

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: May 1, 2020

R. Parekh
C. Filsfils
A. Venkateswaran
Cisco Systems, Inc.
H. Bidgoli
Nokia
D. Voyer
Bell Canada
Z. Zhang
Juniper Networks
October 29, 2019

Multicast VPN with Segment Routing Point-to-Multipoint Trees
draft-parekh-bess-mvpn-sr-p2mp-01.txt

Abstract

A Point-to-Multipoint (P2MP) Tree in a Segment Routing domain efficiently carries traffic from a Root to a set of Leaves. This document describes extensions to BGP encodings and procedures for P2MP trees used in BGP/MPLS IP VPNs.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC 2119](#) [[RFC2119](#)].

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 May 1, 2020.

Copyright Notice

Copyright (c) 2019 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction	3
2.	SR P2MP P-Tunnels for MPVN	3
3.	PMSI Tunnel Attribute for SR P2MP	4
3.1.	MPLS Label	5
3.1.1.	SR MPLS	5
4.	Auto-Discovery and Binding Procedures	5
4.1.	Intra-AS I-PMSI	5
4.1.1.	Originating Intra-AS I-PMSI routes	5
4.1.2.	Receiving Intra-AS I-PMSI A-D routes	6
4.2.	Using S-PMSIs for binding customer flows to P2MP Segments	7
4.2.1.	Originating S-PMSI A-D routes	7
4.2.2.	Receiving S-PMSI A-D routes	7
4.3.	Inter-AS P-tunnels using P2MP Segments	8
4.3.1.	Advertising Inter-AS I-PMSI routes into iBGP	8
4.3.2.	Receiving Inter-AS I-PMSI A-D routes in iBGP	8
4.4.	Leaf A-D routes for P2MP Segment Leaf Discovery	8
4.4.1.	Originating Leaf A-D routes	8
4.4.2.	Receiving Leaf A-D routes	9
5.	Damping of MVPN routes	9
6.	IANA Considerations	10
7.	Security Considerations	10
8.	Contributors	10
9.	References	10
9.1.	Normative References	11
9.2.	Informative References	11
9.3.	URIs	12
	Authors' Addresses	12

1. Introduction

[RFC 6513](#) [[RFC6513](#)] and [RFC 6514](#) [[RFC6514](#)] specify procedures that allow a Service Provider to provide Multicast VPN (MVPN) service to its customers. Multicast traffic from a customer is tunneled across the service provider network over Provider Tunnels (P-Tunnels). P-tunnels can be instantiated via different technologies. A service provider network that uses Segment Routing can use a Point-to-Multipoint (SR P2MP) tree [[I-D.voyer-pim-sr-p2mp-policy](#)] to instantiate P-Tunnels for MVPN.

In a Segment Routing network, a P2MP tree allows efficient delivery of traffic from a Root to set of Leaf nodes. A SR P2MP tree is defined by a SR P2MP Policy and instantiated via a PCE. A P2MP Policy consists of a Root, a Set of Leaf Nodes and a set of candidate paths with optional set of constraints and/or optimization objectives to be satisfied by the P2MP tree. A unique Identifier, called Tree-SID, is associated with a P2MP tree. This Tree-SID can be an MPLS label or an IPv6 address.

This document describes extensions to BGP Auto-Discovery procedures specified in [RFC 6514](#) [[RFC6514](#)] when P-Tunnels are realized by SR P2MP trees. Use of PIM for Auto-Discovery is outside scope of this document. Support for customer BIDIR-PIM is outside the scope of this document.

The reader is expected to be familiar with concepts and terminology of [RFC 6513](#), [RFC 6514](#) and SR P2MP draft.

2. SR P2MP P-Tunnels for MPVN

For MVPN, Provider Edge(PE) routers steer customer multicast traffic into a P-Tunnel instantiated by SR P2MP tree. A SR P2MP tree is defined by a SR P2MP policy [[I-D.voyer-pim-sr-p2mp-policy](#)].

Given a SR P2MP policy, a PCE computes and instantiates the SR P2MP tree on the nodes that are part of the tree using Replication segments and Tree-SID which a unique identifier for the tree [[I-D.voyer-spring-sr-replication-segment](#)]. A Replication segment can be initiated by various methods (BGP, PCEP, others) which are outside the scope of this document.

A PCE provides conceptual APIs, listed below, to define and modify SR P2MP policies. These APIs are invoked by a PCC, which is the root of P2MP tree, using various methods (BGP, PCEP, etc.) which are outside the scope of this document.

CreatePolicy: TBD

DeletePolicy: TBD

AddLeaf: TBD

DeleteLeaf: TBD

The Root of a P2MP tree imposes the Tree-SID to steer the customer payload into the P2MP tree. Provider (P) routers replicate customer payload, using Replication segments, towards the Leaf nodes of the P2MP tree. Leaf nodes of the P2MP tree deliver the customer payload after disposing the Tree-SID.

3. PMSI Tunnel Attribute for SR P2MP

A PMSI Tunnel Attribute (PTA) is defined in [RFC 6514](#) [[RFC6514](#)] to identify the P-Tunnel that is used to instantiate a Provider Multicast Service Interface (PMSI). The PTA is carried in Intra-AS I-PMSI, Inter-AS I-PMSI, Selective PMSI, and Leaf Auto-Discovery routes.

A P2MP tree PTA is constructed as follows:

- o Tunnel Type: The codepoint is set to `[[CREF1: TBD]]` for SR P2MP tree from the "P-Multicast Service Interface Tunnel (PMSI Tunnel) Tunnel Types" registry.
- o Flags: See [Section 4](#) for use of "Leaf Info Required bit".
- o MPLS Label: See [Section 3.1](#)
- o Tunnel Identifier: The SR P2MP P-tunnel is identified by <Tree-ID, Root> where,
 - * Tree-ID is a 32-bit unsigned value that identifies a unique P2MP tree at a Root..
 - * Root is an IP address identifying the Root of a P2MP tree. This can be either an IPv4 or IPv6 address and can be inferred from the PTA length.

When a P-Tunnel is non-segmented, the PTA is created by PE router at the Root of a SR P2MP tree. For segmented P-tunnels, each segment can be instantiated by a different technology. If a segment is instantiated using P2MP tree, the router at the root of a P2MP tree creates the PTA.

3.1. MPLS Label

[RFC6514] allows a PE to aggregate two or more MVPNs onto one P-tunnel by advertising the same P-tunnel in PTA of Auto-Discovery routes of different MVPNs. This section specifies how the "MPLS Label" field of PTA is filled to provide a context bound to a specific MPVN.

3.1.1. SR MPLS

When a SR P2MP P-tunnel, shared across different MVPNs, is instantiated in a SR MPLS domain [[I-D.filsfils-spring-segment-routing-mpls](#)], "MPLS Field" of a PTA advertised in a Auto-Discovery route MUST contain an upstream-assigned MPLS label that the advertising PE has bound to the MVPN.

When a customer payload is steered into a shared SR P2MP P-tunnel, this MPLS label MUST be imposed before the MPLS label representing the Tree-SID.

4. Auto-Discovery and Binding Procedures

[RFC 6514](#) [[RFC6514](#)] defines procedures for discovering PEs participating in a given MVPN and binding customer multicast flows to specific P-Tunnels. This section specifies modifications to these procedures for SR P2MP P-Tunnels.

4.1. Intra-AS I-PMSI

Intra-AS I-PMSI A-D routes are exchanged to discover PEs participating in a MVPN within an AS, or across different ASes when non-segmented P-tunnels for inter-AS MVPNs.

4.1.1. Originating Intra-AS I-PMSI routes

[RFC 6514 Section 9.1.1](#) [[1](#)] describes procedures for originating Intra-AS I-PMSI A-D routes. For SR P2MP P-tunnels, these procedures remain unchanged except as described in the following paragraphs.

When a PE originates an Intra-AS I-PMSI A-D route with a PTA having SR P2MP P-tunnel Type, it MUST create a P2MP policy by invoking CreatePolicy API of the PCE. When the PCE instantiates the P2MP tree on the PE, the Tree-SID MUST be imposed for customer flow(s) steered into the P2MP tree. The Leaf nodes of P2MP tree are discovered using procedures described in [Section 4.1.2](#).

For a PE in "Receiver Sites set", condition (c) is modified to include P2MP tree i.e. such a PE MUST originate an Intra-AS I-PMSI

A-D route when some PEs of the MVPN have VRFs that use SR P2MP tree but MUST NOT create a SR P2MP policy as described above.

When a PE withdraws an Intra-AS I-PMSI A-D route, advertised with a PTA having SR P2MP P-tunnel Type, the Tree-SID imposition state at the PE MUST be removed.

A PE MAY aggregate two or more Intra-AS I-PMSIs from different MVPNs onto the same SR P2MP P-tunnel. When a PE withdraws the last Intra-AS I-PMSI A-D route, advertised with a PTA identifying a SR P2MP P-tunnel, it SHOULD remove the SR P2MP policy by invoking DeletePolicy API of the PCE.

4.1.2. Receiving Intra-AS I-PMSI A-D routes

Procedure for receiving Intra-AS I-PMSI A-D routes, as described in [RFC 6514 Section 9.1.2 \[2\]](#), remain unchanged for SR P2MP P-tunnels except as described in the following paragraphs.

When a PE that advertises a SR P2MP P-tunnel in the PTA of its Intra-AS I-PMSI A-D route, imports an Intra-AS I-PMSI A-D route from some PE, it MUST add that PE as a Leaf node of the P2MP tree. The Originating IP Address of the Intra-AS i-PMSI A-D route is used as the Leaf Address when invoking AddLeaf API of the PCE. This procedure MUST also be followed for all Intra-AS I-PMSI routes that are already imported when the PE advertises a SR P2MP P-tunnel in PTA of its Intra-AS I-PMSI A-D route.

A PE that imports and processes an Intra-AS I-PMSI A-D route from another PE with PTA having SR P2MP P-Tunnel MUST program the Tree-SID of the P2MP tree identified in the PTA of the route for disposition. Note that an Intra-AS I-PMSI A-D route from another PE can be imported before the P2MP tree identified in the PTA of the route is instantiated by the PCE at the importing PE. In such case, the PE MUST correctly program Tree-SID for disposition. A PE in "Sender Sites set" MAY avoid programming the Tree-SID for disposition.

When an Intra-AS I-PMSI A-D route, advertised with a PTA having SR P2MP P-tunnel Type is withdrawn, a PE MUST remove the disposition state of the Tree-SID associated with P2MP tree.

A PE MAY aggregate two or more Intra-AS I-PMSIs from different MVPNs onto the same SR P2MP P-tunnel. When a remote PE withdraws an Intra-AS I-PMSI A-D route from a MVPN, and if this is the last MVPN sharing a SR P2MP P-tunnel, a PE must remove the originating PE as a Leaf from the P2MP tree, by invoking DeleteLeaf API.

4.2. Using S-PMSIs for binding customer flows to P2MP Segments

[RFC 6514](#) [[RFC6514](#)] specifies procedures for binding (C-S,C-G) customer flows to P-tunnels using S-PMSI A-D routes. [RFC 6525](#) [[RFC6625](#)] specifies additional procedures to binding aggregate customer flows to P-tunnels using "wildcard" S-PMSI A-D routes. This section describes modification to these procedures for SR P2MP P-tunnels.

4.2.1. Originating S-PMSI A-D routes

[RFC 6514 Section 12.1](#) [[3](#)] describes procedures for originating S-PMSI A-D routes. For SR P2MP P-tunnels, these procedures remain unchanged except as described in the following paragraphs.

When a PE originates S-PMSI A-D route with a PTA having SR P2MP P-tunnel Type, it MUST set the "Leaf Info Required bit" in the PTA. The PE MUST create a SR P2MP policy by invoking `1` API of the PCE. When the PCE instantiates the P2MP tree on the PE, the Tree-SID MUST be imposed for customer flows steered into the SR P2MP P-tunnel.

The Leaf nodes of P2MP tree are discovered by Leaf A-D routes using procedures described in [Section 4.4.2](#). When a PE originates S-PMSI A-D route with a PTA having SR P2MP P-tunnel Type, it is possible the PE might have imported Leaf A-D routes whose route keys match the S-PMSI A-D route. The PE MUST re-apply procedures of [Section 4.4.2](#) to these Leaf A-D routes.

When a PE withdraws a S-PMSI A-D route, advertised with PTA having P2MP tree P-tunnel type, the Tree-SID imposition state MUST be removed.

A PE MAY aggregate two or more S-PMSIs onto the same SR P2MP P-tunnel. When a PE withdraws the last S-PMSI A-D route, advertised with a PTA identifying a specific SR P2MP P-tunnel, it SHOULD remove the SR P2MP policy by invoking `DeletePolicy` API of the PCE.

4.2.2. Receiving S-PMSI A-D routes

[RFC 6514 Section 12.3](#) [[4](#)] describes procedures for receiving S-PMSI A-D routes. For SR P2MP P-tunnels, these procedures remain unchanged except as described in the following paragraphs.

The procedure to join SR P2MP P-tunnel of S-PMSI A-D route by using a Leaf A-D route is described in [Section 4.4.1](#). If P2MP tree identified in PTA of S-PMSI A-D route is already instantiated by PCE, the PE MUST program Tree-SID for disposition. If the P2MP tree is

instantiated later, the Tree-SID MUST be programmed for disposition at that time.

When a S-PMSI A-D route, whose SR P2MP P-tunnel is joined by a PE, is withdrawn, or when conditions (see [RFC 6514 Section 12.3 \[5\]](#)) required to join that P-Tunnel are no longer satisfied, the PE MUST leave the P-Tunnel. The PE MUST withdraw the Leaf A-D route it had originated and remove the Tree-SID disposition state.

4.3. Inter-AS P-tunnels using P2MP Segments

A segmented inter-AS P-tunnel consists of one or more intra-AS segments, one in each AS, connected by inter-AS segments between ASBRs of different ASes <<https://tools.ietf.org/html/rfc6514#section-9.2>>. These segments are constructed by PEs/ASBRs originating or re-advertising Inter-AS I-PMSI A-D routes. This section describes procedures for instantiating intra-AS segments using SR P2MP trees.

4.3.1. Advertising Inter-AS I-PMSI routes into iBGP

[RFC 6514 Section 9.2.3.2 \[6\]](#) specifies procedures for advertising an Inter-AS I-PMSI A-D route to construct an intra-AS segment. The PTA of the route identifies the type and identifier of the P-tunnel instantiating the intra-AS segment. The procedure for creating SR P2MP P-tunnel for intra-AS segment are same as specified in [Section 4.2.1](#) except that instead of S-PMSI A-D routes, the procedures apply to Inter-AS I-PMSI A-D routes.

4.3.2. Receiving Inter-AS I-PMSI A-D routes in iBGP

[RFC 6514 Section 9.2.3.2 \[7\]](#) specifies procedures for processing an Inter-AS I-PMSI A-D route received via iBGP. If the PTA of the Inter-AS I-PMSI A-D route has SR P2MP P-tunnel Type, the procedures are same as specified in [Section 4.2.2](#) except that instead of S-PMSI A-D routes, the procedures apply to Inter-AS I-PMSI A-D routes. If the receiving router is an ASBR, the Tree-SID is stitched to the inter-AS segments to ASBRs in other ASes.

4.4. Leaf A-D routes for P2MP Segment Leaf Discovery

This section describes procedures for originating and processing Leaf A-D routes used for Leaf discovery of SR P2MP trees.

4.4.1. Originating Leaf A-D routes

The procedures for originating Leaf A-D route in response to receiving a S-PMSI or Inter-AS I-PMSI A-D route with PTA having SR

P2MP P-tunnel Type are same as specified in [RFC 6514 Section 9.2.3.4.1](#) [8].

4.4.2. Receiving Leaf A-D routes

Procedures for processing a received Leaf A-D route are specified in [RFC 6514 Section 9.2.3.5](#) [9]. These procedures remain unchanged for discovering Leaf nodes of P2MP trees except for considerations described in following paragraphs. These procedures apply to Leaf A-D routes received in response to both S-PMSI and Inter-AS I-PMSI A-D routes, shortened to "A-D routes" in this section

A Root PE/ASBR MAY use the same SR P2MP P-tunnel in PTA of two or more A-D routes. For such aggregated P2MP trees, the PE/ASBR MAY receive multiple Leaf A-D routes from a Leaf PE. The P2MP tree for which a Leaf A-D is received can be identified by examining the P2MP tunnel Identifier in the PTA of A-D route that matches "Route Key" field of the Leaf A-D route. When the PE receives the first Leaf A-D route from a Leaf PE, identified by the Originating Router's IP address field, it MUST add that PE as Leaf of the P2MP tree by invoking the AddLeaf API of the PCE.

When a Leaf PE withdraws the last Leaf A-D route for a given SR P2MP P-tunnel, the Root PE MUST remove the Leaf PE from the P2MP tree by invoking DeleteLeaf API of PCE. Note that Root PE MAY remove the P2MP tree, via the DeletePolicyAPI, before the last Leaf A-D is withdrawn. In this case, the Root PE MAY decide to not invoke the DeleteLeaf API.

5. Damping of MVPN routes

When P2MP trees are used as P-Tunnels for S-PMSI A-D routes, change in group membership of receivers connected to PEs has direct impact on the Leaf node set of a P2MP tree. If group membership changes frequently for a large number of groups with a lot of receivers across sites connected to different PEs, it can have an impact on the interaction between PEs and the PCE.

Since Leaf A-D routes are used to discover Leaf PE of a P2MP tree, it is RECOMMENDED that PEs SHOULD damp Leaf A-D routes as described in [Section 6.1 of RFC 7899](#) [RFC7899]. PEs MAY also implement procedures for damping other Auto-Discovery and BGP C-multicast routes as described in [RFC7899].

6. IANA Considerations

IANA to assign a codepoint `[[CREF2: TBD]]` for "P2MP tree" in the "P-Multicast Service Interface Tunnel (PMSI Tunnel) Tunnel Types" registry.

7. Security Considerations

The procedures in this document do not introduce any additional security considerations beyond those mentioned in [\[RFC6513\]](#) and [\[RFC6514\]](#). For general security considerations applicable to P2MP trees, please refer to [\[I-D.voyer-pim-sr-p2mp-policy\]](#) .

8. Contributors

Zafar Ali
Cisco Systems, Inc.
US

Email: zali@cisco.com

Jayant Kotalwar
Nokia
Mountain View
US

Email: jayant.kotalwar@nokia.com

Tanmoy Kundu
Nokia
Mountain View
US

Email: tanmoy.kundu@nokia.com

Clayton Hassen
Bell Canada
Vancouver
CA

Email: clayton.hassen@bell.ca

9. References

9.1. Normative References

- [I-D.voyer-pim-sr-p2mp-policy]
Voyer, D., Filsfils, C., Parekh, R., Bidgoli, H., and Z. Zhang, "Segment Routing Point-to-Multipoint Policy", [draft-voyer-pim-sr-p2mp-policy-00](#) (work in progress), October 2019.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.
- [RFC6513] Rosen, E., Ed. and R. Aggarwal, Ed., "Multicast in MPLS/BGP IP VPNs", [RFC 6513](#), DOI 10.17487/RFC6513, February 2012, <<https://www.rfc-editor.org/info/rfc6513>>.
- [RFC6514] Aggarwal, R., Rosen, E., Morin, T., and Y. Rekhter, "BGP Encodings and Procedures for Multicast in MPLS/BGP IP VPNs", [RFC 6514](#), DOI 10.17487/RFC6514, February 2012, <<https://www.rfc-editor.org/info/rfc6514>>.

9.2. Informative References

- [I-D.filsfils-spring-segment-routing-mpls]
Filsfils, C., Previdi, S., Bashandy, A., Decraene, B., Litkowski, S., Horneffer, M., Milojevic, I., Shakir, R., Ytti, S., Henderickx, W., Tantsura, J., and E. Crabbe, "Segment Routing with MPLS data plane", [draft-filsfils-spring-segment-routing-mpls-03](#) (work in progress), August 2014.
- [I-D.voyer-spring-sr-replication-segment]
Voyer, D., Filsfils, C., Parekh, R., Bidgoli, H., and Z. Zhang, "SR Replication Segment for Multi-point Service Delivery", [draft-voyer-spring-sr-replication-segment-00](#) (work in progress), October 2019.
- [RFC6625] Rosen, E., Ed., Rekhter, Y., Ed., Hendrickx, W., and R. Qiu, "Wildcards in Multicast VPN Auto-Discovery Routes", [RFC 6625](#), DOI 10.17487/RFC6625, May 2012, <<https://www.rfc-editor.org/info/rfc6625>>.
- [RFC7899] Morin, T., Ed., Litkowski, S., Patel, K., Zhang, Z., Kebler, R., and J. Haas, "Multicast VPN State Damping", [RFC 7899](#), DOI 10.17487/RFC7899, June 2016, <<https://www.rfc-editor.org/info/rfc7899>>.

9.3. URIs

- [1] <https://tools.ietf.org/html/rfc6514#section-9.1.1>
- [2] <https://tools.ietf.org/html/rfc6514#section-9.1.2>
- [3] <https://tools.ietf.org/html/rfc6514#section-12.1>
- [4] <https://tools.ietf.org/html/rfc6514#section-12.3>
- [5] <https://tools.ietf.org/html/rfc6514#section-12.3>
- [6] <https://tools.ietf.org/html/rfc6514#section-9.2.3.2>
- [7] <https://tools.ietf.org/html/rfc6514#section-9.2.3.2>
- [8] <https://tools.ietf.org/html/rfc6514#section-9.2.3.4.1>
- [9] <https://tools.ietf.org/html/rfc6514#section-9.2.3.5>

Authors' Addresses

Rishabh Parekh
Cisco Systems, Inc.
170 W. Tasman Drive
San Jose, CA 95134
USA

Email: riparekh@cisco.com

Clarence Filsfils
Cisco Systems, Inc.
Brussels
BE

Email: cfilsfil@cisco.com

Arvind Venkateswaran
Cisco Systems, Inc.
170 W. Tasman Drive
San Jose, CA 95134
USA

Email: arvvenka@cisco.com

Hooman Bidgoli
Nokia
Ottawa
CA

Email: hooman.bidgoli@nokia.com

Daniel Voyer
Bell Canada
Montreal
CA

Email: daniel.voyer@bell.ca

Zhaohui Zhang
Juniper Networks

Email: zzhang@juniper.net

