:E200:0000::/40				
	2406:6400:E300:0000::/40				
	2406:6400:E400:0000::/40				
	2406:6400:E500:0000::/40				
	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

- IPv6 Address 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



APNIC

Training ISP IPV6 Addressing Plan

- IPv6 Address 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 IPV4 Addressing Plan

- Current 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

- Current 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

- Current 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

- Current IPv4 Addressing Plan:

Detail infrastructure block Transport & Loopback IPV4

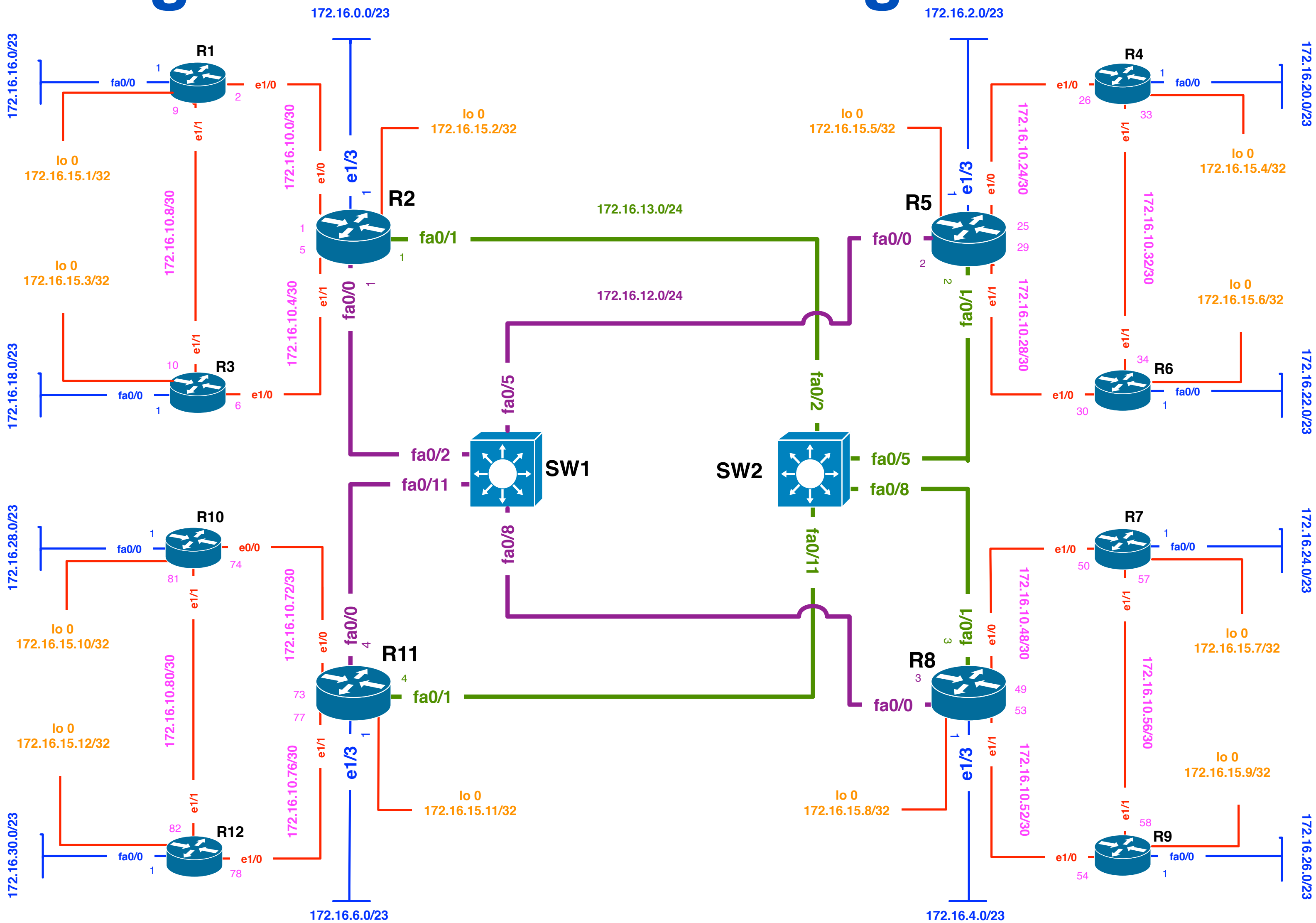
25	172.16.12.0	/24	Transport link PURPLE	No	
26	172.16.13.0	/24	Transport link GREEN	No	
27	172.16.15.0	/24	Loopback	No	

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



Case Study - OSPF

APNIC

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

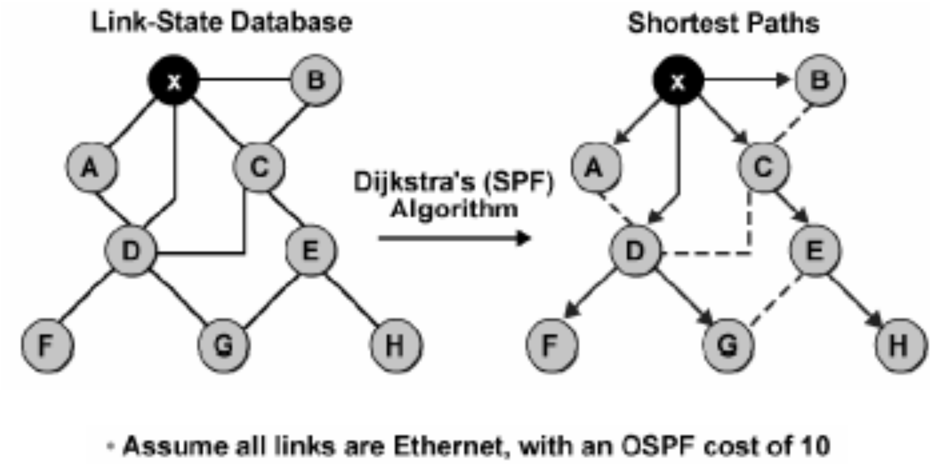
Link State Routing

- 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

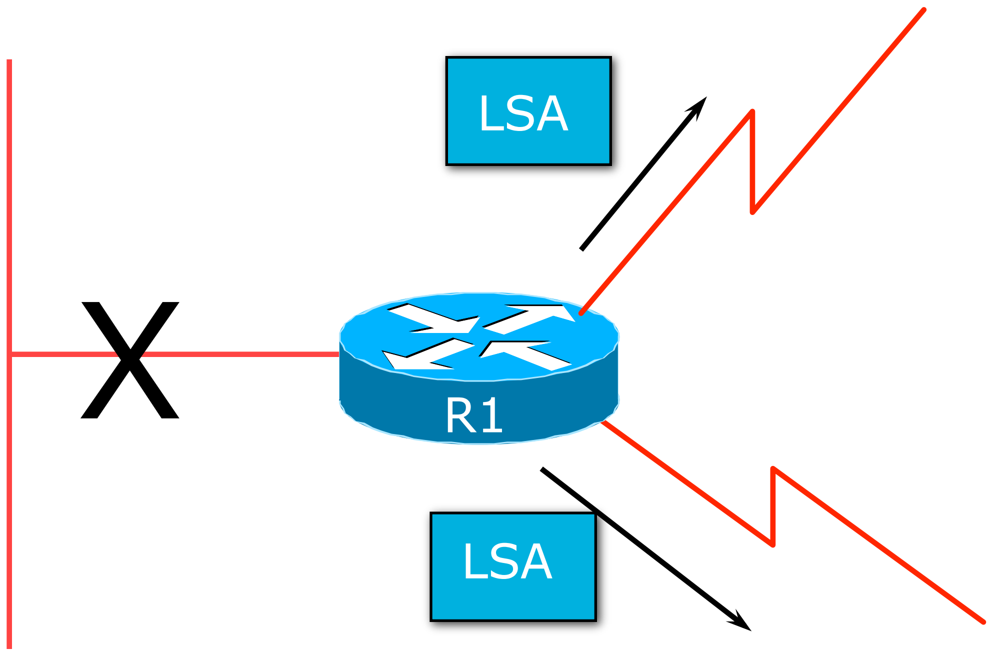
Shortest Path First (SPF) Tree



- 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

Diagram source: www.cisco.com

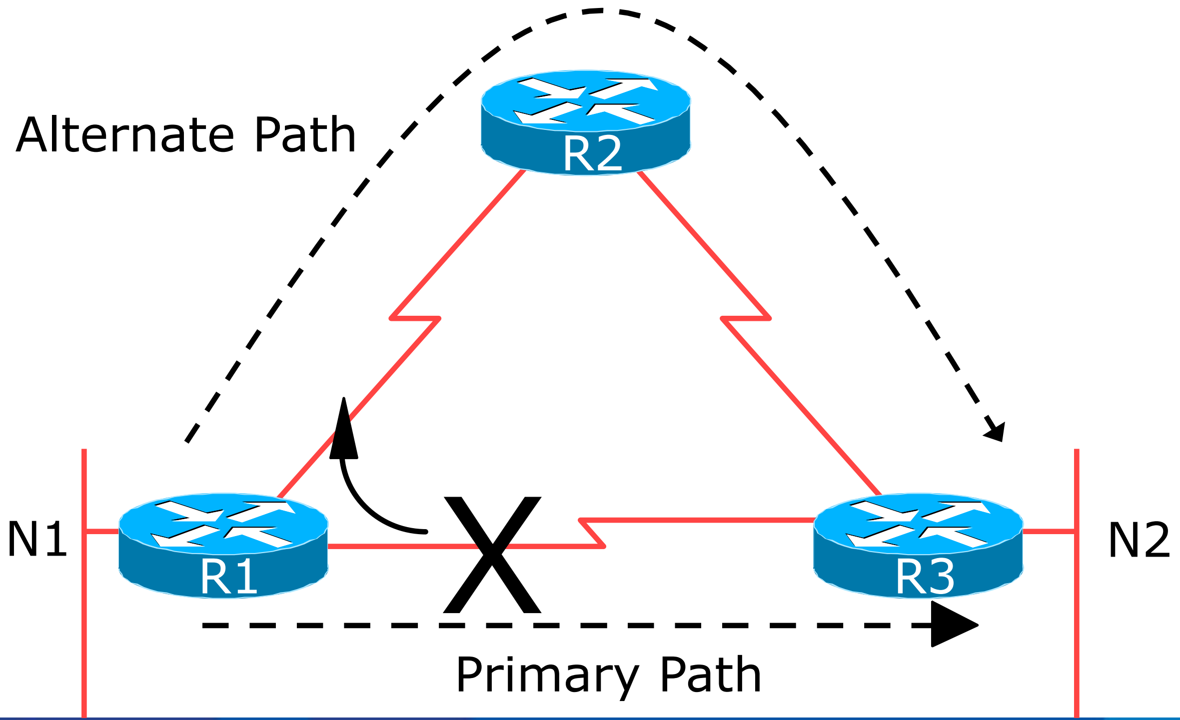
Low Bandwidth Utilisation



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

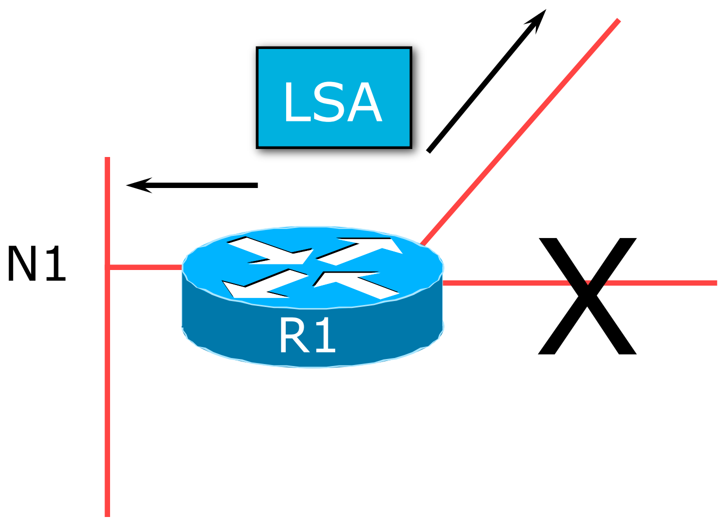
Fast Convergence

- 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

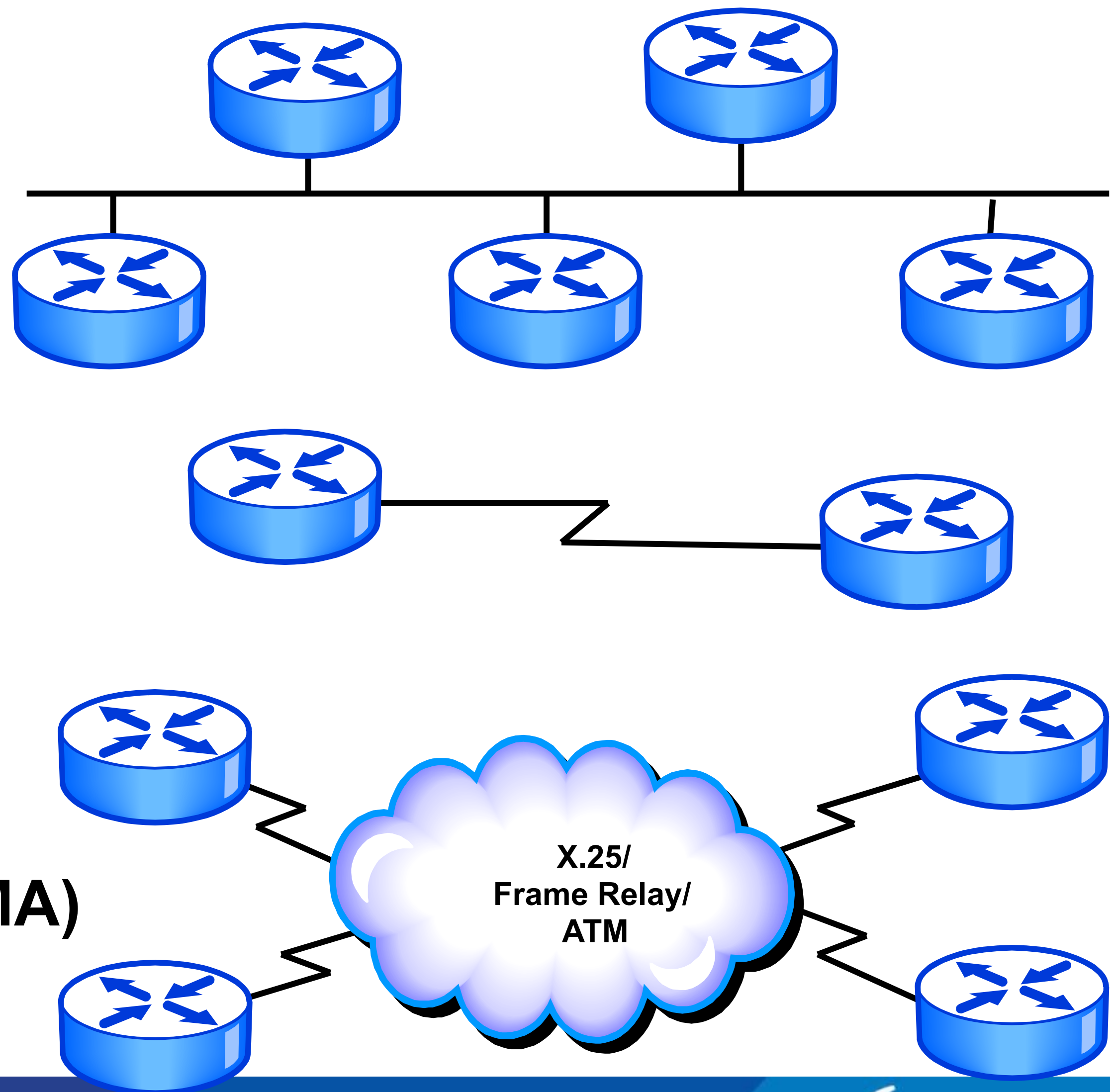
- Neighbor discovery
  - Send L3 multicast message (hello) to discover neighbors
- Exchanging topology table (LSDB)
  - Send L3 multicast message (DBD packets)
- Use SPF algorithm to select best path
  - Each router independently calculates best path from an identical topology database of an OSPF network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

OSPF Network Topology

Broadcast Multi-access

Point-to-Point

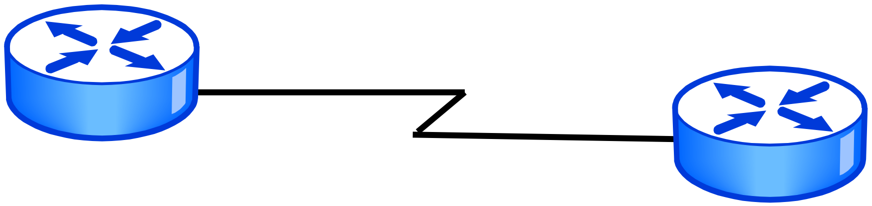
Non Broadcast Multi-access (NBMA)



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

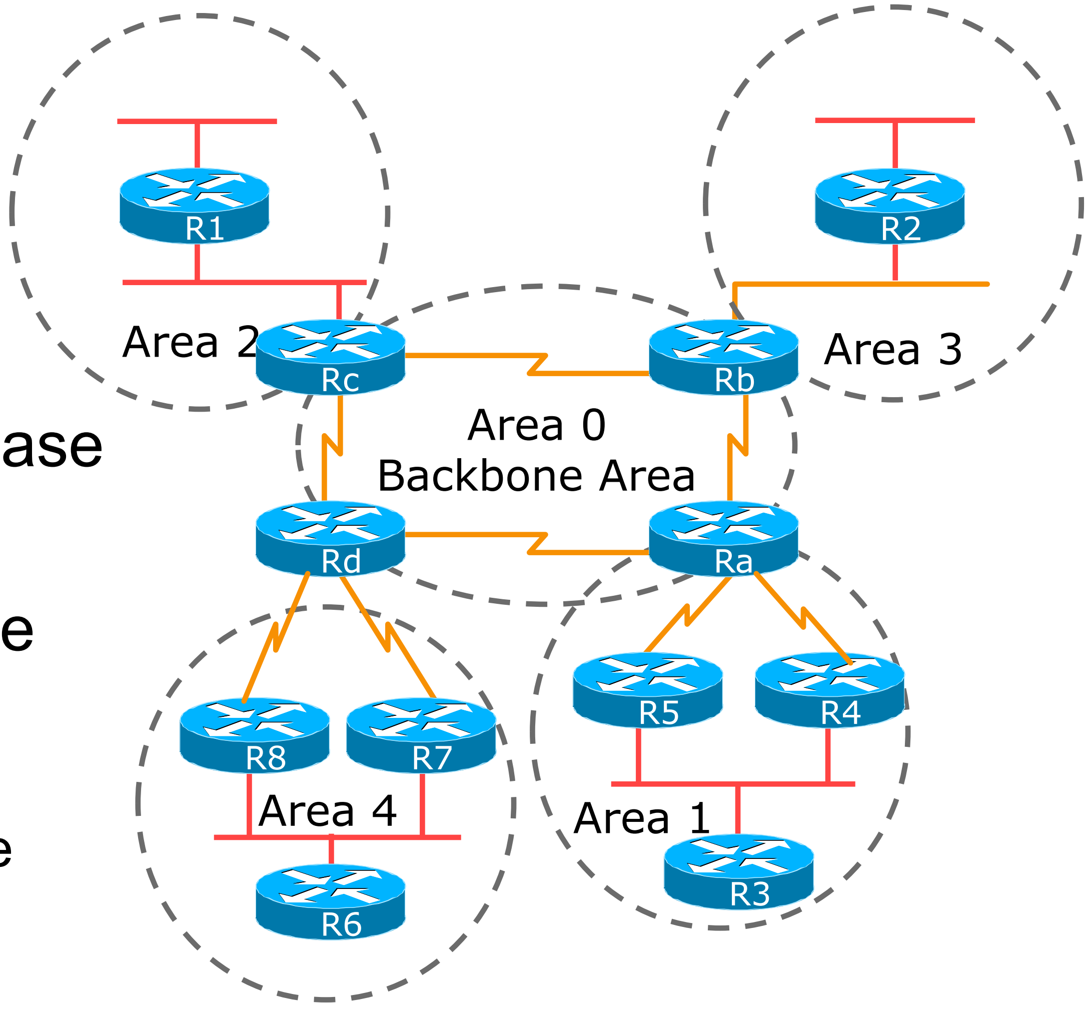
Point-to-Point Network



- 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

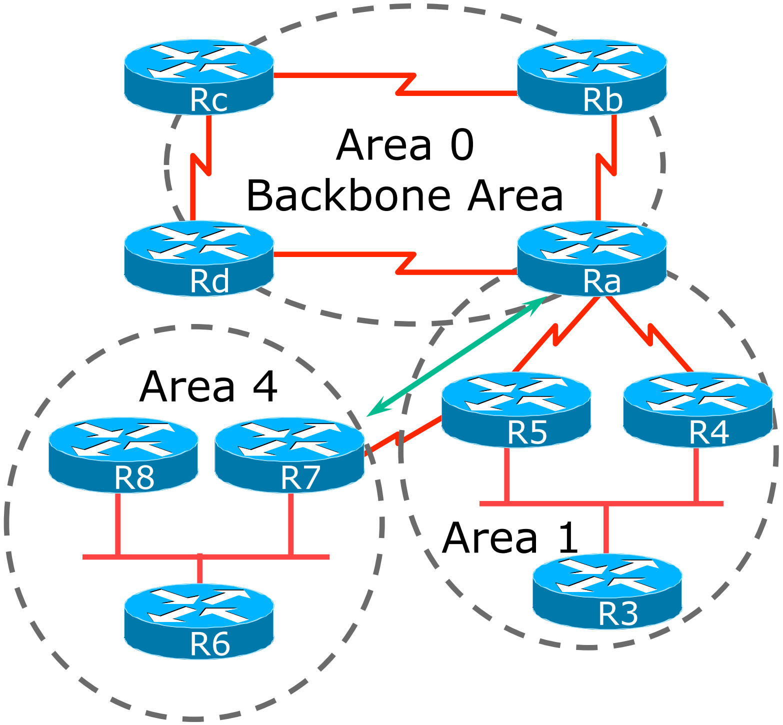
OSPF Areas

- 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

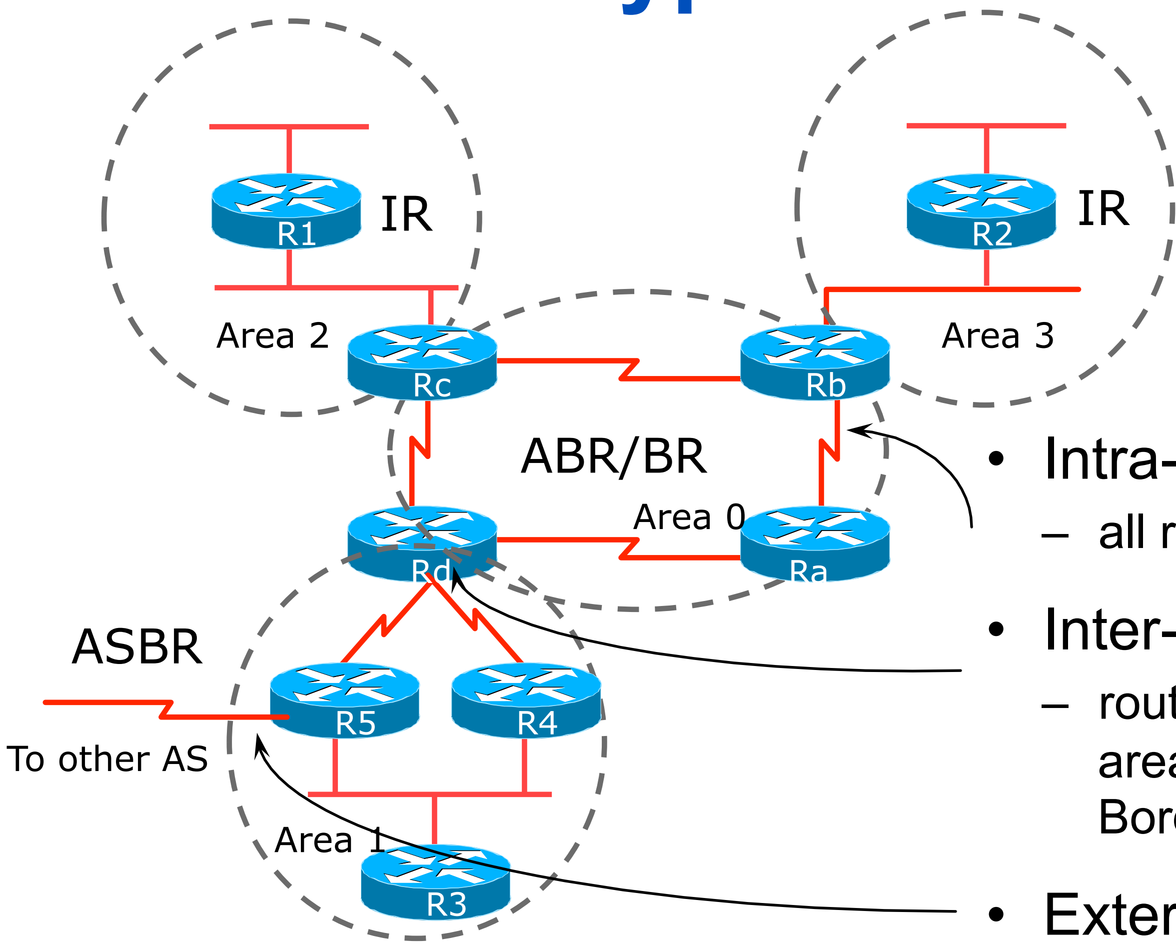


Virtual Links between OSPF Areas

- 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



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

ISP Use of Areas

- ISP networks use:
  - Backbone area
  - Regular area
- Backbone area
  - No partitioning
- Regular area
  - Summarisation of point to point link addresses used within areas
  - Loopback addresses allowed out of regular areas without summarisation (otherwise iBGP won't work)

Configuration of OSPF as IGP

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

Configuration of OSPF as IGP

- Minimum Router OS require for OSPF3:
  - Cisco IOS
  - 12.2(15)T or later (For OSPFv3)
  - 12.2(2)T or later (For IPv6 support)
  - Jun OS
  - JunOS 8.4 or later

Configuration of OSPF as IGP

- Before enabling OSPF3 on an Interface, the following steps must be done on a Router:
  - Enable IPv6 unicast routing
  - Enable IPv6 CEF
  - 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
```

```
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
```

```
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

IPV6 Deployment in IGP

- OSPFv3 or OSPF for IPv6 Overview:
  - OSPFv3 is described in RFC 2740
  - Most of OSPF3 functions are same as OSPFv2
  - In OSPFv3 routing process does not need to be explicitly created. Simply enabling OSPF on an interface will create routing process on a router

IPV6 Deployment in IGP

- OSPFv3 or OSPF for IPv6 Overview:
  - Multiple instances of OSPFv3 can be run on a link which is unlike in OSPFv2
  - OSPFv3 still use 32 bit address as router ID. If no IPv4 address is configured on any interface need to use router-id command to set 32 bit router-id.

IPV6 Deployment in IGP

- OSPFv3 or OSPF for IPv6 Overview:
  - LSA types and functions in OSPF3 are same as OSPF2
  - OSPFv3 use IPv6 address FF02::5 for AllSPF router multicast and IPv6 address FF02::6 for AllD router multicast
  - DR/BDR concepts for Broadcast Multi-access network are same in OSPFv3 as OSPFv2

Questions?

APNIC

Case Study - BGP

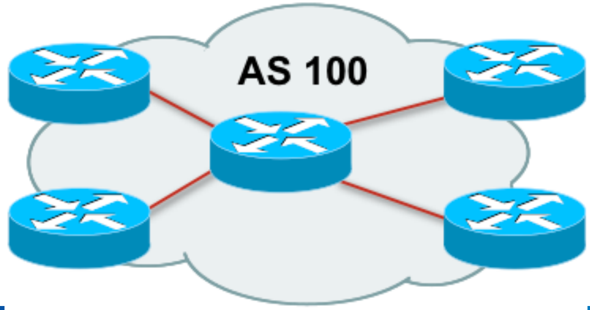
APNIC

What is Border Gateway Protocol?

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

What is an Autonomous System (AS)

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



BGP features

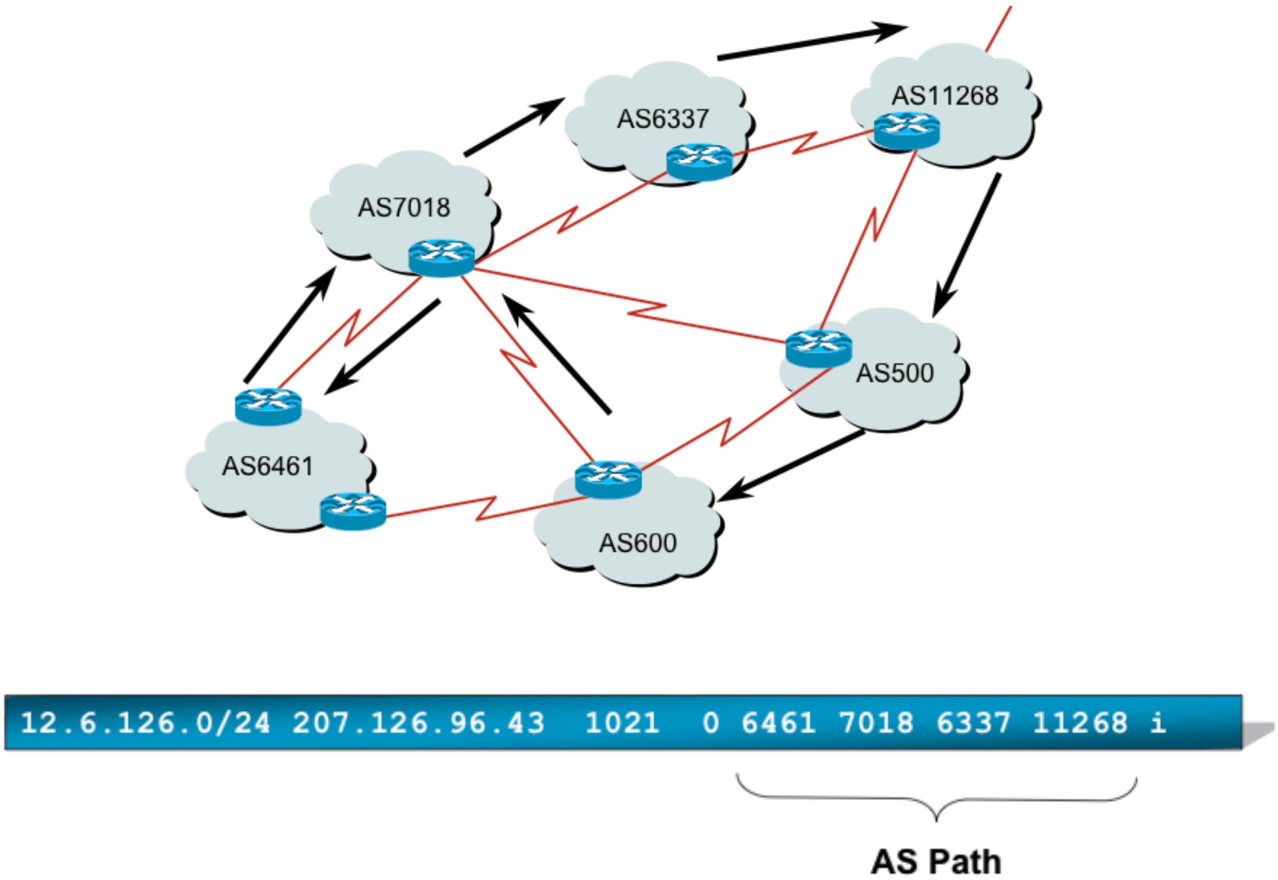
- Path Vector Routing Protocol
- Send incremental updates to peers
- Runs over TCP –Port 179
- Select path based on routing policy/ organization’s business requirement
- Support Classless Inter Domain Routing (CIDR) concept
- Widely used in today’s Internet Backbone
- Current BGP version is MP-BGP

What is Path Vector Routing Protocol

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

What is AS path?

- An AS path example:



BGP Traffic Arrangement Definition

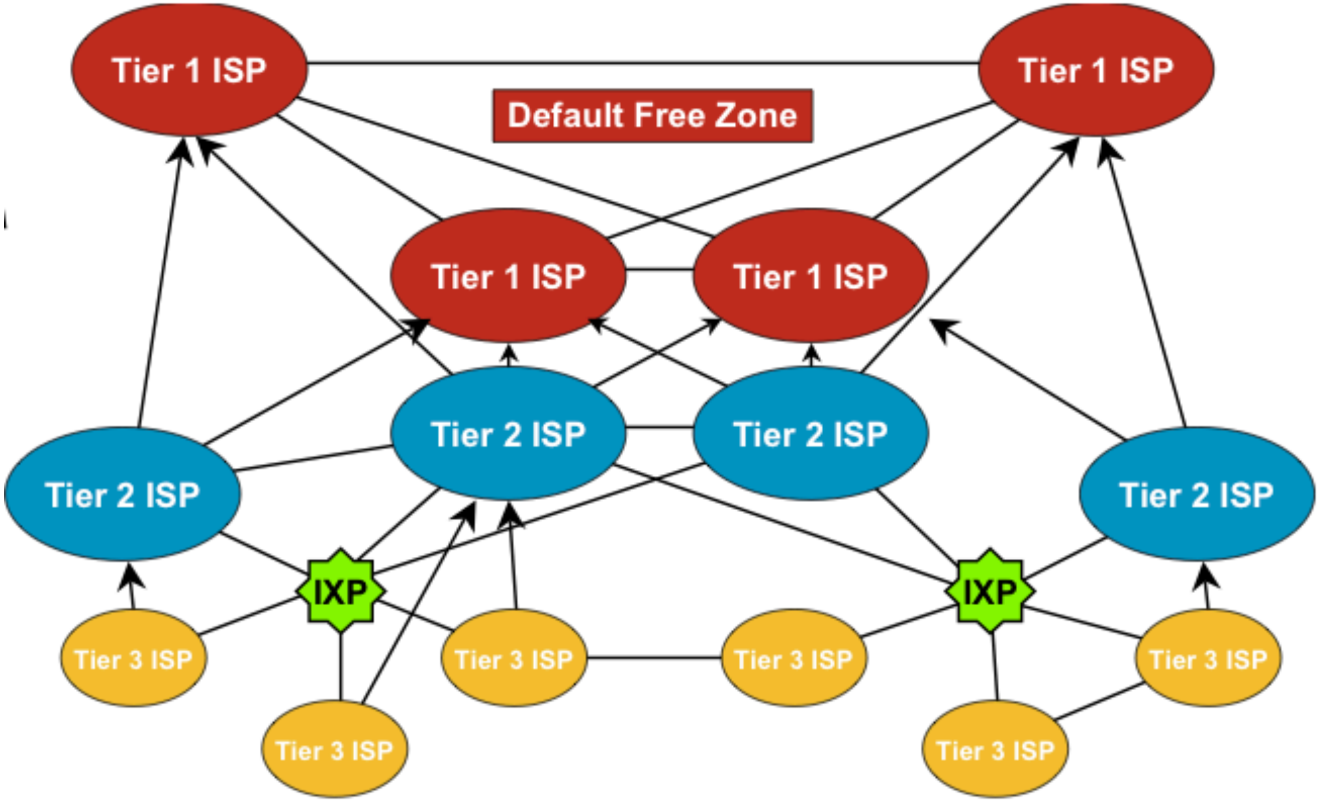
- Transit
  - Forwarding traffic through the network usually for a fee
  - I.e Internet service from upstream ISP
- Peering
  - Exchanging traffic without any fee
  - I.e Connection in an IXP
- Default
  - Where to send traffic if there no explicit route match in the routing table

What is Default Free Zone?

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

ISP Hireracial Connection

- Connectivity Diagram:



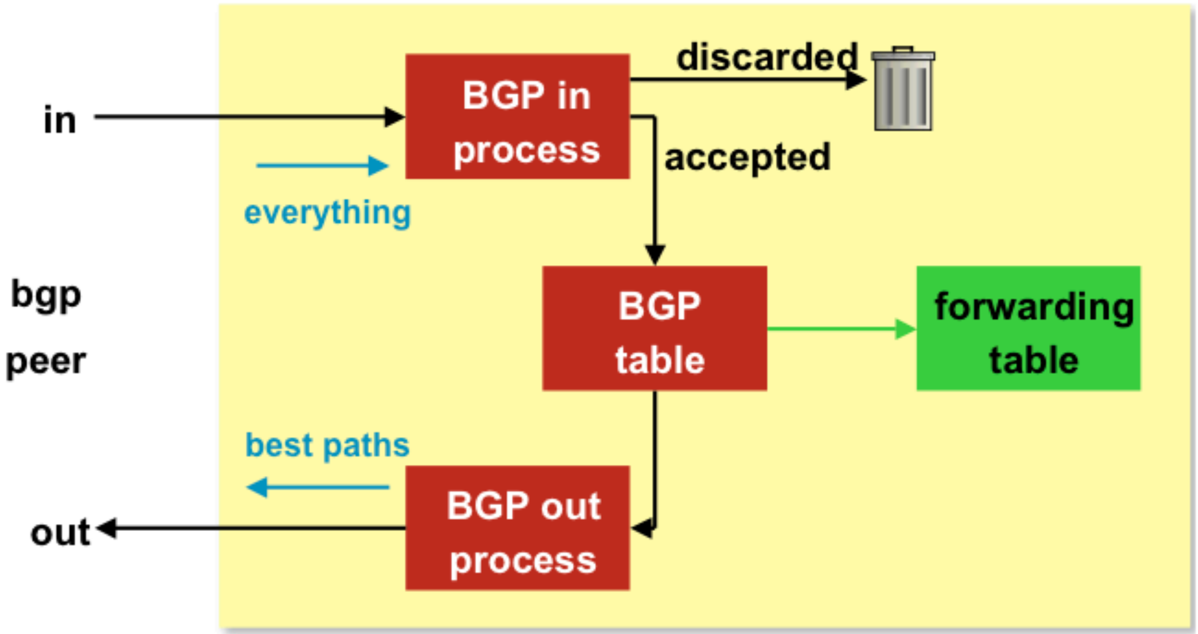
BGP General Operation

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

Constructing the Forwarding Table

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

Constructing the Forwarding Table

- Flowchart:
- 

BGP Terminology

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

Building Neighbor Relationship

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

BGP message type

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

BGP Open message

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

BGP Keepalive Message

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

BGP Update Message

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

BGP Notification message

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

BGP Neighbor Relationship States

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

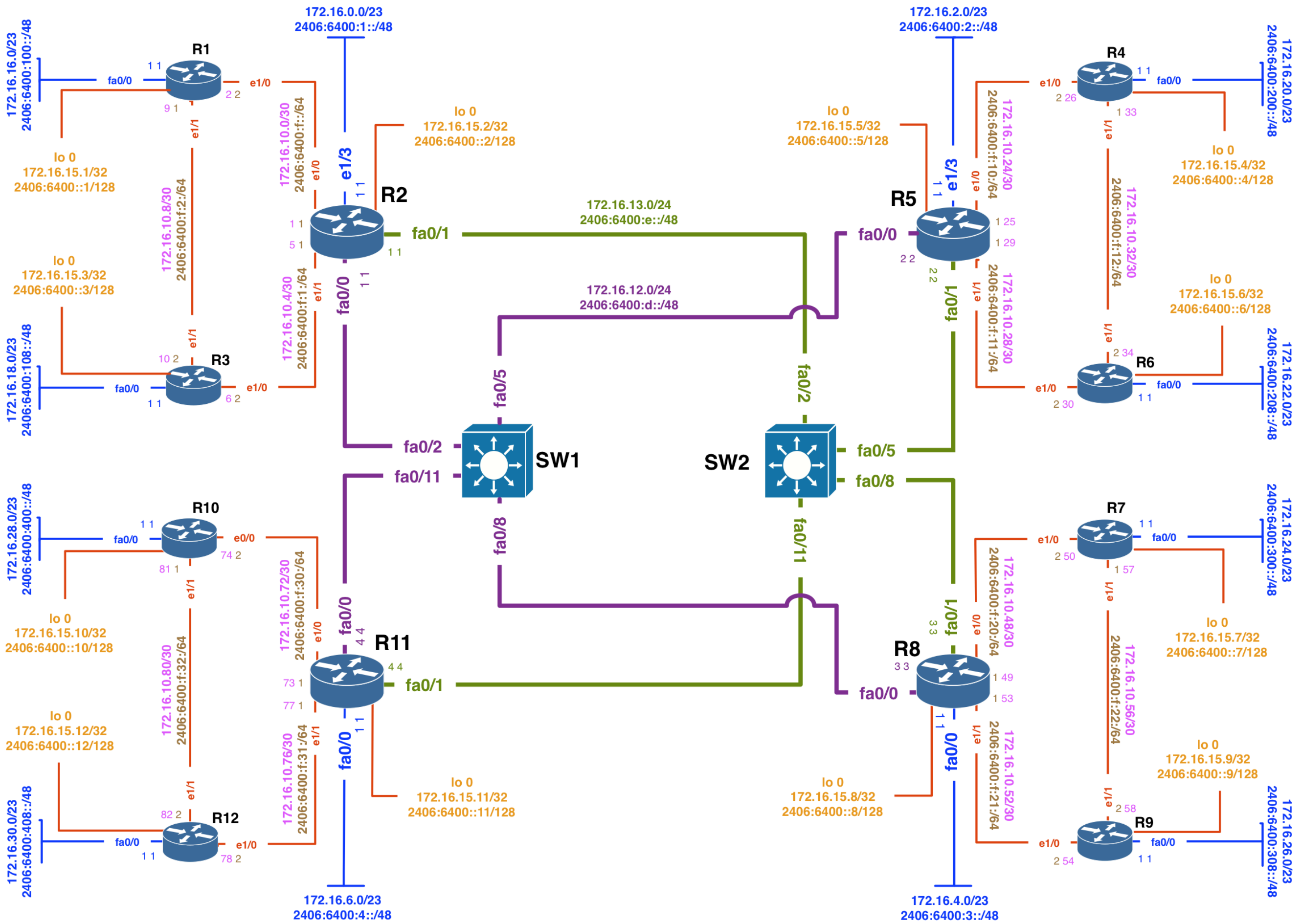
Troubleshoot BGP Neighbor Relation

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

iBGP Peering

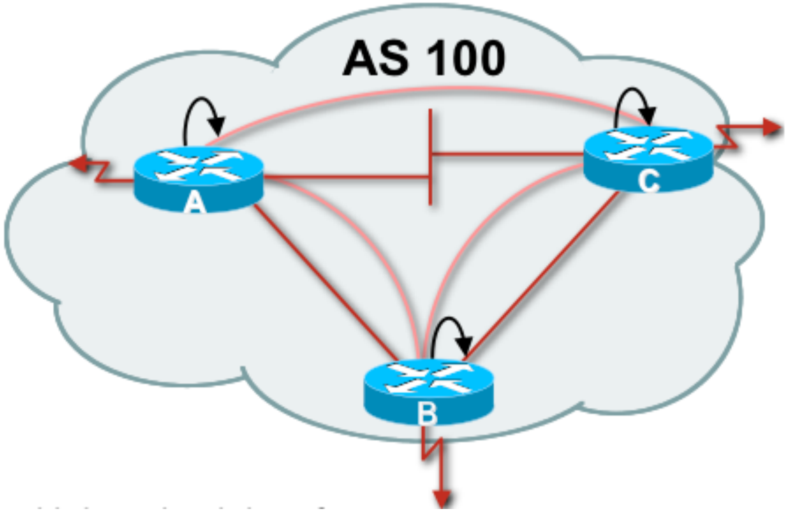
- BGP peer within the same AS
- Not required to be directly connected
- iBGP peering require full mesh peering
  - Within an AS all iBGP speaker must peer with other iBGP speaker
  - They originate connected network
  - Pass on prefixes learned from outside AS
  - They do not forward prefixes learned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

Training ISP IPV6 Addressing Pan



Training ISP IPv6 Address Plan

iBGP Peering with Loopback Interface

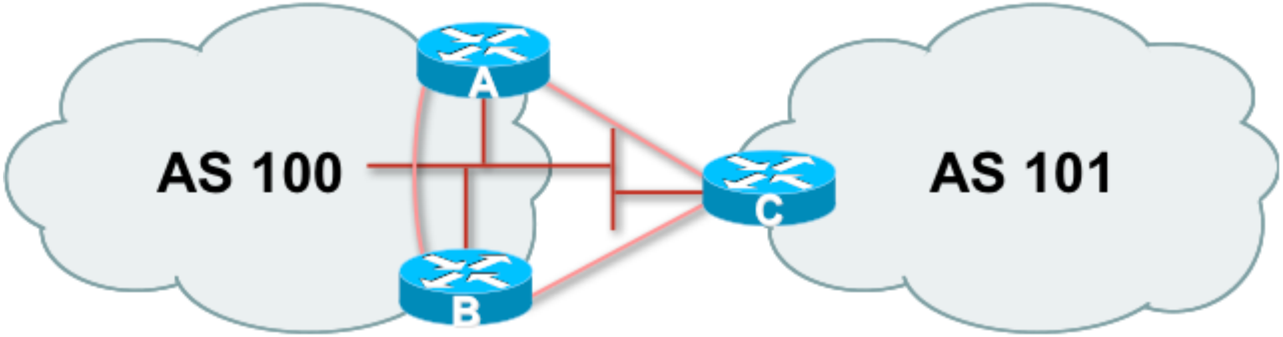


- If iBGP speakers has multiple connection then it is advisable to peer with loopback
- Connected network can go down which might loose iBGP peering
- Loopback interface will never go down

iBGP Neighbor Update Source

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

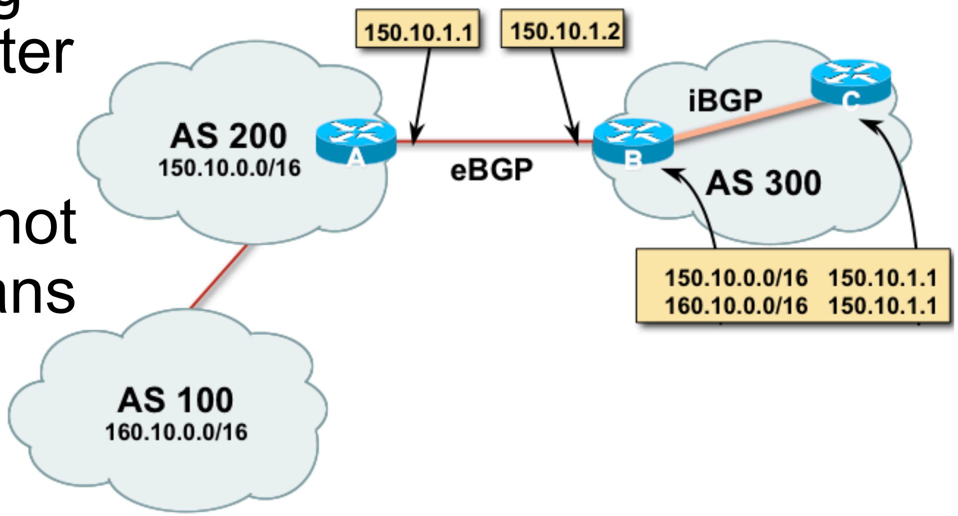
eBGP Peering



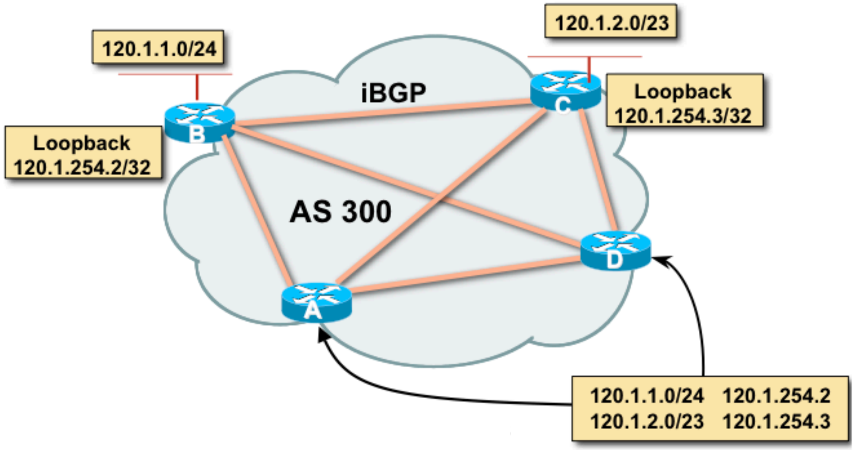
- Peering with BGP speaker in different AS
- Peers should be directly connected and share same WAN link
- eBGP neighbors are usually routed through connected network

BGP Next Hop Behavior

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



iBGP Next Hop



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

BGP Synchronous Rule

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

BGP Attributes

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

Well-known mandatory

- AS-Path
- Next-hop
- Origin

Well-known discretionary

- Local preference
- Atomic aggregate

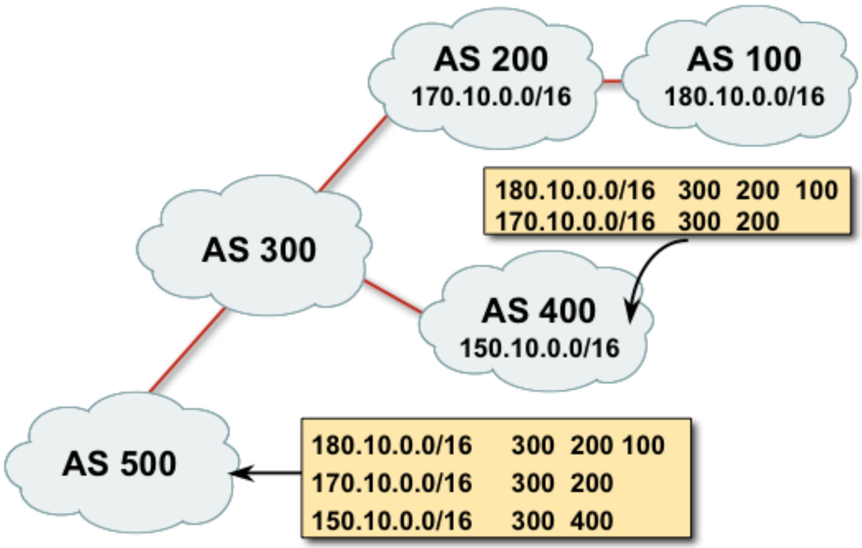
Optional transitive

- Community
- Aggregator

Optional non-transitive

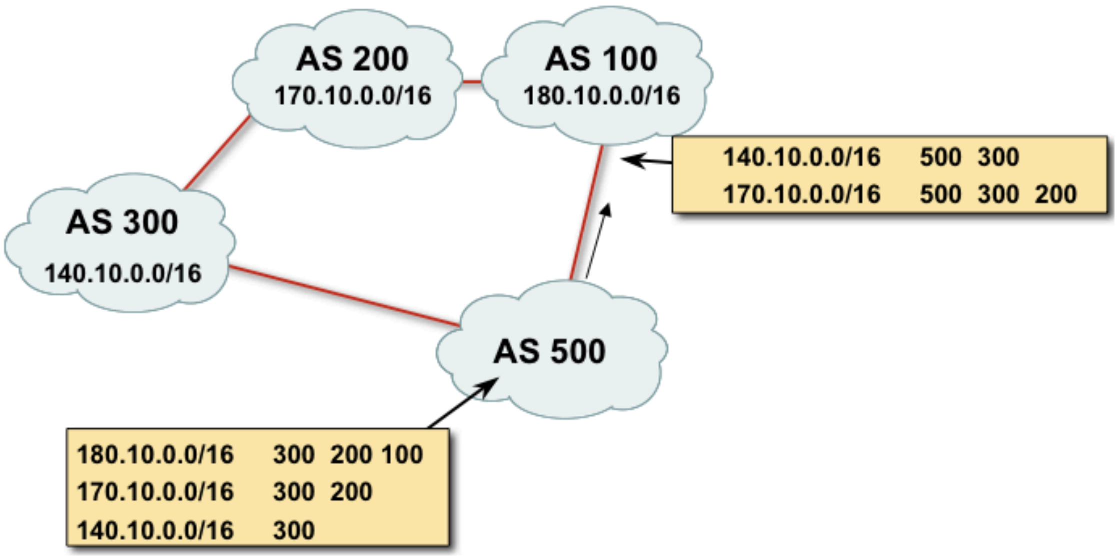
- Multi-exit-discriminator (MED)

AS Path Attribute



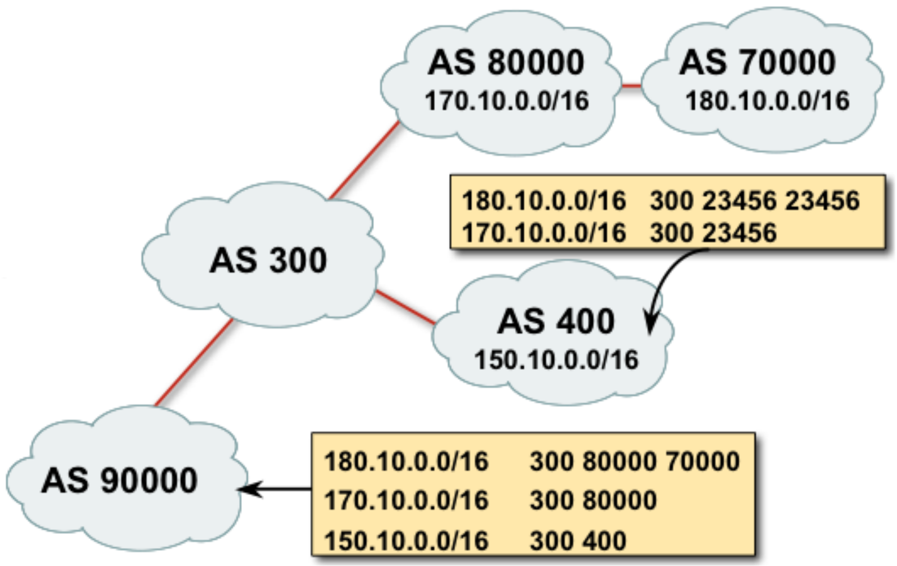
- Sequence of ASes a route has traversed
- Used for
  - Loop detection
  - Path metrics where the length of the AS Path is used as in path selection

AS Path Loop Detection



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

AS Path Attribute (2 byte and 4 byte)



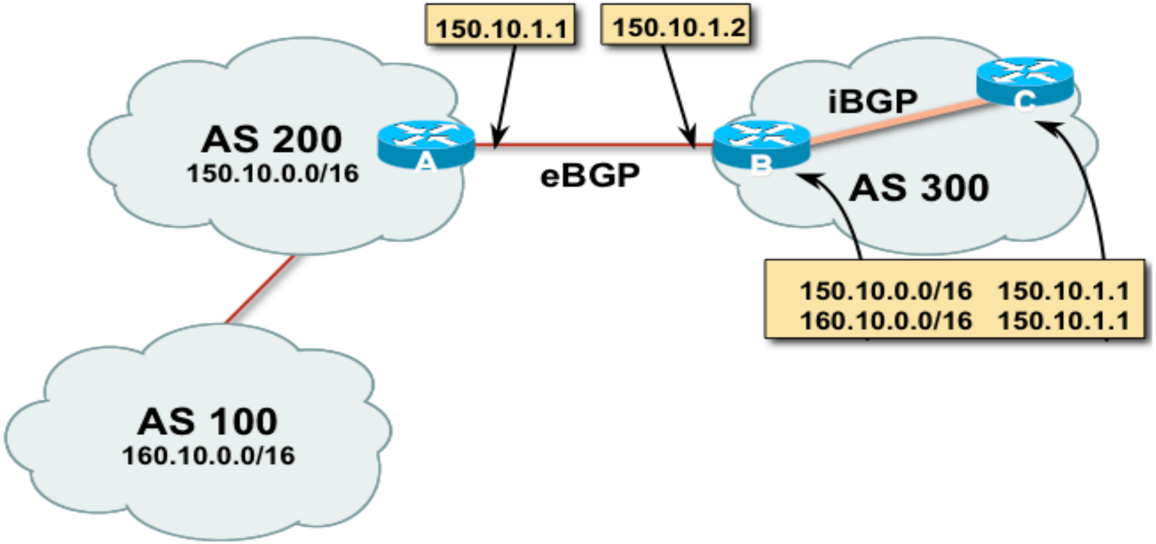
- Internet with 16-bit and 32-bit ASNs
  - 32-bit ASNs are 65536 and above
  - AS-PATH length maintained

AS Path and AS4 Path Example

Router5:

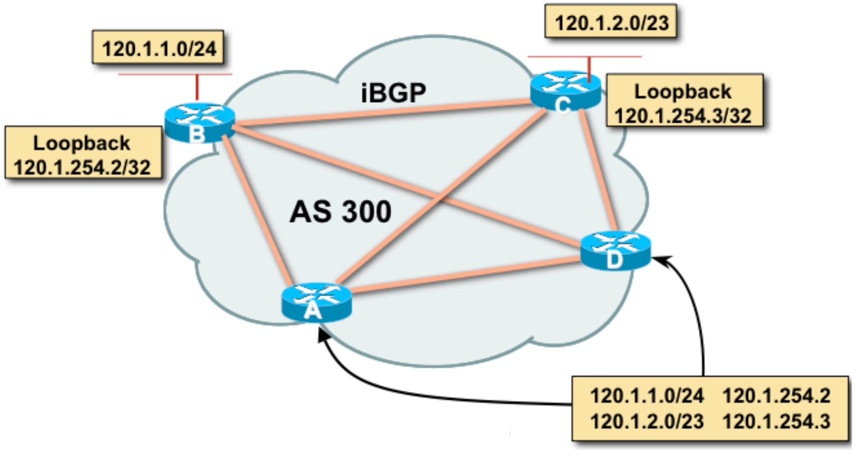
```
Network          Next Hop          Metric  LocPrf Weight        Path
*> 2001::/32     2406:6400:F:41::1
                                           0 23456 38610 6939 I
* i              2406:6400:D::5 0          100   0 45192 4608 4826 6939 i

*> 2001:200::/32    2406:6400:F:41::1
                                           0 23456 38610 6939 2500 i
* i              2406:6400:D::5 0 100 0 45192 4608 4826 6939 2500 i
```

eBGP Next Hop



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

iBGP Next Hop



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

Next Hop Best Practice

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

Next Hop Self Configuration

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

BGP Origin Attribute

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

BGP Origin Attribute Example

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

r RIB-failure, S Stale

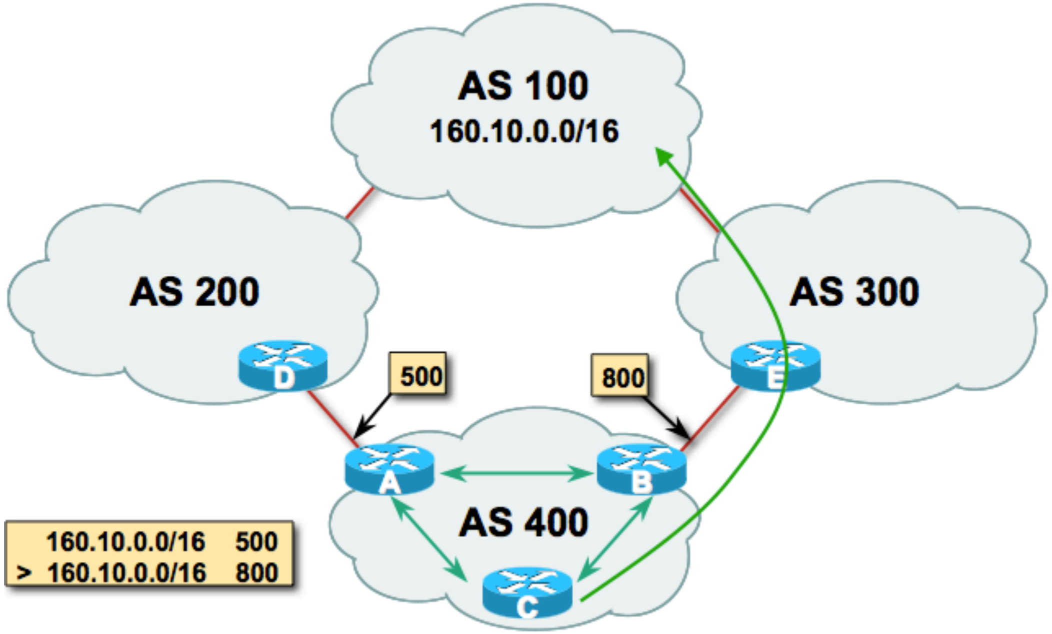
Origin codes: i - IGP, e - EGP, ? – incomplete

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

BGP Local Preference Attribute

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

BGP Local Preference Attribute



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

BGP Local Pref Attribute Example

Network Next Hop Metric LocPrf Weight Path

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

0 23456 38610 6939 i

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

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

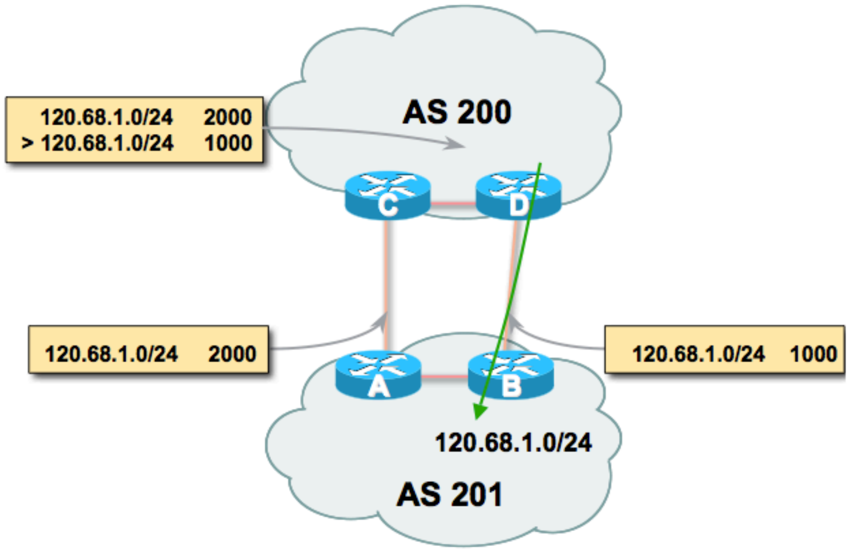
0 23456 38610 6939 2500 i

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

BGP MED Attribute

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

BGP MED Attribute



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

BGP MED Example

Network Next Hop Metric LocPrf Weight Path

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

0 23456 38610 6939 i

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

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

0 23456 38610 6939 2500 i

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

BGP Community Attribute

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

BGP Route Selection Process

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

BGP Peer Group

- Defines a template with parameters set for a group of neighbors instead of individually
- Useful when many neighbors have the same outbound policies
- Members can have a different inbound policy
- Updates generated once per peer group
- Simplifies configuration

BGP Peer Group

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

BGP Peer Group -Advantages

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

BGP Peer Group -BCP

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

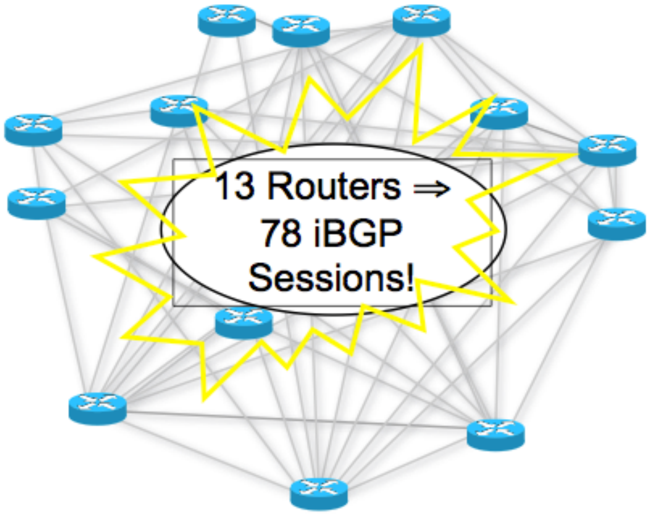
BGP Route Reflector

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

BGP Route Reflector

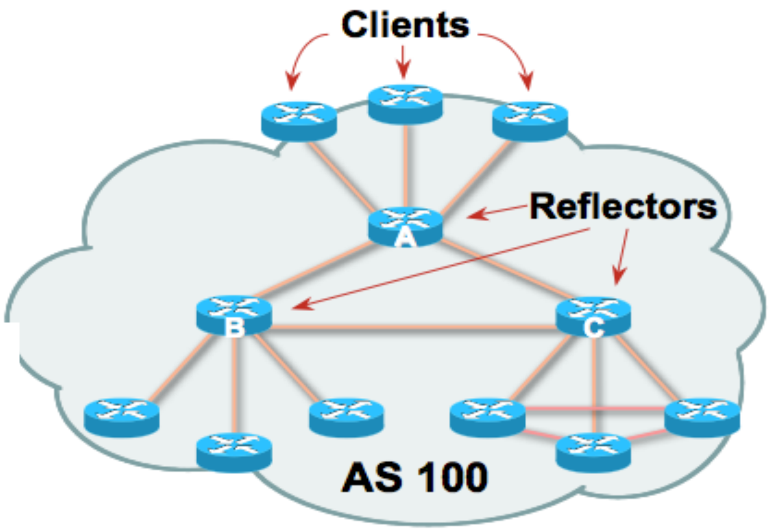
- iBGP use TCP to make reliable delivery of these large routing information across all core router in a transit ISP
- TCP can not broadcast so need to make individual delivery to all iBGP speaker
- To protect routing loop iBGP use split horizon rule so can not send routing information to neighbor learn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

BGP Route Reflector



- Avoid ½n(n-1) iBGP mesh
- n=1000 ⇒ nearly half a million iBGP sessions!
- Solution -Route reflector

BGP Route Reflector



- There will be Reflector, Client and Non-Client
- Reflector receives path from clients and non-clients
- Select the best path then If best path is from client, reflect to other clients and non-clients
- If best path is from non-client, reflect to clients only
- Described in RFC4456

Case study- Deployment IPv6 in EGP

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

Case study- Deployment IPv6 in EGP

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

Case study- Deployment IPv6 in EGP

- Basic BGP Configuration:

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

Case study- Deployment IPv6 in EGP

Adding iBGP Neighbor:

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

iBGP neighbor is always recommended with loopback interface

Case study- Deployment IPv6 in EGP

Announcing IPv6 Prefix:

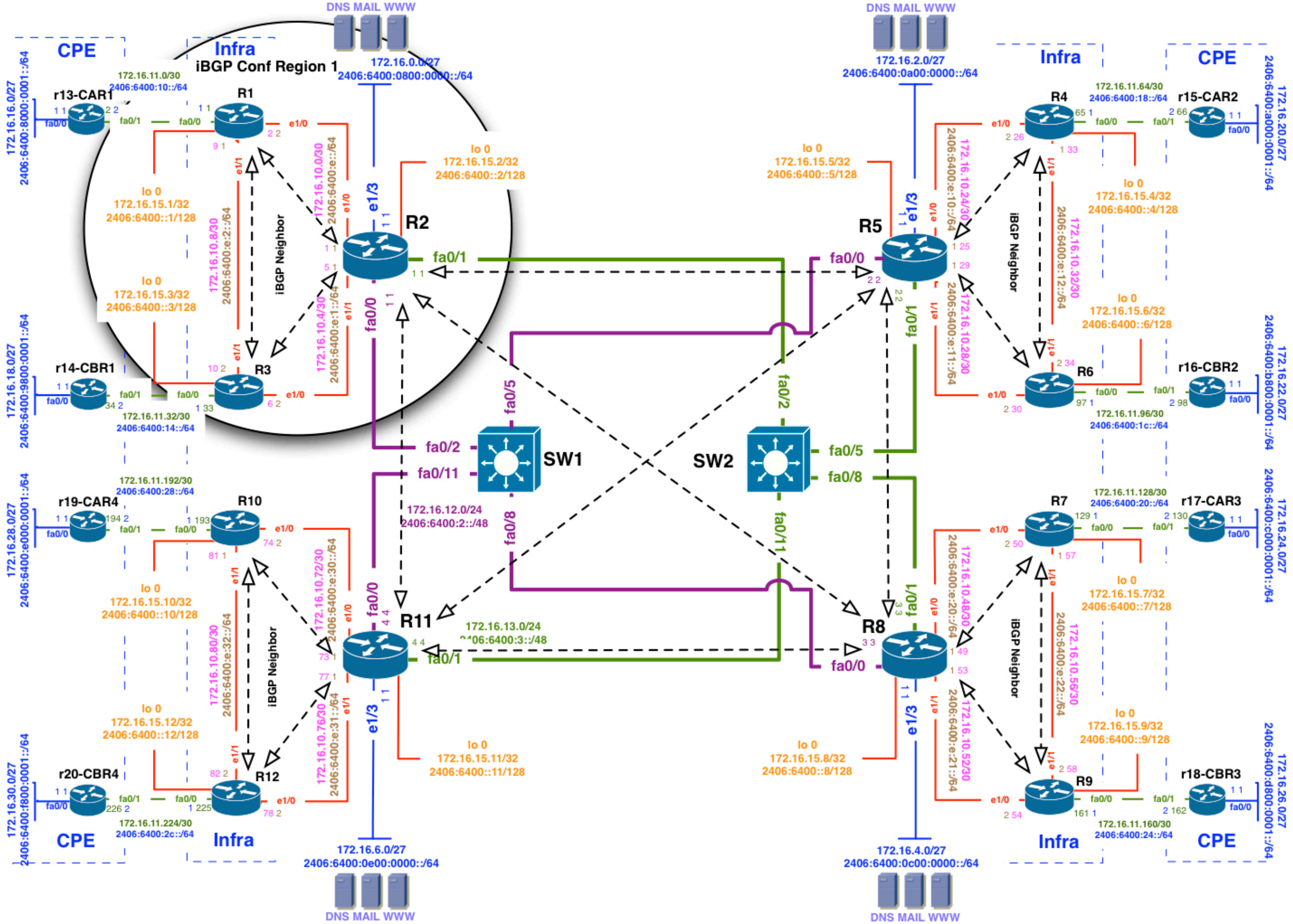
```
router bgp 17821
address-family ipv6
!
neighbor 2406:6400:0000:0000::2 remote-as 17821
neighbor 2406:6400:0000:0000::2 update-source loopback 0
neighbor 2406:6400:0000:0000::2 activate
!
network 2406:6400:0100:0000::/48
```

Case study- Deployment IPv6 in EGP

Add Pull-up route if needed:

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

iBGP Peering For Region 1



IPv4 iBGP Conf POP Router

Router1

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

IPv4 iBGP Configuration Verification

POP Router

sh bgp ipv4 unicast summary

sh bgp ipv4 unicast

sh ip route bgp

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

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

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

IPv6 iBGP Conf POP Router

Router1

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

IPv6 iBGP Configuration Verification

POP Router

sh bgp ipv6 unicast summary

sh bgp ipv6 unicast

sh ipv6 route bgp

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

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

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

IPv4 iBGP Conf Core Router

Router2 Configuration

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

IPv4 iBGP Conf Core Router

Router2 Configuration

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

IPv4 iBGP Configuration Verification

Core Router

sh bgp ipv4 unicast summary

sh bgp ipv4 unicast

sh ip route bgp

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

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

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

IPv6 iBGP Conf Core Router

Router2 Configuration

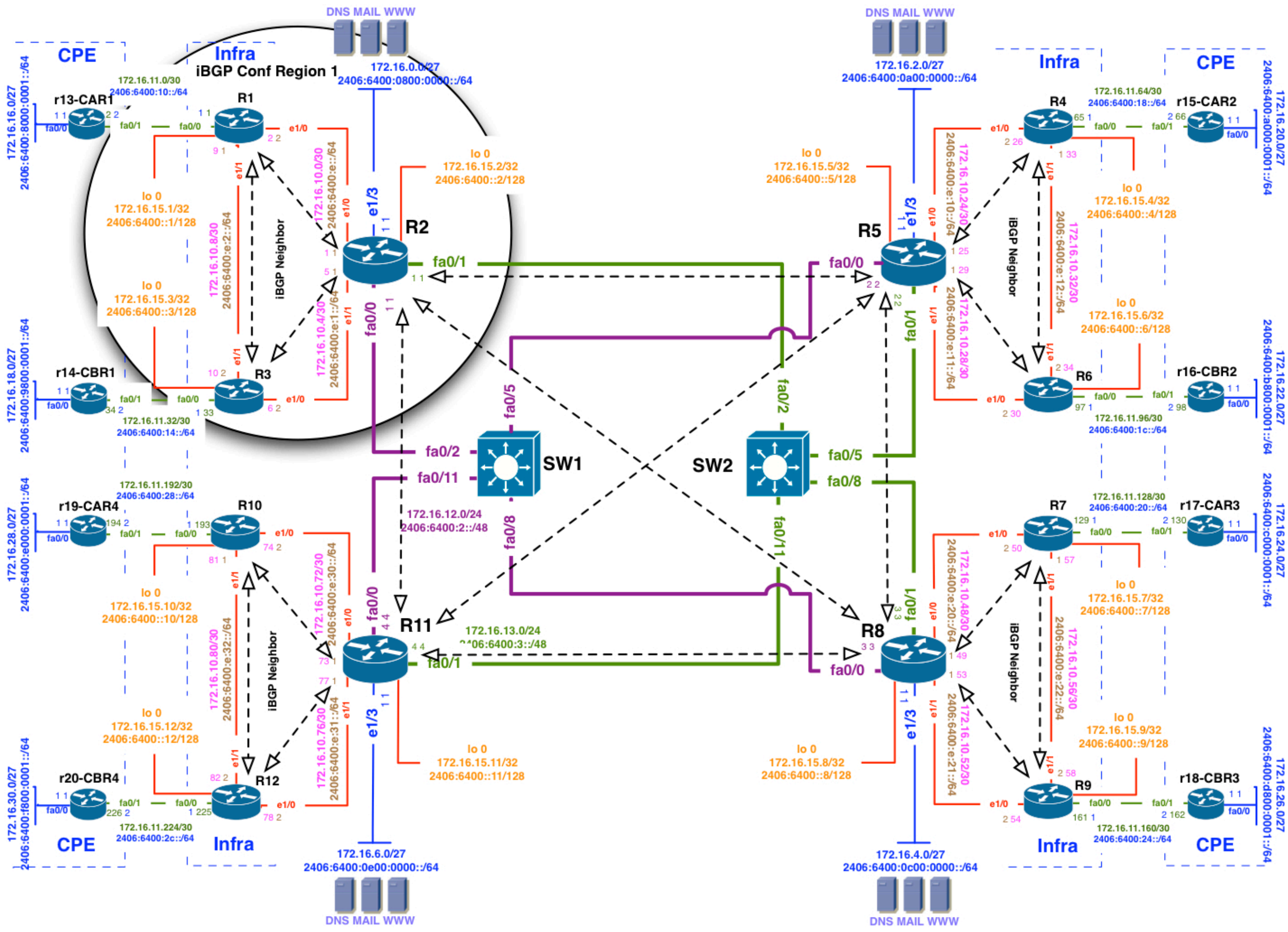
```
config t
router bgp 17821
address-family ipv6
no synchronization
neighbor 2406:6400:0000:0000::1 remote-as 17821
neighbor 2406:6400:0000:0000::1 update-source loopback 0
neighbor 2406:6400:0000:0000::1 activate
neighbor 2406:6400:0000:0000::3 remote-as 17821
neighbor 2406:6400:0000:0000::3 update-source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

IPv6 iBGP Configuration Verification

- Core Router

```
sh bgp ipv6 unicast summary

sh bgp ipv6 unicast

sh ipv6 route bgp

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

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

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

iBGP Full Mesh Issue

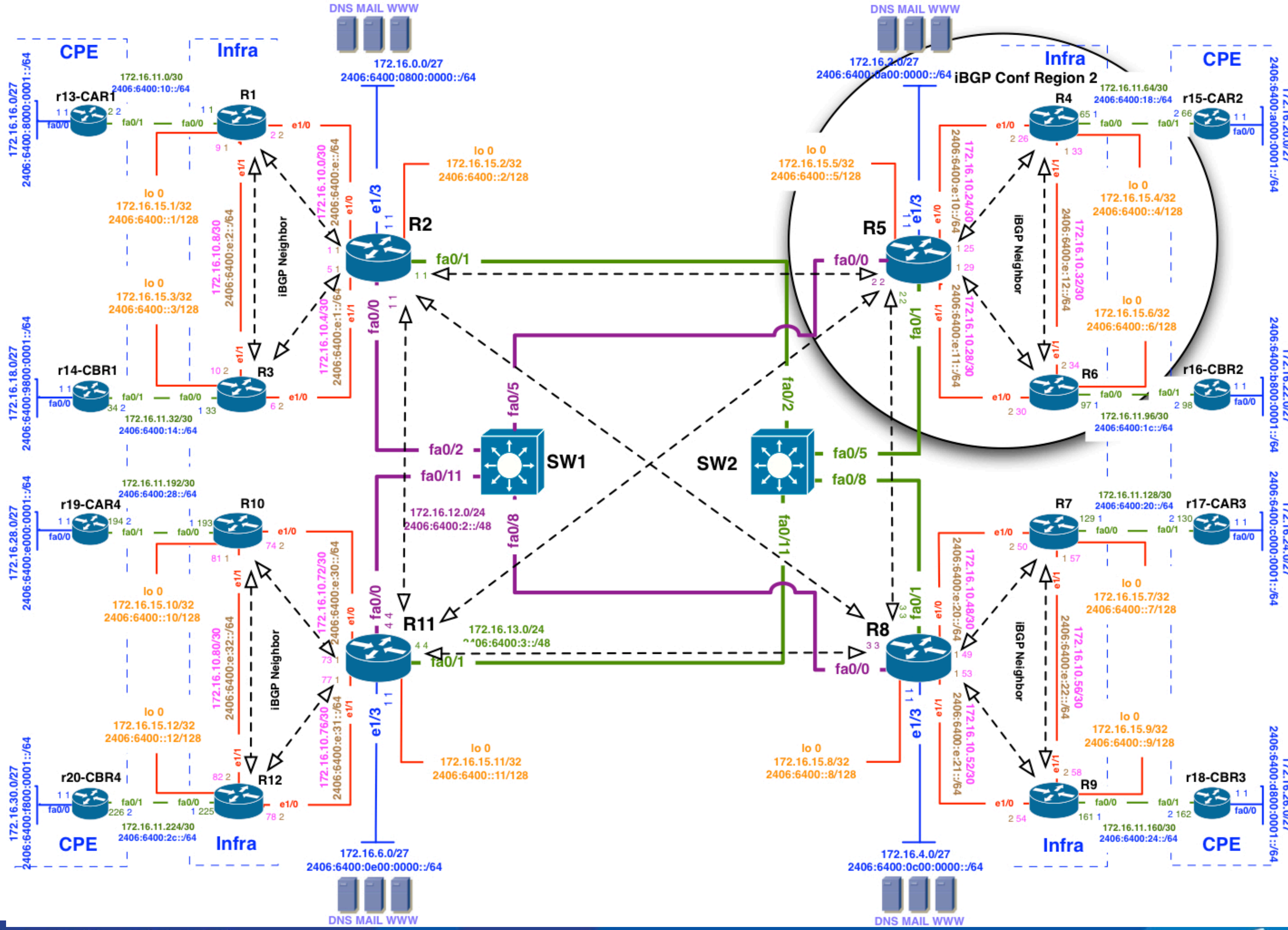


iBGP Full Mesh Issue

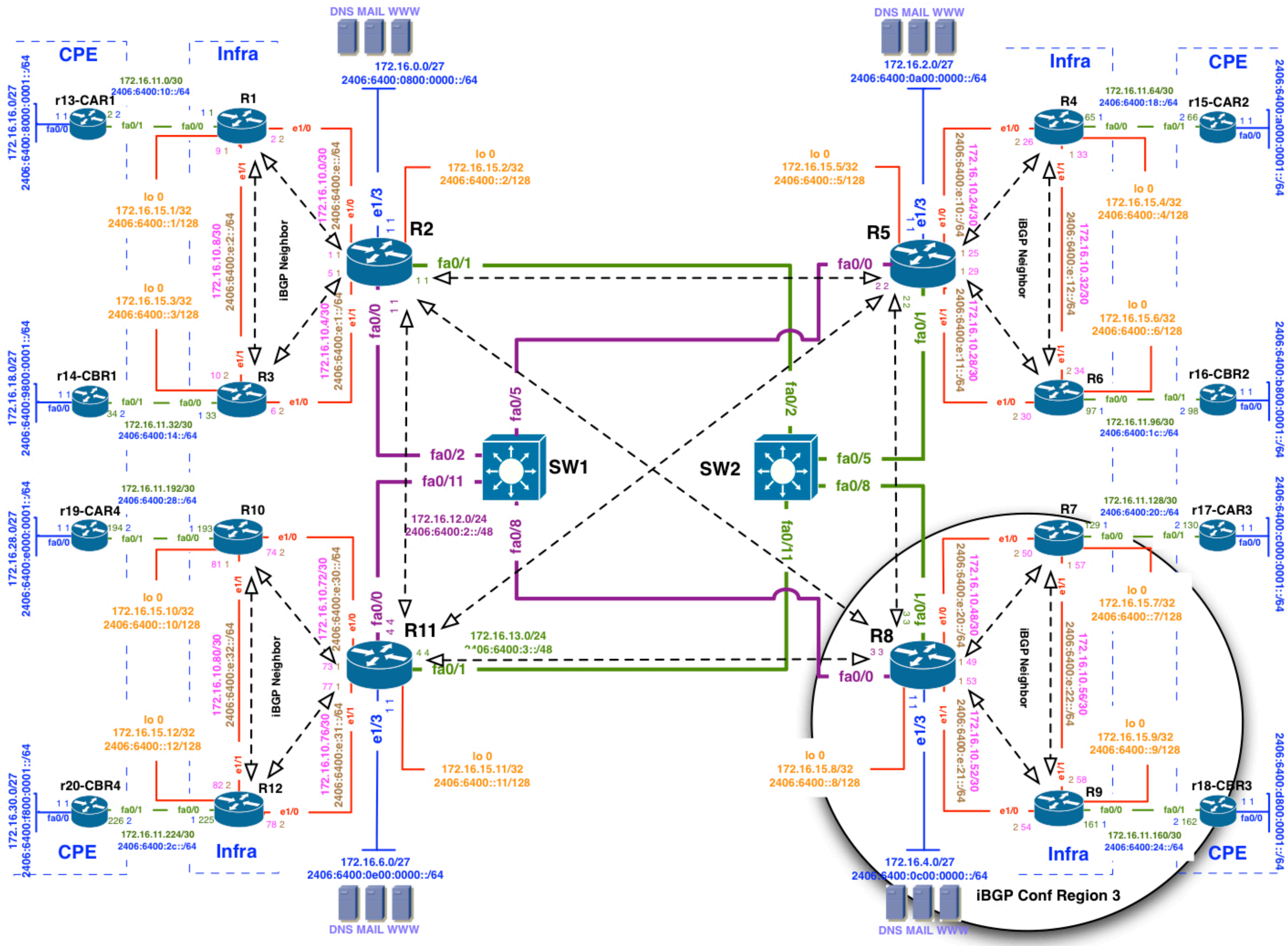
Route reflector configuration:

```
router bgp 17821
address-family ipv6
!
neighbor 2406:6400:0000:0000::1 remote-as 17821
neighbor 2406:6400:0000:0000::1 update-source loopback 0
neighbor 2406:6400:0000:0000::1 activate
!
neighbor 2406:6400:0000:0000::1 route-reflector-client
```

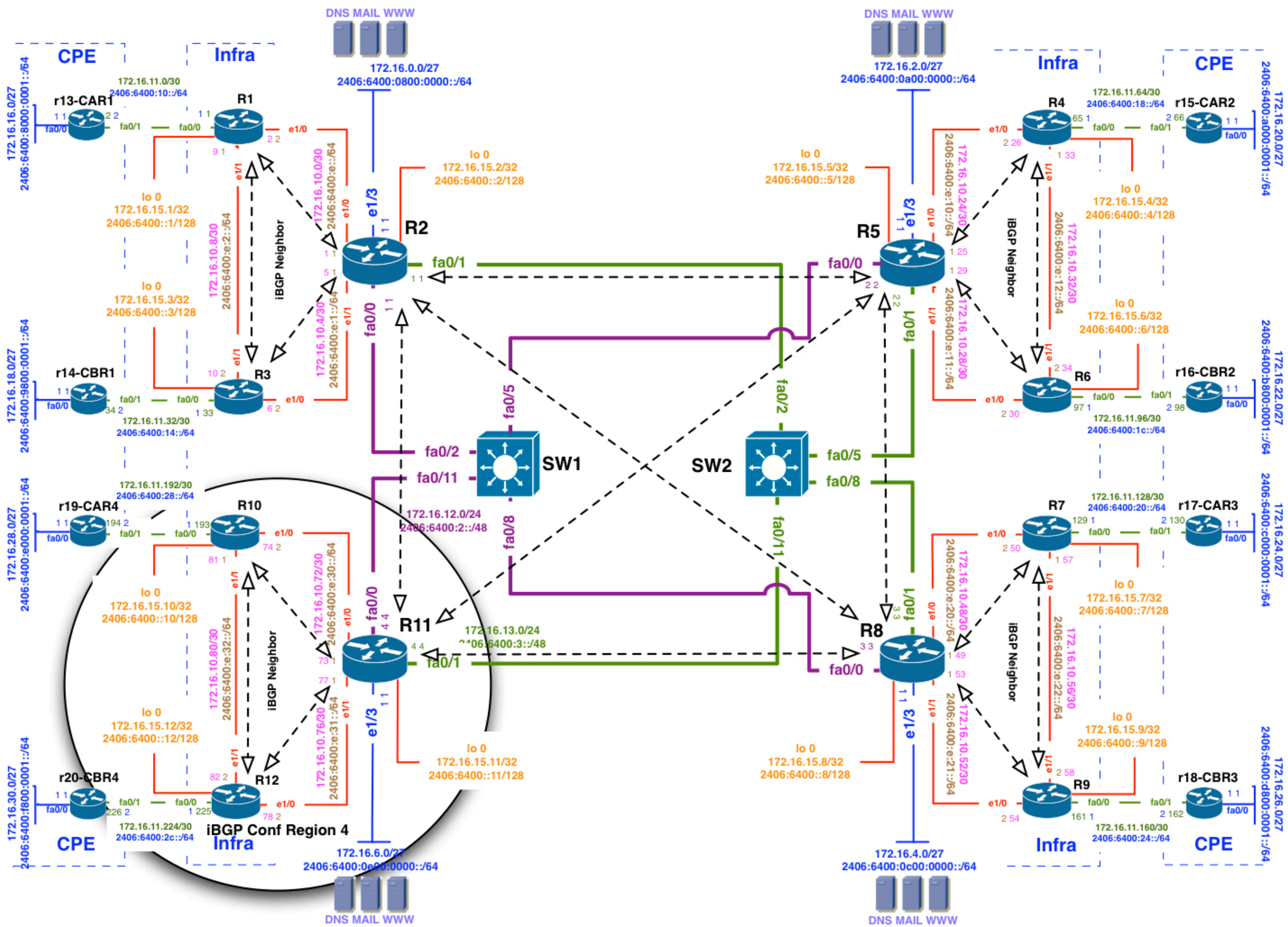
iBGP Peering For Region 2



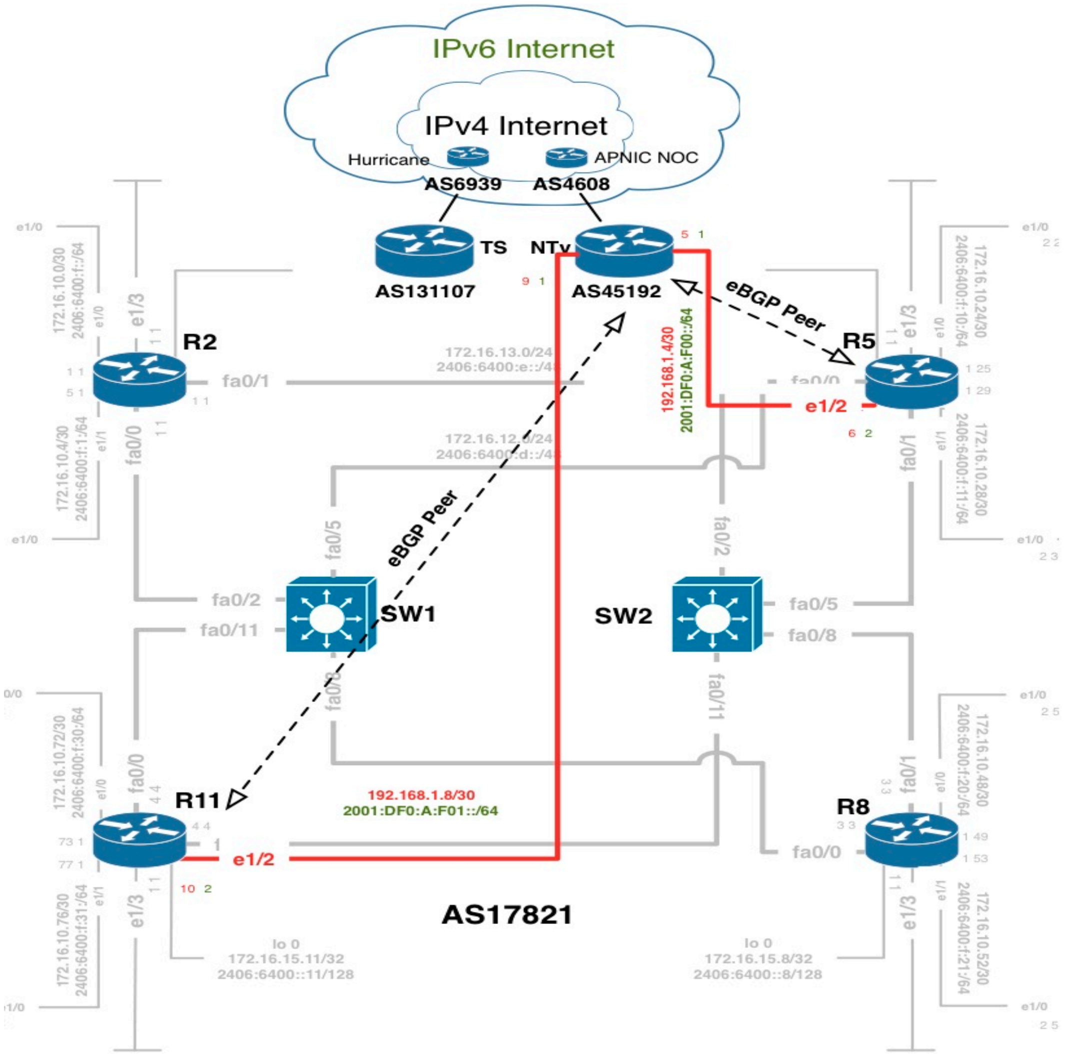
iBGP Peering For Region 3



iBGP Peering For Region 4



IPv6 Native Transit Conf Plan



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

IPv6 Native Transit Configuration

- Router5

```
config t
router bgp 17821
address-family ipv6
neighbor 2406:6400:000D:0000::5 remote-as 45192
neighbor 2406:6400:000D:0000::5 activate
neighbor 2406:6400:000E:0000::5 remote-as 45192
neighbor 2406:6400:000E:0000::5 activate
exit
exit
exit
Wr
```

Questions?

APNIC

IXP Operation

APNIC

IXP Configuration

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

IX Peering Model

- BLPA (Bi-Lateral Peering Agreement)
  – IX will only provide layer two connection/switch port to ISPs
  – Every ISPs will arrange necessary peering arrangement with others by their mutual business understanding.

- MLPA (Multi-Lateral Peering Agreement)
  – IX will provide layer two connection/switch port to ISPs
  – Each ISP will peer with a route server on the IX.
  – Route server will collect and distribute directly connected routes to every peers.

IXP Peering Policy

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

IXP Peering Policy

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

IXP Peering Policy

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

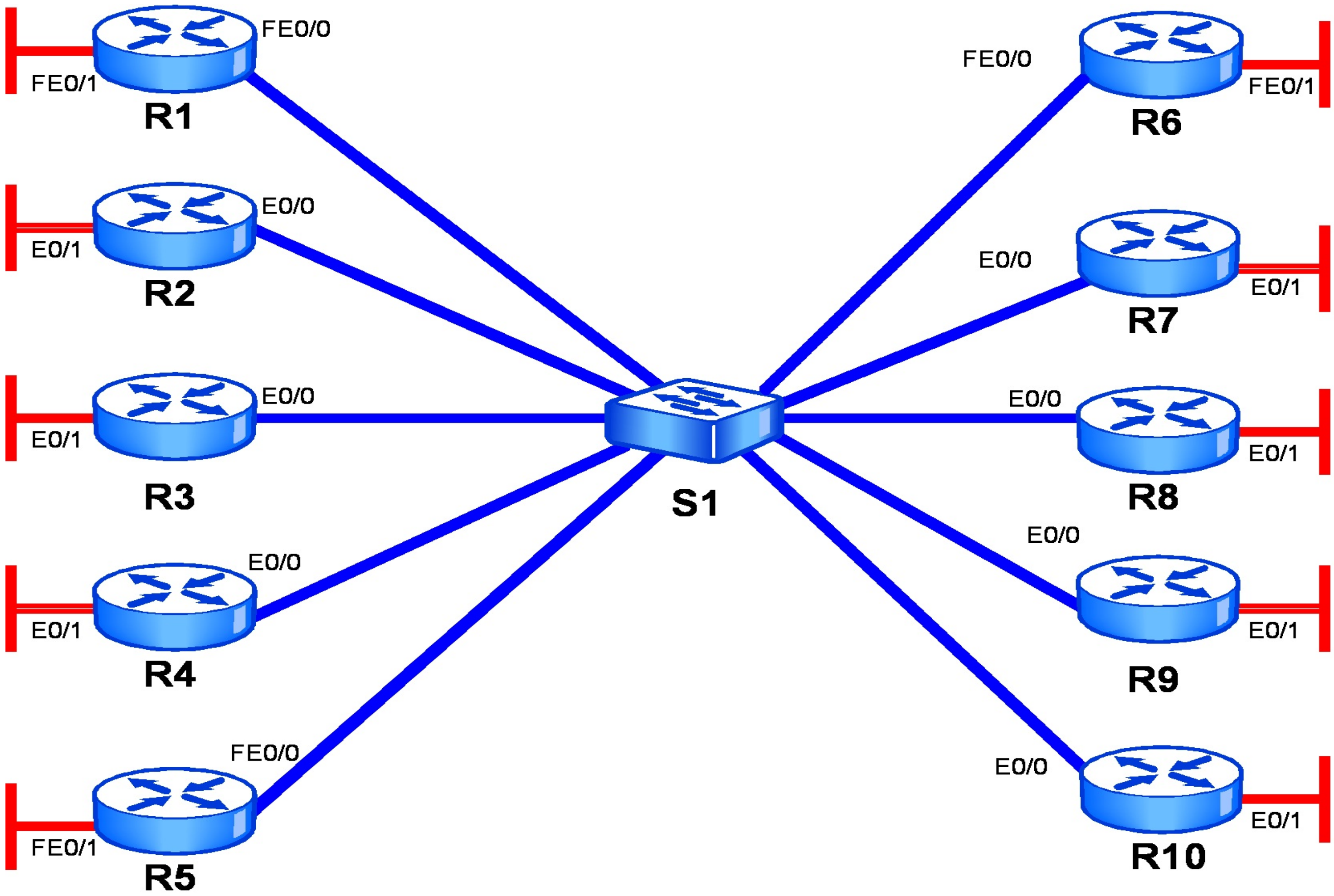
IXP Operating Cost

- Access link
- Link maintenance
- Utility
- Administration

IXP Cost Model

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

IXP Network Diagram



Questions?