SANOG41

innog7

Indian Network Operators Group

April 2024

Some Performance Tweaks for SP Network

Presented By: **Priya Ranjan Karn**

Agenda

- **Challenges Faced By Service Providers**
  - IGP convergence time
  - Increasing Labels within the SP Network
  - Suboptimal Unicast Routing
  - Optimal Performance of RR (Route-Reflector)
  - Increasing size of Global Routing Table
  - Need for IPv6 and dual stack networks
  - Load Balancing and Redundancy in Multicast Network

- **Some Strategies to adopt for Overcoming the Challenges**
  (Based on several tests and implementation within an SP Network)
  - Speeding Your IGP Convergence with LFA
  - Filters to limit labels assignment for IGP routes
  - uRPF
  - BGP Table-Map
  - BGP Route Aggregation
  - 6PE and 6VPE
  - Anycast RP

Loop Free Alternate (LFA)

Challenge

- Even after detecting a failure of the current path, the IGP(OSPF/IS-IS) still have to run the SPF to choose the next available best path.
- This results in delay for packet forwarding to restart.
- IGP's like OSPF/IS-IS doesn't have concept of feasible successors or backup path.

Solution

- LFA helps IGP's to find and install usable backup loop free paths, which helps to quickly switch (within 50 ms) to a backup path when primary path fails.
- LFA helps the Packet Forwarding Engine to correct a path failure before it receives recomputed path from Routing Engine.
- LFA can be configured in two ways for the IGPs.

Per-Link

- IGP calculates a backup next hop for all prefixes that uses the same link (same next-hop).
- Advantage is, it consumes fewer CPU cycles and memory than Per-Prefix LFA.
- Disadvantage however, is that once the primary link fails, suddenly put a lot of burden to the backup link.

Loop Free Alternate (LFA)

```
CISCO(config)# router ospf 1
CISCO(config-ospf)# address-family ipv4
CISCO(config-ospf)# area 0
CISCO(config-ospf-ar)# fast-reroute per-link
```

Per-Prefix

- IGP calculates an LFA path for every individual prefixes.
- Disadvantage is, It requires more CPU cycles and memory.
- Advantage is, offers better load balancing as traffic is spread across to different backup paths.

```
CISCO(config)#router ospf 1
CISCO(config-ospf)#address-family ipv4
CISCO(config-ospf)#area 0
CISCO(config-ospf-ar)#fast-reroute per-prefix
```

Loop Free Alternative (LFA)

Before Enabling LFA

```
RP/0/RP0/CPU0:CISCO#show route 172.20.0.1
Sun Mar 10 10:18:36.965 NPT

Routing entry for 172.20.0.1/32
  Known via "ospf 1", distance 110, metric 4, type intra area
  Installed Mar 10 10:18:28.587 for 00:00:08
  Routing Descriptor Blocks
    172.30.30.13, from 172.20.0.1, via TenGigE0/0/0/0.3
      Route metric is 4
  No advertising protos.
RP/0/RP0/CPU0:CISCO#
```

After Enabling LFA

```
RP/0/RP0/CPU0:CISCO#show route 172.20.0.1
Sun Mar 10 10:19:21.202 NPT

Routing entry for 172.20.0.1/32
  Known via "ospf 1", distance 110, metric 4, type intra area
  Installed Mar 10 10:19:18.060 for 00:00:03
  Routing Descriptor Blocks
    172.30.30.13, from 172.20.0.1, via TenGigE0/0/0/0.3, Protected
      Route metric is 4
    172.30.30.26, from 172.20.0.1, via TenGigE0/0/0/1.5, Backup (Local-LFA)
      Route metric is 0
  No advertising protos.
RP/0/RP0/CPU0:CISCO#
```

Label Allocation

Challenge

- Service Provider's might have some edge devices which are not capable of handling higher number of labels, which can cause scaling issue.
- As we know, by default LDP allocates labels for all IGP learned routes, resulting unwanted label assignments in the SP core.

Solution

- Forwarding in MPLS requires labels only for PE's loopback addresses.
- Label allocation filtering can help SP's for temporary scaling their MPLS network.
- There are two options to accomplish this:

Allocate global **Prefix-list**- Allocates label for only those routes that match the prefix-list
Allocate global **Host-route**- Allocates labels for only for IGP learned route that are **/32**.

```
CISCO(config)#mpls ldp
CISCO(config-ldp)#address-family ipv4 label local allocate for host-routes

                    OR

CISCO(config-ldp)#address-family ipv4 label local allocate for HOST   -----→        where, HOST is a prefix-list  or access-list
```

Label Allocation

Default Label Allocation

```
R1#show mpls forwarding-table
Local      Outgoing   Prefix           Bytes Label   Outgoing    Next Hop
Label      Label      or Tunnel Id     Switched      interface
100        Pop Label  4.4.4.4/32       3700          Gi2/0       10.10.14.2
101        Pop Label  2.2.2.2/32       0             Gi1/0       10.10.12.2
102        Pop Label  10.10.45.0/30    0             Gi2/0       10.10.14.2
           Pop Label  10.10.45.0/30    0             Gi4/0       10.10.15.2
103        Pop Label  10.10.56.0/30    0             Gi4/0       10.10.15.2
104        Pop Label  10.10.25.0/30    0             Gi1/0       10.10.12.2
           Pop Label  10.10.25.0/30    0             Gi4/0       10.10.15.2
105        Pop Label  10.10.23.0/30    0             Gi1/0       10.10.12.2
107        202        3.3.3.3/32       0             Gi1/0       10.10.12.2
108        504        6.6.6.6/32       0             Gi4/0       10.10.15.2
109        209        10.10.36.0/30    0             Gi1/0       10.10.12.2
           508        10.10.36.0/30    0             Gi4/0       10.10.15.2
110        Pop Label  5.5.5.5/32       0             Gi4/0       10.10.15.2
112        503        4.4.4.4 3 [93]   0             Gi4/0       10.10.15.2
R1#
```

After Applying Filter

```
R1#show mpls forwarding-table
Local      Outgoing   Prefix           Bytes Label   Outgoing    Next Hop
Label      Label      or Tunnel Id     Switched      interface
100        Pop Label  4.4.4.4/32       4643          Gi2/0       10.10.14.2
101        Pop Label  2.2.2.2/32       0             Gi1/0       10.10.12.2
107        202        3.3.3.3/32       0             Gi1/0       10.10.12.2
108        515        6.6.6.6/32       0             Gi4/0       10.10.15.2
110        Pop Label  5.5.5.5/32       0             Gi4/0       10.10.15.2
112        503        4.4.4.4 3 [93]   0             Gi4/0       10.10.15.2
R1#
```

Unicast Reverse Path Forwarding (uRPF)

Challenge

- Generally when the router receives a unicast IP packet, it only cares about one thing, the destination address**.**
- It is possible for the attackers to exploit this behavior and spoof the source IP address to send packets that would have otherwise been dropped by the firewall or an access-list.
- Results in Suboptimal Routing and computing overhead for the transit routers most of the times.

Solution

- uRPF feature prevents above mentioned spoofing attacks.
- It helps Routers to drop packets whose source address is unknow to them at the edge of the network.

Unicast Reverse Path Forwarding (uRPF)

- uRPF can be configured to an interface of the router in one of below modes.
- **Strict mode**: Strict mode performs two checks for all incoming packets on an interface before forwarding them
  - Do I have a matching entry for the source in the routing table?

    AND
  - Do I use the same interface to reach this source as which I received this packet?
  - Generally not feasible for SP network, as they have asymmetric traffic flows.

```
CISCO(config)#interface HundredGigE0/0/1/0
CISCO(config-if)#ipv4 verify unicast source reachable-via rx
```

- **Loose Mode**: Loose mode will perform only a single check when it receive an IP packet on an uRPF configured interface.
  - Do I have a matching entry for the source in the routing table?
  - Feasible for SP network as it helps in optimal routing for edge and transit devices.

```
CISCO(config)#interface HundredGigE0/0/1/0
CISCO(config-if)#ipv4 verify unicast source reachable-via any
```

BGP Table-Map

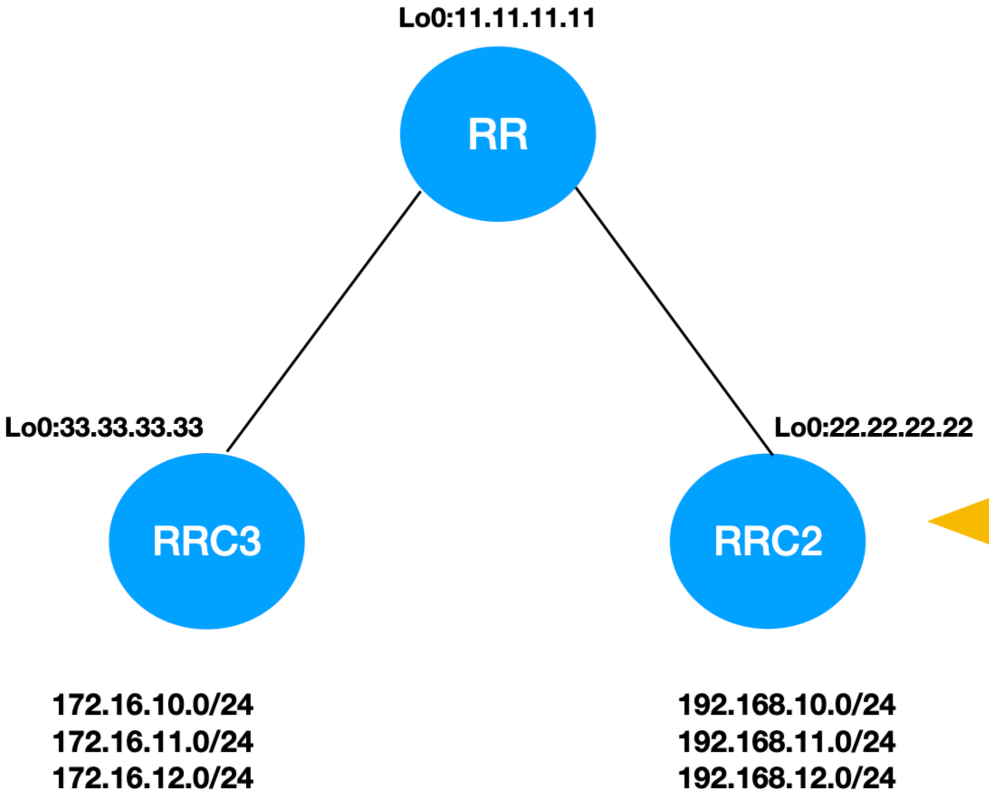
Challenge

- Unnecessary downloading or installing of certain BGP routes to the RIB or FIB on a RR(Route-Reflector),which is not on the data plane.
- Making use of routers that are fit for compute but not for traffic forwarding.

Solution

- BGP table-map is a feature that allows us to filter the BGP routes marked for installation into the RIB.
- A scenario where we could use table-map could be an RR that doesn’t need to be in the data plane but has to host the control plane).
- Optimizes the performance of RR’s.

BGP Table-Map




```
RR
-------------

ip access-list standard ODD
permit 0.0.1.0 255.255.254.255
!

route-map FILTER_ODD deny 10
match ip address ODD
!
route-map FILTER_ODD permit 20
!

router bgp 100

address-family ipv4
table-map FILTER_ODD filter

exit-address-family
```

BGP Table-Map

RR's BGP Table

```
RR#show ip bgp
BGP table version is 7, local router ID is 11.11.11.11
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 172.16.10.0/24   33.33.33.33              0    100      0 ?
 *>i 172.16.11.0/24   33.33.33.33              0    100      0 ?
 *>i 172.16.12.0/24   33.33.33.33              0    100      0 ?
 *>i 192.168.10.0     22.22.22.22              0    100      0 ?
 *>i 192.168.11.0     22.22.22.22              0    100      0 ?
 *>i 192.168.12.0     22.22.22.22              0    100      0 ?
RR#
```

RR's RIB

```
RR#show 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
       + - replicated route, % - next hop override

Gateway of last resort is not set

      172.16.0.0/24 is subnetted, 2 subnets
B        172.16.10.0 [200/0] via 33.33.33.33, 00:27:24
B        172.16.12.0 [200/0] via 33.33.33.33, 00:27:24
B     192.168.10.0/24 [200/0] via 22.22.22.22, 00:27:24
B     192.168.12.0/24 [200/0] via 22.22.22.22, 00:27:24
RR#
```

RRC1 RIB

```
R3#show 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
       + - replicated route, % - next hop override

Gateway of last resort is not set

B     192.168.10.0/24 [200/0] via 22.22.22.22, 00:39:40
B     192.168.11.0/24 [200/0] via 22.22.22.22, 00:39:40
B     192.168.12.0/24 [200/0] via 22.22.22.22, 00:39:40
R3#
```

RRC2 RIB

```
R2#show 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
       + - replicated route, % - next hop override

Gateway of last resort is not set

      172.16.0.0/24 is subnetted, 3 subnets
B        172.16.10.0 [200/0] via 33.33.33.33, 00:39:01
B        172.16.11.0 [200/0] via 33.33.33.33, 00:39:01
B        172.16.12.0 [200/0] via 33.33.33.33, 00:39:01
R2#
```

BGP Route Aggregation

Challenge

- Increasing size of Global routing table.
- Routing instability due to flapping routes.

Solution

- Route Aggregation decreases the size of the global routing table.
- Route Aggregation can help in rapid routing table lookups and prevent TCAM resource exhaustion for network devices.
- Also Route Aggregation positively impacts routing stability.

BGP ROUTE AGGREGATION

- If AS-1 router advertised de-aggregated prefixes to its e-BGP peers.
- And If any of customer link goes down, **AS-1** will advertises the withdrawal of prefix to its eBGP Peers.
- As a result, all internet routers with full BGP table view removes network from their FIB table, which in turn adds to their compute load.
- Again if the customer link comes up, the prefix is re-injected to **AS-1** and re-advertised to its eBGP peers, however connectivity to the internet might not be immediately available for the customer due to BGP propagation delays.
- Now if AS-1 advertises aggregates to its eBGP peers and if any prefix becomes unreachable, the prefix will only be withdrawn from iBGP of **AS-1.**
- The routing tables of the upstream are not impacted as there are no network updates sent from **AS-1** to its eBGP Peers.

AS-1

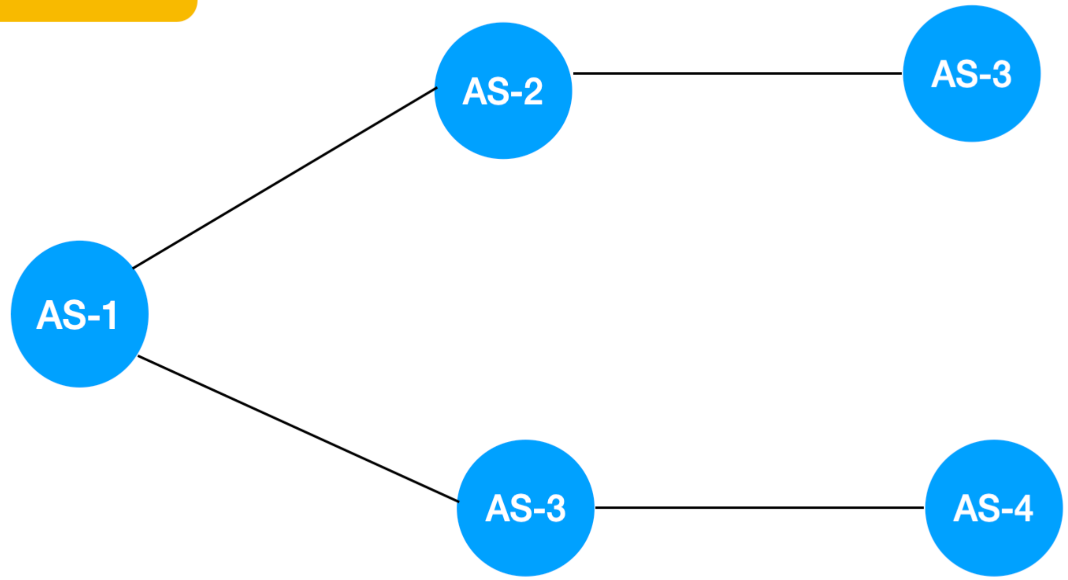
```
router bgp 24550
!
address-family ipv4

aggregate-address 116.90.224.0 255.255.248.0 summary-only
```

Specific Prefixes

116.90.224.0/24
116.90.225.0/24
116.90.226.0/24
116.90.227.0/24
116.90.228.0/24
116.90.229.0/24
116.90.230.0/24
116.90.231.0/24

Aggregated Address

116.90.224.0/21



6PE and 6VPE

Challenge

- Need to deliver IPv6 Services over and IPv4-only MPLS Core.
- One way of doing to would be to move to a dual stack solution.
- It would involve implementing an IPv6 IGP, MP-BGP and IPv6 LDP (or MPLS-TE) for you 'n' number of MPLS boxes.
- Other work around would be supporting IPv6 over the same IPv4 MPLS network with minimal changes.

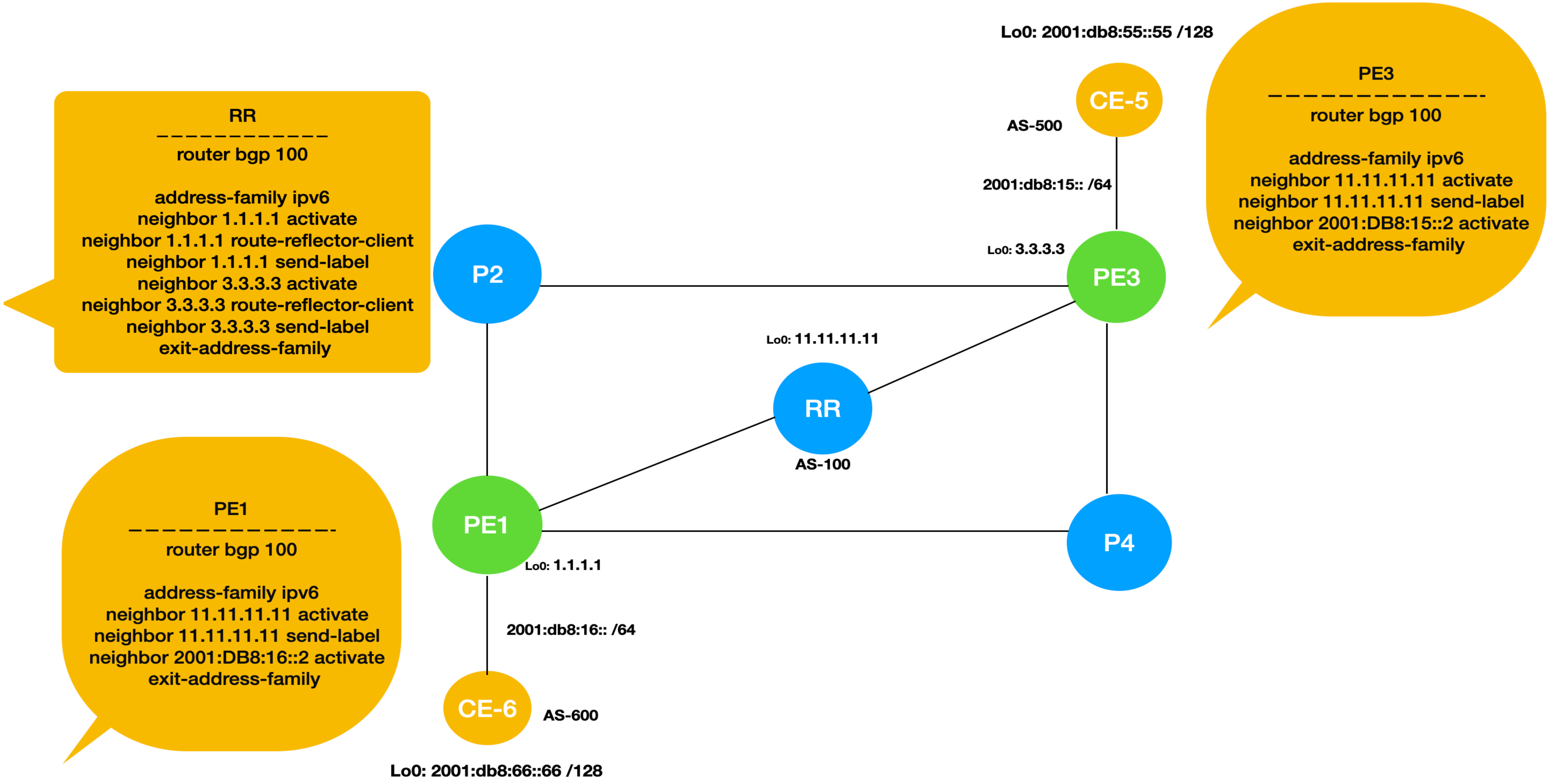
Solution

- 6PE (RFC-4798) and 6VPE (RFC-4659) allows us to run IPv6 over an IPv4-only MPLS core where only the PE's are dual stack.
- Minimal operational, capital cost and risk as no impact on existing IPv4 and MPLS services.
- MPLS Provider Routers (P) doesn't need to be IPv6 aware.
- Only Provider Edge (PE) requires upgrade - as 6PE and 6VPE router can be the existing PE router
- Both 6PE and 6VPE exploit the fact that as long as a packet somehow can be forwarded along an LSP from ingress to egress PE, P routers do not care about anything but the transport label.

6PE (IPv6 Provider Edge)

- 6PE uses the global IPv6 routing table on the PE(i.e. the peering to the CE is under **address-family ipv6 unicast**).
- The neighborship between PE's is also under **address-family ipv6 unicast.**
- However the peers are an IPv4 address with **send-label** (**send NLRI + MPLS label to the peer**) enabled.
- The LSP's between the PE is based on the IPv4 so the next hop addresses are IPv4 addresses.
- When a PE router advertises an IPv6 prefix through MP-BGP to another PE, it embeds the IPv4 address in the IPv6 next hop address.
- This is how a PE knows which IPv4 address (and thus which label) to use to get to the other PE.
- A value of ::FFFF: is prepended to the IPv4 next hop, which is an IPv4-mapped IPv6 address.

6PE (IPv6 Provider Edge)



6PE-PE's Output

PE-1

```
PE1#show bgp ipv6 unicast summary | exclude BGP
4 network entries using 672 bytes of memory
4 path entries using 416 bytes of memory

Neighbor        V           AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
11.11.11.11     4          100      11      10        5    0    0 00:05:31        2
2001:DB8:16::2  4          600      10      10        5    0    0 00:06:13        2

PE1#
```

```
PE1#show bgp ipv6 unicast
BGP table version is 5, local router ID is 1.1.1.1
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 2001:DB8:15::/64 ::FFFF:3.3.3.3           0    100      0 500 ?
 r>  2001:DB8:16::/64 2001:DB8:16::2           0             0 600 ?
 *>i 2001:DB8:55::55/128
                       ::FFFF:3.3.3.3           0    100      0 500 ?
 *>  2001:DB8:66::66/128
                       2001:DB8:16::2           0             0 600 ?
PE1#
```

PE-3

```
PE3#show bgp ipv6 unicast summary | exclude BGP
4 network entries using 672 bytes of memory
4 path entries using 416 bytes of memory

Neighbor        V           AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
11.11.11.11     4          100      13      12        5    0    0 00:07:33        2
2001:DB8:15::2  4          500      12      12        5    0    0 00:08:17        2

PE3#
```

```
PE3#show bgp ipv6 unicast
BGP table version is 5, local router ID is 3.3.3.3
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
 r>  2001:DB8:15::/64 2001:DB8:15::2           0             0 500 ?
 *>i 2001:DB8:16::/64 ::FFFF:1.1.1.1           0    100      0 600 ?
 *>  2001:DB8:55::55/128
                       2001:DB8:15::2           0             0 500 ?
 *>i 2001:DB8:66::66/128
                       ::FFFF:1.1.1.1           0    100      0 600 ?
PE3#
```

6PE-CE's Output

CE-5

```
CE-5#ping 2001:DB8:66::66 source lo0
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 2001:DB8:66::66, timeout is 2 seconds:
Packet sent with a source address of 2001:DB8:55::55
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 60/258/944 ms
CE-5#
CE-5#
CE-5#tracer
CE-5#traceroute 2001:DB8:66::66
Type escape sequence to abort.
Tracing the route to 2001:DB8:66::66

  1 2001:DB8:15::1 92 msec 32 msec 28 msec
  2 ::FFFF:10.10.34.2 [MPLS: Labels 22/24 Exp 0] 128 msec 64 msec 56 msec
  3 2001:DB8:16::1 [AS 600] [MPLS: Label 24 Exp 0] 268 msec 124 msec 112 msec
  4 2001:DB8:16::2 [AS 600] 164 msec 72 msec 132 msec
CE-5#
```

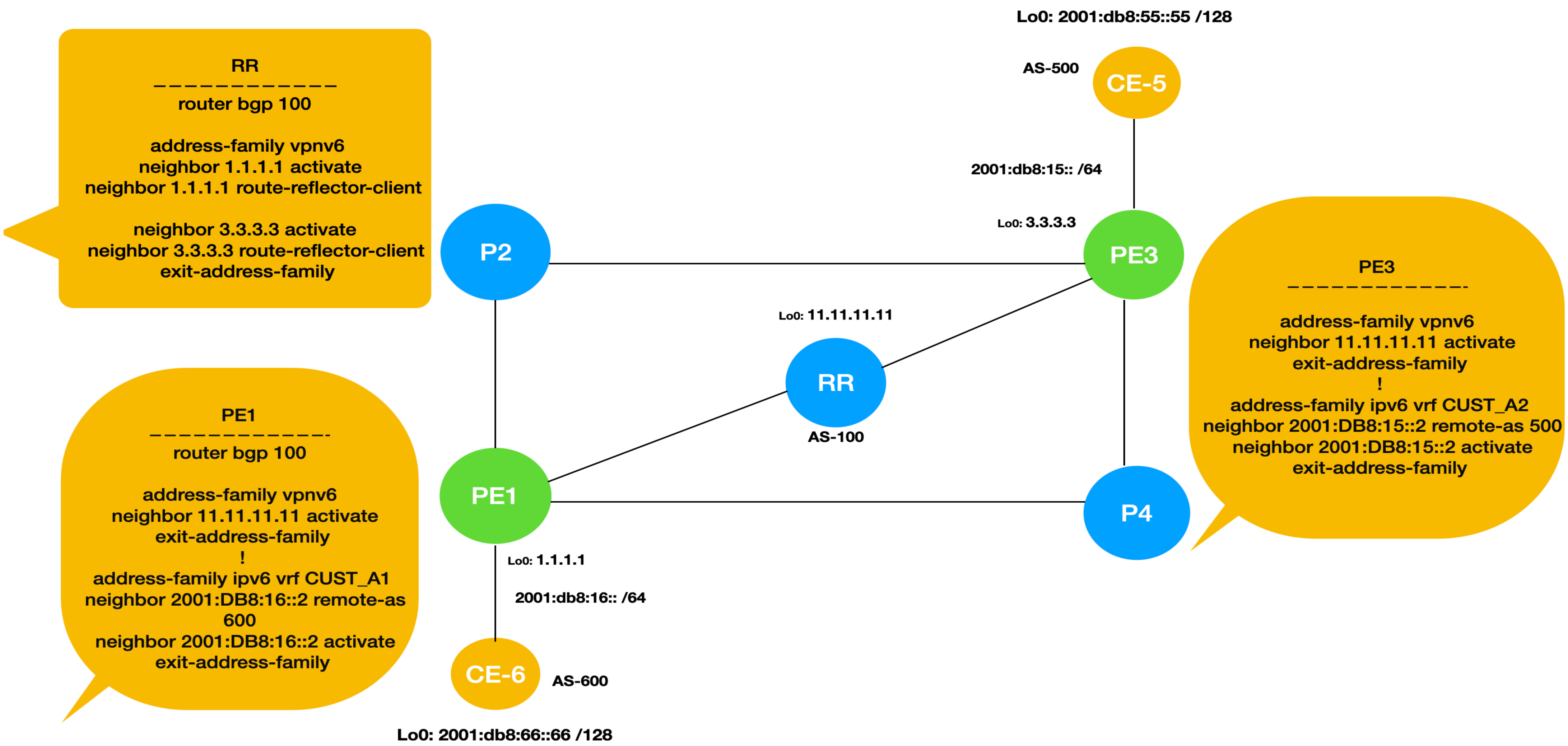
CE-6

```
CE-6#ping 2001:DB8:55::55 so lo0
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 2001:DB8:55::55, timeout is 2 seconds:
Packet sent with a source address of 2001:DB8:66::66
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 60/184/356 ms
CE-6#
CE-6#
CE-6#traceroute 2001:DB8:55::55
Type escape sequence to abort.
Tracing the route to 2001:DB8:55::55

  1 2001:DB8:16::1 168 msec 28 msec 12 msec
  2 ::FFFF:10.10.14.2 [MPLS: Labels 16/24 Exp 0] 272 msec 136 msec 72 msec
  3 2001:DB8:15::1 [AS 500] [MPLS: Label 24 Exp 0] 272 msec 56 msec 96 msec
  4 2001:DB8:15::2 [AS 500] 88 msec 16 msec 152 msec
CE-6#
```

6VPE (IPv6 VPN Provider Edge)

- 6VPE uses VRF's on the PE (i.e. the peering to the CE is under **address-family ipv6 vrf <vrf_name>**).
- The neighborship between PE's is under **address-family vpnv6 unicast** but the peers being an IPv4 address.
- The LSP's between the PE is based on the IPv4 so the next hop addresses are IPv4 addresses.
- When a PE advertises an IPv6 prefix through MP-BGP to another PE , it embeds the IPv4 address in the IPv6 next hop address.
- This is how the PE knows which IPv4 address (and thus which label) to use to get to the other PE.
- A value of ::FFFF: is prepended to the IPv4 next hop, which is an IPv4-mapped IPv6 address.

6VPE (IPv6 VPN Provider Edge)



6VPE-PE's Output

PE-1

```
PE1#show bgp vpnv6 unicast all summary | ex
PE1#show bgp vpnv6 unicast all summary | exclude BGP
6 network entries using 1080 bytes of memory
6 path entries using 648 bytes of memory

Neighbor        V           AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
11.11.11.11     4          100      35      34       11    0    0 00:26:59        2
2001:DB8:16::2  4          600      32      31       11    0    0 00:24:20        2
```

```
PE1#show bgp vpnv6 unicast all
BGP table version is 11, local router ID is 1.1.1.1
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
Route Distinguisher: 1:6 (default for vrf CUST_A1)
 *>i 2001:DB8:15::/64 ::FFFF:3.3.3.3           0    100      0 500 ?
 r>  2001:DB8:16::/64 2001:DB8:16::2           0             0 600 ?
 *>i 2001:DB8:55::55/128
                       ::FFFF:3.3.3.3           0    100      0 500 ?
 *>  2001:DB8:66::66/128
                       2001:DB8:16::2           0             0 600 ?
Route Distinguisher: 3:5
 *>i 2001:DB8:15::/64 ::FFFF:3.3.3.3           0    100      0 500 ?
 *>i 2001:DB8:55::55/128
                       ::FFFF:3.3.3.3           0    100      0 500 ?
```

PE-3

```
PE3#show bgp vpnv6 unicast all summary | exclude BGP
6 network entries using 1080 bytes of memory
6 path entries using 648 bytes of memory

Neighbor        V           AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
11.11.11.11     4          100      49      36        7    0    0 00:27:25        2
2001:DB8:15::2  4          500      39      39        7    0    0 00:31:50        2
```

```
PE3#show bgp vpnv6 unicast all
BGP table version is 7, local router ID is 3.3.3.3
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
Route Distinguisher: 1:6
 *>i 2001:DB8:16::/64 ::FFFF:1.1.1.1           0    100      0 600 ?
 *>i 2001:DB8:66::66/128
                       ::FFFF:1.1.1.1           0    100      0 600 ?
Route Distinguisher: 3:5 (default for vrf CUST_A2)
 r>  2001:DB8:15::/64 2001:DB8:15::2           0             0 500 ?
 *>i 2001:DB8:16::/64 ::FFFF:1.1.1.1           0    100      0 600 ?
 *>  2001:DB8:55::55/128
                       2001:DB8:15::2           0             0 500 ?
 *>i 2001:DB8:66::66/128
                       ::FFFF:1.1.1.1           0    100      0 600 ?
PE3#
```

6VPE-CE's Output

CE-5

```
CE-6#ping 2001:DB8:55::55 so lo0
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 2001:DB8:55::55, timeout is 2 seconds:
Packet sent with a source address of 2001:DB8:66::66
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 92/239/588 ms
CE-6#
CE-6#
CE-6#traceroute 2001:DB8:55::55
Type escape sequence to abort.
Tracing the route to 2001:DB8:55::55

  1 2001:DB8:16::1 140 msec 24 msec 108 msec
  2 ::FFFF:10.10.14.2 [MPLS: Labels 16/24 Exp 0] 272 msec 252 msec 148 msec
  3 2001:DB8:15::1 [AS 500] [MPLS: Label 24 Exp 0] 144 msec 72 msec 72 msec
  4 2001:DB8:15::2 [AS 500] 260 msec 100 msec 124 msec
CE-6#
CE-6#
```

CE-6

```
CE-5#ping 2001:DB8:66::66 source lo0
Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 2001:DB8:66::66, timeout is 2 seconds:
Packet sent with a source address of 2001:DB8:55::55
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 88/193/564 ms
CE-5#
CE-5#
CE-5#
CE-5#traceroute 2001:DB8:66::66
Type escape sequence to abort.
Tracing the route to 2001:DB8:66::66

  1 2001:DB8:15::1 440 msec 228 msec 80 msec
  2 ::FFFF:10.10.34.2 [MPLS: Labels 22/24 Exp 0] 264 msec 108 msec 68 msec
  3 2001:DB8:16::1 [AS 600] [MPLS: Label 24 Exp 0] 260 msec 148 msec 68 msec
  4 2001:DB8:16::2 [AS 600] 84 msec 64 msec 56 msec
CE-5#
```

Anycast RP (Rendezvous Point)

- An **RP (Rendezvous Point)** acts as the meeting place for sources and receivers for multicast data in multicast network.

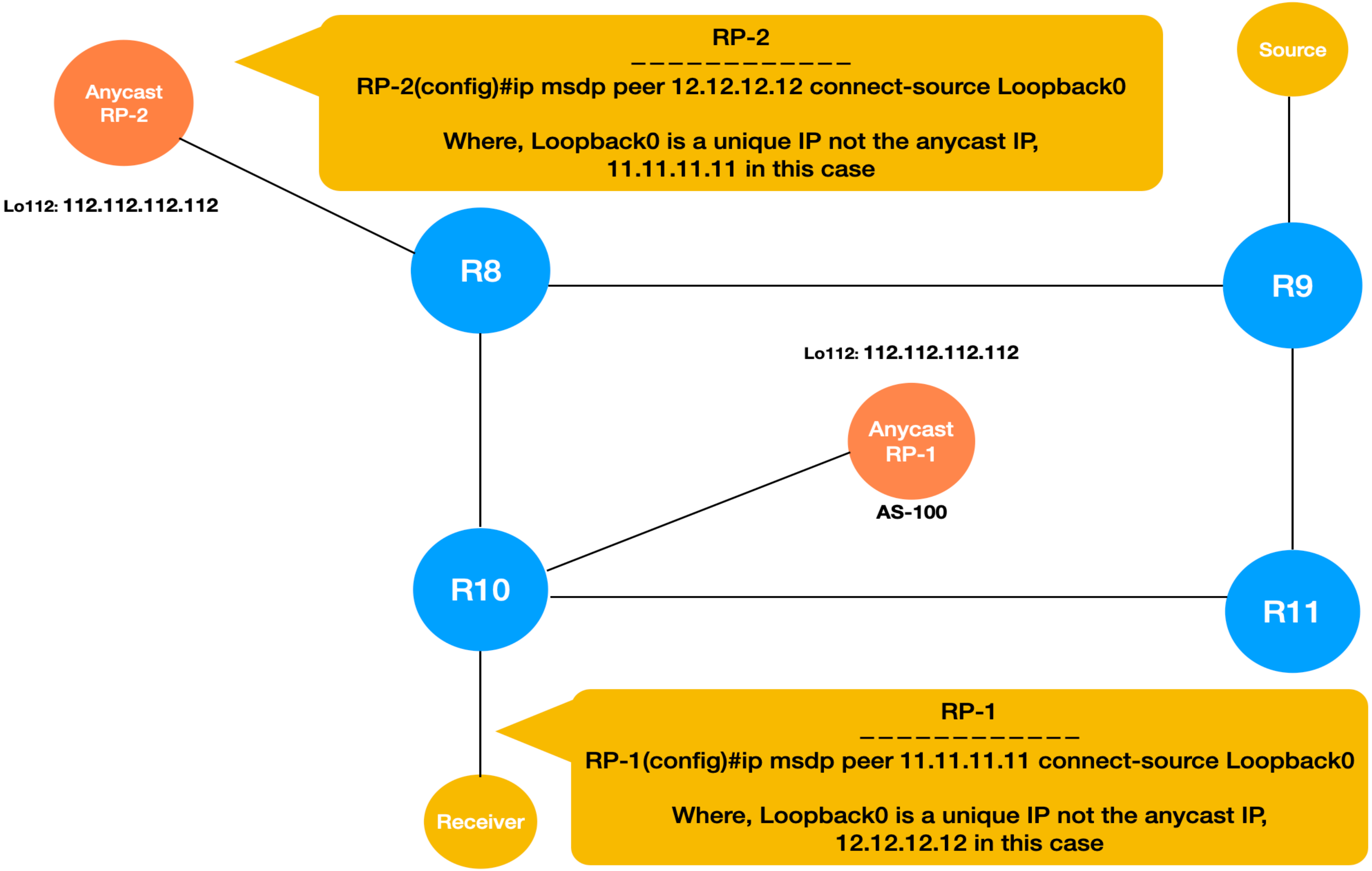
Challenge

- Although there are multiple ways of achieving RP redundancy still the delay in case of failovers is based on the RP/BSR/MA advertisment intervals which are not fast (default upto 60 seconds).

Solution

- Anycast RP is the failover based on the IGP running in the multicast domain which be really fast (especially when used with BFD and LFA).
- Anycast RP solution also provides shared-tree load balancing among any number of active RPs in a multicast domain.
- RP's in the multicast domain share the same unicast IP address.
- PIM join/prune as well the source registration message are sent to the closest RP based on the unicast routing table.
- In order for all the RP's in the multicast domain to be synchronized with each other regarding PIM join/prune and source registration infromation, Anycast RP in conjunction with **MSDP (Multicast Source Discovery Protocol)** defined in RFC 3618 is to be used.

Anycast RP



RP-2
RP-2(config)#ip msdp peer 12.12.12.12 connect-source Loopback0
Where, Loopback0 is a unique IP not the anycast IP, 11.11.11.11 in this case
Source
Anycast RP-2
Lo112: 112.112.112.112
R8
R9
Lo112: 112.112.112.112
Anycast RP-1
AS-100
R10
R11
RP-1
RP-1(config)#ip msdp peer 11.11.11.11 connect-source Loopback0
Where, Loopback0 is a unique IP not the anycast IP, 12.12.12.12 in this case
Receiver

QUESTIONS?

Thank You !!