draft-agrawal-spring-srv6-mpls-interworking-05

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)
Summary

• The draft describes solution for the SRv6 and MPLS interworking.
• The initial version was posted in October 2018; the rev 5 provides the additional details.
• The draft describes the data plane and the associated control plane procedures.
• For data plane, End.DTM and decap variant End.BM behaviors are used.
• For control plane, both SR-PCE based and BGP based solutions are detailed.
Interworking (IW) scenarios

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

   1.1 **SRv6 over SR-MPLS-IPv4 (6oM)**
   Tunnel traffic destined to SRv6 Service SID bound to SRv6 locator of egress PE over SR-MPLS-IPv4 domain.

   1.2 **SR-MPLS-IPv4 over SRv6 (Mo6)**
   Tunnel MPLS LSP bound to IPv4 loopback address of egress PE over SRv6 C domain.

2. **Service IW** - Service discontinuity over a different intermediate transport i.e. BGP SRv6 VPN PE interworking with BGP MPLS VPN PE for L3/L2 service connectivity.
End.DTM SRv6 SID behavior

- The "Endpoint with decapsulation and MPLS table lookup" behavior.
- This behavior is executed on IW routers between the SRv6 and 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. }
SRv6 Headend Behaviors

• H.Encaps.M


• H.Encaps.M.Red

The draft enhances two well-known solutions to provide PE locator/IPv4 PE loopback LSP tunneling:

• 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.

(RED,10)SRv6 TE Policy
- B:2:End::
- B:4:BM-Red-7::
- B:8:End::
- B:10:End::

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 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.

VPN prefixes with SRv6 Service SID
B:10:DT4::, color: RED
SR-PCE (Mo6)

(Red, 10)SR-MPLS TE Policy
16002
16004
30007(Binding SID for SRv6 Policy at 4)
16008
16010

SRv6 TE Policy (bound to MPLS BSID 30007)
- B:5:End::
- B:7:DTM:: (behavior: decap and MPLS lookup)

VPN prefixes with VPN Label NH: 10.10.10.10
Color: RED

MPLS
30007
16008
16010
VPN Label

IPv6
SA = A4::
DA = B:5:End::
SRH, NH = 137
IPv6 (B:7:DTM::, SL=1)
MPLS
SA = A4::
DA = B:7:DTM::
NH=137
MPLS
16008
16010
VPN Label

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)

Intuitive solution for an MPLS-minded operator

- 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
- Draft introduces a new TLV called "SRv6 label route tunnel" TLV of the BGP Prefix-SID Attribute to signal SRv6 SID of behavior DTM that tunnel MPLS packet with label in NLRI at the top of its label stack through SRv6/IPv6 domain.
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.
Next Steps

• The authors would like the WG to review and adopt the document.