BGP fundamentals

Lwin Lwin Aung

CCIE #53660 (R&S), PMP, CISSP

Sai Nyan Lynn Swe

CCIE #38501 (R&S, SP and DC), PMP, CISSP

Border Gateway Protocol

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

BGP

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

Path Vector Protocol

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

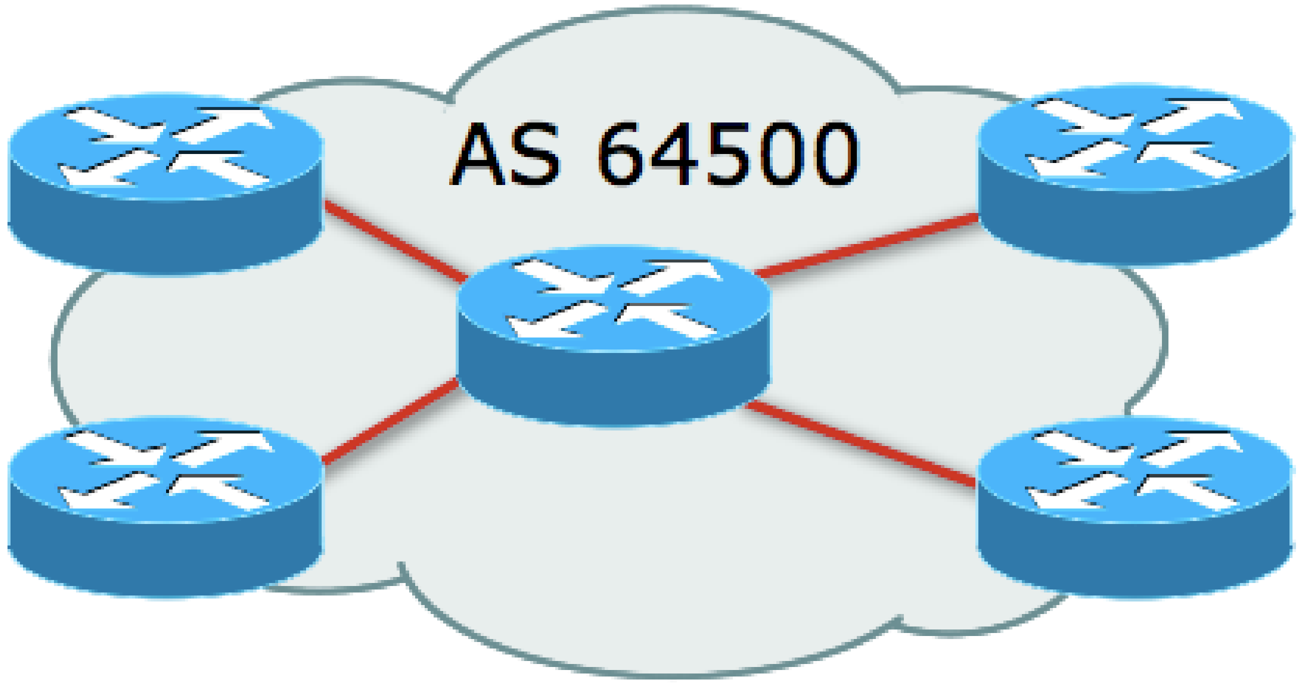
```
12.6.126.0/24   207.126.96.43    1021  0    6461 7018 6337 11268  i
```

AS Path

Definitions

- Transit – carrying traffic across a network
  - (Commercially: for a fee)
- Peering – exchanging routing information and traffic
  - (Commercially: between similar sized networks, and for no fee)
- Default – where to send traffic when there is no explicit match in the routing table

Autonomous System (AS)



- Collection of networks with same routing policy
- Single routing protocol
- Usually under single ownership, trust and administrative control
- Identified by a unique 32-bit integer (ASN)

Autonomous System Number

Range:	
0-4294967295	(32-bit range – RFC6793)
	(0-65535 was original 16-bit range)
Usage:	
0 and 65535	(IANA Reserved)
1-64495	(public Internet)
64496-64511	(documentation – RFC5398)
64512-65534	(private use only)
23456	(represent 32-bit range in 16-bit world)
65536-65551	(documentation – RFC5398)
65552-131071	(IANA Reserved)
131072-458751	(public Internet)
458752-4199999999	(IANA Reserved/Unallocated)
4200000000-4294967294	(private use only – RFC6996)
4294967295	(IANA Reserved – RFC7300)

- 32-bit range representation specified in RFC5396
  - Defines "asplain" (traditional format) as standard notation

Autonomous System Number (ASN)

- ASNs are distributed by the Regional Internet Registries
  - They are also available from upstream ISPs who are members of one of the RIRs
- The entire 16-bit ASN pool has been assigned to the RIRs
  - Around 39400 16-bit ASNs are visible on the Internet
    - (this number is dropping slightly as 32-bit ASN announcements increase)
- Each RIR has also received a block of 32-bit ASNs
  - Out of 44500 assignments, around 36500 are visible on the Internet (May 2024)
- See www.iana.org/assignments/as-numbers

Configuring BGP in Cisco IOS

- This command enables BGP in Cisco IOS:

```
router bgp 64500
```

- For ASNs > 65535, the AS number can be entered in either plain or dot notation:

```
router bgp 131076
```

  - Or

```
router bgp 2.4
```

- IOS displays ASNs in plain notation by default
  - Dot notation is optional (and NOT recommended):

```
router bgp 2.4
 bgp asnotation dot
```

Configuring BGP in JunOS

- This command sets the local autonomous system number

```
set routing-options autonomous-system 131076
```

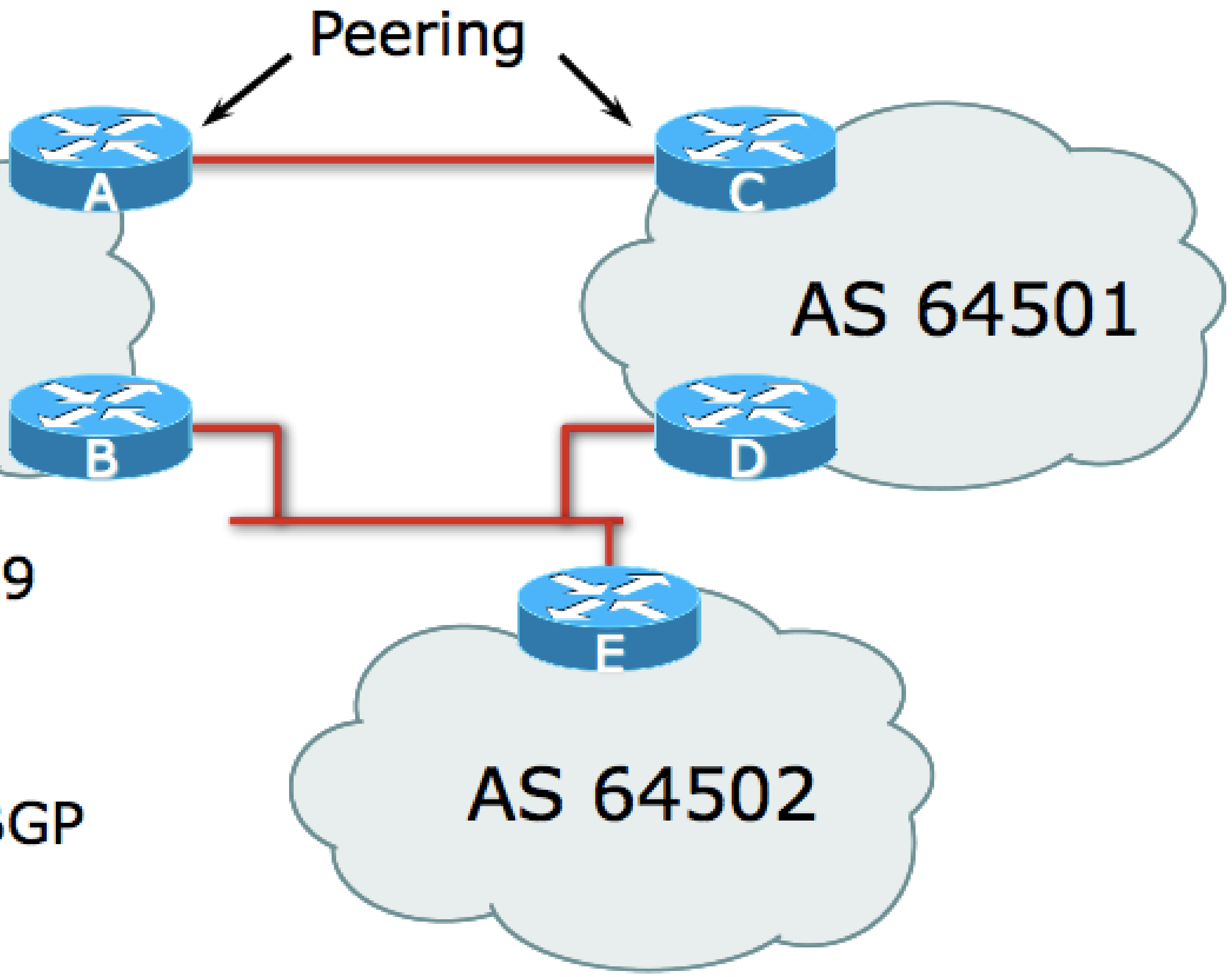
- All BGP configuration is then carried out under:

```
edit protocols bgp
```

- JunOS displays ASNs in plain notation by default
  - Dot notation is optional (and NOT recommended):

```
set routing-options autonomous-system asdot-notation 2.4
```

BGP Basics



- Runs over TCP – port 179
- Path vector protocol
- Incremental updates
- “Internal” & “External” BGP
- Supports IPv4 & IPv6

BGP General Operation

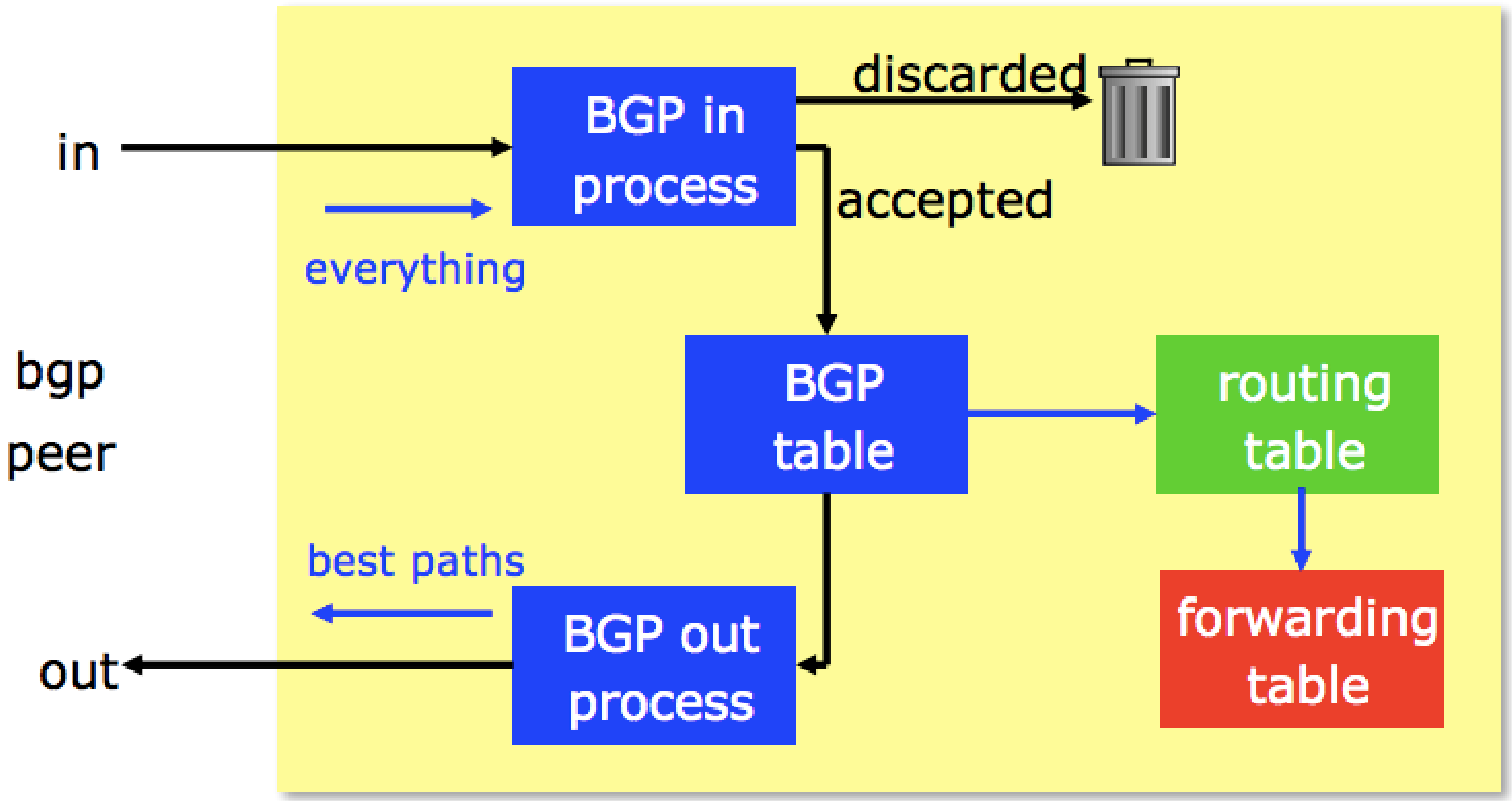
- Learns multiple paths via internal and external BGP speakers
- Picks the best path and installs it in the routing table (RIB)
- Best path is sent to external BGP neighbours
- Policies are applied by influencing the best path selection

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
  - Announces "best path" information to peers
- Best path stored in Routing Table (RIB) if:
  - Prefix and prefix length are unique (after best path selection)

    and
  - Lowest "protocol distance"
- Best paths in the RIB are installed in forwarding table (FIB)

Constructing the Forwarding Table

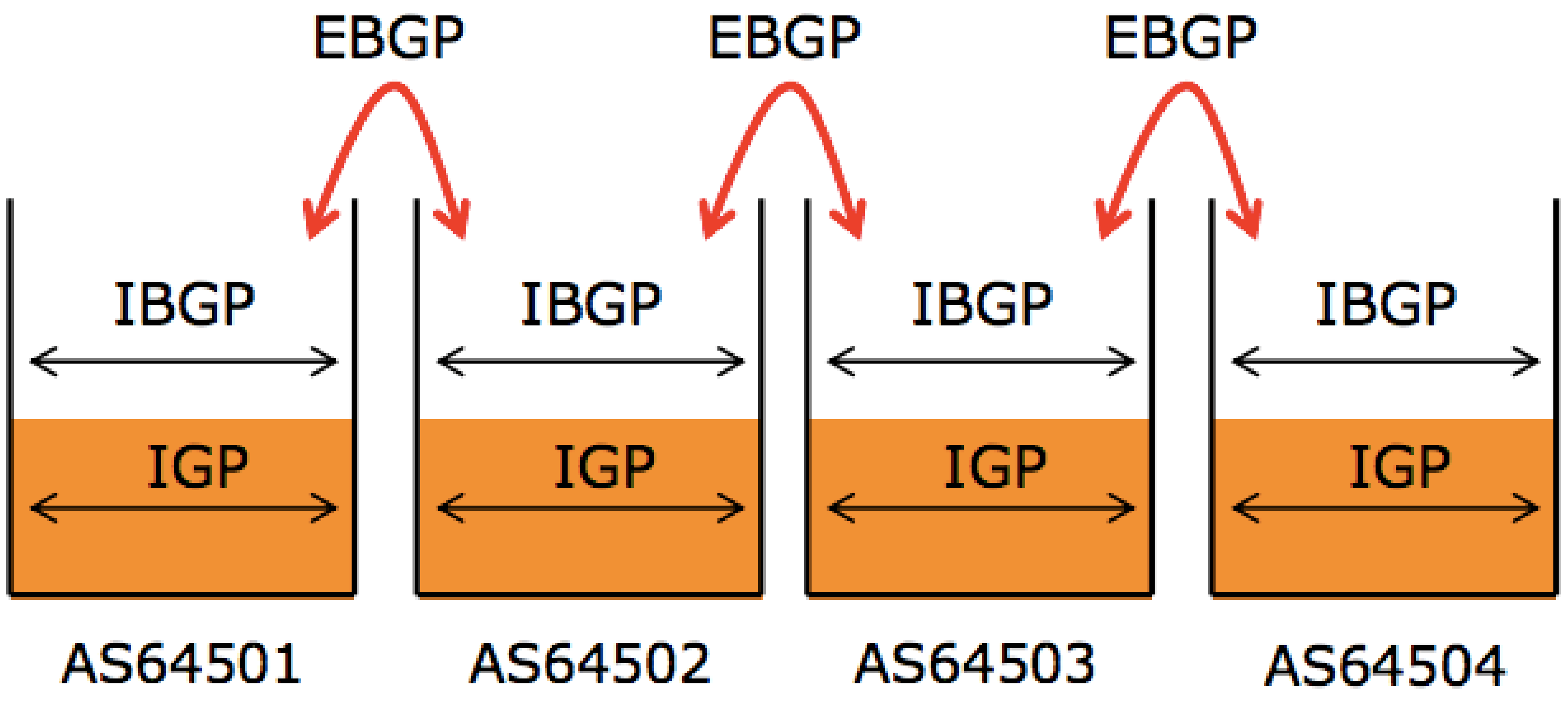


EBGP & IBGP

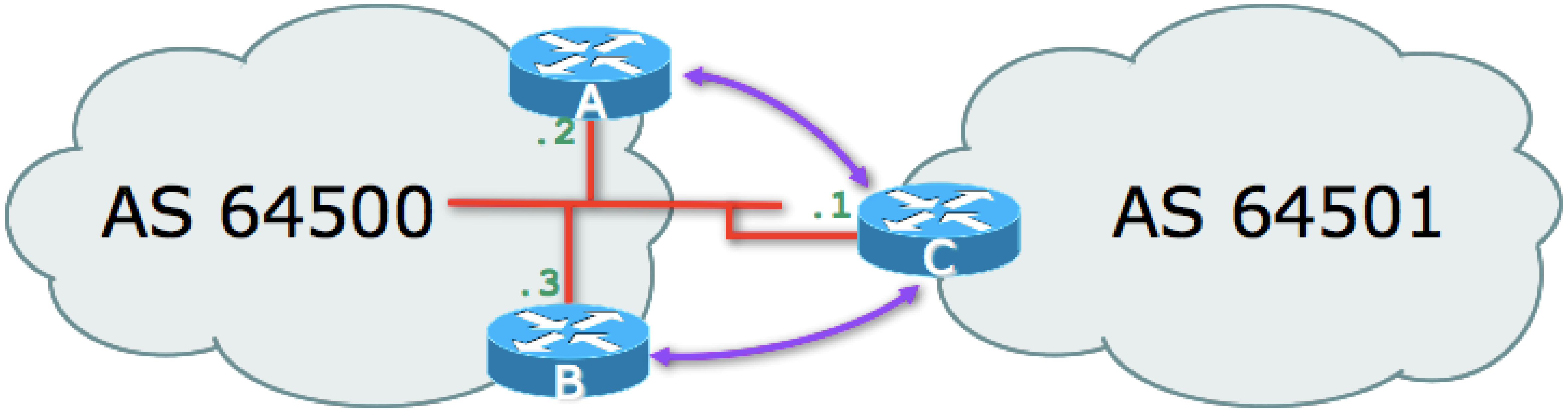
- BGP is used
  - Internally (IBGP)
  - Externally (EBGP)
- IBGP used to carry
  - Some/all Internet prefixes across network operator backbone
  - ISP’s customer prefixes
- EBGP used to
  - Exchange prefixes with other ASes
  - Implement routing policy

BGP/IGP model used in service provider networks

- Model representation



External BGP Peering (EBGP)



- Between BGP speakers in different AS
- Should be directly connected
- **Never** run an IGP between EBGP peers

Configuring External BGP

- Router A in AS64500

ip address on ethernet interface

```
interface FastEthernet 5/0
 ip address 102.102.10.2 255.255.255.240
!
router bgp 64500
 address-family ipv4
  network 100.100.8.0 mask 255.255.252.0
  neighbor 102.102.10.1 remote-as 64501
  neighbor 102.102.10.1 prefix-list RouterC-in in
  neighbor 102.102.10.1 prefix-list RouterC-out out
  neighbor 102.102.10.1 activate
!
```

Local ASN

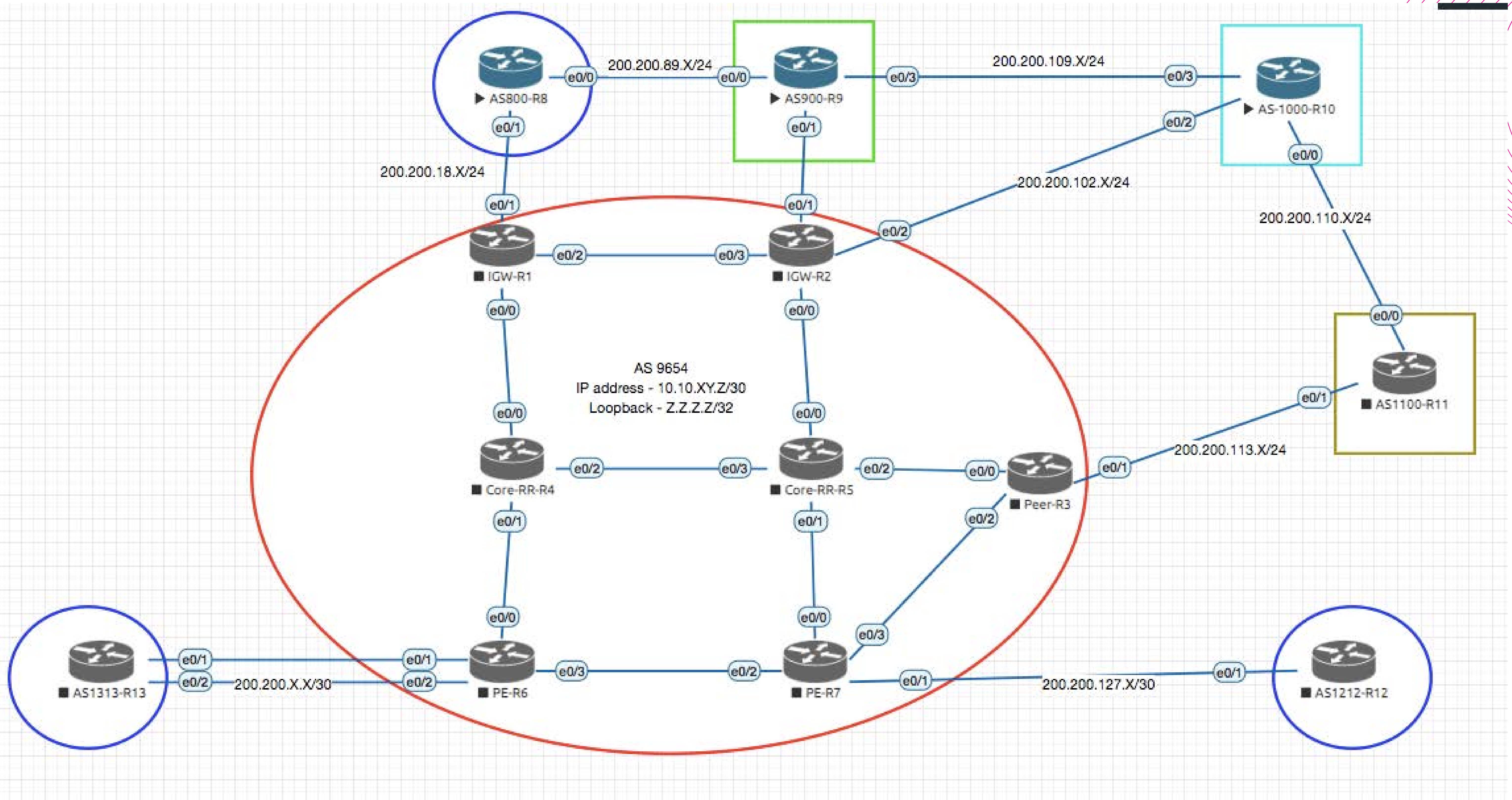
Select IPv4 or IPv6

Remote ASN

Inbound and outbound filters

ip address of Router C ethernet interface

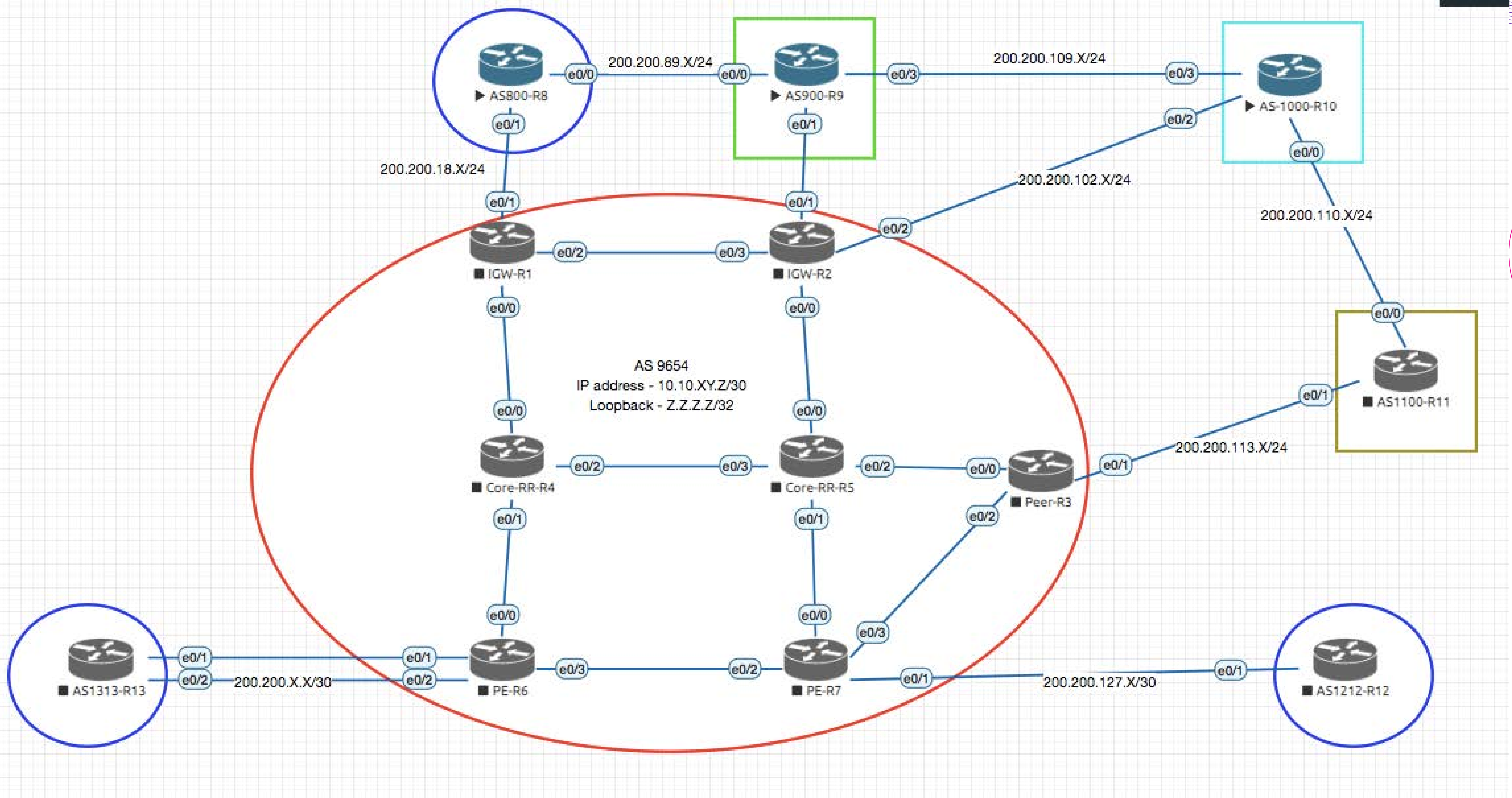
BGP Foundation Lab



Lab Access guide

- Each students can login to individual lab

Username(WebUI)	Password(WebUI)	IP Address
admin	Optimity@123	103.103.194.131 - 220



- **Check the IP address assignment for each router**

AS800-R8 — -telnet 103.103.194.75 32776 — 102×24

```
AS800-R8#
AS800-R8#
AS800-R8#sh ip int brief | e una
Interface              IP-Address      OK? Method Status                Protocol
Ethernet0/0            200.200.89.8    YES manual up                    up
Ethernet0/1            200.200.18.8    YES manual up                    up
```

- **Configure EBGP config at Relevant Routers**

```
router bgp 800
 bgp router-id 8.8.8.8
 bgp log-neighbor-changes
 neighbor 200.200.18.1 remote-as 9654
 neighbor 200.200.18.1 description MMIX
 neighbor 200.200.89.9 remote-as 900
 neighbor 200.200.89.9 description AS900



Wr mem
```

Check EBGP status

sh ip int brief | e una

Sh run | s router bgp
Sh ip bgp summary
Sh ip bgp

```
AS800-R8#sh ip bgp summary
BGP router identifier 8.8.8.8, local AS number 800
BGP table version is 1, main routing table version 1

Neighbor        V           AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
200.200.18.1    4         9654       0       0        1    0    0 never    Idle
200.200.89.9    4          900       0       0        1    0    0 never    Idle
```

```
AS800-R8#sh ip bgp
AS800-R8#sh ip route bgp
Codes: L - local, C - connected, S - static, 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
       i - IS-IS, su - IS-IS summary, 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, H - NHRP, l - LISP
       a - application route
       + - replicated route, % - next hop override

Gateway of last resort is not set
```

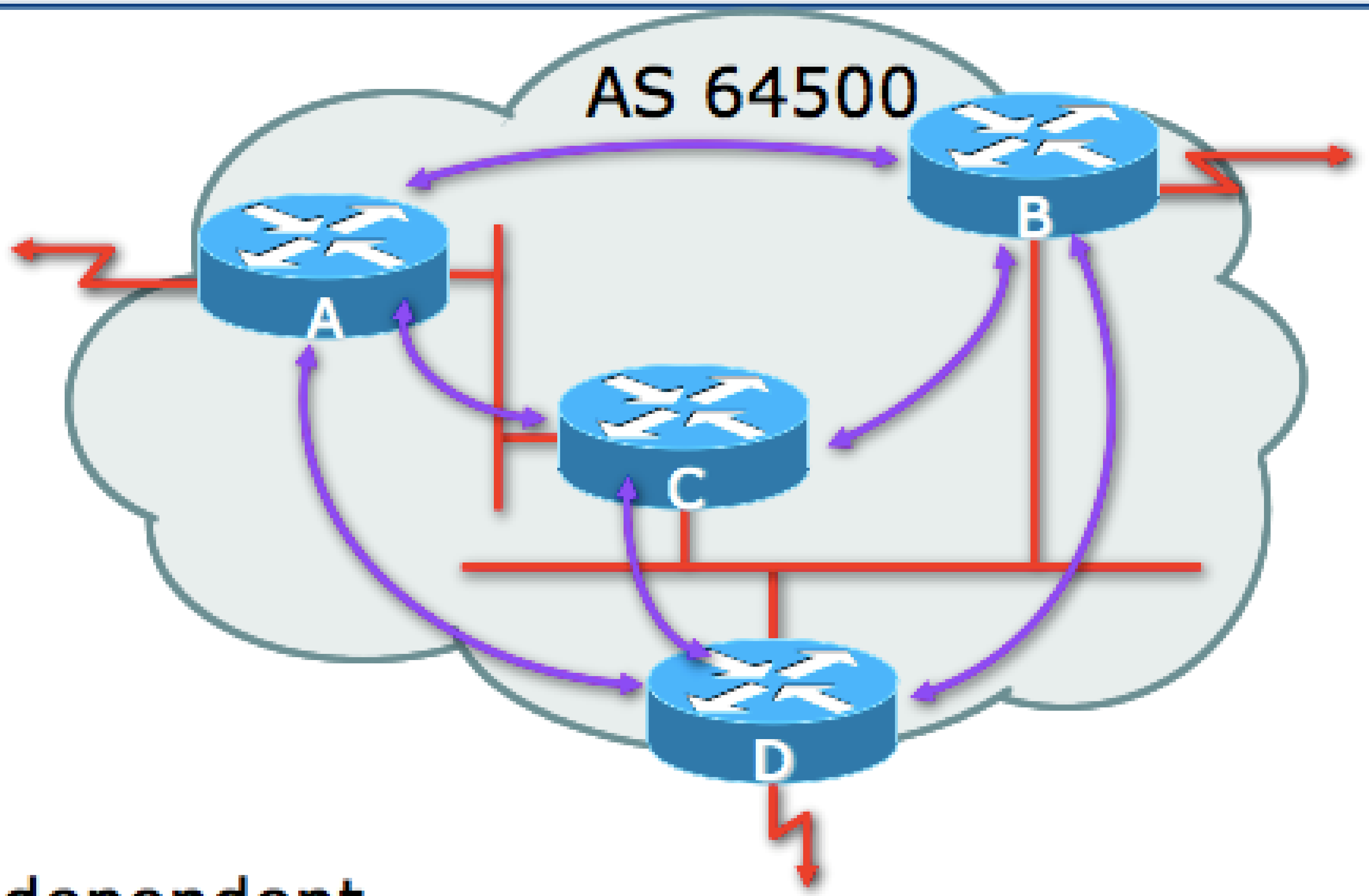
Internal BGP (IBGP)

- BGP peer within the same AS
- Not required to be directly connected
  - IGP takes care of inter-BGP speaker connectivity
- IBGP speakers must be fully meshed:
  - They originate connected networks
  - They pass on prefixes learned from outside the AS
  - They do not pass on prefixes learned from other IBGP speakers

Internal BGP (IBGP)

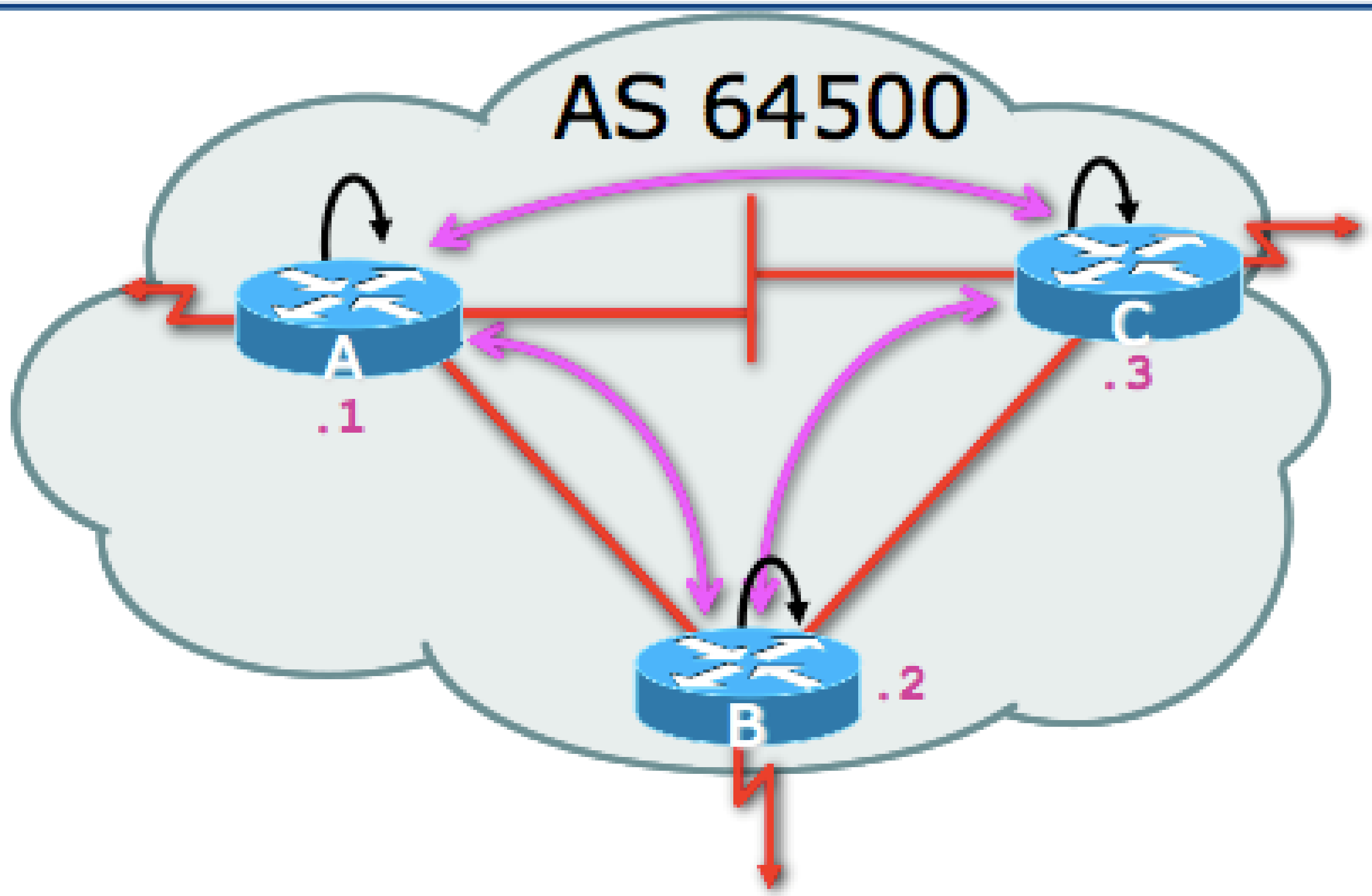
- BGP peer within the same AS
- Not required to be directly connected
  - IGP takes care of inter-BGP speaker connectivity
- IBGP speakers must be fully meshed:
  - They originate connected networks
  - They pass on prefixes learned from outside the AS
  - They do not pass on prefixes learned from other IBGP speakers

Internal BGP Peering (IBGP)



- Topology independent
- Each IBGP speaker must peer with every other IBGP speaker in the AS as per ⟷

Peering between Loopback Interfaces



- Peer with loop-back interface
  - Loop-back interface does not go down – ever!
- Do not want IBGP session to depend on state of a single interface or the physical topology

Configuring Internal BGP

- Router A in AS64500

ip address on
loopback interface

```
interface loopback 0
 ip address 105.3.7.1 255.255.255.255
!
router bgp 64500
 address-family ipv4
  network 100.100.1.0 mask 255.255.255.0
  neighbor 105.3.7.2 remote-as 64500
  neighbor 105.3.7.2 update-source loopback0
  neighbor 105.3.7.2 activate
  neighbor 105.3.7.3 remote-as 64500
  neighbor 105.3.7.3 update-source loopback0
  neighbor 105.3.7.3 activate
!
```

Local ASN

Local ASN

ip address of Router B
loopback interface

Configuring Internal BGP

- Router B in AS64500

ip address on loopback interface

```
interface loopback 0
 ip address 105.3.7.2 255.255.255.255
!
router bgp 64500
 address-family ipv4
  network 100.100.1.0 mask 255.255.255.0
  neighbor 105.3.7.1 remote-as 64500
  neighbor 105.3.7.1 update-source loopback0
  neighbor 105.3.7.1 activate
  neighbor 105.3.7.3 remote-as 64500
  neighbor 105.3.7.3 update-source loopback0
  neighbor 105.3.7.3 activate
!
```

Local ASN

Local ASN

ip address of Router A loopback interface

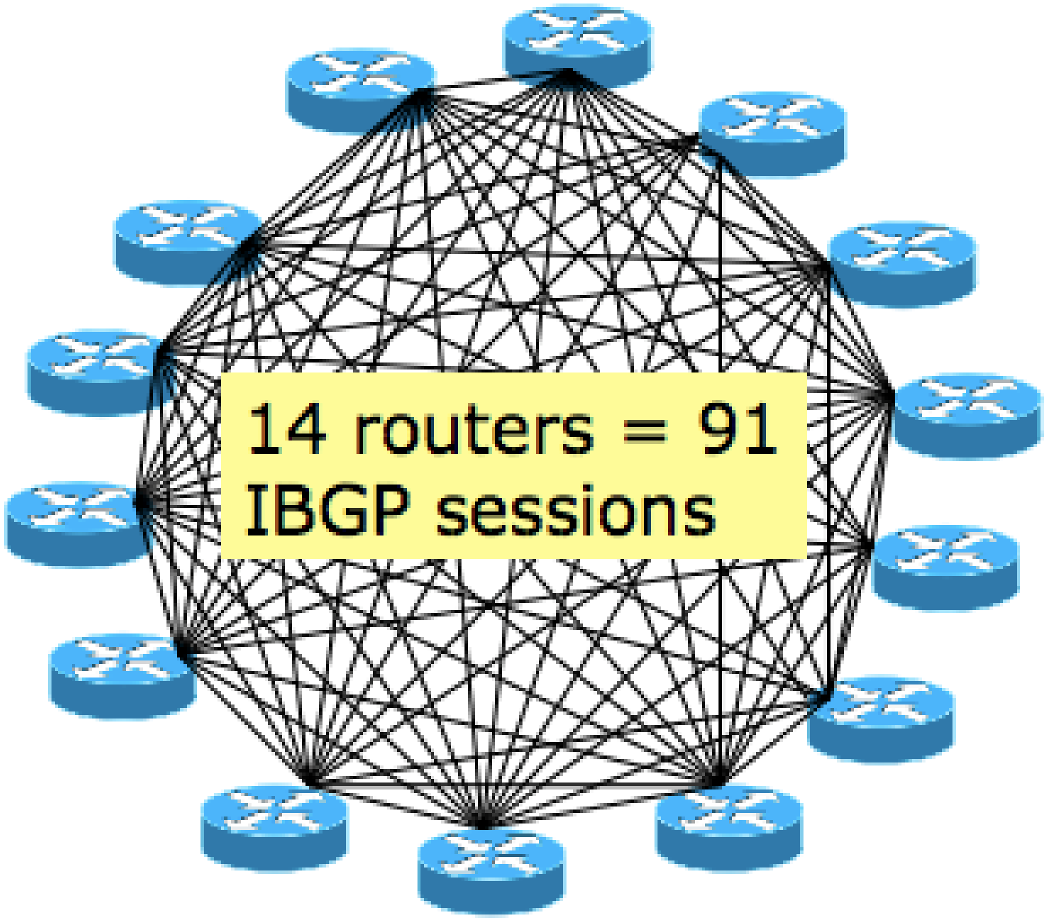
Route Reflectors

Scaling the IBGP mesh

Scaling the IBGP mesh

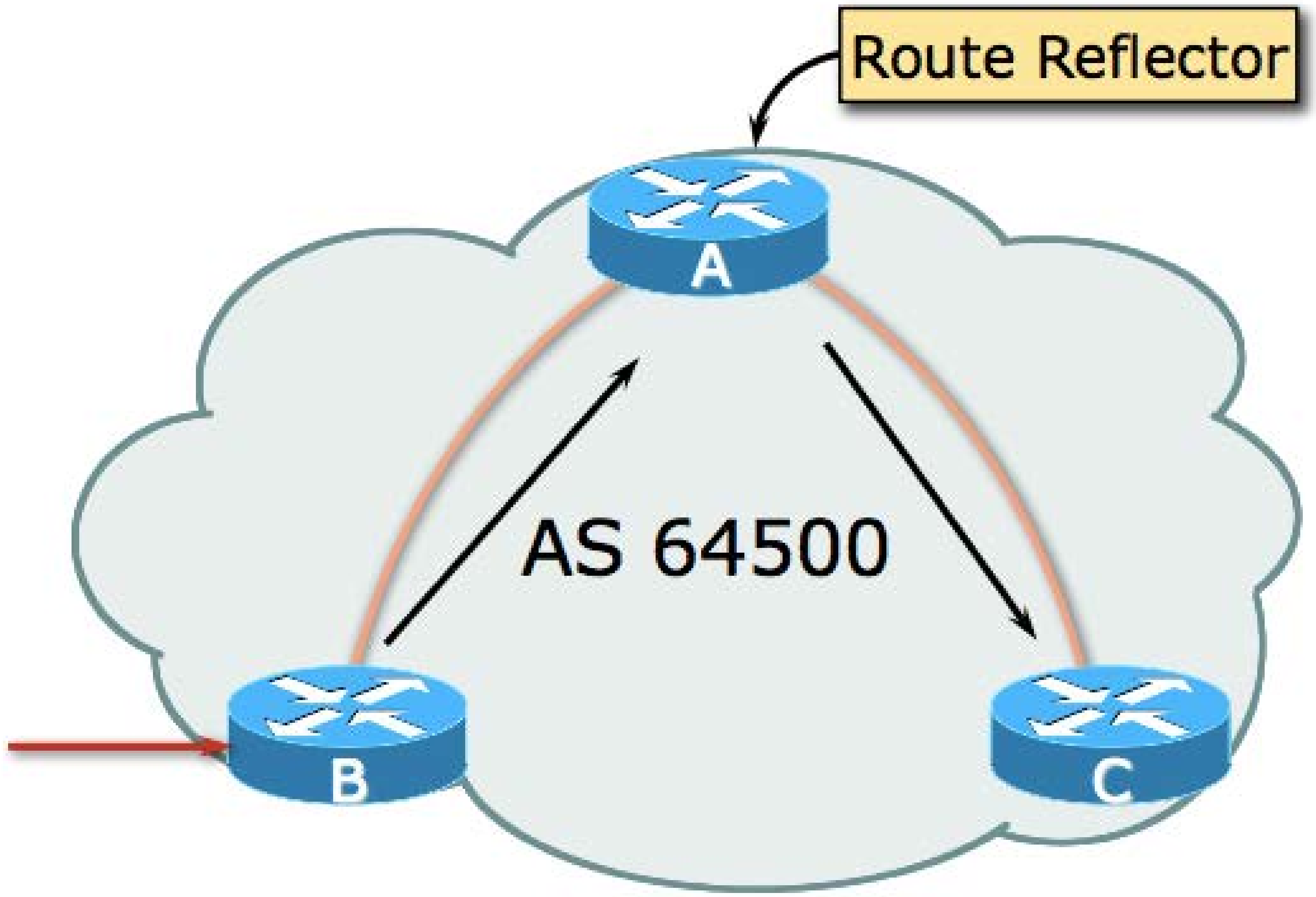
- Avoid ½n(n-1) IBGP mesh

n=1000 ⇒ nearly
half a million
IBGP sessions!



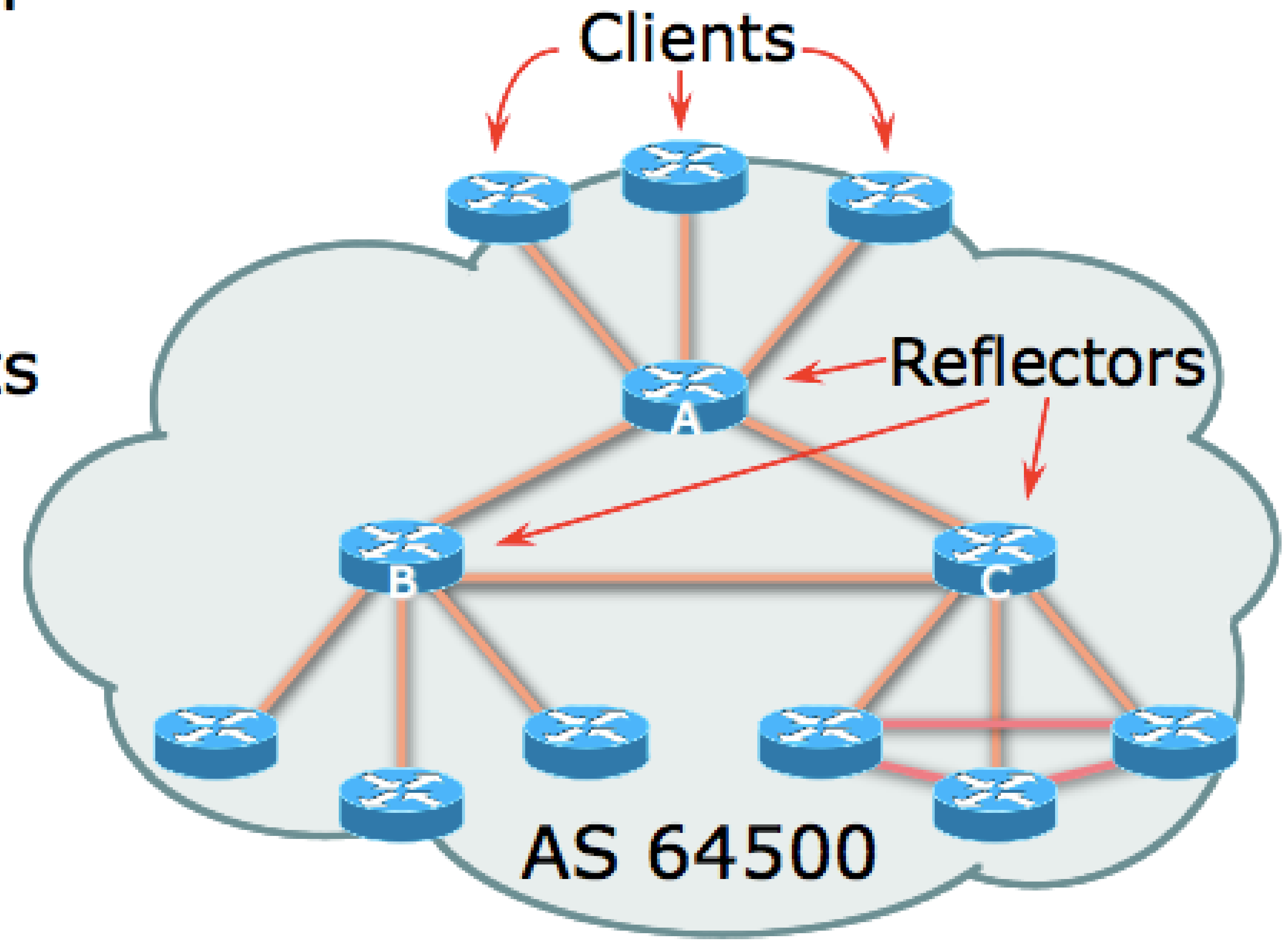
- Two solutions
  - Route reflector: simpler to deploy and run
  - BGP Confederation: more complex, has corner case advantages

Route Reflector: Principle



Route Reflector: Rules

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



Route Reflector: Benefits

- Solves IBGP mesh problem
- Packet forwarding is not affected
- Normal BGP speakers co-exist
- Multiple reflectors for redundancy
- Easy migration
- Multiple levels of route reflectors

Route Reflector: Cisco IOS Configuration

- Router D configuration:

```
router bgp 64500
 address-family ipv4
 ...
  neighbor 100.64.3.4 remote-as 64500
  neighbor 100.64.3.4 route-reflector-client
  neighbor 100.64.3.5 remote-as 64500
  neighbor 100.64.3.5 route-reflector-client
  neighbor 100.64.3.6 remote-as 64500
  neighbor 100.64.3.6 route-reflector-client
 ...
```

Inserting prefixes into BGP

- Two ways to insert prefixes into BGP
  - **`redistribute static`**
  - **`network`** command

Inserting prefixes into BGP – redistribute static

- Configuration Example:

```
router bgp 64500
 address-family ipv4
  redistribute static
ip route 100.64.32.0 255.255.254.0 serial0
```

- Static route must exist before redistribute command will work
- Forces origin to be “incomplete”
- Care required!

Inserting prefixes into BGP – network command

- Configuration Example

```
router bgp 64500
 address-family ipv4
  network 100.64.32.0 mask 255.255.254.0
ip route 100.64.32.0 255.255.254.0 serial0
```

- A matching route must exist in the routing table before the network is announced
- Forces origin to be “IGP”

Summary
BGP neighbour status (Cisco IOS IPv4)

```
Router6>show ip bgp summary
BGP router identifier 10.0.15.246, local AS number 10
BGP table version is 16, main routing table version 16
7 network entries using 819 bytes of memory
14 path entries using 728 bytes of memory
2/1 BGP path/bestpath attribute entries using 248 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 1795 total bytes of memory
BGP activity 7/0 prefixes, 14/0 paths, scan interval 60 secs

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
10.0.15.241     4    10       9       8       16    0    0 00:04:47        2
10.0.15.242     4    10       6       5       16    0    0 00:01:43        2
10.0.15.243     4    10       9       8       16    0    0 00:04:49        2
...
```

BGP Version

Updates sent and received

Updates waiting

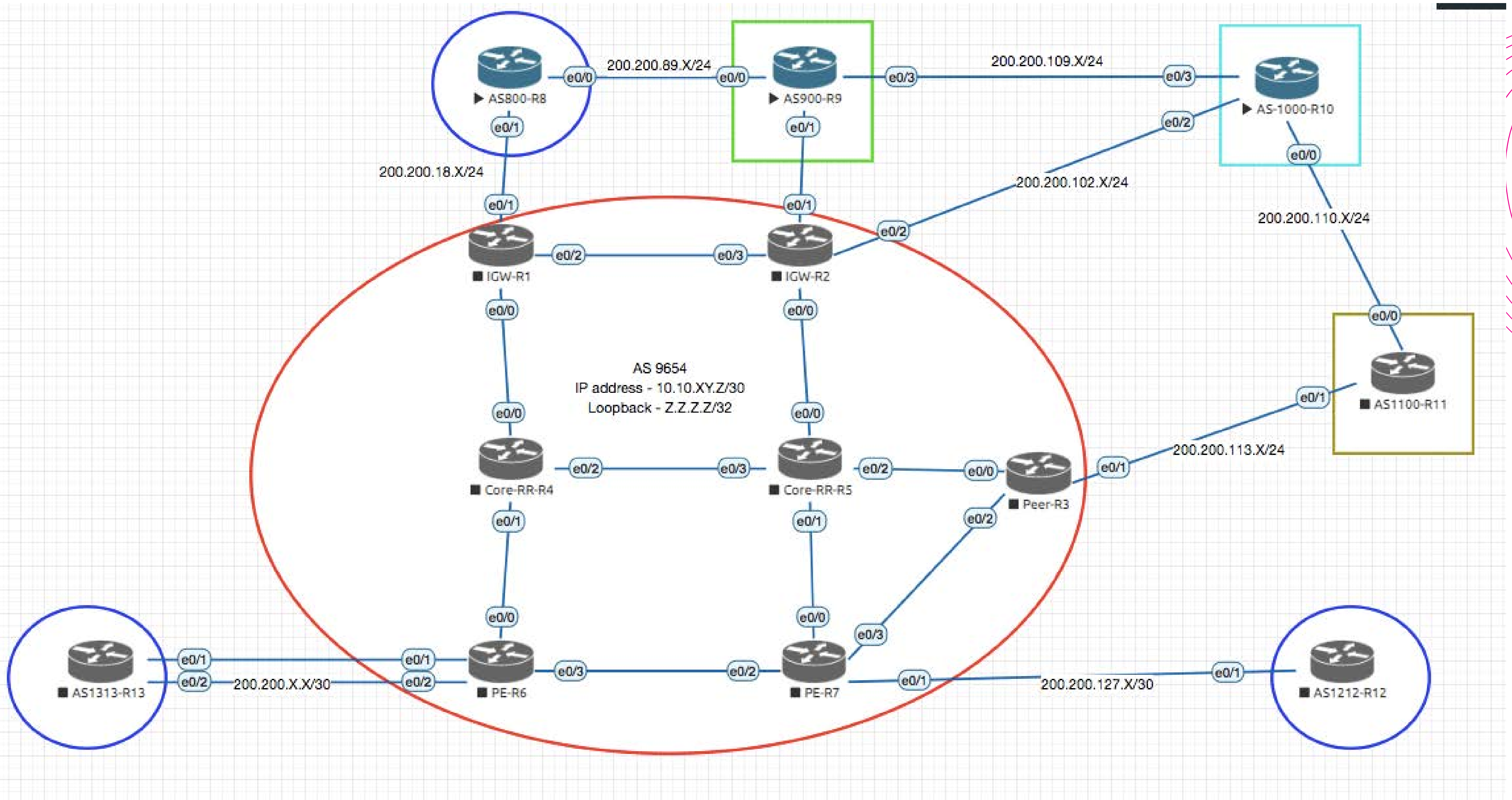
Summary
BGP Table (Cisco IOS IPv4)

```
Router6>sh ip bgp
BGP table version is 18, local router ID is 10.0.15.246
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
              x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

     Network          Next Hop            Metric LocPrf Weight Path
 *>i 10.0.0.0/26      10.0.15.241              0    100      0 i
 *>i 10.0.0.64/26     10.0.15.242              0    100      0 i
 *>i 10.0.0.128/26    10.0.15.243              0    100      0 i
 *>i 10.0.0.192/26    10.0.15.244              0    100      0 i
 *>i 10.0.1.0/26      10.0.15.245              0    100      0 i
 *>  10.0.1.64/26     0.0.0.0                  0         32768 i
 *>i 10.0.1.128/26    10.0.15.247              0    100      0 i
 *>i 10.0.1.192/26    10.0.15.248              0    100      0 i
 *>i 10.0.2.0/26      10.0.15.249              0    100      0 i
 *>i 10.0.2.64/26     10.0.15.250              0    100      0 i
 *>i 10.0.2.128/26    10.0.15.251              0    100      0 i
 *>i 10.0.2.192/26    10.0.15.252              0    100      0 i
 *>i 10.0.3.0/26      10.0.15.253              0    100      0 i
 *>i 10.0.3.64/26     10.0.15.254              0    100      0 i
```

Summary

- BGP – path vector protocol
- Multi-protocol (IPv4 & IPv6)
- IBGP versus EBGP
- Stable IBGP – peer with loopbacks
- Announcing prefixes & aggregates

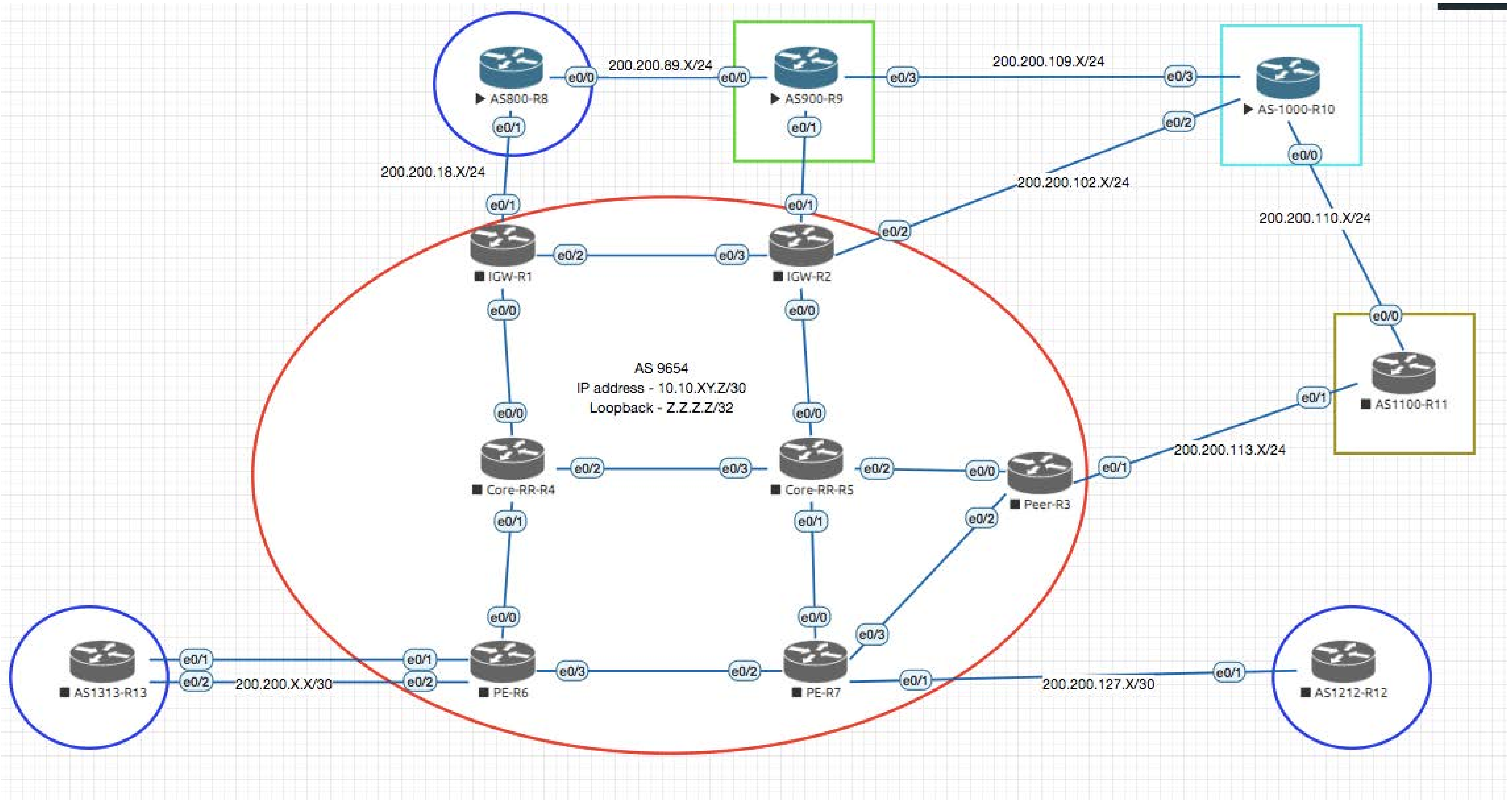
Configure IBGP config at Relevant Routers



IGW-R1

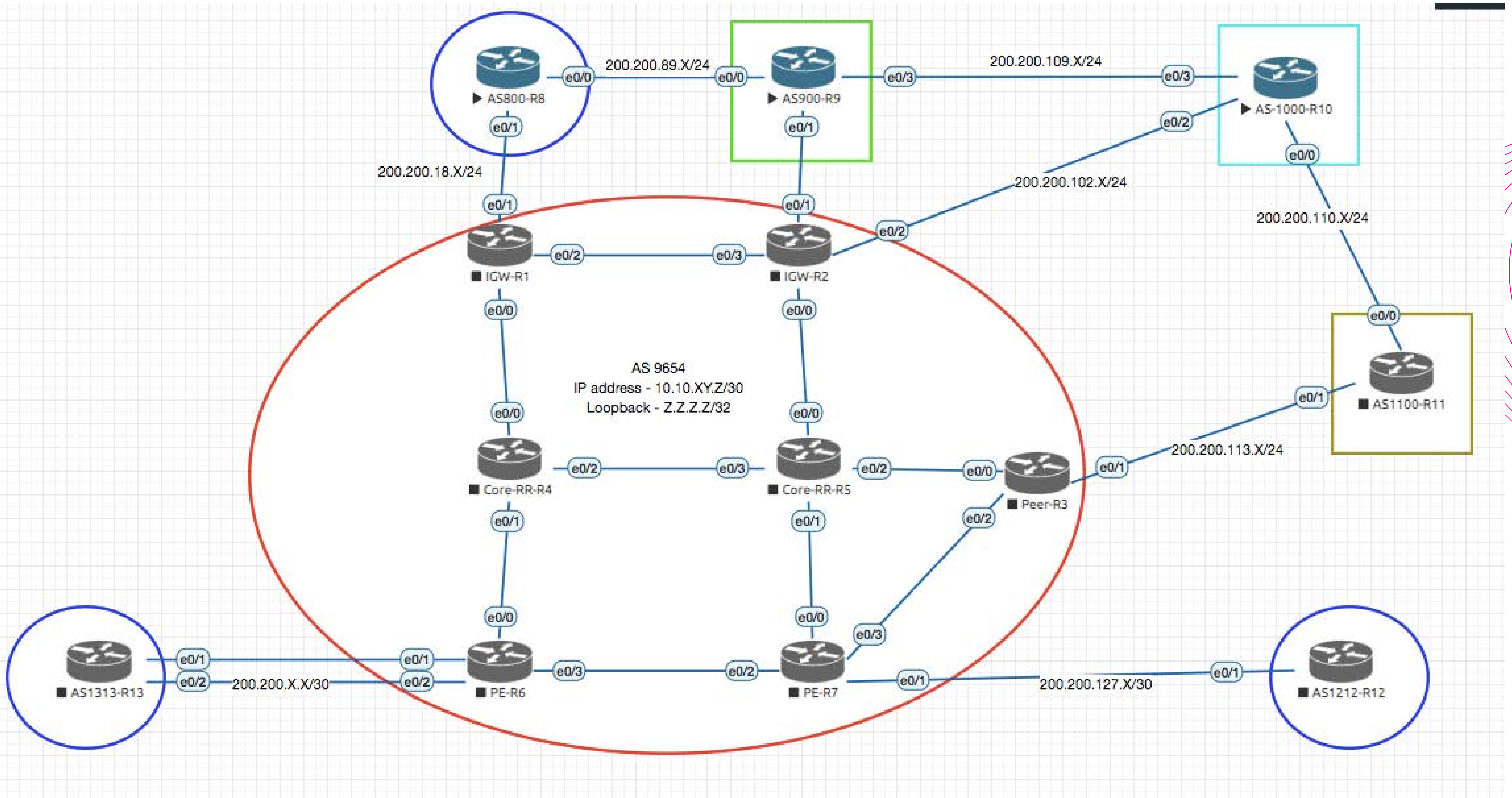
```
router bgp 9654
 bgp router-id 1.1.1.1
 bgp log-neighbor-changes
 neighbor 4.4.4.4 remote-as 9654
 neighbor 4.4.4.4 description RR
 neighbor 4.4.4.4 update-source Loopback0
 neighbor 4.4.4.4 next-hop-self
 neighbor 5.5.5.5 remote-as 9654
 neighbor 5.5.5.5 description RR
 neighbor 5.5.5.5 update-source Loopback0
 neighbor 5.5.5.5 next-hop-self
 neighbor 200.200.18.8 remote-as 800
 neighbor 200.200.18.8 description AS800
```

```
Sh ip int brief | e una
Sh ip ospf neigh
Sh ip ospf int brief

sh run  | s router bgp
sh ip bgp summ
```




```
IGW-R2
router bgp 9654
 bgp router-id 2.2.2.2
 bgp log-neighbor-changes
 neighbor 4.4.4.4 remote-as 9654
 neighbor 4.4.4.4 description RR
 neighbor 4.4.4.4 update-source Loopback0
 neighbor 4.4.4.4 next-hop-self
 neighbor 5.5.5.5 remote-as 9654
 neighbor 5.5.5.5 description RR
 neighbor 5.5.5.5 update-source Loopback0
 neighbor 5.5.5.5 next-hop-self
```

```
Sh ip int brief | e una
Sh ip ospf neigh
Sh ip ospf int brief

sh run  | s router bgp
sh ip bgp summ
```




```
IGW-R2
router bgp 9654
 bgp router-id 2.2.2.2
 bgp log-neighbor-changes
 neighbor 4.4.4.4 remote-as 9654
 neighbor 4.4.4.4 description RR
 neighbor 4.4.4.4 update-source Loopback0
 neighbor 4.4.4.4 next-hop-self
 neighbor 5.5.5.5 remote-as 9654
 neighbor 5.5.5.5 description RR
 neighbor 5.5.5.5 update-source Loopback0
 neighbor 5.5.5.5 next-hop-self
```

```
Sh ip int brief | e una
Sh ip ospf neigh
Sh ip ospf int brief

sh run  | s router bgp
sh ip bgp summ
```

- Configure IBGP config at Route Reflector (R4 & R5)

```
Core-R4
router bgp 9654
 bgp router-id 4.4.4.4
 bgp cluster-id 96.54.96.54
 bgp log-neighbor-changes
 neighbor 1.1.1.1 remote-as 9654
 neighbor 1.1.1.1 description IGW-1
 neighbor 1.1.1.1 update-source Loopback0
 neighbor 1.1.1.1 route-reflector-client

 neighbor 2.2.2.2 remote-as 9654
 neighbor 2.2.2.2 description IGW-2
 neighbor 2.2.2.2 update-source Loopback0
 neighbor 2.2.2.2 route-reflector-client

 neighbor 3.3.3.3 remote-as 9654
 neighbor 3.3.3.3 description PEER-R3
 neighbor 3.3.3.3 update-source Loopback0
 neighbor 3.3.3.3 route-reflector-client

 neighbor 6.6.6.6 remote-as 9654
 neighbor 6.6.6.6 description PE-6
 neighbor 6.6.6.6 update-source Loopback0
 neighbor 6.6.6.6 route-reflector-client

 neighbor 7.7.7.7 remote-as 9654
 neighbor 7.7.7.7 description PE-7
 neighbor 7.7.7.7 update-source Loopback0
 neighbor 7.7.7.7 route-reflector-client

 neighbor 5.5.5.5 remote-as 9654
 neighbor 5.5.5.5 description RR
 neighbor 5.5.5.5 update-source Loopback0
```

```
Sh ip int brief | e una
Sh ip ospf neigh
Sh ip ospf int brief

sh run  | s router bgp
sh ip bgp summ
```

Task – Create Loopback 100 with IP address of 100.100.100/24 at Core-R4 router.
Task - Advertised the loopback interface by BGP
Verify – Check BGP status and advertised status

```
Core-R4(config-router)#do sh run int lo 100
!
interface Loopback100
 ip address 100.100.100.100 255.255.255.0
end

=========
Core-R4(config-router)#
router bgp 9654
 network 100.100.100.0 mask 255.255.255.0
```

```
Core-R4#sh ip bgp
BGP table version is 2, local router ID is 4.4.4.4
Status codes: s suppressed, d damped, h history, * valid, > best, i
              r RIB-failure, S Stale, m multipath, b backup-path, f
              x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

     Network          Next Hop            Metric LocPrf Weight Path
 *>  100.100.100.0/24 0.0.0.0                  0         32768 i
```

```
Core-R4#sh ip bgp neighbors 1.1.1.1 advertised-routes
BGP table version is 2, local router ID is 4.4.4.4
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
              x best-external, a additional-path, c RIB-compressed,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

     Network          Next Hop            Metric LocPrf Weight Path
 *>  100.100.100.0/24 0.0.0.0                  0         32768 i

Total number of prefixes 1
```

Task – Create Loopback 101 with IP address of 101.101.101/24 at AS1100 router.
Task - Advertised the loopback interface by BGP
Verify – Check BGP status and advertised status

Task – Create Loopback 88 with IP address of 88.88.88.88/16 at AS800 router.
Task - Advertised the loopback interface by BGP
Verify – Check BGP status and advertised status

OPTIMITY