Workgroup: mpls
Internet-Draft:
draft-zzhang-mpls-mldp-rsvp-p2mp-srv6-01
Published: 30 December 2022
Intended Status: Standards Track
Expires: 3 July 2023
Authors: Z. Zhang P. Beeram R. Parekh
Juniper Networks Juniper Networks Cisco
I. Wijnands R. Chen
Arrcus ZTE

mLDP/RSVP-TE P2MP Tunnel with SRv6 SID

Abstract

This document specifies extensions to mLDP and RSVP-TE P2MP protocols to set up mLDP and RSVP-TE P2MP tunnels with SRv6 SIDs intead of MPLS labels.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 3 July 2023.

Copyright Notice

Table of Contents

1. Introduction

[I-D.ietf-spring-sr-replication-segment] specifies an SR replication segment as a logical construct which connects a Replication Node to a set of Downstream Nodes. A replicaiton segment is identified by <replication-id, node-id> in control plane.

SR replication segments are building blocks of SR-P2MP replication trees. As specified in [I-D.ietf-pim-sr-p2mp-policy], an SR-P2MP tree is the concatenation of replication segments installed on tree nodes (the packets carried by the tree do not carry a concatenation of replication SIDs for those segments). A controller calculates the P2MP tree, and signals individual replication segments onto tree nodes via PCEP [I-D.ietf-pce-sr-p2mp-policy], BGP [I-D.ietf-idr-sr-p2mp-policy] [I-D.ietf-bess-bgp-multicast-controller], Netconf [RFC6241] or other means.

Each tree is identified by a <root-id, tree-id> tuple and has a corresponing SR P2MP Policy, which may have multiple candidate paths. As such, the replication-id of a replication segment is a <root-id, tree-id, candidate-path-id> tuple.

With MPLS data plane, the forwarding state for a replication segment is identical to the forwarding state on mLDP/RSVP-TE P2MP tree nodes ([RFC6388], [RFC4875]), i.e., in the form of "incoming label -> (labeled) replication branches". In other words, the only difference between mLDP/RSVP-TE P2MP and SR-MPLS P2MP is in the control plane -

instead of hop-by-hop signaling, SR-P2MP signaling is from a controller and with a different control plane identification.

With SRv6 data plane, while SRv6 SID instead of MPLS labels are used, the FUNCT bits in the LOC:FUNCT:ARG format of SID encoding [RFC8986] are the equivalent of MPLS label, while the LOC bits get the packet to local or downstream nodes. Nonetheless, for operators who does not use MPLS data palne, SRv6 P2MP is a natural choice.

However, even an SRv6-only operator may want to use another option to set up its P2MP trees, instead of using controller-based signaling with <root-id, tree-id, candidate-path-id> identification.

Consider an existing MVPN deployment with PE-PE mLDP or RSVP-TE P2MP tunnels and the provider network is being transitioned from MPLS to SRv6 part by part incrementally. Considering the following three factors:

- The MVPN PE-PE tunnel is mLDP/RSVP-TE P2MP so during the transition it is ideal to keep using mLDP FEC or RSVP-TE P2MP Session Object to identify the tunnel in the control plane

- The are some border nodes between the MPLS part and SRv6 part of the network to do MPLS-SRv6 interworking

- Even after the entire network is converted to SRv6, hop-by-hop mLDP/RSVP-TE signaling may still be preferred because controller-based tree calculation and signaling may not be needed or desired for certain reasons

Therefore, it is desired to have P2MP trees identified by mLDP FEC or RSVP Session Object in the control plane but with SRv6 data plane, and there are two options for that:

- Use controller to signal mLDP/RSVP-TE trees with SRv6 SIDs

- Extend mLDP/RSVP-TE P2MP protocol to support SRv6 SIDs

The first option will be specified in [I-D.ietf-bess-bgp-multicast-controller] and [I-D.li-pce-multicast] for BGP and PCEP signaling respectively, and this document specifies the second option.

2. mLDP P2MP Procedures

There are two options to use mLDP protocol and procedures to signal mLDP tunnels for SRv6 data plane, as specified in the following two sections.

2.1. SRv6 SIDs Constructed from Signaled Labels

In this simpliest option, no protocol extension is needed. MPLS labels in various mLDP messages are treated as the FUNCT bits of the LOC:FUNCT:ARG format of SRv6 SID encoding.

All tree nodes MUST have the following provisioned:

- *Whether the signaled labels to/from all neighbors are real MPLS labels or FUNCT bits of SRv6 SIDs, or,

- *Whether the signaled labels to/from each neighbor are real MPLS labels or FUNCT bits of SRv6 SIDs. This allows a node to interwork between MPLS and SRv6 parts of the network.

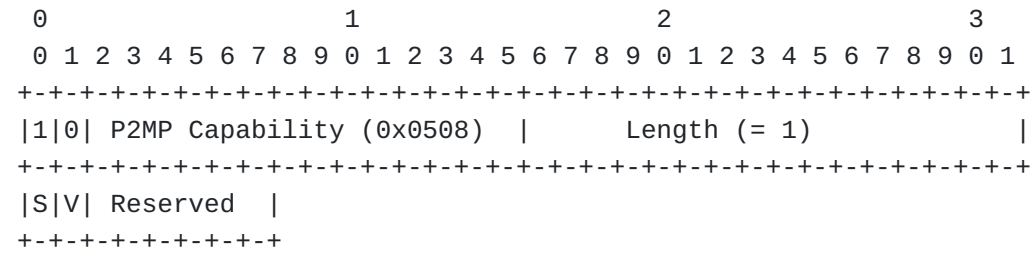
- *If the FUNCT bits of SRv6 SIDs are more than 20-bit, each node MUST be provisioned with a consitent FUNCT "prefix" to be combined with signaled "label".

With the above provisioning, a node determines if a signaled label is a real MPLS label for MPLS data planes, or is to be treated as the FUNCT bits of an SRv6 SID, and installs forwarding state accordingly.

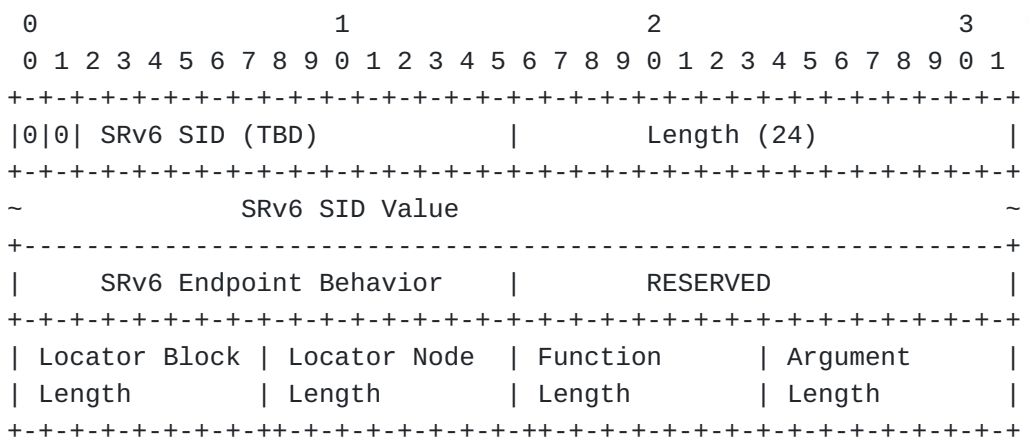
2.2. Explicitly Signaled SRv6 SIDs

With this options, mLDP signaling is extended as following.

A new V-bit is defined in the P2MP Capability TLV to indicate that this node uses SRv6 SIDs:



An SRv6 SID TLV is defined to signal the SRv6 SID instead of a label:



The SRv6 SID TLV is used in place of Generic Label TLV if and only if the neighbor has indicated via the V-bit in the P2MP Capability TLV that it uses SRv6 SIDs.

SRv6 SID Value (16 octets): This field encodes an SRv6 SID, as defined in [RFC8986].

SRv6 Endpoint Behavior (2 octets): This field encodes the SRv6 Endpoint Behavior codepoint value that is associated with the SRv6 SID. The codepoints used are from IANA's "SRv6 Endpoint Behaviors" subregistry under the "Segment Routing" registry that was introduced by [RFC8986]. The opaque SRv6 Endpoint Behavior (i.e., value 0xFFFF) MAY be used when the advertising router wishes to abstract the actual behavior of its locally instantiated SRv6 SID.

RESERVED (2 octet): This field MUST be set to 0 by the sender and ignored by the receiver.

Locator Block Length (1 octet): This field contains the length of the SRv6 SID Locator Block in bits.

Locator Node Length (1 octet): This field contains the length of the SRv6 SID Locator Node in bits.

Function Length (1 octet): This field contains the length of the SRv6 SID Function in bits.

Argument Length (1 octet): This field contains the length of the SRv6 SID Argument in bits.

The choice of SRv6 Endpoint Behavior of the SRv6 SID is entirely up to the originator of the TLV. While this document expects End.Replicate [I-D.ietf-spring-sr-replication-segment], the reception of other SRv6 Endpoint Behaviors (e.g., new behaviors that may be introduced in the future) is not considered an error. An unrecognized SRv6 Endpoint Behavior MUST NOT be considered invalid

by the receiver. An implementation MAY log a rate-limited warning when it receives an unexpected behavior.

2.3. mLDP over Targeted Sessions

To be added.

2.4. Multi-topology and FlexAlgo Considerations

To be added.

3. RSVP-TE P2MP Procedures

Similarly, there are two options to use RSVP-TE P2MP protocol and procedures to signal RSVP-TE P2MP tunnels for SRv6 data plane, as specified in the following two sections.

3.1. SRv6 SIDs Constructed from Signaled Labels

This is the same as Section 2.1.

3.2. Explicitly Signaled SRv6 SIDs

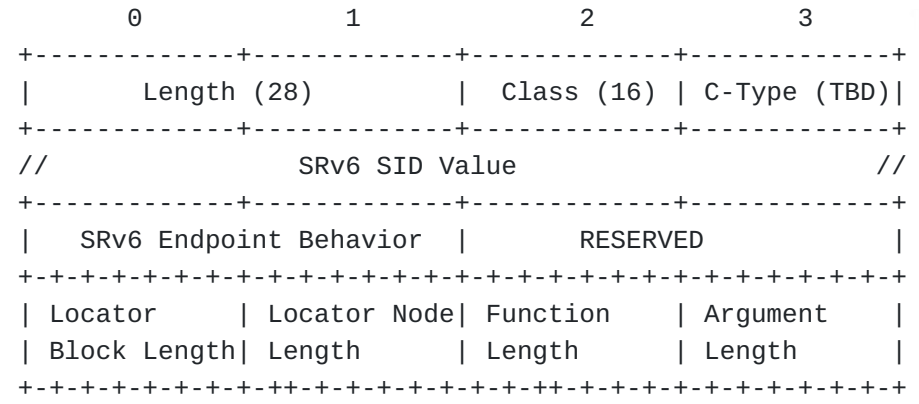
Similar to Section 2.2, RSVP-TE P2MP signaling is extended as following:

3.2.1. Hello Extension

This is to indicate a node uses SRv6 SIDs. To be expanded.

3.2.2. Label Object for SRv6 SID

A new C-type (TBD) is defined for the Label Object to indicate an IPv6 address as SRv6 SID:



The C-Type TBD Label Object is used in place of C-Type 1 Label Object if and only if the neighbor has indicated via Hello that it uses SRv6 SIDs.

SRv6 SID Value (16 octets):
This field encodes an SRv6 SID, as defined in [RFC8986].

SRv6 Endpoint Behavior (2 octets): This field encodes the SRv6 Endpoint Behavior codepoint value that is associated with the SRv6 SID. The codepoints used are from IANA's "SRv6 Endpoint Behaviors" subregistry under the "Segment Routing" registry that was introduced by [RFC8986]. The opaque SRv6 Endpoint Behavior (i.e., value 0xFFFF) MAY be used when the advertising router wishes to abstract the actual behavior of its locally instantiated SRv6 SID.

RESERVED (2 octet): This field MUST be set to 0 by the sender and ignored by the receiver.

Locator Block Length (1 octet): This field contains the length of the SRv6 SID Locator Block in bits.

Locator Node Length (1 octet): This field contains the length of the SRv6 SID Locator Node in bits.

Function Length (1 octet): This field contains the length of the SRv6 SID Function in bits.

Argument Length (1 octet): This field contains the length of the SRv6 SID Argument in bits.

The choice of SRv6 Endpoint Behavior of the SRv6 SID is entirely up to the originator of the TLV. While this document expects End.Replicate [I-D.ietf-spring-sr-replication-segment], the reception of other SRv6 Endpoint Behaviors (e.g., new behaviors that may be introduced in the future) is not considered an error. An unrecognized SRv6 Endpoint Behavior MUST NOT be considered invalid by the receiver. An implementation MAY log a rate-limited warning when it receives an unexpected behavior.

4. Security Considerations

To be added.

5. IANA Considerations

To be added.

6. References

6.1. Normative References

[I-D.ietf-spring-sr-replication-segment]

Voyer, D., Filsfils, C., Parekh, R., Bidgoli, H., and Z. J. Zhang, "SR Replication Segment for Multi-point Service Delivery", Work in Progress, Internet-Draft, draft-ietf-spring-sr-replication-segment-10, 20 October 2022, <https://www.ietf.org/archive/id/draft-ietf-spring-sr-replication-segment-10.txt>.

[RFC3471] Berger, L., Ed. and RFC Publisher, "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, DOI 10.17487/RFC3471, January 2003, <https://www.rfc-editor.org/info/rfc3471>.

[RFC4875] Aggarwal, R., Ed., Papadimitriou, D., Ed., Yasukawa, S., Ed., and RFC Publisher, "Extensions to Resource Reservation Protocol - Traffic Engineering (RSVP-TE) for Point-to-Multipoint TE Label Switched Paths (LSPs)", RFC 4875, DOI 10.17487/RFC4875, May 2007, <https://www.rfc-editor.org/info/rfc4875>.

[RFC6388] Wijnands, IJ., Ed., Minei, I., Ed., Kompella, K., Thomas, B., and RFC Publisher, "Label Distribution Protocol Extensions for Point-to-Multipoint and Multipoint-to-Multipoint Label Switched Paths", RFC 6388, DOI 10.17487/RFC6388, November 2011, <https://www.rfc-editor.org/info/rfc6388>.

[RFC8986] Filsfils, C., Ed., Camarillo, P., Ed., Leddy, J., Voyer, D., Matsushima, S., Li, Z., and RFC Publisher, "Segment Routing over IPv6 (SRv6) Network Programming", RFC 8986, DOI 10.17487/RFC8986, February 2021, <https://www.rfc-editor.org/info/rfc8986>.

6.2. Informative References

[I-D.ietf-bess-bgp-multicast-controller] Zhang, Z. J., Raszuk, R., Pacella, D., and A. Gulko, "Controller Based BGP Multicast Signaling", Work in Progress, Internet-Draft, draft-ietf-bess-bgp-multicast-controller-09, 11 April 2022, <https://www.ietf.org/archive/id/draft-ietf-bess-bgp-multicast-controller-09.txt>.

[I-D.ietf-idr-sr-p2mp-policy] Bidgoli, H., Voyer, D., Stone, A., Parekh, R., Krier, S., and S. Agrawal, "Advertising p2mp policies in BGP", Work in Progress, Internet-Draft, draft-ietf-idr-sr-p2mp-policy-00, 27 May 2022, <https://

www.ietf.org/archive/id/draft-ietf-idr-sr-p2mp-policy-00.txt>.

[I-D.ietf-pce-sr-p2mp-policy] Bidgoli, H., Voyer, D., Rajarathinam, S., Hemmati, E., Saad, T., and S. Sivabalan, "PCEP extensions for p2mp sr policy", Work in Progress, Internet-Draft, draft-ietf-pce-sr-p2mp-policy-01, 7 July 2022, <https://www.ietf.org/archive/id/draft-ietf-pce-sr-p2mp-policy-01.txt>.

[I-D.ietf-pim-sr-p2mp-policy] Voyer, D., Filsfils, C., Parekh, R., Bidgoli, H., and Z. J. Zhang, "Segment Routing Point-to-Multipoint Policy", Work in Progress, Internet-Draft, draft-ietf-pim-sr-p2mp-policy-05, 2 July 2022, <https://www.ietf.org/archive/id/draft-ietf-pim-sr-p2mp-policy-05.txt>.

[I-D.li-pce-multicast] Li, H., Wang, A., Zhang, Z. J., Chen, H., and R. Chen, "Multicast Tree Setup via PCEP", Work in Progress, Internet-Draft, draft-li-pce-multicast-02, 20 September 2022, <https://www.ietf.org/archive/id/draft-li-pce-multicast-02.txt>.

[RFC6241] Enns, R., Ed., Bjorklund, M., Ed., Schoenwaelder, J., Ed., Bierman, A., Ed., and RFC Publisher, "Network Configuration Protocol (NETCONF)", RFC 6241, DOI 10.17487/RFC6241, June 2011, <https://www.rfc-editor.org/info/rfc6241>.

Authors' Addresses

Zhaohui Zhang
Juniper Networks

Email: zzhang@juniper.net

Vishnu Pavan Beeram
Juniper Networks

Email: vbeeram@juniper.net

Rishabh Parekh
Cisco

Email: riparekh@cisco.com

IJsbrand Wijnands
Arrcus

Email: ice@braindump.be

Ran Chen
ZTE

Email: chen.ran@zte.com.cn