IPv6-Only PE Design – ALL SAFI

draft-ietf-bess-ipv6-only-pe-design-00
&
draft-ietf-bess-ipv6-only-pe-design-all-safi-00

Gyan Mishra (Verizon)
Jeff Tantsura (Microsoft)
Mankamana Mishra (Cisco)
Sudha Madhavi (Juniper)
Mohana Sundari (Juniper)
Adam Simpson (Nokia)
Qing Yang (Arista)
Shaunglong Chen (Huawei)

IETF-113 March 21st 2022

IPv6-Only PE Design - ALL SAFI

IPv6-Only PE Design (BCP): Adopted April 28th 2021 (Original Draft)

https://datatracker.ietf.org/doc/draft-ietf-bess-ipv6-only-pe-design/

Focus of this draft is on the Proof of Concept testing BCP for vendor implementation & operator deployment

IPv6-Only PE Design PE-CE Edge Peering – Supports IPv4-Unicast 1/1, IPv6-Unicast 2/1 over IPv6 Next Hop

This draft had VPN & MVPN in the original versions so we are just adding back in – so net-net no change (Vendor Testing Draft)

IPv4-Unicast 1/1, IPv6-Unicast 2/1, IPv4-VPN 1/128 IPv4-MVPN 1/129 IPv6-VPN 2/128 IPv6-MVPN 2/129

IPv6-Only PE Design All SAFI (Standards Track): (New draft)

https://datatracker.ietf.org/doc/draft-mishra-bess-ipv4nlri-all-safi-ipv6nh/ (Name change)

This draft is a super set of the original draft above.

Focus of this draft is on the IPv6-Only PE Design from a design change perspective and normative language pertaining to the new design procedural technological as well as paradigm change from traditional “Dual Stacking” to new IPv6-Only PE design for all AFI/SAFI

IPv6-Only PE Design ALL SAFI – Supports All AFI/SAFI over IPv6 Next Hop.

IPv4-Unicast 1/1, IPv6-Unicast 2/1, IPv4-VPN 1/128 IPv4-MVPN 1/129 IPv6-VPN 2/128 IPv6-MVPN 2/129

Replacement Draft ⇔ This draft defines the new procedure for Edge & Inter-AS peering optimization to traditional Dual Stacking providing the identical functionality with the IPV4 addressing & IPv4 BGP peering savings, OPEX Saving and its design applicability to all AFI & SAFI.

https://datatracker.ietf.org/doc/draft-ietf-bess-ipv6-only-pe-design-all-safi/

IPv6-Only PE ALL SAFI Design

https://datatracker.ietf.org/doc/draft-ietf-bess-ipv6-only-pe-design-all-safi/

IPv6-Only PE Design ALL SAFI

- IPv6-Only PE design **ALL** SAFI
- Advertise both IPv4 and IPv6 routes (NLRI) ANY SAFI over a single IPv6 BGP peer
- Supports PE-CE Edge, Inter-AS peering & every type of peering relationships

Functional use cases for ALL AFI/SAFI Design:

- Edge Customer NLRI IPv4 or IPV6 related AFI/SAFI (PE-CE)
  - 1/1 2/1 (Unicast), 1/2 2/2 (Multicast)
- Inter-AS Customer NLRI IPv4 or IPV6 related AFI/SAFI (ASBR-ASBR)
  - 1/1 2/1 (Unicast), 1/2 2/2 (Multicast), 1/128 2/128 (VPN), 1/129 2/129 (MVPN), 1/4 2/4 BGP-LU (6PE/4PE), 1/140 2/140 (BGP VPN Auto Discovery)
- Inter-AS Multicast NLRI IPv4 or IPV6 related AFI/SAFI (ASBR-ASBR)
  - 1/5 2/5 (MCAST-VPN) , 1/8 2/8 (MCAST-VPLS), 1/66 2/66 (BGP MDT-SAFI), 1/78 2/78 (MCAST-TREE)
- Inter-AS PE to Controller NLRI IPv4 or IPV6 related AFI/SAFI (ASBR-ASBR)
  - 1/71 2/71 (BGP-LS), 1/72 2/72 (BGP-LS VPN), 1/75 2/75 (Routing Policy SAFI), 1/80 2/80 BGP-LS-SPF
- Inter-AS L1 VPN, L2 VPN NLRI IPv4 or IPV6 related AFI/SAFI
  - 1/65 2/65 (VPLS), 1/70 2/70 (BGP EVPN), 1/69 2/69 (L1 VPN)
- Inter-AS BGP FlowSpec, Optimizations & SFC NLRI IPv4 or IPV6 related AFI/SAFI
  - 1/132 2/132 (RTC), 1/133 2/133 (BGP FlowSpec), 1/134 2/134 (VPN BGP FlowSpec), 1/9 2/9 (BGP SFC)
- Inter-AS BGP Policy - SR-TE Policy, SD-WAN Policy NLRI IPv4 or IPV6 related AFI/SAFI
  - 1/73 2/73 (SR-TE), 1/74 2/74 (SD-WAN Capabilities)

IPv6-Only PE ALL SAFI Design

IANA BGP AFI SAFI ⇔ ALL SAFI's listed below support the IPv6-Only PE Design framework
https://www.iana.org/assignments/safi-namespace/safi-namespace.xhtml

Value	Description	Reference
0	Reserved	[RFC4760]
1	Network Layer Reachability Information used for unicast forwarding	[RFC4760]
2	Network Layer Reachability Information used for multicast forwarding	[RFC4760]
3	Reserved	[RFC4760]
4	Network Layer Reachability Information (NLRI) with MPLS Labels	[RFC8277]
5	MCAST-VPN	[RFC6514]
6	Network Layer Reachability Information used for Dynamic Placement of Multi-Segment Pseudowires	[RFC7267]
7	Encapsulation SAFI (OBSOLETE)	[RFC9012]
8	MCAST-VPLS	[RFC7117]
9	BGP SFC	[RFC9015]
10-63	Unassigned	
64	Tunnel SAFI	[Gargi_Nalawade][draft-nalawade-kapoor-tunnel-safi-01]
65	Virtual Private LAN Service (VPLS)	[RFC4761][RFC6074]
66	BGP MDT SAFI	[RFC6037]
67	BGP 4over6 SAFI	[RFC5747]
68	BGP 6over4 SAFI	[Yong_Cui]
69	Layer-1 VPN auto-discovery information	[RFC5195]
70	BGP EVPNs	[RFC7432]
71	BGP-LS	[RFC7752]
72	BGP-LS-VPN	[RFC7752]

IPv6-Only PE ALL SAFI Design

IANA BGP AFI SAFI (Continued)

https://www.iana.org/assignments/safi-namespace/safi-namespace.xhtml

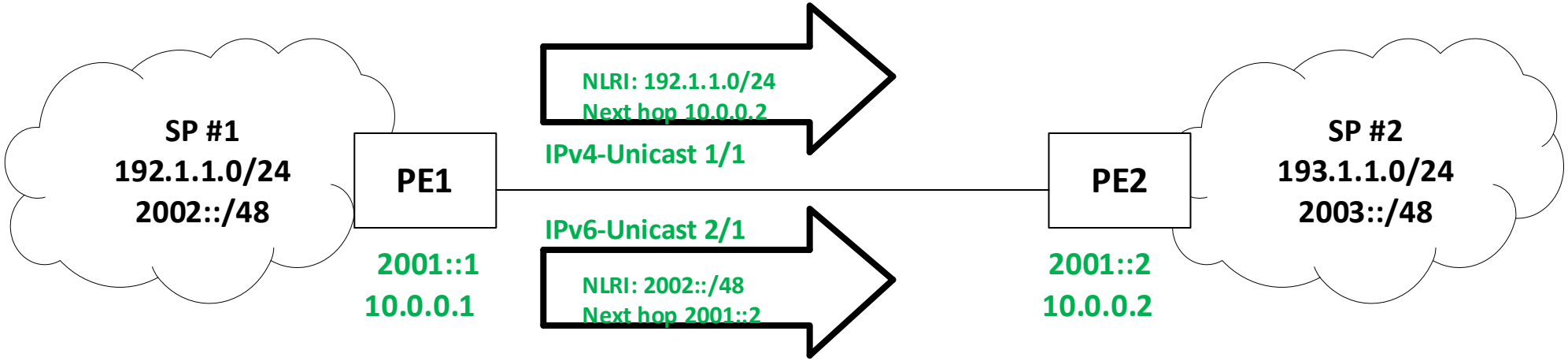
Value	Description	Reference
73	SR TE Policy SAFI	[draft-previdi-idr-segment-routing-te-policy]
74	SD-WAN Capabilities	[draft-dunbar-idr-sdwan-port-safi]
75	Routing Policy SAFI	[draft-ietf-idr-rpd-02]
76	Classful-Transport SAFI	[draft-kaliraj-idr-bgp-classful-transport-planes-00]--→ Transport Only
77	Tunneled Traffic Flowspec	[draft-ietf-idr-flowspec-nvo3-10]
78	MCAST-TREE	[draft-ietf-bess-bgp-multicast-03]
79	BGP-DPS (Dynamic Path Selection)	[https://eos.arista.com/eos-4-26-2f/dps-vpn-scaling-using-bgp][Venkit_Kasiviswanathan]
80	BGP-LS-SPF	[draft-ietf-lsvr-bgp-spf-15][Victor_Kuarsingh]
81-127	Unassigned	
128	MPLS-labeled VPN address	[RFC4364][RFC8277]
129	Multicast for BGP/MPLS IP Virtual Private Networks (VPNs)	[RFC6513][RFC6514]
130-131	Reserved	[RFC4760]
132	Route Target constrains	[RFC4684]
133	Dissemination of Flow Specification rules	[RFC8955]
134	L3VPN Dissemination of Flow Specification rules	[RFC8955]
135-139	Reserved	[RFC4760]
140	VPN auto-discovery	[draft-ietf-l3vpn-bgpvpn-auto]
141-240	Reserved	[RFC4760]
241-254	Reserved for Private Use	[RFC4760]
255	Reserved	[RFC4760]

Traditional PE-CE Dual Stacked Peering with Separate IPv4 & IPv6 Peer

IXP Peering has a separate IPv4 & IPv6 peer with IPv4 NH peer carrying IPv4 NLRI & IPv6 NH peer carrying IPv6 NLRI

PE-CE Native IP (GRT) Routing

Traditional Dual Stack ⇔ Separate IPv4 & IPv6 peering



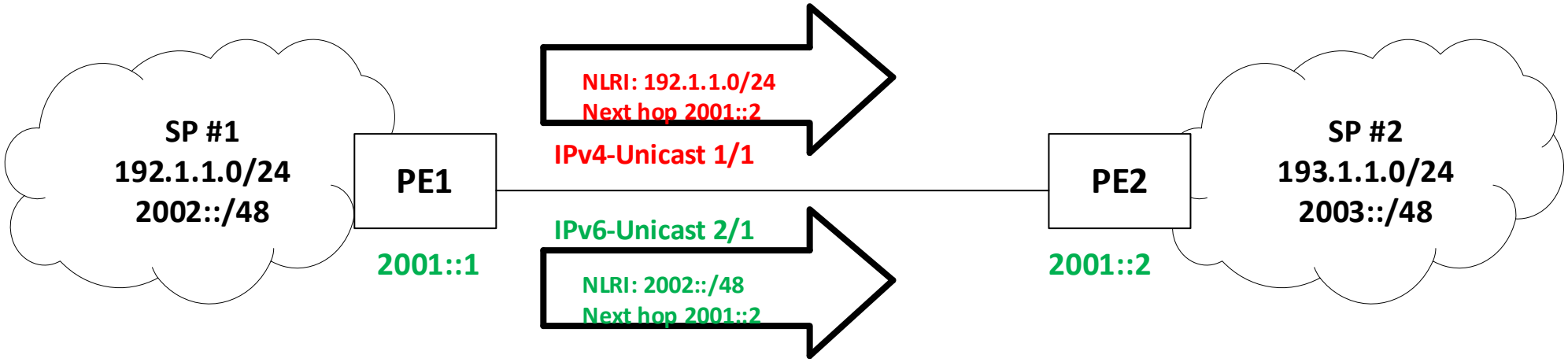
IPv6-Only PE Design ALL SAFI ⇔ PE-CE Edge Single IPv6 Peer carrying IPv4 & IPv6 NLRI (Same Dual Stack Functionality)

So now with RFC 8950 NH encoding schema of 16 / 32 byte IPv6 next hop both IPv4 & IPv6 NLRI can be advertised using a single IPv6 peer.

This basic concept can eliminate all IPv4 peering at the Edge and within the Core.

PE-CE Native IP (GRT) Routing

IPv6-Only PE Design ⇔ Single IPv6-Only Pure Transport Peer to carry both IPv4 & IPv6 NLRI



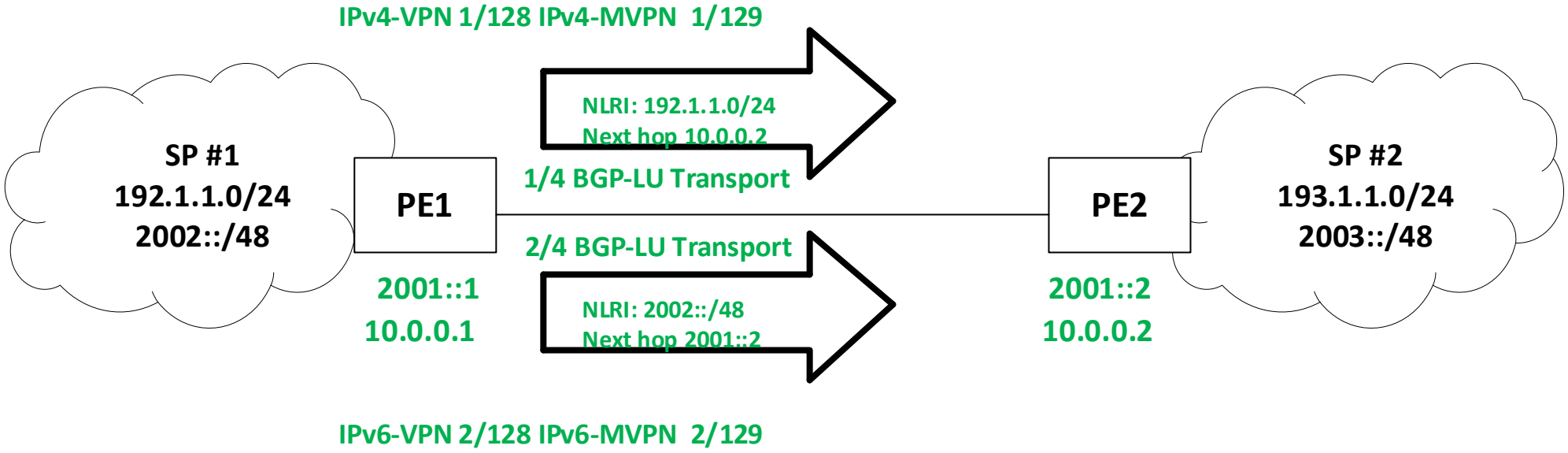
IXP Peering of all SPs would now be able to use
a single IPv6 peer per SP adjacency
Eliminate IPv4 Address depletion issues

Traditional PE-PE Inter-AS Dual Stacked Peering with Separate IPv4 & IPv6 Peer

IXP Peering has a separate IPv4 & IPv6 peer with IPv4 NH peer carrying IPv4 NLRI & IPv6 NH peer carrying IPv6 NLRI

Inter-AS Opt A,B,C,AB / SR-TE Binding SID

Traditional Dual Stack ⇔ Separate IPv4 & IPv6 peering

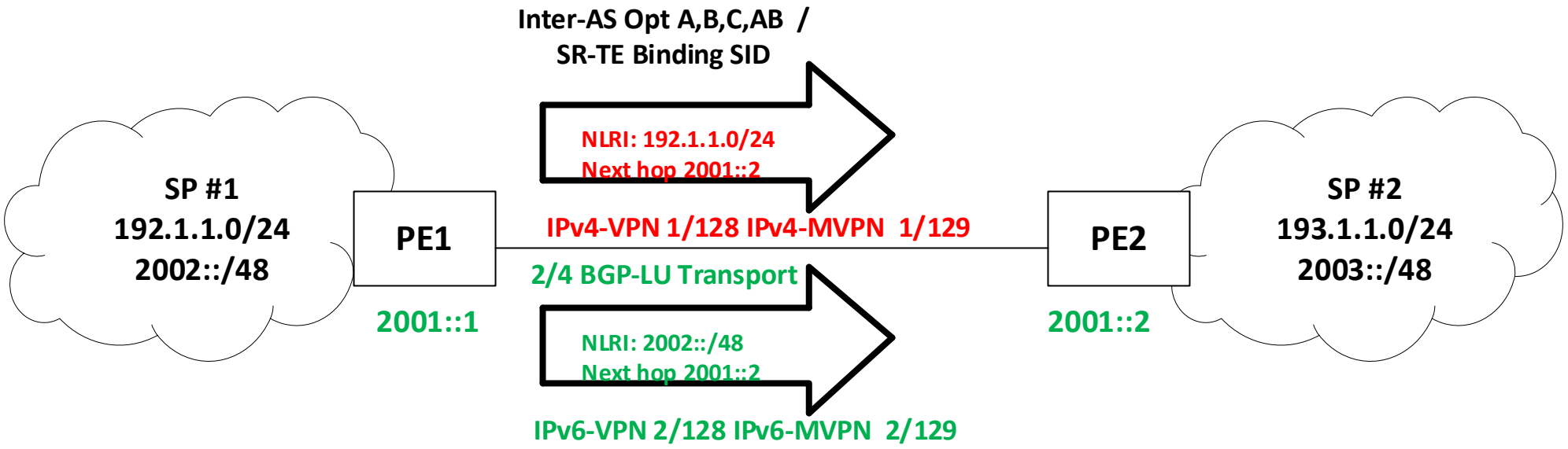


IPv6-Only PE Design ALL SAFI ⇔ Inter-AS - Single IPv6 Peer carrying IPv4 & IPv6 NLRI (Same Dual Stack Functionality)

So now with RFC8950- NH encoding schema of 24/48 byte IPv6 next hop both IPv4 & IPv6 NLRI can be advertised using a single IPv6 peer.

This basic concept can eliminate all IPv4 peering at the Edge and within the Core.

IPv6-Only PE Design ⇔ Single IPv6-Only Pure Transport Peer to carry both IPv4 & IPv6 NLRI



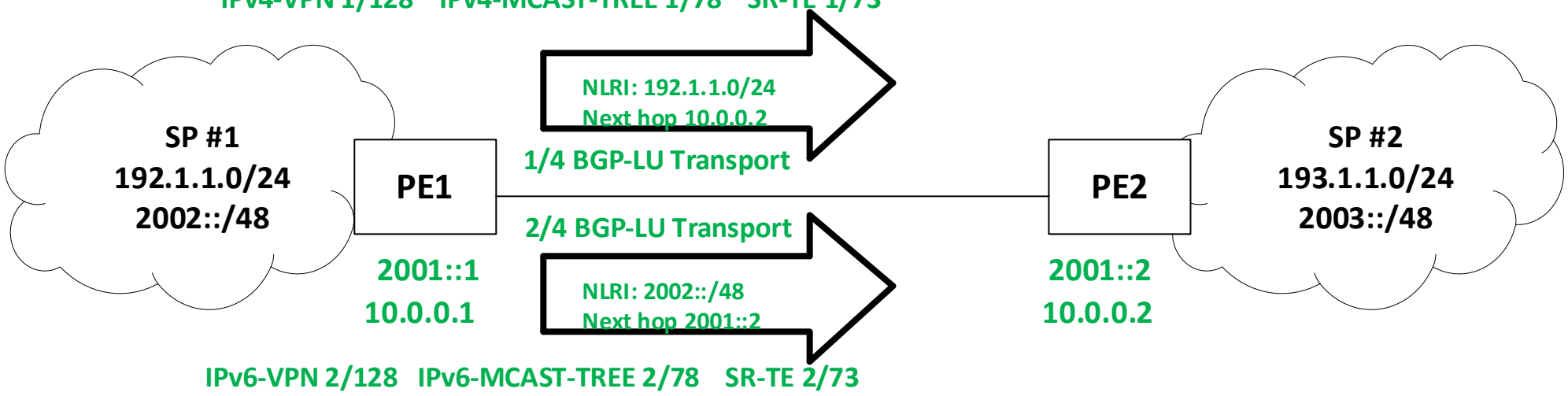
IXP Peering of all SPs would now be able to use
a single IPv6 peer per SP adjacency
Eliminate IPv4 Address Depletion Issues

Traditional Inter-AS Dual Stacked Peering with Separate IPv4 & IPv6 Peer

IXP Peering has a separate IPv4 & IPv6 peer with IPv4 NH peer carrying IPv4 NLRI & IPv6 NH peer carrying IPv6 NLRI

Inter-AS Opt A,B,C,AB / SR-TE Binding SID

Traditional Dual Stack ⇔ Separate IPv4 & IPv6 peering



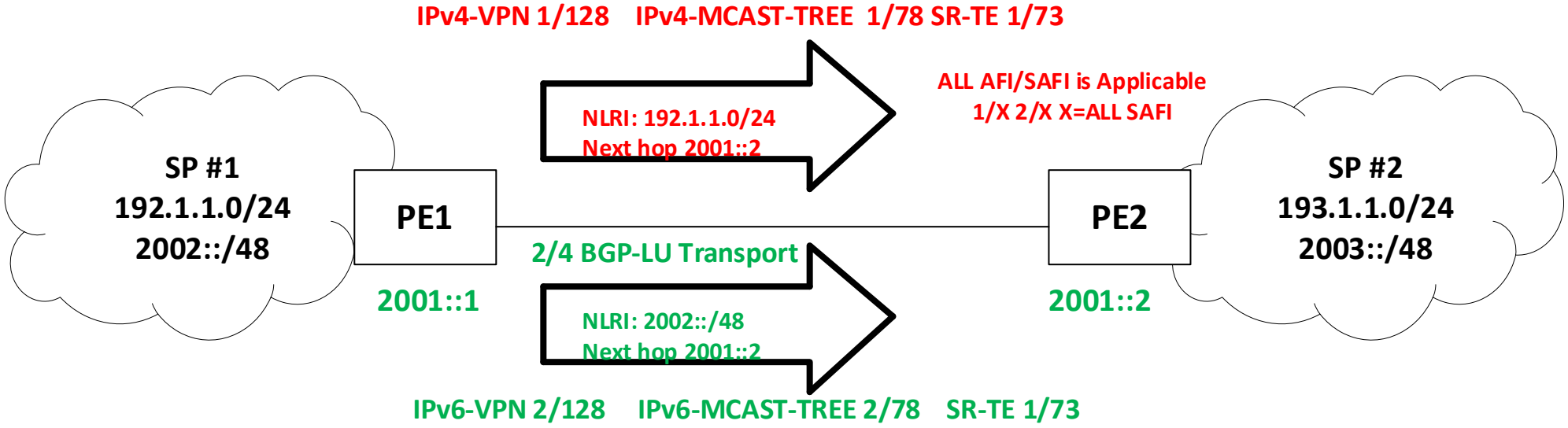
IPv6-Only PE Design ALL SAFI ⇔ Inter-AS - Single IPv6 Peer carrying IPv4 & IPv6 NLRI (Same Dual Stack Functionality)

So now with RFC8950- NH encoding schema of 24/48 byte IPv6 next hop both IPv4 & IPv6 NLRI can be advertised using a single IPv6 peer.

This basic concept can eliminate all IPv4 peering at the Edge and within the Core.

IPv6-Only PE Design ⇔ Single IPv6-Only Pure Transport Peer to carry both IPv4 & IPv6 NLRI

Inter-AS Opt A,B,C,AB /
SR-TE Binding SID



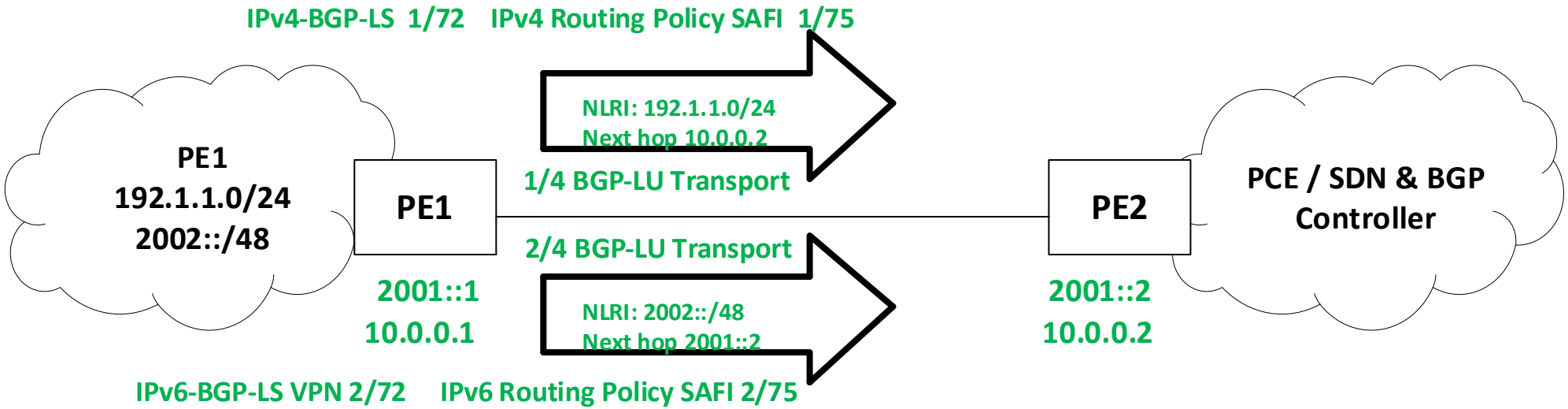
IXP Peering of all SPs would now be able to use a single IPv6 peer per SP adjacency
Eliminate IPv4 Address Depletion Issues

Traditional PE to Controller Dual Stacked Peering with Separate IPv4 & IPv6 Peer

PE to PCE/SD or BGP Controller has a separate
IPv4 & IPv6 peer with IPv4 NH peer carrying
IPv4 NLRI & IPv6 NH peer carrying IPv6 NLRI

Inter-AS Opt A,B,C,AB /
SR-TE Binding SID

Traditional Dual Stack ⇔ Separate IPv4 & IPv6 peering

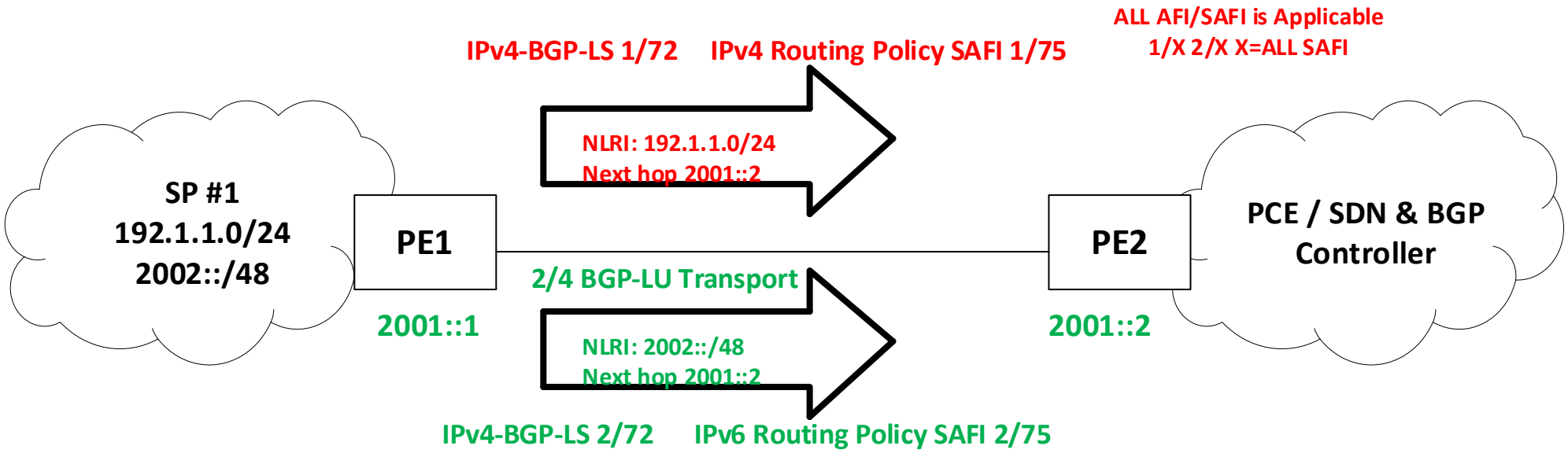


IPv6-Only PE Design ALL SAFI ⇔ PE to Controller - Single IPv6 Peer carrying IPv4 & IPv6 NLRI (Same Dual Stack Functionality)

PE to PCE/SDN or BGP Controller single IPv6 Peer
So now with RFC8950- NH encoding schema of 24/48 byte IPv6 next hop both IPv4 & IPv6 NLRI can be advertised using a single IPv6 peer.

This basic concept can eliminate all PE to Controller peering at the Edge and within the Core.

IPv6-Only PE Design ⇔ Single IPv6-Only Pure Transport Peer to carry both IPv4 & IPv6 NLRI



Vendors: Cisco, Juniper, Nokia, Arista, Huawei

Hardware Platforms & Router Code Revision
on the IPv6-Only PE to source the ICMP packets.

Vendors: Cisco, Juniper, Nokia, Arista, Huawei
Code & Platform chosen for testing platform by Vender

Cisco: Edge Router- XR ASR 9910 IOS XR 7.4.1, Core Router- NCS 6000 7.2.2, CRS-X 6.7.4

Juniper: Edge Router- MX platform MX480, MX960, Core Router- PTX Platform PTX5000, PTC10K8 (JUNOS and EVO) Release 20.4R2

Nokia: Edge and Core-7750 Service Router, Release R21

Huawei: Edge and Core-VRPv8, Release VRP-V800R020C10

Arista: XXX

IPv6-Only Edge Peering Design Vendor POC Testing

Core - MPLS underlay IGP-ISIS

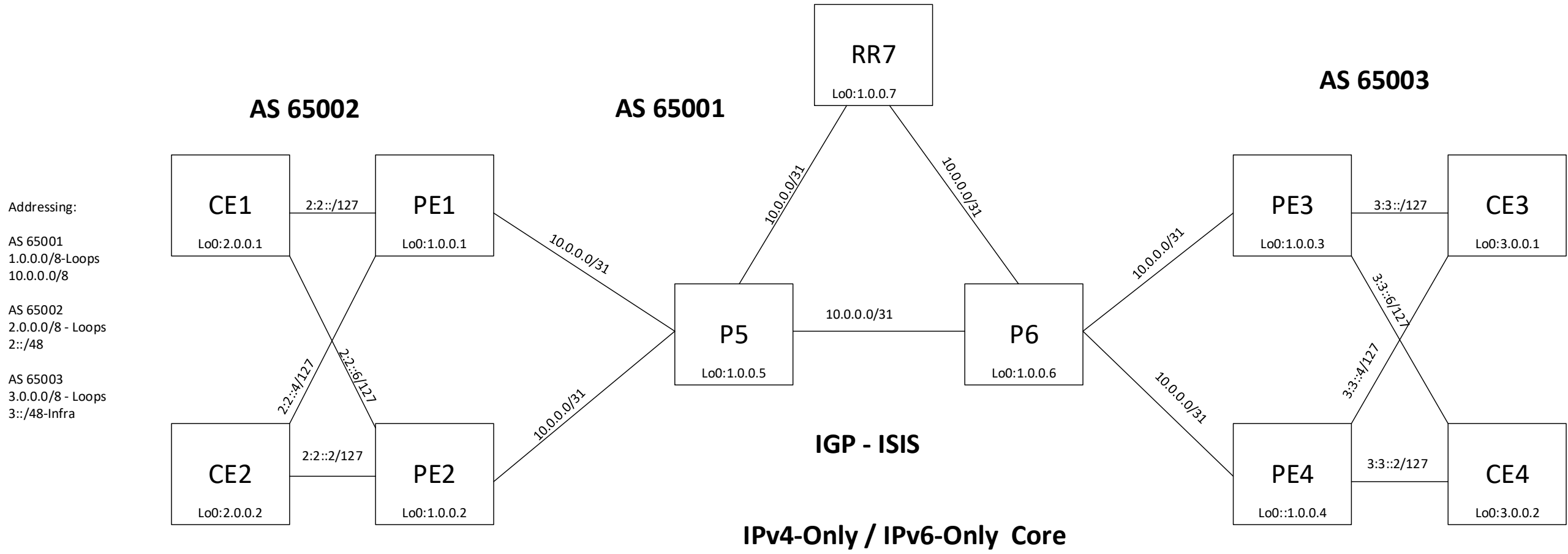
- ECMP
- PE-CE(FRR)
- BGP PIC Edge (PE)
- PE – next-hop-self
- Label Allocation mode - prefix based, per-CE, per-VRF (Note each vendors nomenclature is different)
- RR (Route Reflector) - Add Path & Next-Hop-Unchanged (Inter-AS Option-C)
- PE Link and node failover & route withdrawn – Ping & Traceroute operational testing slide 24

Edge & Inter-AS testing:

- BFD – Single Hop – Enabled BFD on Interface (IPv6) with BGP session bootstraps only to IPv6 as the peer is IPv6 (test)
- BFD – BFD LAG – RFC 7730 Micro BFD (Optional)
- IPFIX
- QOS marking traffic classification & scheduling
- CE Link & Node failover test & route withdrawn – Ping & Traceroute operational testing end to end

Common Vendor test bed “Intra-AS” used by all vendors – Cisco, Juniper, Nokia, Arista, Huawei

IPv4-Only Core Global & VPN Test #1-4



Addressing:

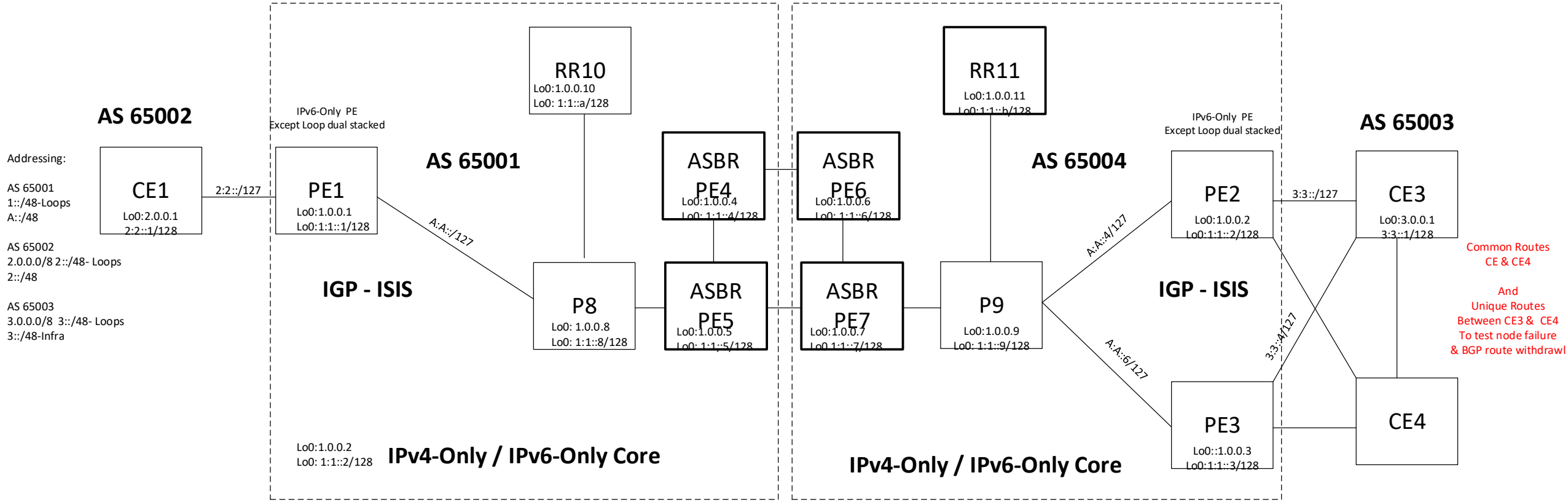
AS 65001
1.0.0.0/8-Loops
10.0.0.0/8

AS 65002
2.0.0.0/8 - Loops
2::/48

AS 65003
3.0.0.0/8 - Loops
3::/48-Infra

Common Vendor “Inter-AS” test bed used by all vendors – Cisco, Juniper, Nokia, Arista, Huawei
(Updated with Dual ASBRs for PE-PE failover and only one side Dual CE for failover test)

IPv6-Only Core Global & VPN Test #5-8



Inter-AS Opt
A, B, AB, C
VRF & Global 4PE

IPv6-Only PE ALL SAFI DESIGN – Test Scenario's

IPv6-Only PE ALL SAFI Design Test Scenario's:

1. Unicast SAFI=1 (Original IPv6-Only PE Design)
IPv4 Core:
Test #1 Global table 6PE
Test #2 L3 VPN

Global table IPv6
IPv6 Core:
Test #3 Global table 4PE
Test #4 L3 VPN

2. SAFI VPN=128 & MVPN=129 (Updated tests)

MVPN SAFI testing:
MVPN we will keep it very simple using RFC 6388 X-PMSI PTA Tunnel Type 2 mLDP P2MP BGP-AD w/ PIM C-Signaling.

Inter-AS Options testing:
This test will cover the common IXP POP L3 VPN Inter-AS peering w/ IPv6-Only PE over IPv4 Core & IPv6 Core and Inter-AS options B & C. (Opt-A is excluded as that is similar to PE-CE and Opt-AB is a combo of A & B so that is excluded as well)

IPv4 Core
Test #5 Global table Option B (6PE)
Test #6 Global table Option C (6PE)
Test #7 L3 VPN Inter AS Option B
Test #8 L3 VPN Inter AS Option C

IPv6 Core
Test #9 Global table Option B (4PE)
Test #10 Global table Option C (4PE)
Test #11 L3 VPN Inter AS Option B
Test #12 L3 VPN Inter AS Option C

Juniper Updates:
Code & Platform chosen for testing platform by Vender

Juniper: Edge Router- MX platform MX480, MX960, Core Router- PTX
Platform PTX5000, PTC10K8 (JUNOS and EVO) Release 20.4R2
PE-PE core different release inter-as

Juniper Test :

- Tested v4 edge over v6 core in a virtual setup using vMX platform and 20.4R2 and LDPv6 as underlay, but there were some data plane forwarding issues. Tested same setup on latest release 21.4 and it worked. Investigating what the minimum version is for this setup to work.
- Tested on physical hardware above using the Juniper POC diagram.
- Completed IPv6-Only PE design functionality test with PE-CE IPv6 peer carrying IPv4 & IPv6 prefixes control plane validation & data plane forwarding plane validation and verified end to end reachability CE to CE forwarding plane with Default Per-CE label allocation mode. Tested with IPv4-Only Core & IPv6-Only Core and proved that the IPv6-Only PE design solution works. Both IPv4 & IPv6 packets were forwarded identical functionality of “Dual Stack” without having IPv4 address configured.


Q&A

THANK YOU