Considerations for Protection of SRv6 Networks

draft-liu-rtgwg-srv6-protection-considerations-01

Yisong Liu (China Mobile)(Presenter)

Weiqiang Cheng (China Mobile)

Changwang Lin (New H3C Technologies)

Xuesong Geng (Huawei Technologies)

Yao Liu (ZTE Corp.)

IETF 113

Outline

- Forwarding over SRv6 Network
  - SRv6 BE path
  - SRv6 TE path (with Compressed Segment List encoding)
- Protection of SRv6 Newtork
  - Path Protection
    - Local Protection
    - End-to-End Protection
  - Service Protection
  - Coexistence of Service Protection and Path Protection
- Running Code Status
- Implementation Recommendations
  - SRv6 BE Deployments and Examples
  - SRv6 TE Deployments and Examples

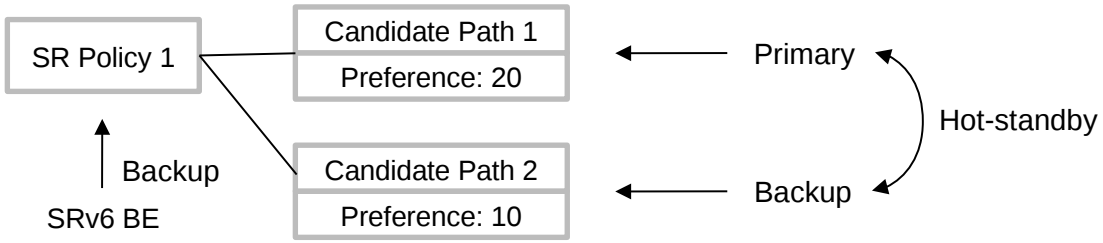
Path Protection

- Local Proctection (performed by the node adjacent to the failed component):

TI-LFA can provide a loop free backup irrespective of the topologies. If G-SRv6 Compressed encoding is enabled, the repair node should try to use G-SIDS to encode the TI-LFA repair path.

- End-to-End Protection (the ingress PE node in charge of the failure recovery):

In SR Policy for SRv6 TE path, the candidate path with second highest preference can be selected as the hot-standby backup. If all the candidate paths fail, SRv6 BE path may be used.



- Liveness Check:

BFD for interface, BFD/S-BFD between neighbors, BFD/S-BFD for SR Policy (at the level of segment list) , Other OAM methods, such as Ping, TWAMP or STAMP, etc.

Service Protection

If egress PE fails, packets should be forwarded to another egress PE of the same service.

- Local Repair:

[I-D.ietf-rtgwg-srv6-egress-protection] provides a method for the adjacent node to encode repair path with Mirror SID.

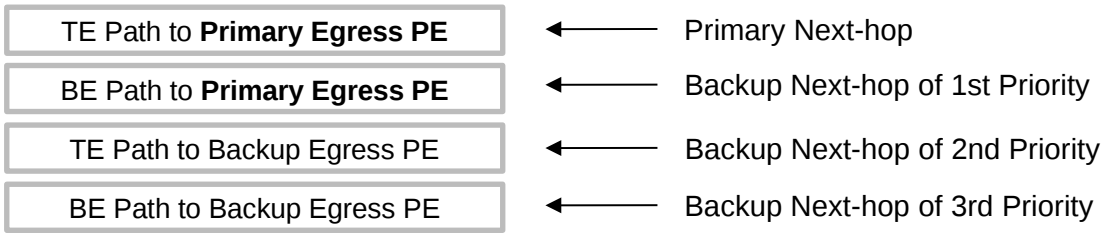
- Ingress Node Switchover:

The ingress node steers the flow to the path belonging to another egress PE node for protection.

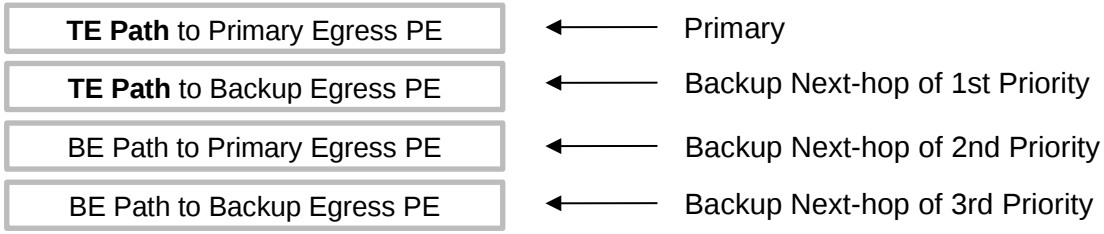
Liveness Check: BFD for Service SID/Locator/PE address

- Coexistence of Service Protection and Path Protection:

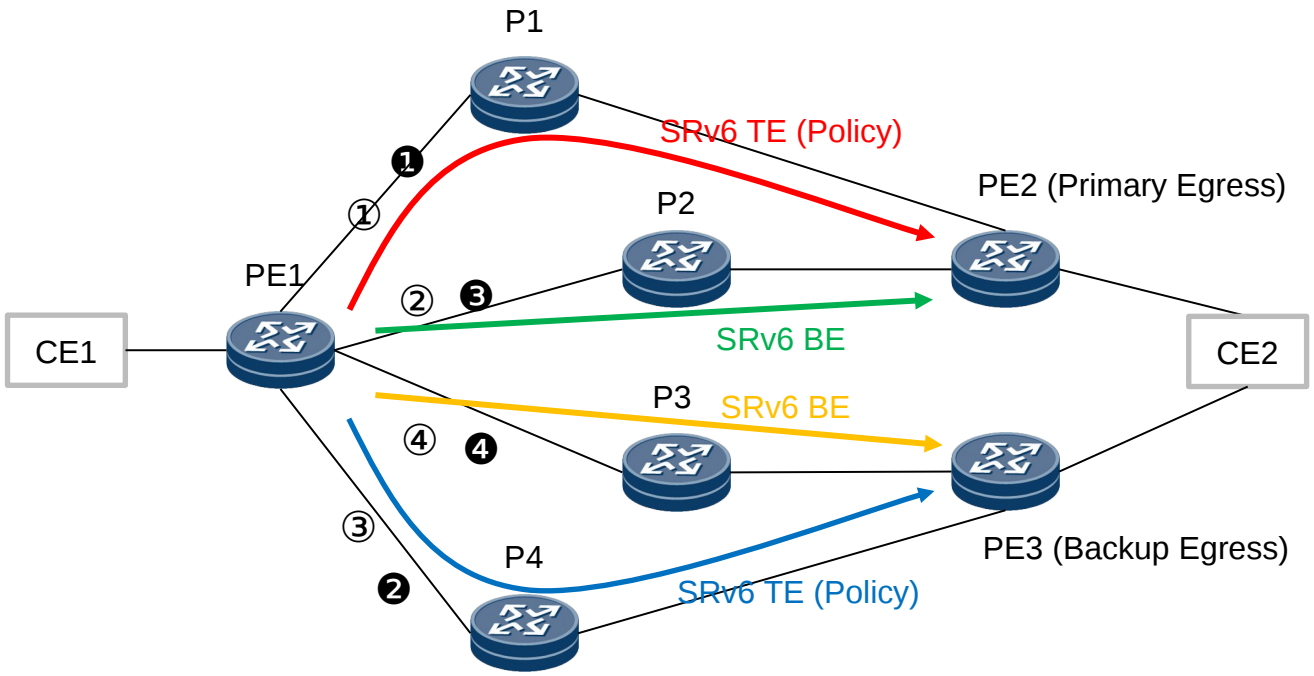
Egress-node-first strategy:



TE-first strategy:



Service Protection Strategy



	Egress-node-first strategy	TE-first strategy
Primary Next-hop:	①	❶
Backup Next-hop of 1st Priority:	②	❷
Backup Next-hop of 2nd Priority:	③	❸
Backup Next-hop of 3rd Priority:	④	❹

Running Code Status

Lab Interop-test Status

Hardware devices and software implementations which have passed G-SRv6 protection interoperability tests hosted by China Mobile in 2021:

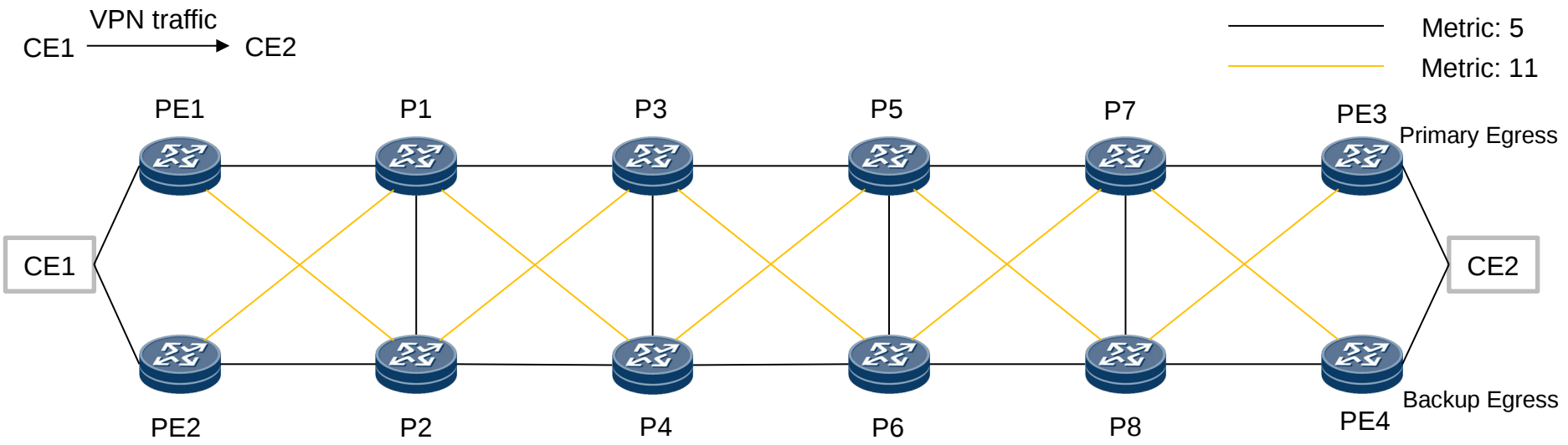
- China Unitechs Unified Controller
- Huawei NE40E and NE5000E
- H3C CR16010H-FA and CR19000
- ZTE M6000-8S Plus and M6000-3S

Deployment Status

Trials of G-SRv6 protection in three province branch networks of China Mobile in 2021:

- Huawei devices with a China Unitechs Unified Controller, Henan Province.
- H3C devices with a China Unitechs Unified Controller, Zhejiang Province.
- ZTE devices with a China Unitechs Unified Controller, Fujian Province.

Reference Topology



SID Assignments:

NodeID:	AX for PEX; BX for PX
Locator:	AA:NodeID::/32
End SID:	Locator:1::
End SID with COC:	Locator:2::
End DT:	Locator:100::
End.X SID:	Locator:NeigborNodeID + F1::
End.X SID with COC :	Locator:NeigborNodeID + F2::

Take PE1 as an example:

Locator:	AA:A1::/32
End SID:	AA:A1:1::
End SID with COC:	AA:A1:2::
End DT:	AA:A1:100::
For PE1->P1:	
End.X SID:	AA:A1:B1F1::
End.X SID with COC:	AA:A1:B1F2::
For PE1->P2:	
End.X SID:	AA:A1:B2F1::
End.X SID with COC:	AA:A1:B2F2::

SR Policy Configurations on PE1:

SR Policy 1 (Strict Path to PE3)

Candidate Path 1 (Preference: 20, **PE1->P1->P3->P5->P7->PE3**)

Segment List: AA:A1:B1F2::, AA:B1:B3F2::, AA:B3:B5F2::, AA:B5:B7F2::, AA:B7:A3F1::

Candidate Path 2 (Preference: 10, **PE1->P2->P4->P6->P8->PE3**)

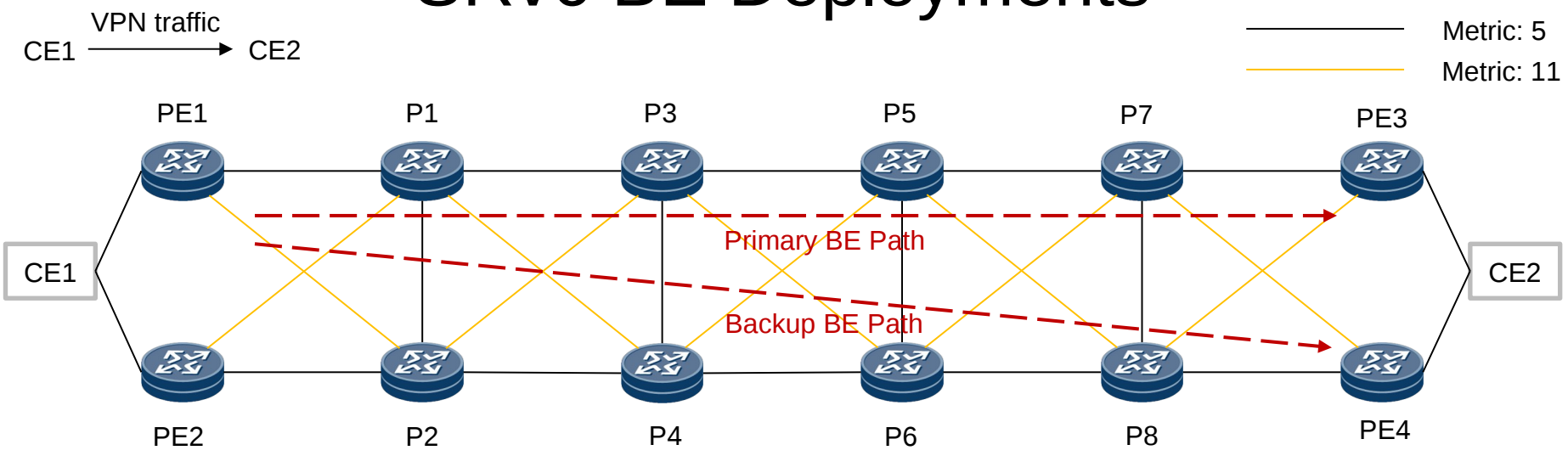
Segment List: AA:A1:B2F2::, AA:B2:B4F2::, AA:B4:B6F2::, AA:B6:B8F2::, AA:B8:A3F1::

SR Policy 2 (Loose Path to PE4)

Candidate Path 1 (Preference: 20, **PE1->…->P4->…->P8->…->PE4**)

Segment List: AA:B4:2::, AA:B8:2::, AA:A4:1::

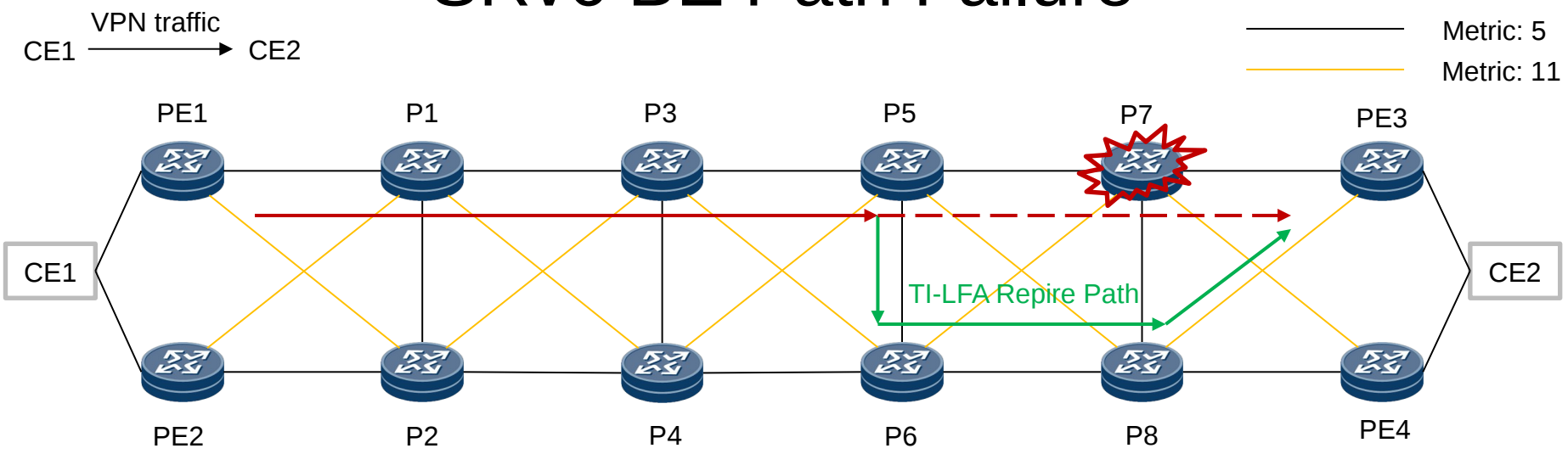
SRv6 BE Deployments



Deployments:

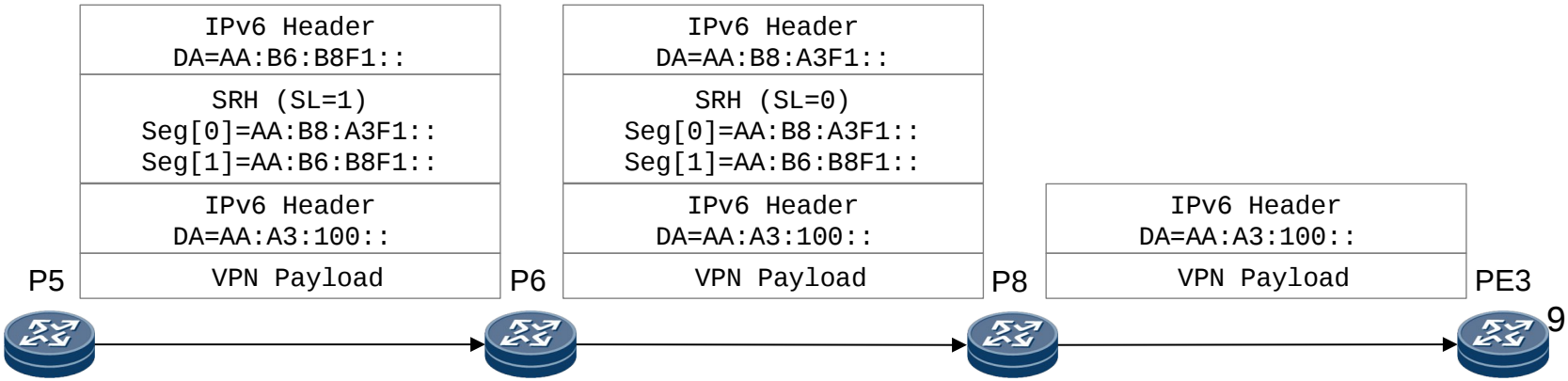
- Next-hop for VPN traffic on PE1:
  - Primary: SRv6 BE path to PE3 ← BFD for locator of PE3
  - Backup: SRv6 BE path to PE4 ← BFD for locator of PE4
- All nodes enable TI-LFA for local protection, BFD for links and neighbors.

SRv6 BE Path Failure

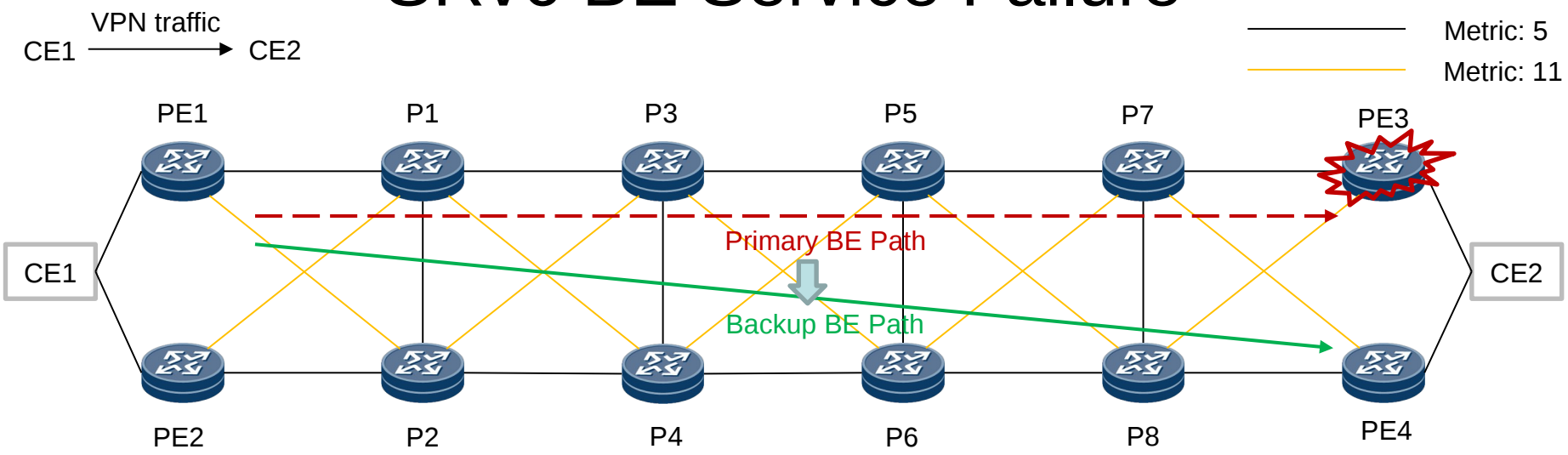


When P7 fails:

- The fail-timer of BFD from P5 to P7 expires, so P5 perceives the failure.
- The TI-LFA repair path is P5->P6->P8->PE3.
- P5 encapsulates the packet in an outer IPv6 Header with SRH carrying the TI-LFA repair list.



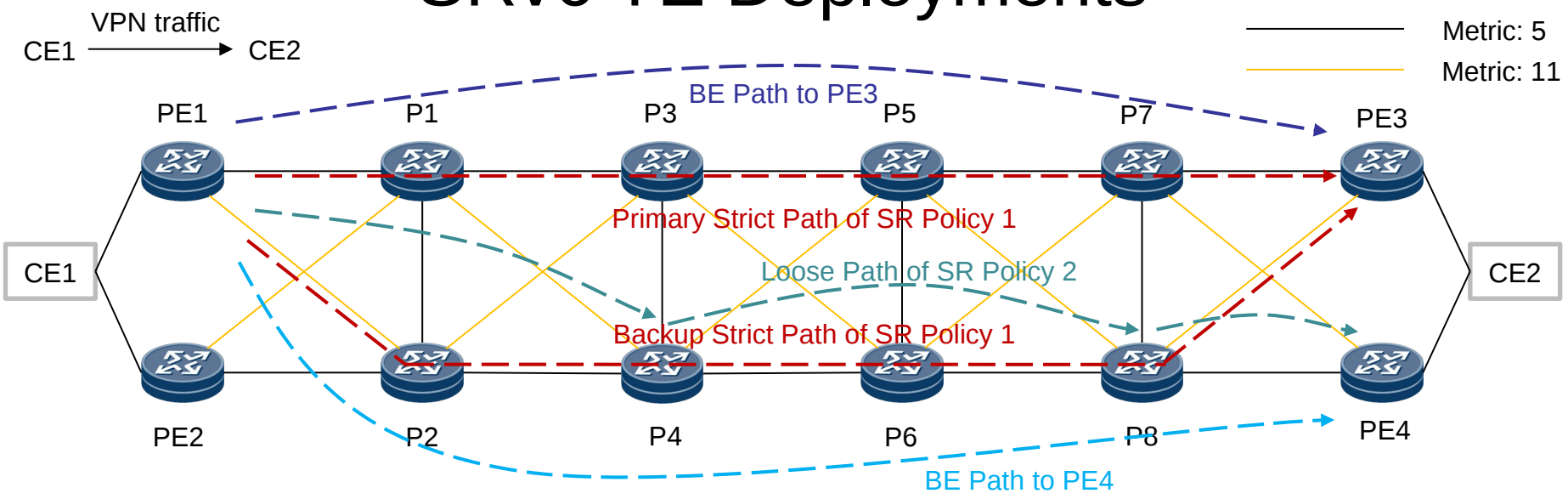
SRv6 BE Service Failure



When PE3 fails:

- The BFD session from PE1 to locator AA:A3::/32 is down.
- PE1 triggers the switchover to the SRv6 BE path to PE4.

SRv6 TE Deployments



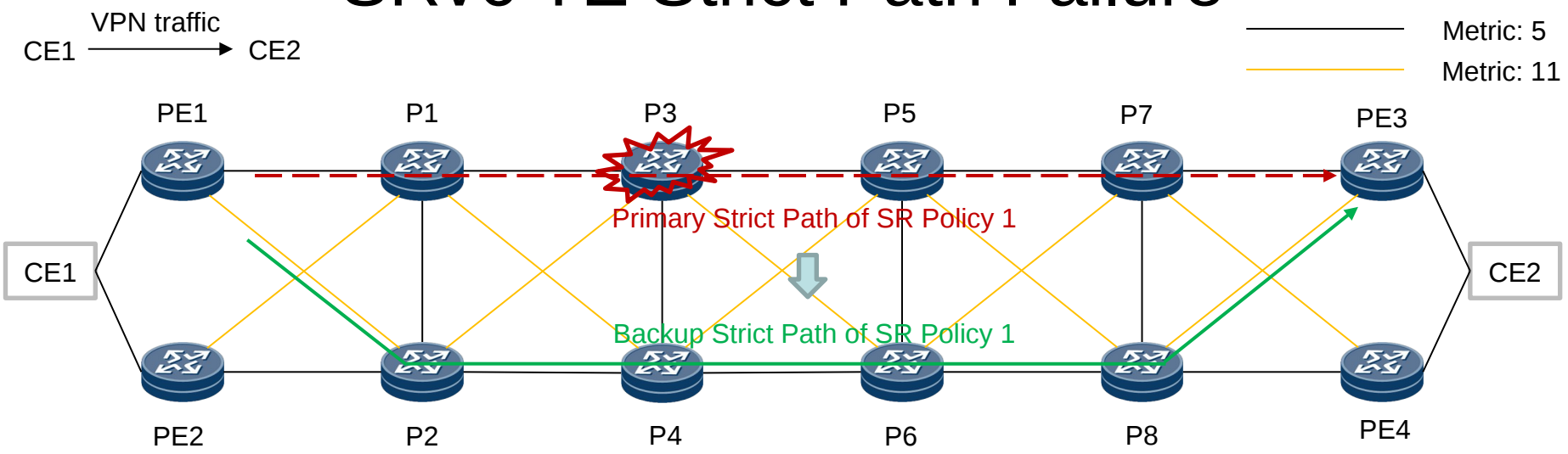
Deployments:

- Next-hop for VPN traffic on PE1 (using egress-node-first strategy):
  - Primary: SR Policy 1 (Strict TE path)
    - Primary Candidate Path: PE1->P1->P3->P5->P7->PE3
    - Backup Candidate Path: PE1->P2->P4->P6->P8->PE3
  - Backup (1st priority): SRv6 BE path to PE3
  - Backup (2nd priority): SR Policy 2 (Loose TE path:)
    - Candidate Path: PE1->…->P4->…->P8->…->PE4
  - Backup (3rd priority): SRv6 BE path to PE4

← BFD for SR Policy 1
← BFD Reverse Path: PE3->P7->P5->P3->P1->PE1
← BFD Reverse Path: PE3->P8->P6->P4->P2->PE1
← BFD for locator of PE3
← BFD for SR Policy 2
← BFD for locator of PE4

- All nodes enable TI-LFA for local protection, BFD for links and neighbors.

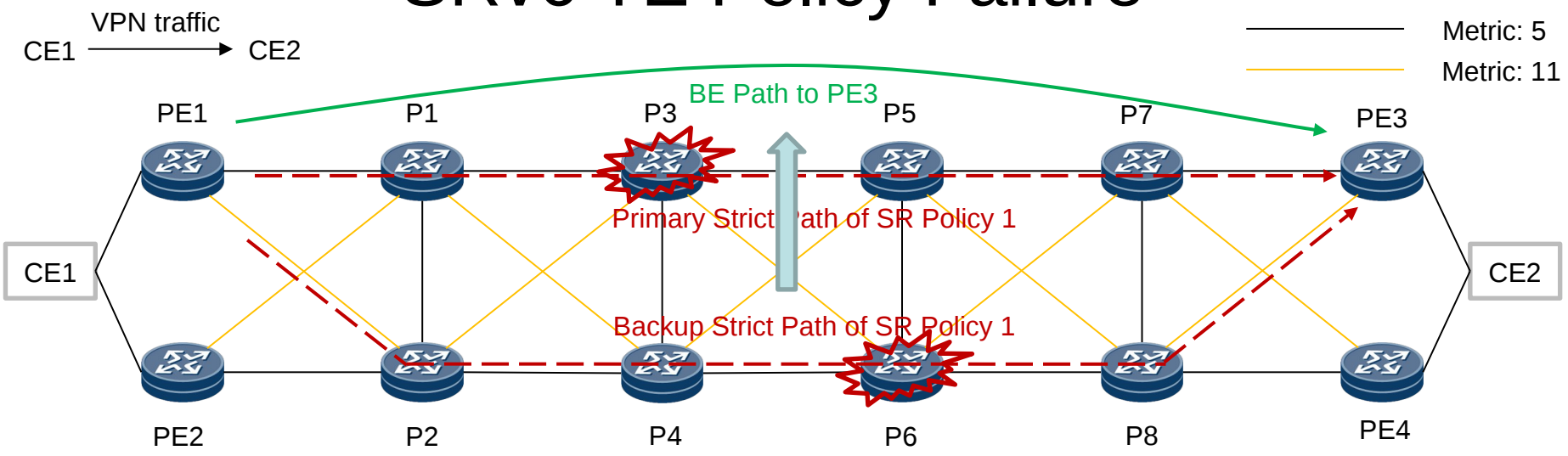
SRv6 TE Strict Path Failure



When P3 fails:

- The BFD session of the segment list in the primary candidate path of SR Policy 1 is down.
- PE1 triggers the switchover to the backup candidate path of SR Policy 1.

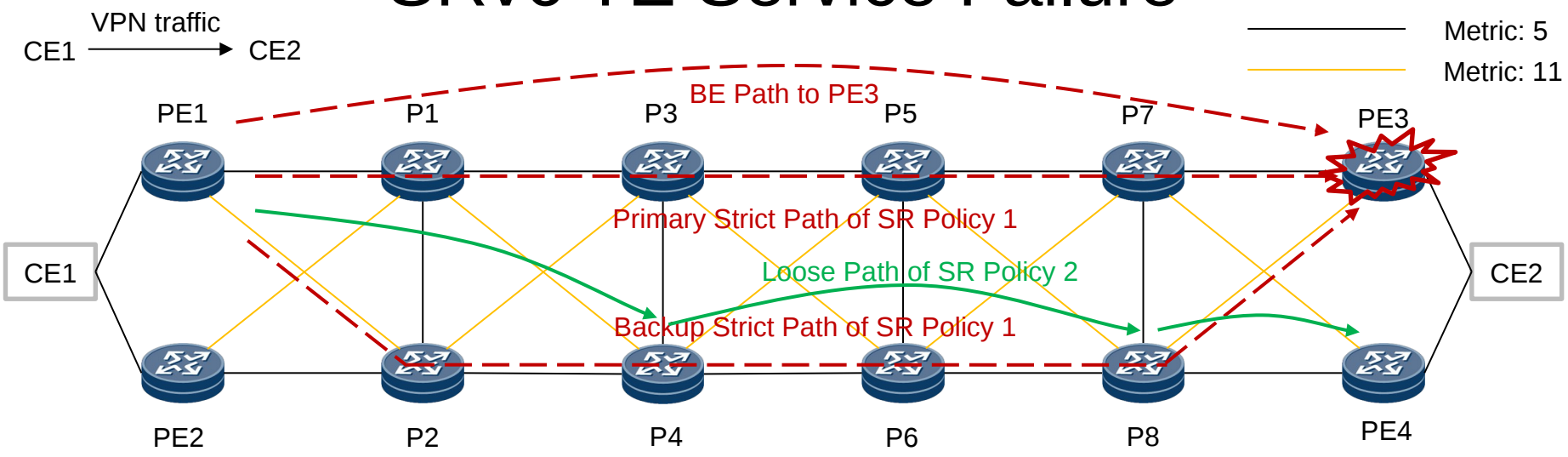
SRv6 TE Policy Failure



Before the recovery of P3, P6 also fails:

- The BFD session of the segment list in the backup candidate path of SR Policy 1 is also down.
- PE1 triggers the switchover to the 1st priority backup next-hop which is the SRv6 BE path to PE3.

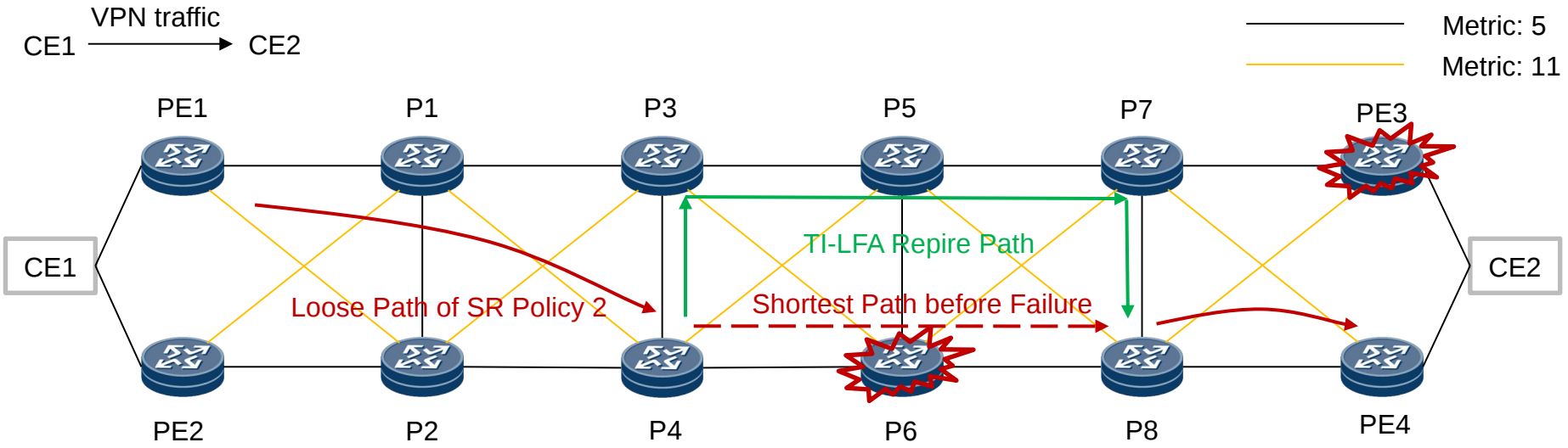
SRv6 TE Service Failure



When PE3 fails:

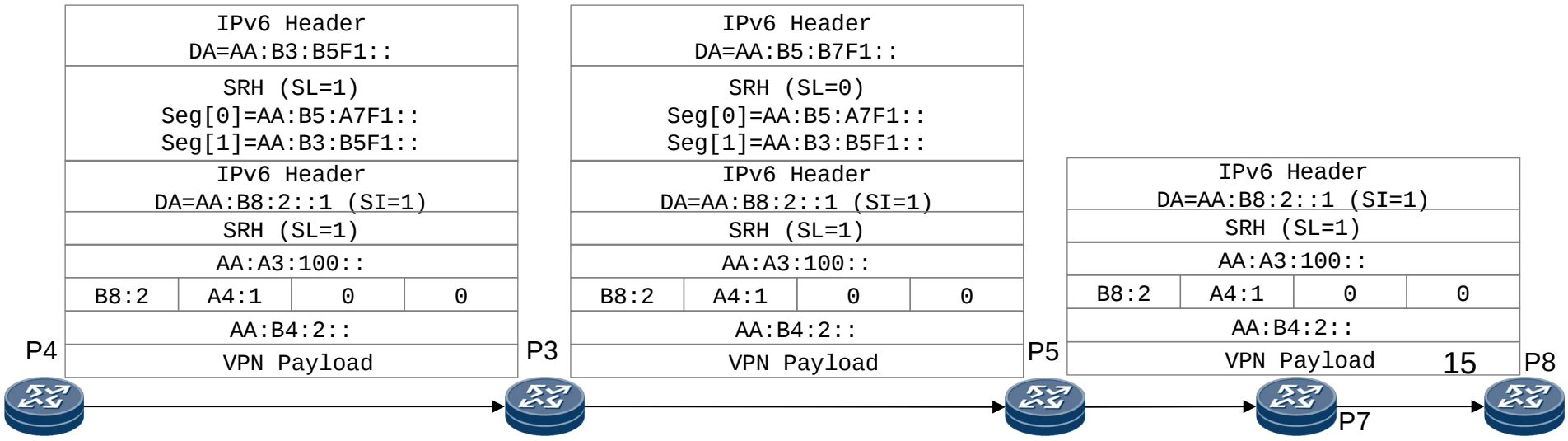
- Both the BFD sessions of SR Policy 1 and locator AA:A3::/32 are down, which means the primary next-hop and the 1st priority backup next-hop are both down.
- PE1 triggers the switchover to the 2nd priority backup next-hop which is the SRv6 TE loose path to PE4.

SRv6 TE Loose Path Failure(Non-Endpoint)

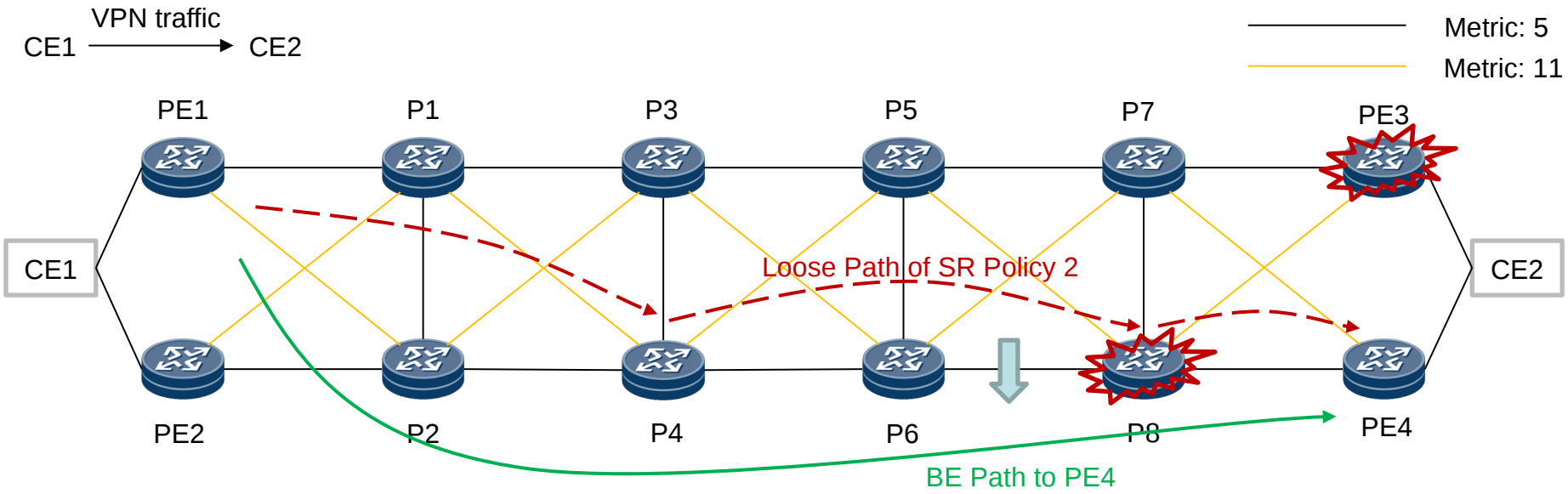


Before the recovery of PE3, assume that P6 fails:

- The fail-timer of BFD from P4 to P6 expires, so P4 perceives the failure.
- The TI-LFA repair path is P4->P3->P5->P7->P8.
- P4 encapsulates the packet in an outer IPv6 Header with SRH carrying the TI-LFA repair list.



SRv6 TE Loose Path Failure (Endpoint)



Before the recovery of PE3, assume that P8 fails:

- The TI-LFA on P6 does not work, since the failed endpoint node P8 is the destination.
- The BFD session of SR Policy 2 is down.
- PE1 triggers the switchover to the 3rd priority backup next-hop which is the SRv6 BE path to PE4.

Next Step

- Any questions or comments are Welcomed
- Seeking for feedback from WG