IPv4 NLRI with IPv6 NH
Use cases

draft-ietf-bess-deployment-guide-ipv4nlri-ipv6nh-02

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-111, July 27th 2021

Latest Draft Updates during WG Adoption Poll:

- The draft was adopted by the WG on April 28th 2021
- Update authors list Juniper contact Lili wang and replaced with Sudha Madhavi & Mohana Sundari
- Updated the draft from feedback during the WG adoption clarifying the new IPv6-only PE-CE peering architecture & design framework where IPv4 and IPv6 forwarding occurs just as it did with a IPv4/IPv6 Dual stacked PE-CE design, but now without and IPv4 address defined on the PE-CE attachment circuit. This new IPv6-Only PE-CE edge peering design can be applied ubiquitously to any use case.
- Co-authors meeting in June with each vendor deep dive into the IPv6-Only PE-CE peering architecture and created 4 test cases added to the draft.
- Created a detailed test plan of what protocols and functions would be tested as well as a common lab diagram layout of devices in the test bed.
- During the meeting we decided to investigate Interoperability testing through EANTC Denmark and I would take the lead through Verizon engagement.
- Draft has been updated with the vendor platforms and code revisions & latest on initial proof of concept testing completed by juniper.

IPv6-Only PE-CE Peering Design Interoperability testing status:

- We discussed feasibility and practicality of interoperability testing as well as logistics with shipping hardware and determined that it would be logistically impossible to ship hardware to a central site for vendor interoperability testing.
- We discussed virtual VM OVA testing with QCOW2 image in a virtual environment and determined that due to the history of hardware caveats related to testing that the testing must be done on physical hardware to make the test results accurate.
- We investigated possibility of using EANTC in Denmark for Interoperability testing in July 2021, however due to the short timeframe as well as initial prerequisite testing to be completed first this was not possible.
- We decided that we can accomplish the goals of the draft as a BCP for IPv6-Only peering architecture proliferation worldwide by providing the vendor specific detailed test results of the IPv6-Only peering architecture “4 E2E Use Cases” so that all vendors that are not part of the testing will implement this design as well. We do not think that by not doing the vendor interoperability testing diminishes the BCP draft in any way, and as vendors looking to deploy will now be able to take draft test results and code to their vendors that were included in this draft or even vendors not included in this draft, and can now go to the next step on vendor interoperability testing as necessary for their environments.

Vendors: Cisco, Juniper, Nokia, Arista, Huawei

IPv6-Only PE-CE Peering Design E2E Test Vendor test Cases:

- Test-1 E2E IPv6-Only PE-CE, Global table routing over IPv4-Only Core
- Test-2 E2E IPv6-Only PE-CE, VPN Overlay over IPv4-Only Core
- Test-3 E2E IPv6-Only PE-CE, Global table routing over IPv6-Only Core
- Test-4 E2E IPv6-Only PE-CE, VPN Overlay over IPv6-Only Core
- IPv6-Only operational considerations testing related to IPv4 Ping & traceroute & all ICMP types & code points. Scenario where IPv4 route withdrawn and ICMP Destination Unreachable has to be sent by the PE, however the PE does not have an IPv4 address. The solution the authors came up with was that we would add a Loopback 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

Juniper Test Results:

Tested v4 edge over v6 core in a virtual setup using vMX platforrm 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.

IPv6-Only Edge Peering Design Vendor POC

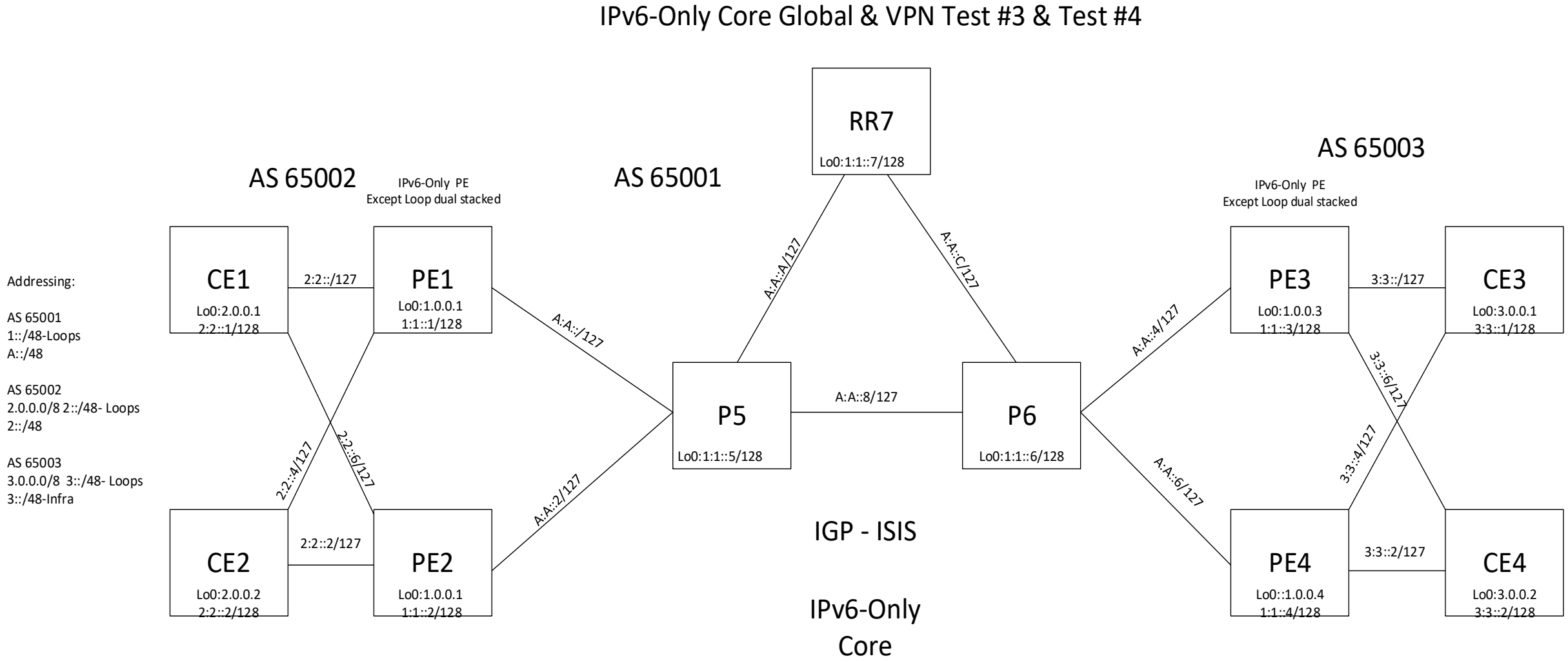
Core - MPLS underlay IGP-ISIS

- 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)

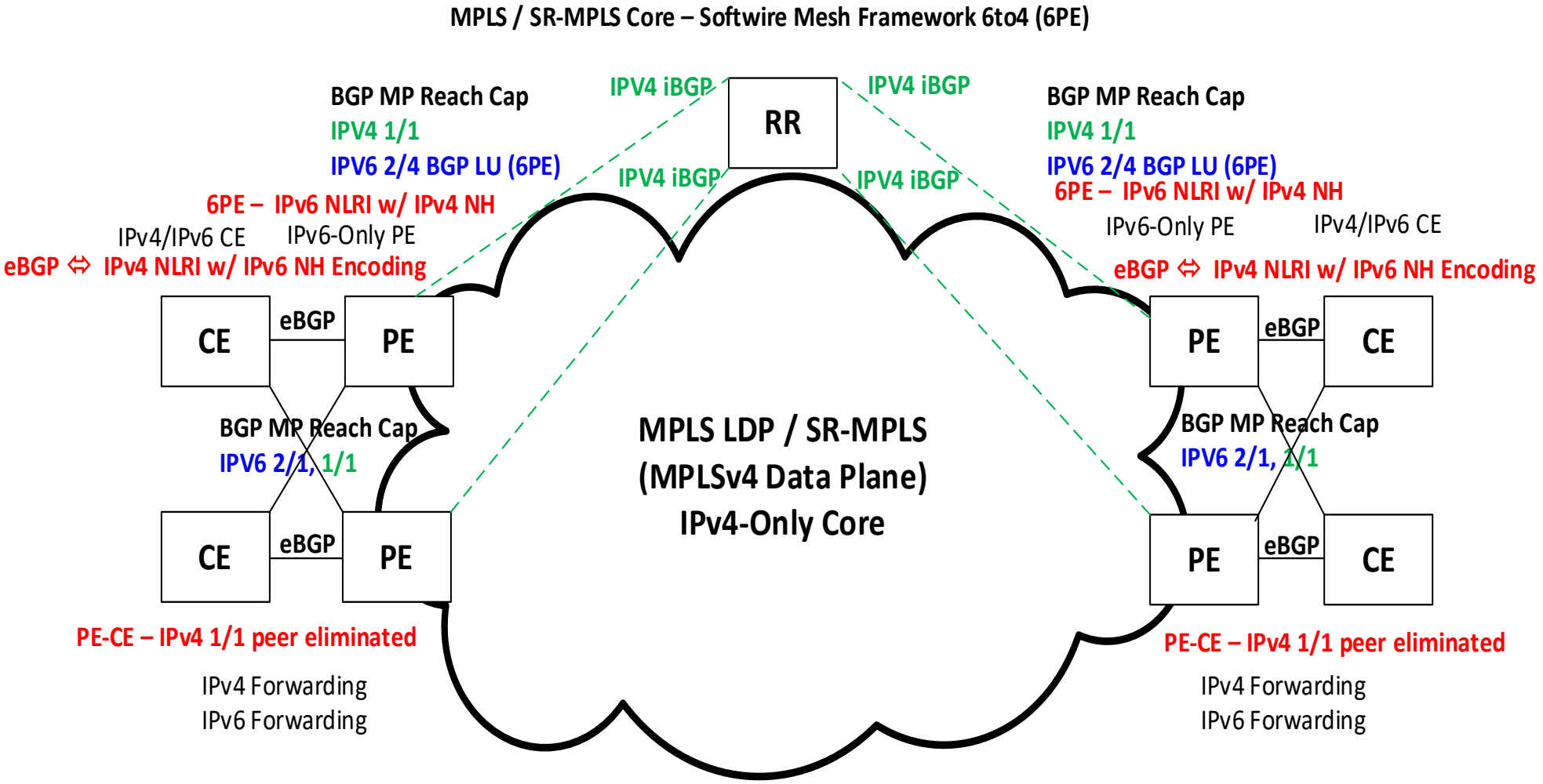
Edge testing:

- QOS
- IPFIX
- RR
- Add Path

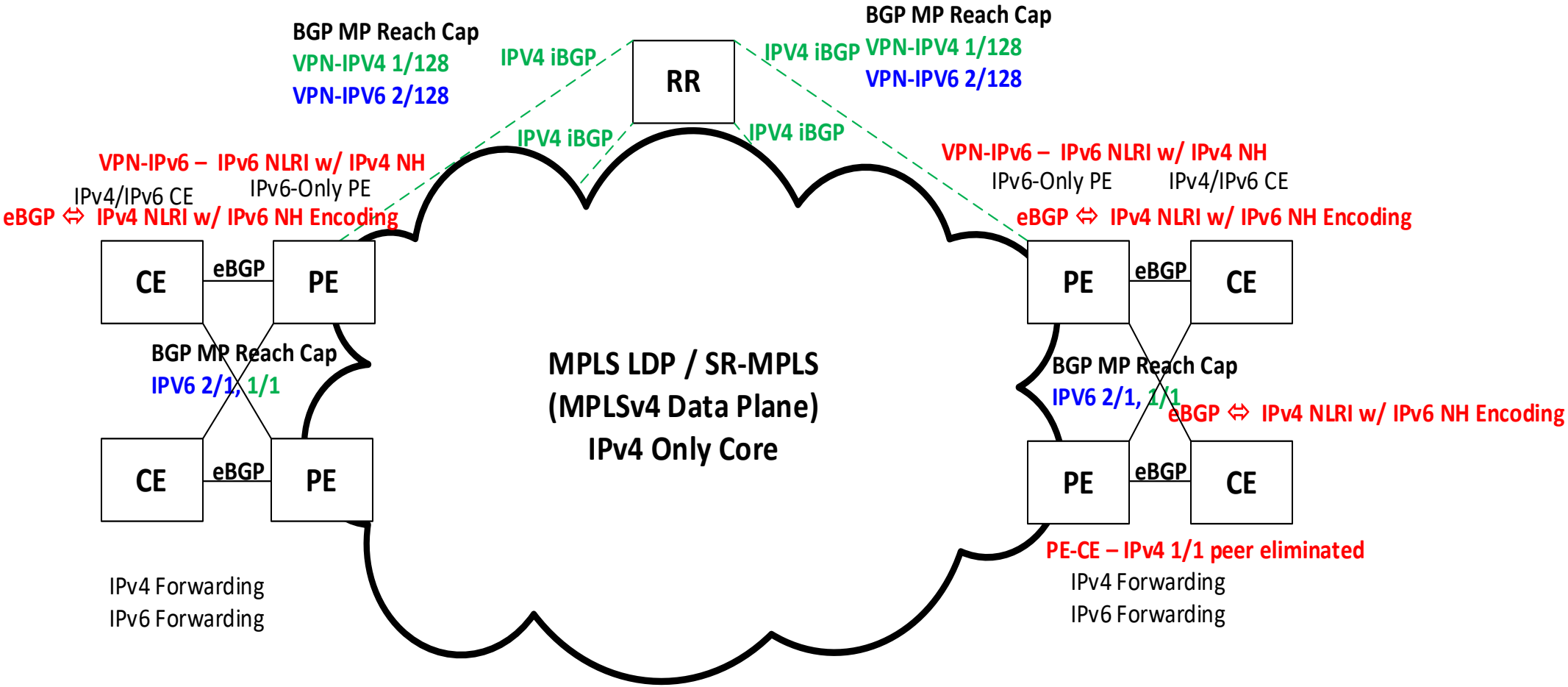
Common Vendor test bed used by all vendors – Cisco, Juniper, Nokia, Arista, Huawei



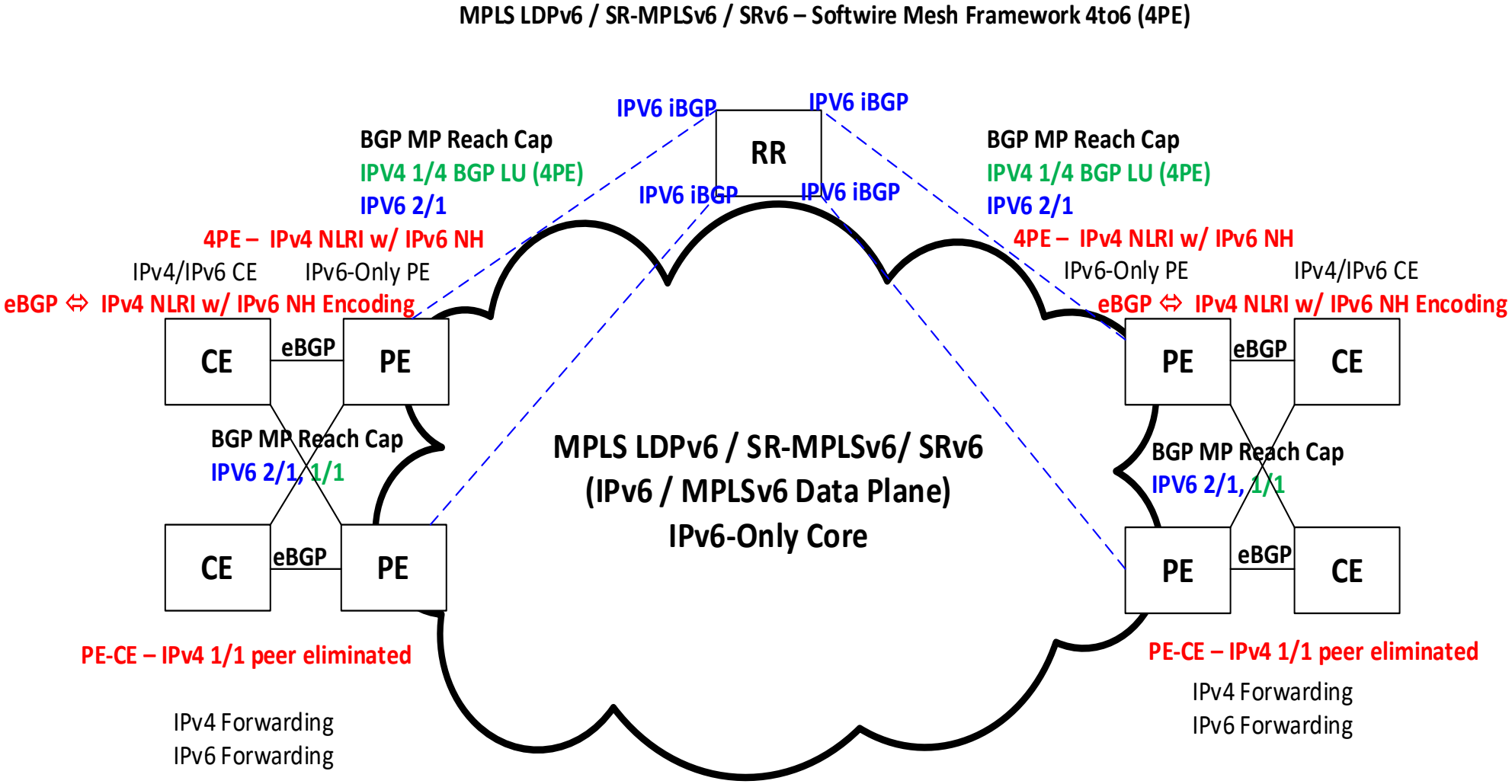
Test Case #1 E2E IPv6-Only PE-CE, Global Table over IPv4-Only Core (6PE)
Softwire mesh framework 6to4 softwire



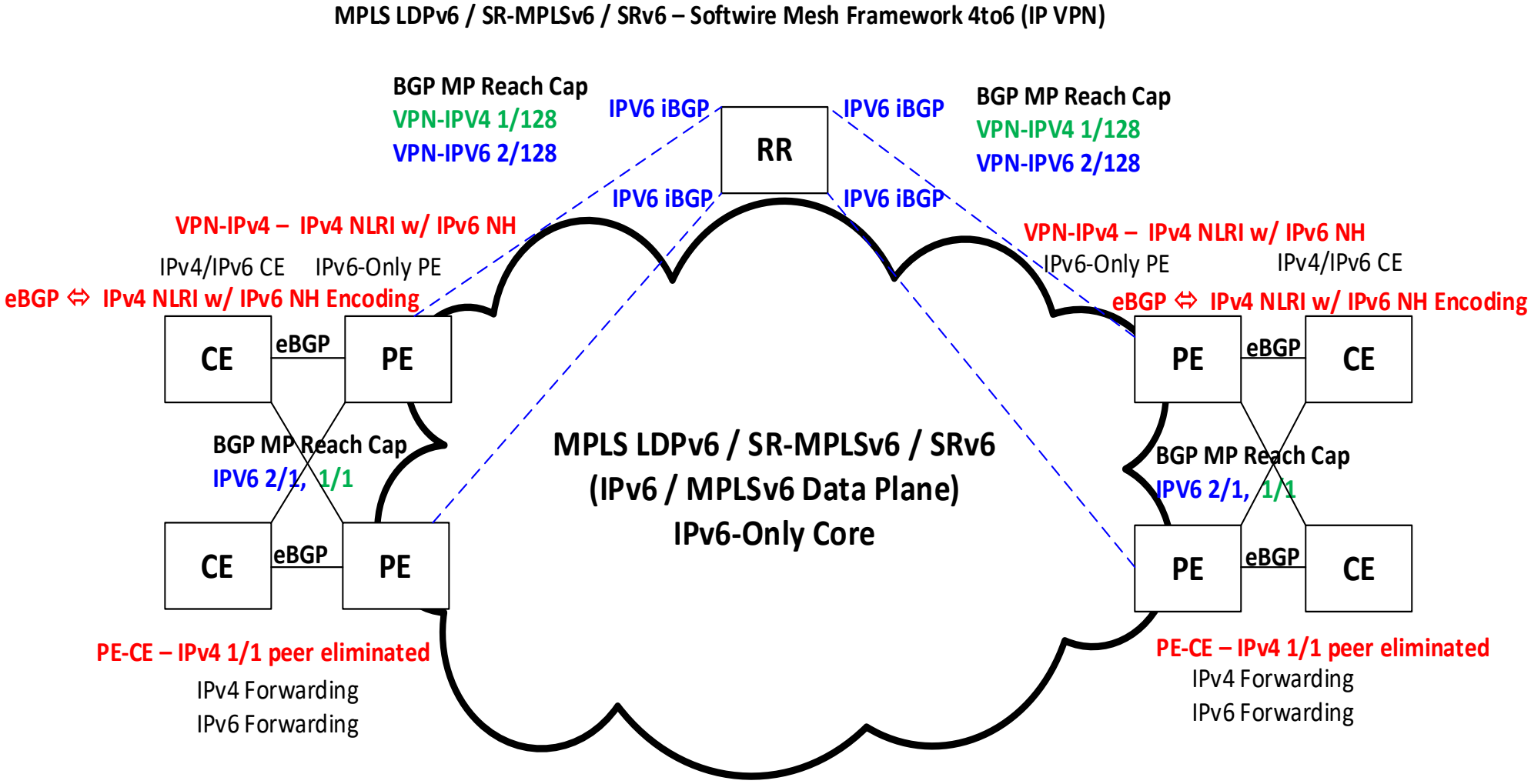
Test Case #2 E2E IPv6-Only PE-CE, VPN over IPv4-Only Core (IP VPN)
Softwire mesh framework 6to4 softwire



Test Case #3 E2E IPv6-Only PE-CE, Global Table over IPv6-Only Core (4PE)
Softwire mesh framework 4to6 softwire



Test case #4 E2E IPv6-Only PE-CE, VPN over IPv6-Only Core (IP VPN)
Softwire mesh framework 4to6 softwire



New Drafts parity as a result of this draft:

4to6 Softwire mesh

<Current name> “this presentations draft” ⇔ PE-CE AFI/SAFI 1/1 2/1 Unicast SAFI Only

draft-ietf-bess-deployment-guide-ipv4nlri-ipv6nh-02

<New name>

draft-ietf-bess-ipv6-only-pe-design

! Ubiquitous AFI/SAFI applicability –Same IPv6-Only peering design control plane IPv4 NLRI over IPv6 NH and IPv4 forwarding with IPv6-Only Peering⇔ Apply now to all AFI/SAFI

draft-mishra-bess-all-safi-ipv4nlri-ipv6nh

6to4 Softwire mesh

Single AFI/SAFI⇔ PE-CE AFI/SAFI 1/1 2/1 Unicast SAFI Only (BCP test cases identical to other draft)

draft-mishra-bess-ipv4-only-pe-design

! Ubiquitous AFI/SAFI applicability –Same IPv4-Only peering design control plane IPv6 NLRI over IPv4 NH and IPv4 forwarding with IPv4-Only Peering⇔ Apply now to all AFI/SAFI

draft-mishra-bess-all-safi-ipv6nlri-ipv4nh


Q&A

THANK YOU

With RFC 5549 NH encoding of IPv4 NLRI with IPv6 next hop used for iBGP PE-RR peering we can apply the same concept to all eBGP "Core" & "Edge" peering and utilize a "single protocol only" on all peering, IPv6 AFI=2.

Tremendous OPEX Saving to eliminate IPv4 eBGP peering. This basic concept will help alleviate IPv4 address depletion issues on IXP NAPs as well as both public & private peering pints for both Service Providers & Enterprises. This concept will also help pave the way as a stepping stone to eventual IPv6 Only customers as well as the eventual complete elimination of IPv4.

IPv4 AFI=1 eBGP Peering can be eliminated from all Edge PE-CE for both Enterprise and Service Provider networks.

BGP is a transport and so just as we stack SAFI's on an AFI on our PE-RR iBGP peering in a typical MPLS or SR environment, we can do the same with stacking both AFI=1 & AFI=2 capability onto a **single IPv6 eBGP** peer. So basically the **IPv6 eBGP** peering has in the MP Reach capability exchange has advertised & received both AFI=1 & AFI=2 capabilities, and so now the IPv6 peer can advertise both IPv4 NLRI & IPV6 NLRI with IPv6 next hop encoding as an IPv6 address as defined in RFC 5549 update.

With IPv4 address depletion issues with Internet IXP exchange points which have come up in recent at NANOG 65 Montreal 2015, have proposed using RFC 5594 IPv4 NLRI encoding in IPv6 NH as a way to eliminate all IPv4 peering at the IXP POP's to save on address space as well as OPEX expenditure in maintaining both IPV4 & IPV6 peering.

Softwire mesh framework RFC 5565 tunneling v6 over v4 transport data plane using IPv4 Next Hop encoding over MPLS LDPv4 core(6PE scenario)

In an Enterprise Dual stacked customers edge environment over IPv4 MPLS "BGP free" core, where 6PE is used to connect IPv6 islands over an IPv4 core, PE-RR core peering AFI/SAFI 1/1 and 6PE BGP-LU IPv6 labeled unicast AFI/SAFI 1/4. PE-RR 6PE requires IPv6 NLRI to be carried with IPv4 Next Hop encoding.

Softwire mesh framework RFC 5565 tunneling v6 over v4 transport data plane using IPv4 Next Hop encoding over MPLS LDPv4 core(IP VPN scenario)

In a Service provider Dual stacked customers edge environment over IPv4 MPLS "BGP free" core at the edge, we maintain the Dual stacked PE-CE IPv4 peering AFI/SAFI 1/1(IPV4) & 2/1 (IPV6). In the core PE-RR peering we have AFI/SAFI 1/128 (VPN-IPV4) 1/129 (MVPN) 2/128 (VPN-IPV6) 2/129 (MVPN). PE-RR peering requires 2/128 (VPN-IPV6) 2/129 (MVPN) NLRI to be carried over IPv4 Next Hop encoding.

With Enterprise or Service provider Dual stacked customers edge environment in Green & Brown field deployments of IPv6 transport Core using MPLS LDPv6, SR-MPLSv6 or SRv6 IPv6 data plane with “BGP free” core, we can now eliminate separate v4 & v6 Edge PE-CE peering so all SAFI related to AFI=1 can now be carried over AFI=2.

Softwire mesh framework RFC 5565 tunneling v4 over v6 transport data plane using RFC 5549 NH encoding over MPLS LDPv6 or SRv6 core(4PE scenario)

With Enterprise or Service provider Dual stacked customers edge environment in Green & Brown field deployments of IPv6 transport Core using MPLS LDPv6, SR-MPLSv6 or SRv6 IPv6 data plane with “BGP free” core, we can now eliminate separate v4 & v6 peering so all SAFI related to AFI=1 can now be carried over AFI=2.

In an Enterprise Dual stacked customers edge environment over an IPv6 MPLS LDPv6, SR-MPLSv6 or SRv6 IPv6 core, where 4PE is used to connect IPv4 islands over a IPv6 core, PE-RR peering ⇔ BGP-LU IPV4 labeled unicast (4PE) AFI/SAFI 1/4 to connect IPv4 islands over IPv6 core, carrying AFI/SAFI 2/1 would now also carry AFI/SAFI 1/1 & 1/4 4PE IPv4 Labeled unicast with RFC 5549 IPv6 Next Hop encoding.

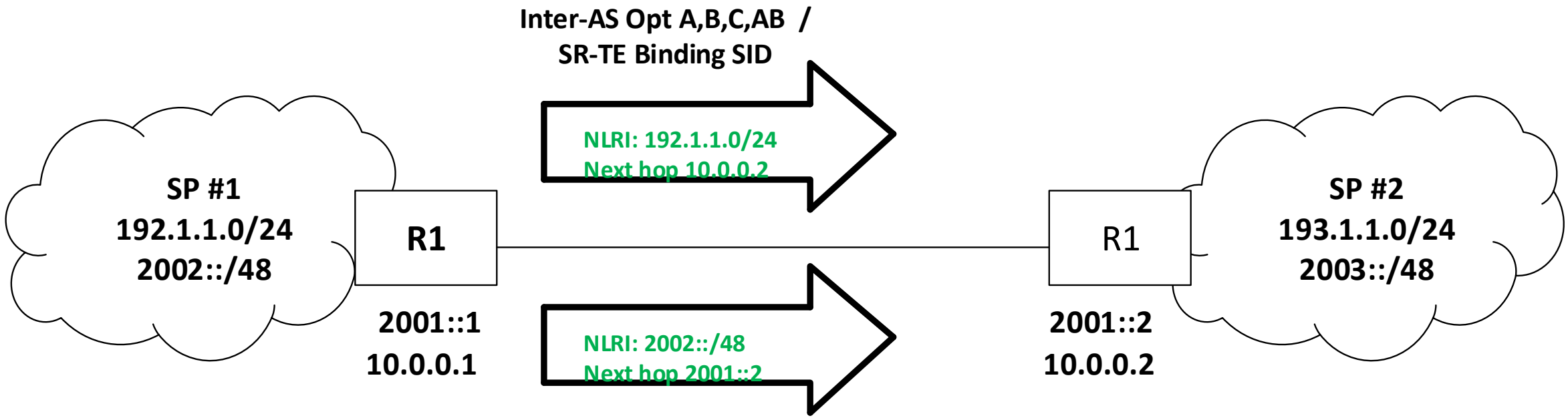
At the PE-CE **eBGP** customer edge now AFI/SAFI 1/1 & 2/1 would now both be carried by a single protocol **IPv6**, with IPv6 Next Hop defined in **RFC 5549 draft update for the IPv6 Next Hop encoding**.

Softwire mesh framework RFC 5565 tunneling v4 over v6 transport data plane using RFC 5549 NH encoding over MPLS LDPv6 or SRv6 core(IP VPN scenario)

In a Service provider Dual stacked customers edge environment over an IPv6 MPLS LDPv6, SR-MPLSv6 or SRv6 IPv6 core, with core PE-RR peering, IPv6 peering PE-RR already carrying IPv6 NLRI - 2/128 (VPN-IPV6) 2/129 (MVPN) would now carry as well 1/128 (VPN-IPV6) 1/129 (MVPN) with **RFC 5549 draft updated IPv6 NH encoding**.

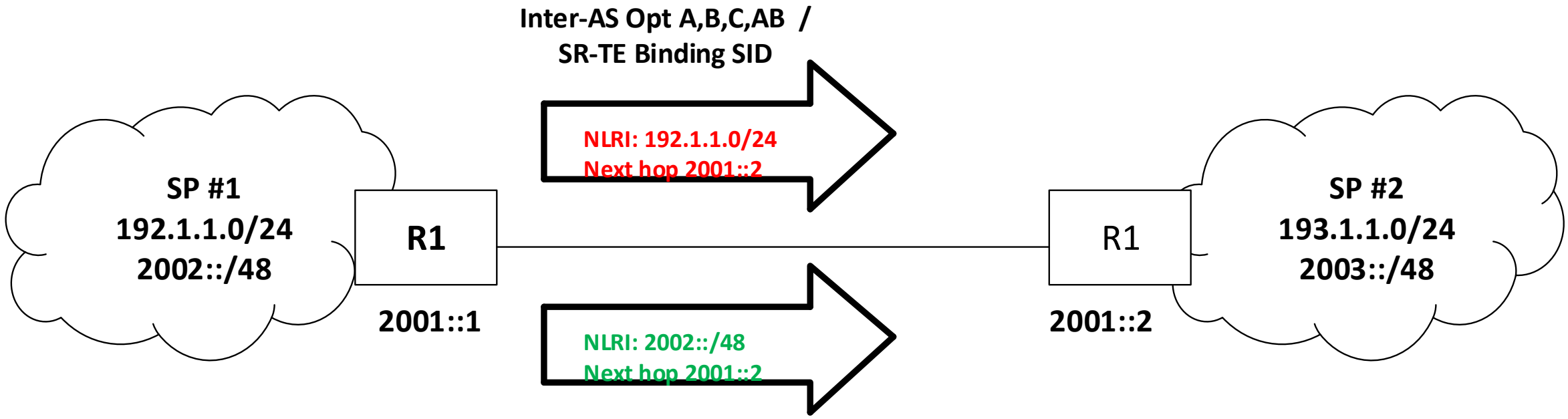
At the PE-CE **eBGP** customer edge now AFI/SAFI 1/1 & 2/1 would now both be carried by the single protocol **IPv6,** with IPv6 Next Hop defined in **RFC 5549 draft update for the IPv6 Next Hop Encoding**.

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



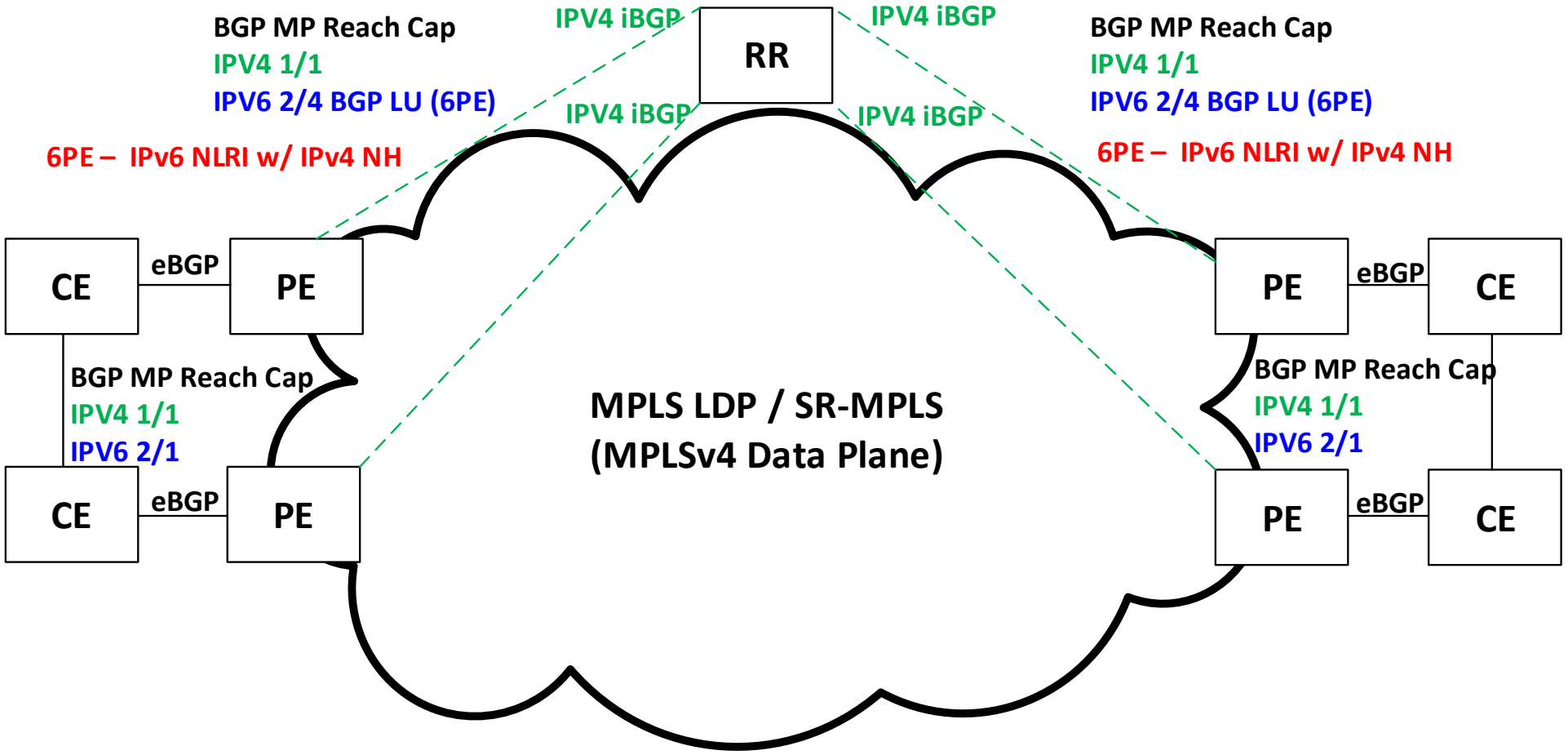
So now with RFC 5549 NH encoding schema of 16 / 32 byte IPv6 next hop both IPv4 & IPv6 NLRI can be advertised using a single IPv6 peer.

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

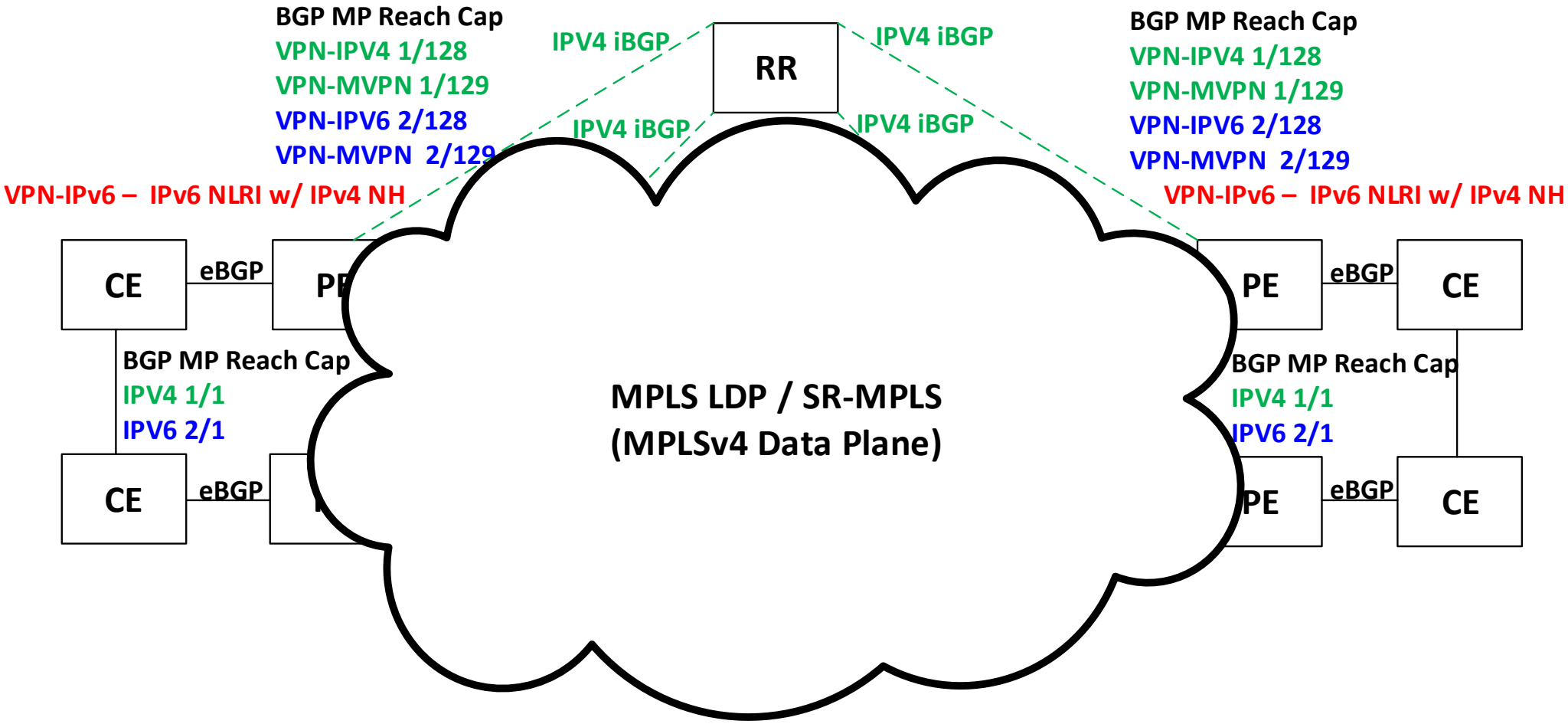


IXP Peering of all SPs would now be able to use a single IPv6 peer per SP adjacency

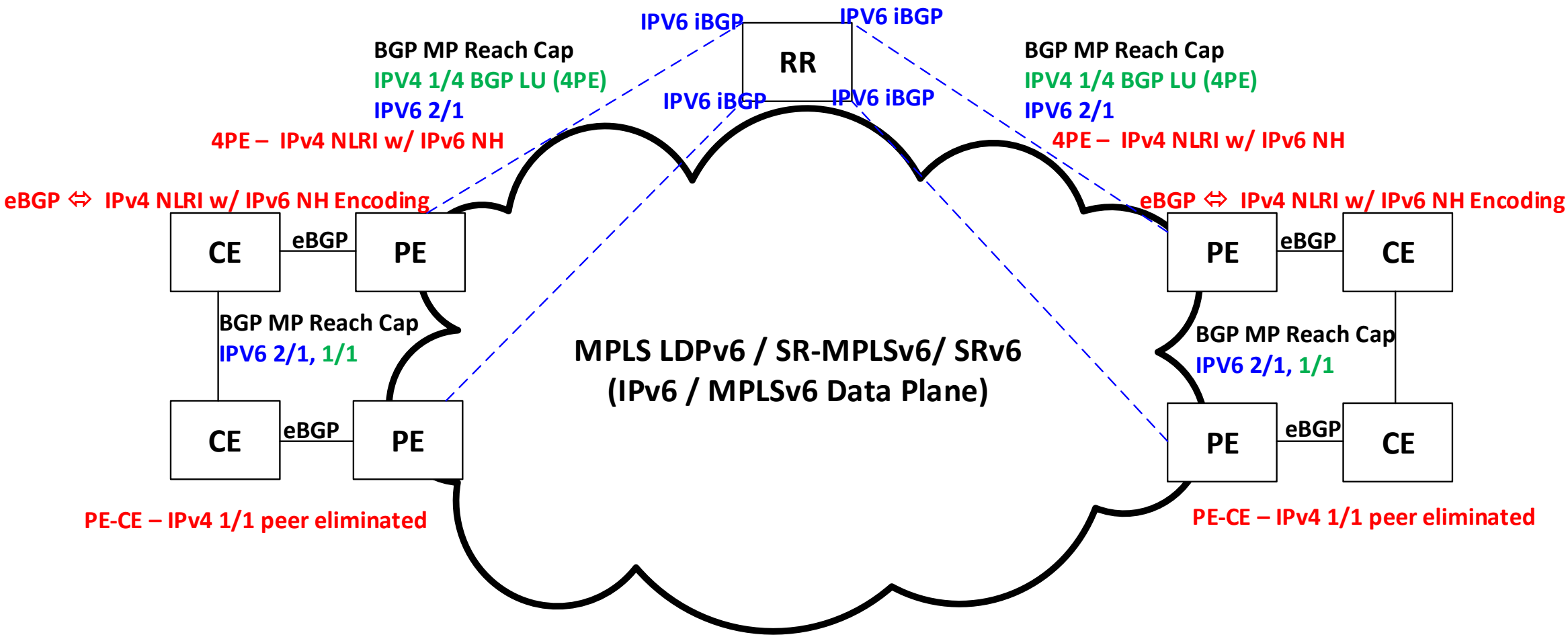
MPLS / SR-MPLS Core– SOFTWIRE MESH FRAMEWORK 6to4 (6PE)



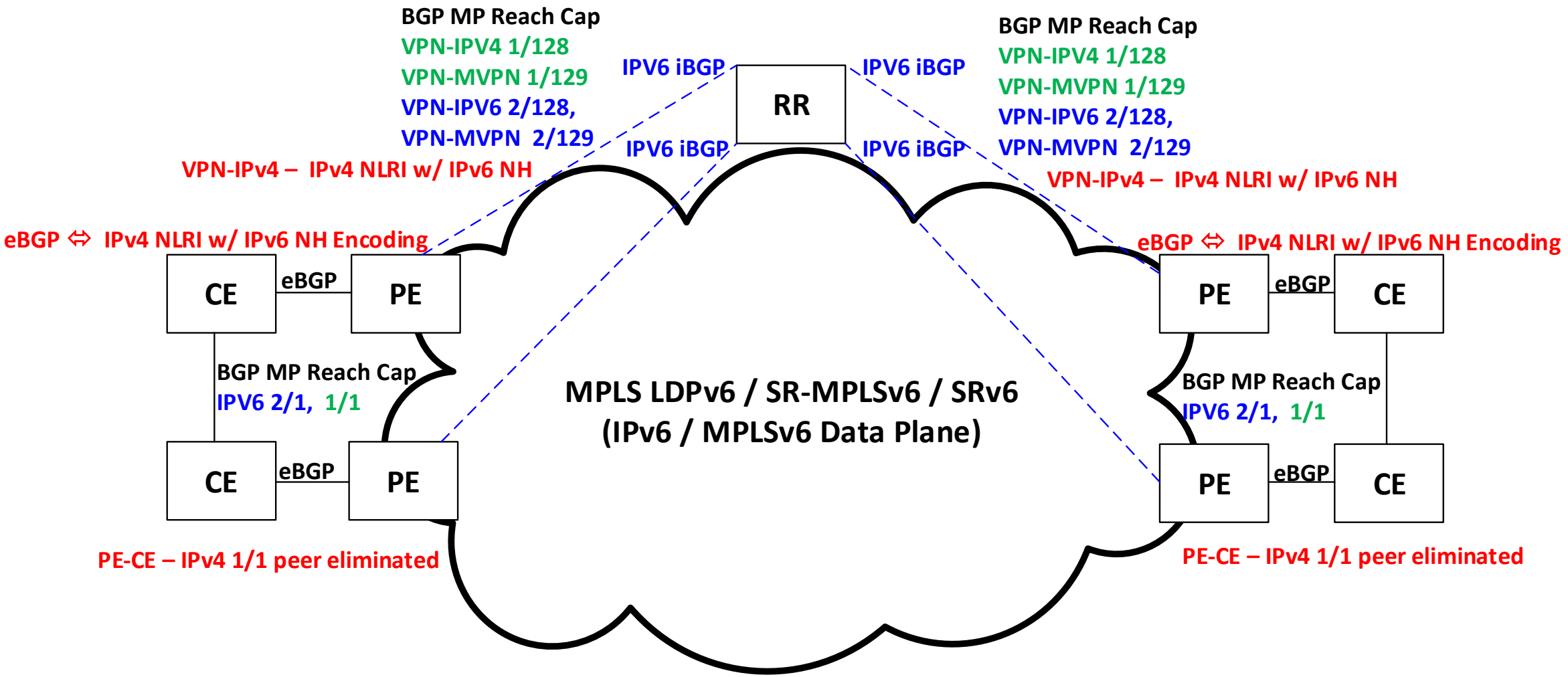
MPLS / SR-MPLS Core – SOFTWIRE MESH FRAMEWORK 6to4 (IP VPN)



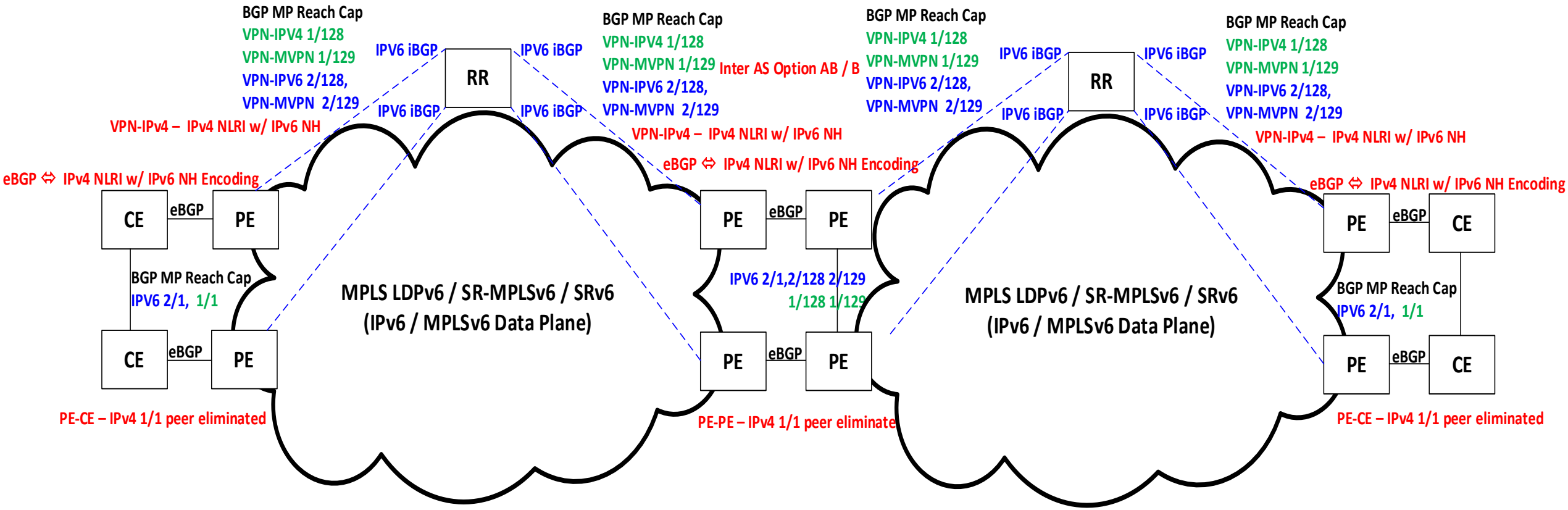
MPLS LDPv6 / SR-MPLSv6 / SRv6 Core – SOFTWIRE MESH FRAMEWORK 4to6 (4PE)



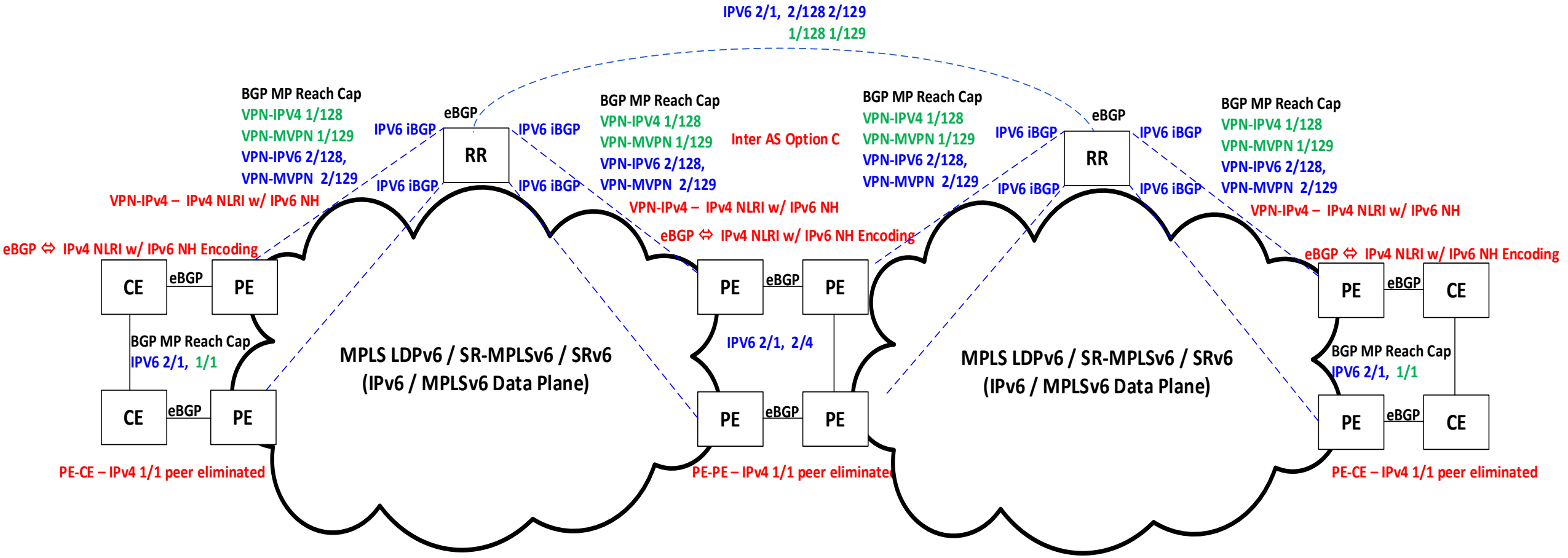
MPLS LDPv6 / SR-MPLSv6 / SRv6 Core – SOFTWIRE MESH FRAMEWORK 4to6 (IP VPN)



MPLS LDPv6 / SR-MPLSv6 / SRv6 Core – SOFTWIRE MESH FRAMEWORK 4to6 (IP VPN)
Inter-AS Option AB / B



MPLS LDPv6 / SR-MPLSv6 / SRv6 Core – SOFTWIRE MESH FRAMEWORK 4to6 (IP VPN)
Inter-AS Option C



Appendix A: IPv4 NLRI IPv6 Next Hop Vendor Testing

Appendix A.1 Router and Switch Vendor Support and Interoperability Test Results

Vendor	Support	Interoperability
Cisco	***	
Juniper	***	
Nokia	***	
Arista	***	
Huawei	***	