Multicast VPN with Segment Routing Point-to-Multipoint Segment
draft-parekh-bess-mvpn-sr-p2mp-00.txt

Abstract

A Point-to-Multipoint (P2MP) Segment 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 segments 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 September 12, 2019.
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. P2MP Segment P-Tunnels for MPVN .......................... 3
3. PMSI Tunnel Attribute for P2MP Segment ................... 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 .................... 9
      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 ................................ 10
   9.2. Informative References ................................. 11
   9.3. URIs ................................................... 11
10. 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 Segment [I-D.voyer-spring-sr-p2mp-policy] (Tree-SID P2MP segment) to instantiate P-Tunnels for MVPN.

In a Segment Routing network, a P2MP segment allows efficient delivery of traffic from a Root to set of Leaf nodes. A P2MP segment is defined by a P2MP segment Policy and instantiated via a P2MP policy engine. A P2MP segment Policy consists of a Root, a Set of Leaf Nodes and an optional set of constraints to be satisfied by the P2MP segment. A unique P2MP segment Identifier (P2MP SID) is associated with a P2MP segment. This P2MP 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 P2MP segments (via P2MP segment Policy). 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. P2MP Segment P-Tunnels for MVPN

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

A P2MP policy engine provides conceptual APIs, listed below, to define and modify P2MP segment policies. These APIs can be invoked by different methods (BGP, PCEP, etc.) which are outside the scope of this document.

CreatePolicy: TBD

DeletePolicy: TBD

AddLeaf: TBD

DeleteLeaf: TBD
For a SR P2MP segment policy, the P2MP Policy Engine computes and instantiates the Tree-SID P2MP segment on the nodes that are part of the segment using an Identifier (TreeSID) and a SR Replication [I-D.voyer-spring-sr-p2mp-policy]. A TreeSID segment can be initiated by various methods (BGP, PCEP, others) which are outside the scope of this document.

the Root of a P2MP segment imposes the P2MP SID to steer the customer payload into the P2MP segment. Provider (P) routers replicate customer payload, using the P2MP SID, towards the Leaf nodes of the P2MP segment. Leaf nodes of the P2MP segment deliver the customer payload after dispoing the P2MP SID.

3. PMSI Tunnel Attribute for P2MP Segment

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 segment PTA is constructed as follows:

- **Tunnel Type**: The codepoint is set to [[CREF1: TBD]] for P2MP segment from the "P-Multicast Service Interface Tunnel (PMSI Tunnel) Tunnel Types" registry.
- **Flags**: See Section 4 for use of "Leaf Info Required bit".
- **MPLS Label**: See Section 3.1
- **Tunnel Identifier**: The P2MP segment P-Tunnel is identified by <Tree-ID, Root> where,
  * Tree-ID is a 32-bit unsigned value that identifies a unique P2MP segment at a Root.
  * Root is an IP address identifying the Root of a P2MP segment. 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 P2MP segment. For segmented P-tunnels, each segment can be instantiated by a different technology. If a segment is instantiated using P2MP segment, the router at the root of a P2MP segment 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 MVPN.

3.1.1. SR MPLS

When a P2MP segment P-tunnel, shared across different MVPNs, is instantiated in a SR MPLS domain [I-D.filips-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 P2MP segment P-tunnel, this MPLS label MUST be imposed before the MPLS label representing the P2MP SID.

4. Auto-Discovery and Binding Procedures

RFC 6514 [RFC6514] defines procedures for discovering PE's participating in a given MVPN and binding customer multicast flows to specific P-Tunnels. This section specifies modifications to these procedures when P-Tunnels are instantiated by P2MP segments.

4.1. Intra-AS I-PMSI

Intra-AS I-PMSI A-D routes are exchanged to discover PE's 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 P2MP segment 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 P2MP segment Tunnel Type, it MUST create a P2MP segment policy by invoking CreatePolicy API of the P2MP Policy Engine. When the P2MP Policy Engine instantiates the P2MP segment on the PE, the P2MP SID MUST be imposed for customer flow(s) steered into the P2MP segment. The Leaf nodes of P2MP segment are discovered using procedures described in Section 4.1.2.
For a PE in "Receiver Sites set", condition (c) is modified to include P2MP segment 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 P2MP segment but MUST NOT create a P2MP segment policy as described above.

When a PE withdraws an Intra-AS I-PMSI A-D route, advertised with a PTA having P2MP segment Tunnel Type, the P2MP 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 P2MP segment P-Tunnel. When a PE withdraws the last Intra-AS I-PMSI A-D route, advertised with a PTA identifying a P2MP segment P-Tunnel, it SHOULD remove the P2MP segment policy by invoking DeletePolicy API of the P2MP policy engine.

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 P2MP segment P-tunnels except as described in the following paragraph.

When a PE that advertises a P2MP segment 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 segment. 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 P2MP Policy Engine. This procedure MUST also be followed for all Intra-AS I-PMSI routes that are already imported when the PE advertises a P2MP segment 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 Tunnel Type as P2MP segment MUST program the P2MP SID of the P2MP segment 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 segment identified in the PTA of the route is instantiated by the P2MP policy engine at the importing PE. In such case, the PE MUST correctly program P2MP SID for disposition. A PE in "Sender Sites set" MAY avoid programming the P2MP SID for disposition.

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

A PE MAY aggregate two or more Intra-AS I-PMSIs from different MVPNs onto the same P2MP segment 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 P2MP segment P-tunnel, a PE must remove the originating PE as a Leaf from the P2MP segment, 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 [RFC6525] 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 when P2MP segment 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 P2MP segment 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 P2MP segment Tunnel Type, it MUST set the "Leaf Info Required bit" in the PTA. The PE MUST create a P2MP segment policy by invoking API of the P2MP Policy Engine. When the P2MP Policy Engine instantiates the P2MP segment on the PE, the P2MP SID MUST be imposed for customer flow(s) steered into the P2MP segment P-Tunnel.

The Leaf nodes of P2MP segment 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 P2MP segment 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 segment P-tunnel type, the P2MP SID imposition state MUST be removed.

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

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 P2MP segment P-tunnels, these procedures remain unchanged except as described in the following paragraphs.
The procedure to join P2MP segment P-Tunnel of S-PMSI A-D route by using a Leaf A-D route is described in Section 4.4.1. If P2MP segment identified in PTA of S-PMSI A-D route is already instantiated by P2MP policy engine, the PE MUST program P2MP SID for disposition. If the P2MP segment is instantiated later, the P2MP SID MUST be programmed for disposition at that time.

When a S-PMSI A-D route, whose P2MP segment 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 P2MP 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 P2MP segments.

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 P2MP segment 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 P2MP segment 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 P2MP 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 P2MP segment P-tunnels.
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 P2MP segment 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 segments 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 aggregate two or more A-D routes on the same P2MP segment P-Tunnel. For such aggregated P2MP segments, the PE/ASBR MAY receive multiple Leaf A-D routes from a Leaf PE. The P2MP segment 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 segment by invoking the AddLeaf API of the P2MP policy engine.

When a Leaf PE withdraws the last Leaf A-D route for a given P2MP segment, the Root PE MUST remove the Leaf PE from the P2MP segment by invoking DeleteLeaf API of P2MP policy engine. Note that Root PE MAY remove the P2MP segment, 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 segments 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 segment. 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 P2MP policy engine.

Since Leaf A-D routes are used to discover Leaf PE of a P2MP segment, 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 segment" 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 segments, please refer to [I-D.voyer-spring-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

9. References

9.1. Normative References

[I-D.voyer-spring-sr-p2mp-policy]
daniel.voyer@bell.ca, d., Hassen, C., Gillis, K.,
9.2. Informative References

[I-D.filsfils-spring-segment-routing-mpls]


9.3. URIs

[8] https://tools.ietf.org/html/rfc6514#section-9.2.3.4.1

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
Clayton Hassen
Bell Canada
Vancouver
CA

Email: clayton.hassen@bell.ca