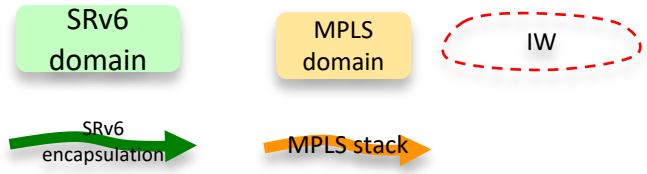
IETF 115
Nov 2022

draft-agrawal-spring-srv6-mpls-interworking-10

Swadesh Agrawal, Cisco Systems (swaagraw@cisco.com) - Presenter
Zafar Ali, Cisco Systems (zali@cisco.com)
Clarence Filsfils, Cisco Systems (cfilsfil@cisco.com)
Daniel Voyer, Bell Canada (daniel.voyer@bell.ca)
Gaurav dawra, LinkedIn (gdawra.ietf@gmail.com)
Zhenbin Li, Huawei (lizhenbin@huawei.com)

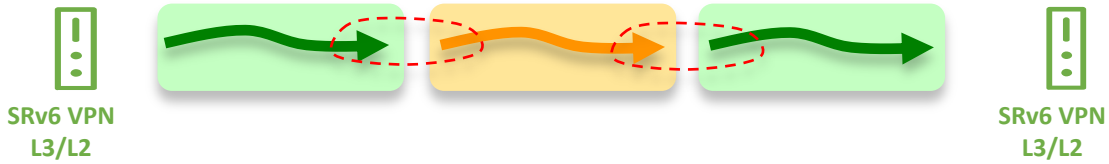
Interworking (IW) scenarios



1. **Transport IW** - L3/L2 service continuity over a different intermediate transport

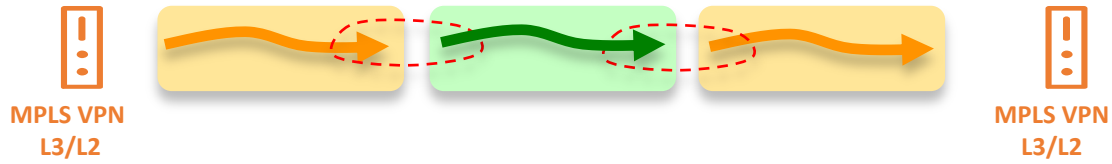
1.1 SRv6 over SR-MPLS-IPv4 (6oM)

Forward SRv6 encapsulated traffic destined to egress PE over MPLS domain.

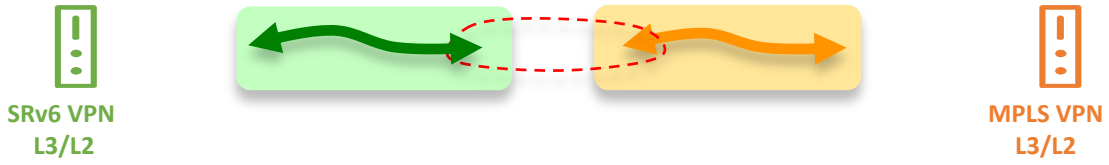


1.2 SR-MPLS-IPv4 over SRv6 (Mo6)

Forward encapsulated label stack to egress PE over SRv6 domain.



2. **Service IW** - L3/L2 service signaling discontinuity i.e. SRv6 service SID based PE interworks with BGP MPLS based PE for service connectivity.



End.DTM SRv6 SID behavior

- The "Endpoint with decapsulation and MPLS table lookup" behavior.
- This behavior is executed on IW router (SRv6 domain to MPLS domain).

Pseudo Code

When N receives a packet destined to S and S is a local End.DTM SID, N does:

S01. When an SRH is processed {

S02. If (Segments Left != 0) {

S03. Send an ICMP Parameter Problem to the Source Address, Code 0 (Erroneous header field encountered), Pointer set to the Segments Left field, interrupt packet processing and discard the packet.

S04. }

S05. Proceed to process the next header in the packet

S06. } When processing the Upper-layer header of a packet matching a FIB entry locally instantiated as an End.DTM SID, N does:

S01. If (Upper-Layer Header type == 137(MPLS)) {

S02. Remove the outer IPv6 Header with all its extension headers

S03. Set the packet's associated FIB table to T

S04. Submit the packet to the MPLS FIB lookup for transmission according to the lookup result.

S05. } Else {

S06. Process as per [ietf-spring-srv6-network-programming] section 4.1.1

S07. }

End.DPM SRv6 SID behavior

- The "Endpoint with decapsulation and MPLS label push" behavior.
- This behavior is executed on IW router (SRv6 domain to MPLS domain).

Pseudo Code

When N receives a packet destined to S and S is a local End.DPM SID, N does:

S01. When an SRH is processed {

S02. If (Segments Left != 0) {

S03. Send an ICMP Parameter Problem to the Source Address, Code 0 (Erroneous header field encountered), Pointer set to the Segments Left field, interrupt packet processing and discard the packet.

S04. }

S05. Proceed to process the next header in the packet

S06. } When processing the Upper-layer header of a packet matching a FIB entry locally instantiated as an End.DTM SID, N does:

S01. Remove the outer IPv6 Header with all its extension headers

S02. Push the MPLS label stack associated with S

S03. Submit the packet to the MPLS engine for transmission

SRv6 Headend Behaviors

- H.Encaps.M

H.Encaps applied to MPLS label stack. The H.Encaps.M behavior encapsulates a received MPLS Label stack [RFC3032] packet in an IPv6 header with an SRH. Together MPLS label stack and its payload becomes the payload of the new IPv6 packet. The Next Header field of the SRH MUST be set to 137 [RFC4023].

- H.Encaps.M.Red

H.Encaps.Red applied to MPLS label stack. The H.Encaps.M.Red behavior is an optimization of the H.Encaps.M behavior.

Interconnecting Binding SIDs

- Binding Segment (BSID) is bound to SR policy [RFC8402].
- An SR-MPLS label can be bound to an SRv6 Policy and an SRv6 SID can be bound to an SR-MPLS Policy.
- These BSIDs in segment list of SR policy on headend represent intermediate domain of different dataplane type and act as interconnecting BSIDs.

Transport IW

The draft enhances two well-known solutions to create IW state in network:

- SR-PCE (SDN Controller) procedure provides a path that satisfies the intent (e.g. low latency), across multiple domains. SR PCE detects the data plane discontinuity.
- BGP Inter-Domain routing procedure advertises PE locator/IPv4 Loopback address LSP for best effort end to end connectivity.

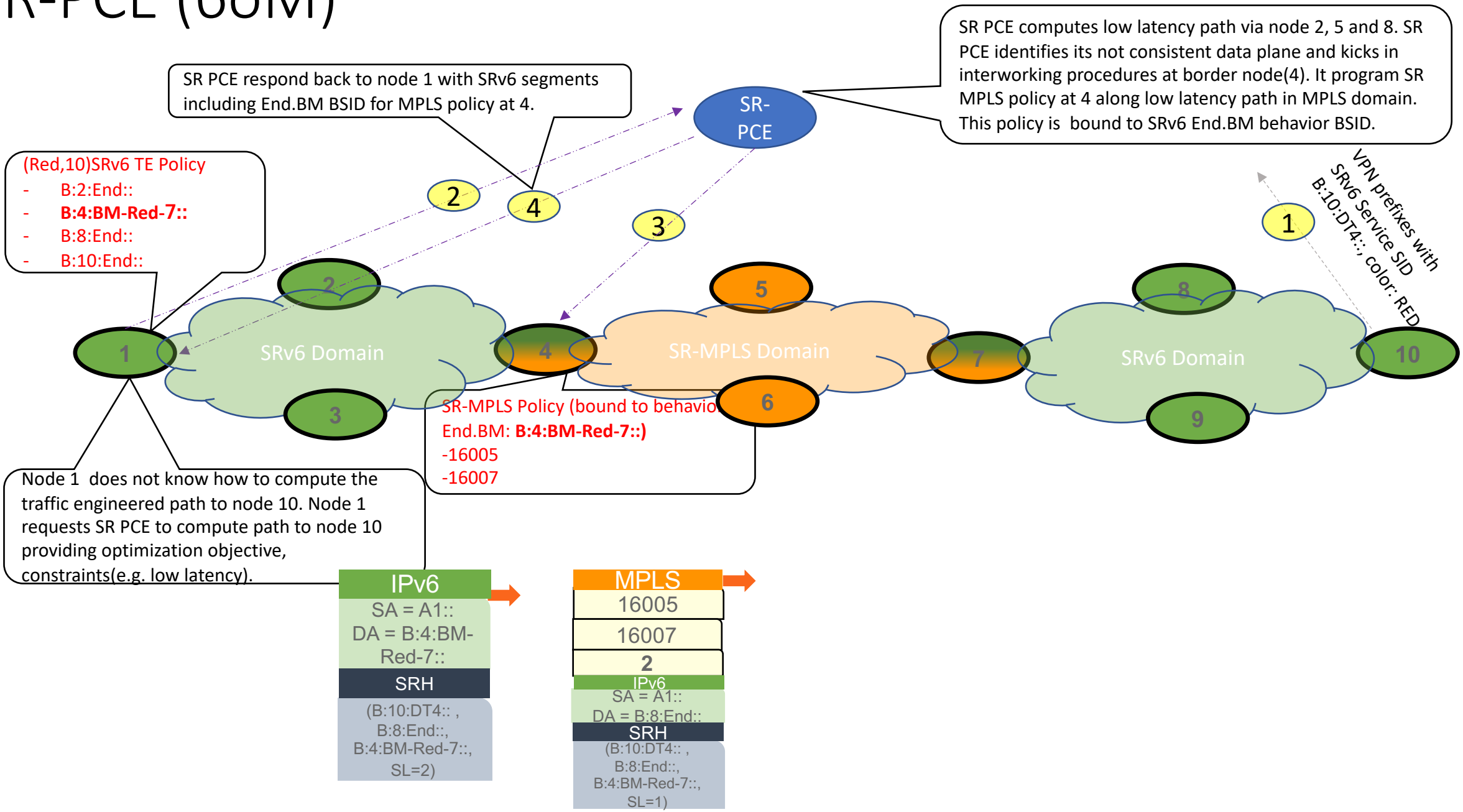
Legends

SRv6 Capable (F function on node B:k:F::, Router-id: Ak::)

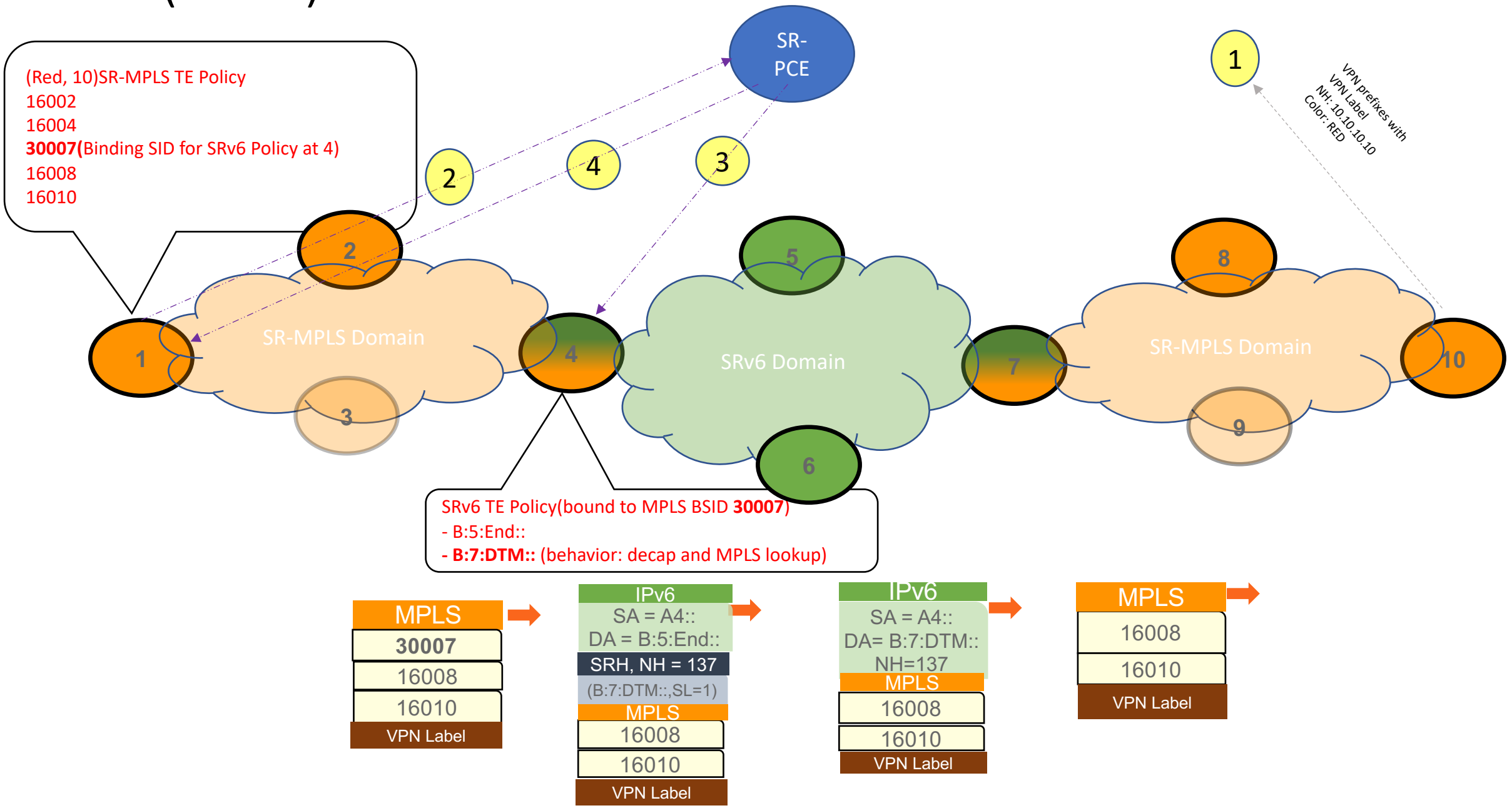
SR-MPLS IPv4 Node (Prefix SID label: 1600k, Router-id: 1.1.1.k)

SR-MPLS IPv4 and SRv6 Capable Node

SR-PCE (6oM)

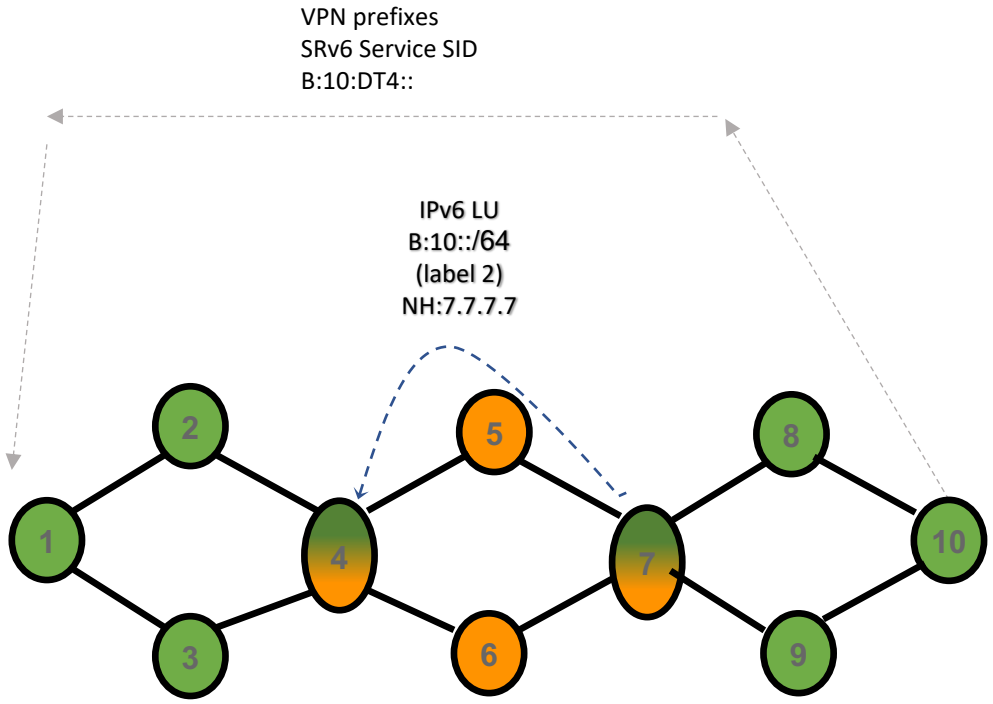

SR PCE respond back to node 1 with SRv6 segments including End.BM BSID for MPLS policy at 4.
SR PCE computes low latency path via node 2, 5 and 8. SR PCE identifies its not consistent data plane and kicks in interworking procedures at border node(4). It program SR MPLS policy at 4 along low latency path in MPLS domain. This policy is bound to SRv6 End.BM behavior BSID.
SR-PCE
(Red,10)SRv6 TE Policy
- B:2:End::
- B:4:BM-Red-7::
- B:8:End::
- B:10:End::
VPN prefixes with SRv6 Service SID B:10:DT4::, color: RED
SRv6 Domain
SR-MPLS Domain
SRv6 Domain
SR-MPLS Policy (bound to behavio
End.BM: B:4:BM-Red-7::)
-16005
-16007
Node 1 does not know how to compute the traffic engineered path to node 10. Node 1 requests SR PCE to compute path to node 10 providing optimization objective, constraints(e.g. low latency).
IPv6
SA = A1::
DA = B:4:BM-Red-7::
SRH
(B:10:DT4:: , B:8:End::, B:4:BM-Red-7::, SL=2)
MPLS
16005
16007
2
IPv6
SA = A1::
DA = B:8:End::
SRH
(B:10:DT4:: , B:8:End::, B:4:BM-Red-7::, SL=1)

SR-PCE (Mo6)


(Red, 10)SR-MPLS TE Policy
16002
16004
30007(Binding SID for SRv6 Policy at 4)
16008
16010
SR-PCE
1
VPN prefixes with
VPN Label
NH: 10.10.10.10
Color: RED
2
4
3
SR-MPLS Domain
SRv6 Domain
SR-MPLS Domain
5
8
7
10
9
6
SRv6 TE Policy(bound to MPLS BSID 30007)
- B:5:End::
- B:7:DTM:: (behavior: decap and MPLS lookup)
MPLS
30007
16008
16010
VPN Label
IPv6
SA = A4::
DA = B:5:End::
SRH, NH = 137
(B:7:DTM::,SL=1)
MPLS
16008
16010
VPN Label
IPv6
SA = A4::
DA= B:7:DTM::
NH=137
MPLS
16008
16010
VPN Label
MPLS
16008
16010
VPN Label

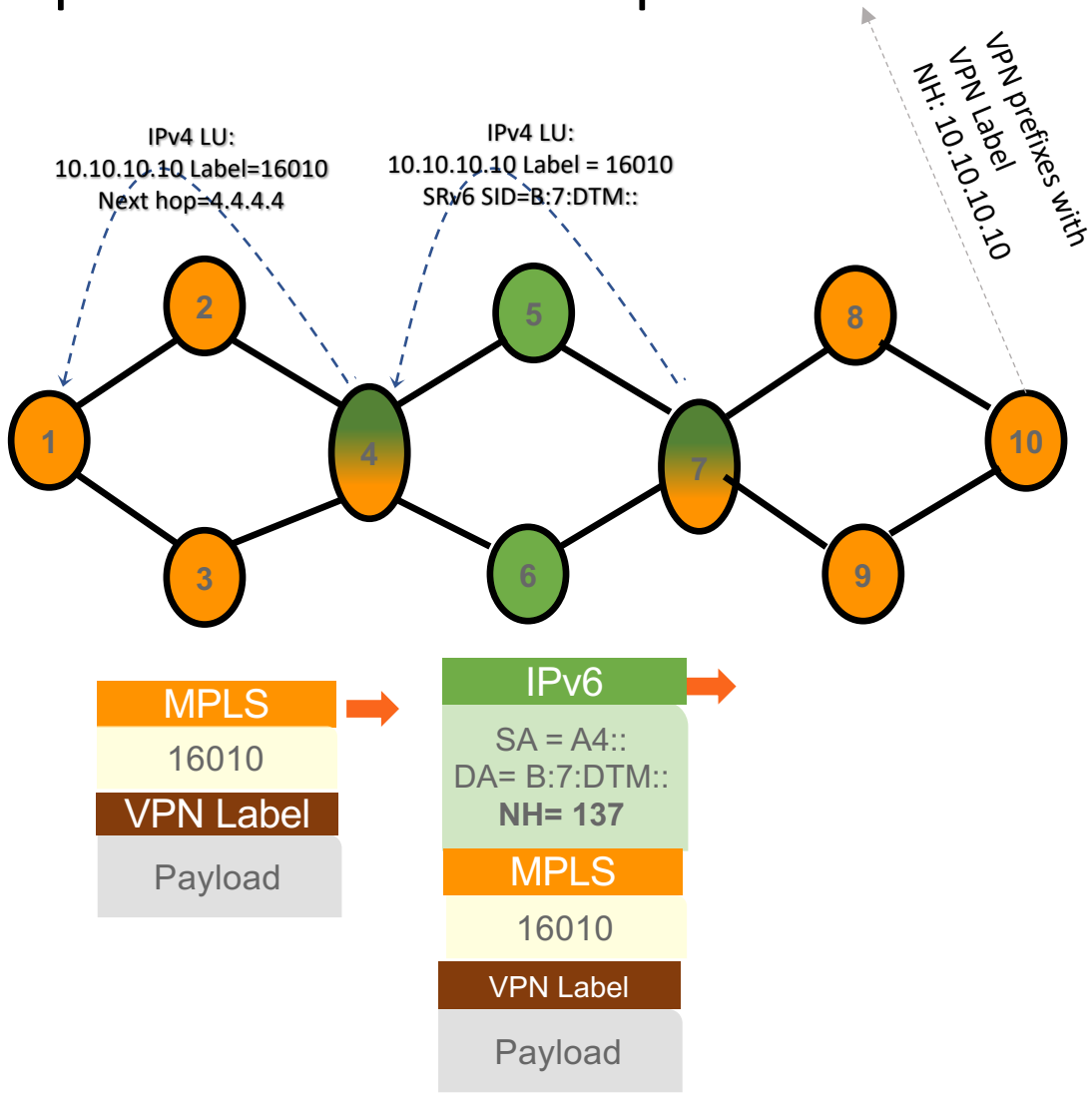
BGP (6oM)

- Advertise PE locators i.e. node 10.
- Its classic 6PE on IW nodes (4 & 7) over SR-MPLS-IPv4 domain. (RFC 4798)
- Leak locator in left domain IGP or advertise locator to ingress PE (node 1) in IPv6 BGP with SRv6 SID of node 4 End behavior.



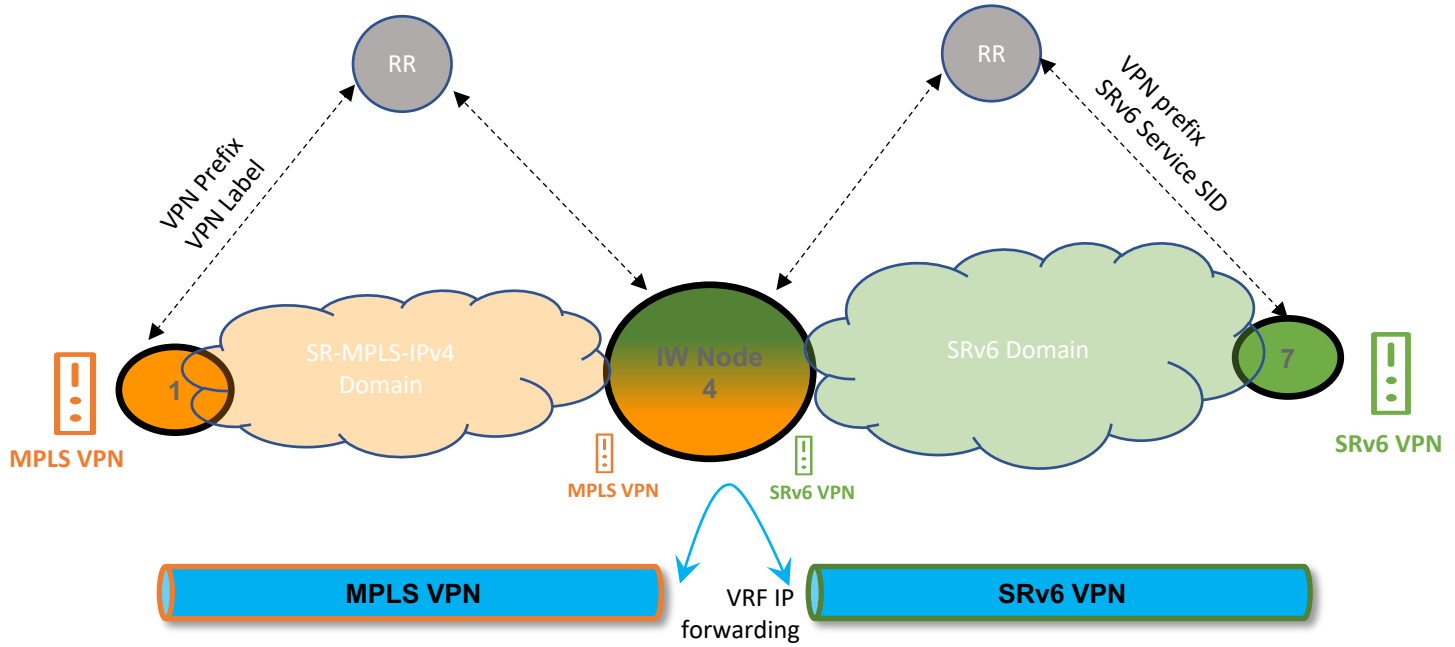
BGP (Mo6): BGP LU LSP IPv6 encaped to next hop

- Existing BGP 3017 label cross-connect on border routers for each PE IPv4 loopback address
- The lookups at the ingress border router are based on BGP 3107 label as usual
- Just the SR-MPLS IPv4 LSP to next hop is replaced by an IPv6 tunnel with DA = SRv6 SID associated with End.DTM behavior of Egress IW node i.e. node 7.
- Ingress border router forwarding perform 3107 label swap and H.Encaps.M with DA = SRv6 SID associated with DTM behavior



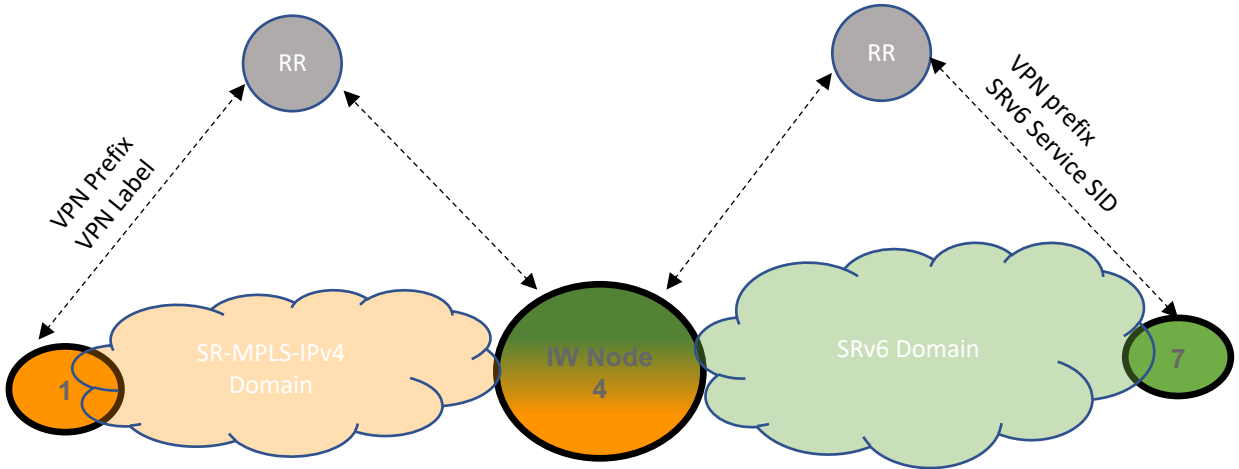
Mto6 <-> 6toM Service IW: Gateway solution

Gateway is router which supports both BGP SRv6 based L2/L3 services and BGP MPLS based L2/L3 services for a service instance (e.g. L3 VRF). It terminates service encapsulation and perform L2/ L3 destination lookup in service instance. Lookup result in re-encapsulation with service information of destination domain.



Mto6 <-> 6toM Service IW: Translation option

This is like Inter-as option B procedures described in [RFC4364] just that service label cross-connect on ASBR router is replaced with service label to SRv6 service SID or vice versa translation on IW node.



Translation	action
VPN label	H.Encap: B:7:DT4::
B:4:DPM_PE1VRF: :	Push VPN label and LSP to PE1

Summary

- The initial version was posted in October 2018;
- The draft describes the data plane and the associated control plane procedures.
- For data plane, new End.DTM and End.DPM behaviors are defined.
- Interconnecting Binding SIDs usage to traverse heterogenous data plane.
- For control plane, both SR-PCE based and BGP based solutions are detailed.

Updates from IETF 113

- Moved BGP protocol extensions in separate draft (draft-agrawal-bess-bgp-srv6-mpls-interworking-00) in BESS WG.
- This is done to independently state BGP protocol extensions, future applicability of them for other use cases and review by BESS.
- Expanded on Service Interworking using translation between SRv6 service SID and VPN label

Next Steps

The authors request review and adoption of the draft.