Internet Engineering Task Force (IETF)
Request for Comments: 7524
Category: Standards Track
ISSN: 2070-1721

Y. Rekhter
E. Rosen
Juniper Networks
R. Aggarwal
Arktan
T. Morin
I. Grosclaude
Orange
N. Leymann
Deutsche Telekom AG
S. Saad
AT&T

May 2015

Inter-Area Point-to-Multipoint (P2MP) Segmented Label Switched Paths (LSPs)

Abstract

This document describes procedures for building inter-area point-to-multipoint (P2MP) segmented service label switched paths (LSPs) by partitioning such LSPs into intra-area segments and using BGP as the inter-area routing and Label Distribution Protocol (LDP). Within each IGP area, the intra-area segments are either carried over intra-area P2MP LSPs, using P2MP LSP hierarchy, or instantiated using ingress replication. The intra-area P2MP LSPs may be signaled using P2MP RSVP-TE or P2MP multipoint LDP (mLDP). If ingress replication is used within an IGP area, then (multipoint-to-point) LDP LSPs or (point-to-point) RSVP-TE LSPs may be used in the IGP area. The applications/services that use such inter-area service LSPs may be BGP Multicast VPN, Virtual Private LAN Service (VPLS) multicast, or global table multicast over MPLS.

Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc7524.

Copyright Notice

Table of Contents

1. Introduction

This document describes procedures for building inter-area point-to-multipoint (P2MP) segmented service LSPs by partitioning such LSPs into intra-area segments and using BGP as the inter-area routing and label distribution protocol. Within each IGP area, the intra-area segments are either carried over intra-area P2MP LSPs, potentially using P2MP LSP hierarchy, or instantiated using ingress replication. The intra-area P2MP LSPs may be signaled using P2MP RSVP-TE [RFC4875] or P2MP mLDP [RFC6388]. If ingress replication is used in an IGP area, then (multipoint-to-point) LDP LSPs [RFC5036] or (point-to-point) RSVP-TE LSPs [RFC3209] may be used within the IGP area. The applications/services that use such inter-area service LSPs may be BGP Multicast VPN (BGP MVPN), VPLS multicast, or global table multicast over MPLS.

The primary use case of such segmented P2MP service LSPs is when the Provider Edge (PE) routers are in different areas but in the same Autonomous System (AS) and thousands or more of PEs require P2MP connectivity. For instance, this may be the case when MPLS is pushed further to the metro edge and the metros are in different IGP areas. This may also be the case when a service provider's network comprises multiple IGP areas in a single AS, with a large number of PEs. Seamless MPLS is the industry term to address this case [SEAMLESS-MPLS]. Thus, one of the applicabilities of this document is that it describes the multicast procedures for seamless MPLS.

It is to be noted that [RFC6514] and [RFC7117] already specify procedures for building segmented inter-AS P2MP service LSPs. This document complements those procedures, as it extends the segmented P2MP LSP model such that it is applicable to inter-area P2MP service LSPs as well. In fact, an inter-AS deployment could use inter-AS segmented P2MP LSPs as specified in [RFC6514] and [RFC7117] where each intra-AS segment is constructed using inter-area segmented P2MP LSPs, as specified in this document.

2. Specification of Requirements

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 [RFC2119].

3. General Assumptions and Terminology

The reader is assumed to be familiar with MVPN procedures and terminology [RFC6513] [RFC6514] and VPLS procedures and terminology [RFC7117].

This document allows Area Border Routers (ABRs), acting as Route Reflectors, to follow the procedures specified in [SEAMLESS-MPLS] when handling the BGP Next Hop of the routes to the PEs. Specifically, when reflecting such routes from the non-backbone areas into the backbone area, the ABRs MUST set the BGP Next Hop to their own loopback addresses (next-hop-self), while when reflecting such routes from the backbone area into the non-backbone areas, the ABRs SHOULD NOT change the BGP Next Hop addresses (next-hop-unchanged).

While this document allows ABRs to follow the procedures specified in [SEAMLESS-MPLS], procedures specified in this document are applicable even when ABRs do not follow the procedures specified in [SEAMLESS-MPLS].

This document specifies a particular way of supporting the global table multicast service. Although the document refers to this approach simply as "global table multicast", it does not mean to imply that there are no other ways to support global table multicast.

An alternative way to support global table multicast is to use the procedures for MVPN that are specified in [RFC6514] and in this document. That alternative is discussed in more detail in [GTM]. However, that alternative is not further considered in the current document.

This document assumes that, in the context of global table multicast, ABRs do not carry routes to the destinations external to their own AS. Furthermore, in the context of global table multicast, this document assumes that an Autonomous System Border Router (ASBR), when re-advertising into Internal BGP (IBGP) routes received from an external speaker (received via External BGP (EBGP)), may not change the BGP Next Hop to self.

Within an AS, a P2MP service LSP is partitioned into three segments: ingress area segment, backbone area segment, and egress area segment. Within each area, a segment is carried over an intra-area P2MP LSP or instantiated using ingress replication.

When intra-area P2MP LSPs are used to instantiate the intra-area segments, there could be either 1:1 or n:1 mapping between intra-area segments of the inter-area P2MP service LSP and a given intra-area P2MP LSP. The latter is realized using P2MP LSP hierarchy with upstream-assigned labels [RFC5331]. For simplicity of presentation, we assume that P2MP LSP hierarchy is used even with 1:1 mapping; in which case, an Implicit NULL is used as the upstream-assigned label.

When intra-area segments of the inter-area P2MP service LSP are instantiated using ingress replication, multiple such segments may be carried in the same P2P RSVP-TE or MP2P LDP LSP. This can be achieved using downstream-assigned labels alone.

The ingress area segment of a P2MP service LSP is rooted at a PE (or at an ASBR in the case where the P2MP service LSP spans multiple ASes). The leaves of this segment are other PEs/ASBRs and ABRs in the same area as the root PE.

The backbone area segment is rooted at either an ABR that is connected to the ingress area (ingress ABR), an ASBR if the ASBR is present in the backbone area, or a PE if the PE is present in the backbone area. The backbone area segment has its leaf ABRs that are connected to the egress area(s) or PEs in the backbone area, or ASBRs in the backbone area.

The egress area segment is rooted at an ABR in the egress area (egress ABR), and has its leaf PEs and ASBR in that egress area (the latter covers the case where the P2MP service LSP spans multiple ASes). For a given P2MP service LSP, note that there may be more than one backbone segment, each rooted at its own ingress ABR, and more than one egress area segment, each rooted at its own egress ABR.

This document uses the term "A-D routes" for "auto-discovery routes".

An implementation that supports this document MUST implement the procedures described in the following sections to support inter-area P2MP segmented service LSPs.

4. Inter-Area P2MP Segmented Next-Hop Extended Community

This document defines a new Transitive IPv4-Address-Specific Extended Community Sub-Type: "Inter-Area P2MP Next-Hop". This document also defines a new BGP Transitive IPv6-Address-Specific Extended Community Sub-Type: "Inter-Area P2MP Next-Hop".

A PE, an ABR, or an ASBR constructs the Inter-Area P2MP Segmented Next-Hop Extended Community as follows:

- The Global Administrator field MUST be set to an IP address of the PE, ABR, or ASBR that originates or advertises the route carrying the P2MP Next-Hop Extended Community. For example this address may be the loopback address or the PE, ABR, or ASBR that advertises the route.

- The Local Administrator field MUST be set to 0.

  If the Global Administrator field is an IPv4 address, the IPv4-Address-Specific Extended Community is used; if the Global Administrator field is an IPv6 address, the IPv6-Address-Specific Extended Community is used.

  The detailed usage of these Extended Communities is described in the following sections.

5. Discovering P2MP FEC of Inter-Area P2MP Service LSP

Each inter-area P2MP service LSP has associated with it P2MP Forwarding Equivalence Class (FEC). The egress PEs need to learn this P2MP FEC in order to initiate the creation of the egress area segment of the P2MP inter-area service LSP.

The P2MP FEC of the inter-area P2MP LSP is learned by the egress PEs either by configuration or based on the application-specific procedures (e.g., MVPN-specific procedures or VPLS-specific procedures).

5.1. BGP MVPN

Egress PEs and/or ASBRs discover the P2MP FEC of the service LSPs used by BGP MVPN using the Inclusive Provider Multicast Service Interface (I-PMSI) or Selective PMSI (S-PMSI) A-D routes that are originated by the ingress PEs or ASBRs following the procedures of [RFC6514], along with modifications as described in this document. The Network Layer Reachability Information (NLRI) of such routes encodes the P2MP FEC.

The procedures in this document require that at least one ABR in a given IGP area act as a Route Reflector for MVPN A-D routes. Such a Router Reflector is responsible for re-advertising MVPN A-D routes across area boundaries. When re-advertising these routes across area boundaries, this Route Reflector MUST follow the procedures in this document. Note that such a Route Reflector may also re-advertise MVPN A-D routes within the same area; in which case, it follows the plain BGP Route Reflector procedures [RFC4456].

5.1.1. Routes Originated by PE or ASBR

The "Leaf Information Required" flag MUST be set in the PMSI Tunnel attribute carried in the MVPN A-D routes, when originated by the ingress PEs or ASBRs, except for the case where (a) as a matter of policy (provisioned on the ingress PEs or ASBRs) there is no aggregation of ingress area segments of the service LSPs and (b) mLDP is used as the protocol to establish intra-area transport LSPs in the ingress area. Before any Leaf A-D route is advertised by a PE or ABR in the same area, as described in the following sections, an I-PMSI/S-PMSI A-D route is advertised either with an explicit Tunnel Type and Tunnel Identifier in the PMSI Tunnel attribute, if the Tunnel Identifier has already been assigned, or with a special Tunnel Type of "No tunnel information present" otherwise.

5.1.2. Routes Re-advertised by PE or ASBR

When the I-PMSI/S-PMSI A-D routes are re-advertised by an ingress ABR, the "Leaf Information Required" flag MUST be set in the PMSI Tunnel attribute present in the routes, except for the case where (a) as a matter of policy (provisioned on the ingress ABR) there is no aggregation of backbone area segments of the service LSPs and (b) mLDP is used as the protocol to establish intra-area transport LSPs in the backbone area. Likewise, when the I-PMSI/S-PMSI A-D routes are re-advertised by an egress ABR, the "Leaf Information Required" flag MUST be set in the PMSI Tunnel attribute present in the routes, except for the case where (a) as a matter of policy (provisioned on the egress ABR) there is no aggregation of egress area segments of the service LSPs and (b) mLDP is used as the protocol to establish intra-area transport LSPs in the egress area.

Note that the procedures in the above paragraph apply when intra-area segments are realized by either intra-area P2MP LSPs or by ingress replication.

5.1.3. Inter-Area Routes

When BGP MVPN I-PMSI or S-PMSI A-D routes are advertised or propagated to signal inter-area P2MP service LSPs, to indicate that these LSPs should be segmented using the procedures specified in this document, these routes MUST carry the Inter-Area P2MP Segmented Next-Hop Extended Community. This Extended Community MUST be included in the I-PMSI/S-PMSI A-D route by the PE that originates such a route, or an ASBR that re-advertises such a route into its own

AS. The Global Administrator field in this Extended Community MUST be set to the advertising PE or ASBR's IP address. This Extended Community MUST also be included by ABRs as they re-advertise such routes. An ABR MUST set the Global Administrator field of the Inter-Area P2MP Segmented Next-Hop Extended Community to its own IP address. Presence of this Extended Community in the I-PMSI/S-PMSI A-D routes indicates to ABRs and PEs/ASBRs that they have to follow the procedures in this document when these procedures differ from those in [RFC6514].

If an ASBR receives from an IBGP peer an I-PMSI or S-PMSI A-D route that carries the Inter-Area P2MP Segmented Next-Hop Extended Community, then before re-advertising this route to an EBGP peer, the ASBR SHOULD remove this Extended Community from the route.

Suppose an ASBR receives an I-PMSI/S-PMSI A-D route from an EBGP peer, and this route carries the Inter-Area P2MP Segmented Next-Hop Extended Community. If the inter-area P2MP service LSP signaled by this route should not be segmented, then before re-advertising this route to its IBGP peers, the ASBR MUST remove this Extended Community from the route.

To avoid requiring ABRs to participate in the propagation of C-multicast routes, this document requires that ABRs MUST NOT modify the BGP Next Hop when re-advertising Inter-AS I-PMSI A-D routes. For consistency, this document requires that ABRs MUST NOT modify the BGP Next Hop when re-advertising either Intra-AS or Inter-AS I-PMSI/S-PMSI A-D routes. The egress PEs may advertise the C-multicast routes to RRs that are different than the ABRs. However, ABRs can still be configured to be the Route Reflectors for C-multicast routes; in which case, they will participate in the propagation of C-multicast routes.

5.2. LDP VPLS with BGP Auto-discovery or BGP VPLS

Egress PEs discover the P2MP FEC of the service LSPs used by VPLS, using the VPLS A-D routes that are originated by the ingress PEs [RFC4761] [RFC6074] or VPLS S-PMSI A-D routes that are originated by the ingress PEs [RFC7117]. The NLRI of such routes encodes the P2MP FEC.

5.2.1. Routes Originated by PE or ASBR

The "Leaf Information Required" flag MUST be set in the PMSI Tunnel attribute carried in the VPLS A-D routes or VPLS S-PMSI A-D routes, when originated by the ingress PEs or ASBRs, except for the case where (a) as a matter of policy (provisioned on the ingress PEs or ASBRs) there is no aggregation of ingress area segments of the service LSPs and (b) mLDP is used as the protocol to establish intra-area transport LSPs in the ingress area. Before any Leaf A-D route is advertised by a PE or ABR in the same area, as described in the following sections, a VPLS/S-PMSI A-D route is advertised either with an explicit Tunnel Type and Tunnel Identifier in the PMSI Tunnel attribute, if the Tunnel Identifier has already been assigned, or with a special Tunnel Type of "No tunnel information present" otherwise.

5.2.2. Routes Re-advertised by PE or ASBR

When the VPLS/S-PMSI A-D routes are re-advertised by an ingress ABR, the "Leaf Information Required" flag MUST be set in the PMSI Tunnel attribute present in the routes, except for the case where (a) as a matter of policy (provisioned on the ingress ABR) there is no aggregation of backbone area segments of the service LSPs and (b) mLDP is used as the protocol to establish intra-area transport LSPs in the backbone area. Likewise, when the VPLS/S-PMSI A-D routes are re-advertised by an egress ABR, the "Leaf Information Required" flag MUST be set in the PMSI Tunnel attribute present in the routes, except for the case where (a) as a matter of policy (provisioned on the egress ABR) there is no aggregation of egress area segments of the service LSPs and (b) mLDP is used as the protocol to establish intra-area transport LSPs in the egress area.

5.2.3. Inter-Area Routes

When VPLS A-D routes or S-PMSI A-D routes are advertised or propagated to signal inter-area P2MP service LSPs, to indicate that these LSPs should be segmented using the procedures specified in this document, these routes MUST carry the Inter-Area P2MP Segmented Next-Hop Extended Community. This Extended Community MUST be included in the A-D route by the PE or ASBR that originates such a route, and the Global Administrator field MUST be set to the advertising PE or ASBR's IP address. This Extended Community MUST also be included by ABRs as they re-advertise such routes. An ABR MUST set the Global Administrator field of the Inter-Area P2MP Segmented Next-Hop Extended Community to its own IP address. Presence of this Extended Community in the I-PMSI/S-PMSI A-D routes indicates to ABRs and PEs/ASBRs that they have to follow the procedures in this document when these procedures differ from those in [RFC7117].

Note that the procedures in the above paragraph apply when intra-area segments are realized by either intra-area P2MP LSPs or by ingress replication.

The procedures in this document require that at least one ABR in a given area act as a Route Reflector for VPLS A-D routes. Such a Router Reflector is responsible for re-advertising VPLS A-D routes across areas boundaries. When re-advertising these routes across areas boundaries, this Route Reflector MUST follow the procedures in this document. Note that such a Route Reflector may also re-advertise VPLS A-D routes within the same area; in which case, it follows plain BGP Route Reflector procedures [RFC4456].

When re-advertising VPLS A-D routes, a Route Reflector MUST NOT modify the BGP Next Hop of these routes.

5.3. Global Table Multicast over MPLS

This section describes how the egress PEs discover the P2MP FEC when the application is global table multicast over an MPLS-capable infrastructure. In the rest of the document, we will refer to this application as "global table multicast".

When Protocol Independent Multicast - Sparse Mode (PIM-SM) is used for non-bidirectional ASM ("Any Source Multicast") group addresses, this document refers to this as "PIM-SM in ASM mode".

In the case where global table multicast uses PIM-SM in ASM mode, the following assumes that an inter-area P2MP service LSP could be used to carry traffic either on a shared (*,G) or a source (S,G) tree.

An egress PE learns the (S/*,G) of a multicast stream as a result of receiving IGMP or PIM messages on one of its IP multicast interfaces. This (S/*,G) forms the P2MP FEC of the inter-area P2MP service LSP. For each such P2MP FEC, there MAY exist a distinct inter-area P2MP service LSP, or multiple such FECs MAY be carried over a single P2MP service LSP using a wildcard (*,*) S-PMSI [RFC6625].

Note that this document does not require the use of (*,G) inter-area P2MP service LSPs when global table multicast uses PIM-SM in ASM mode. In fact, PIM-SM in ASM mode may be supported entirely by using only (S,G) inter-area P2MP service LSPs.

6. Egress PE/ASBR Signaling Procedures

This section describes the egress PE/ASBR procedures for constructing segmented inter-area P2MP LSPs. The procedures in this section apply irrespective of whether the egress PE/ASBR is in a leaf IGP area, the backbone area, or even in the same IGP area as the ingress PE/ASBR.

An egress PE/ASBR applies procedures specified in this section to MVPN I-PMSI or S-PMSI A-D routes only if these routes carry the Inter-Area P2MP Segmented Next-Hop Extended Community. An egress PE applies procedures specified in this section to VPLS A-D routes or VPLS S-PMSI A-D routes only if these routes carry the Inter-Area P2MP Segmented Next-Hop Extended Community.

In order to support global table multicast, an egress PE MUST be auto-configured to import routes that carry an AS-specific Route Target Extended Community ([RFC4360]) with the Global Administrator field set to the AS of the PE and the Local Administrator field set to 0.

Once an egress PE/ASBR discovers the P2MP FEC of an inter-area segmented P2MP service LSP, it MUST propagate this P2MP FEC in BGP in order to construct the segmented inter-area P2MP service LSP. This propagation uses BGP Leaf A-D routes.

6.1. Determining the Upstream ABR/PE/ASBR (Upstream Node)

An egress PE/ASBR discovers the P2MP FEC of an inter-area P2MP segmented service LSP as described in Section 5. Once the egress PE/ASBR discovers this P2MP FEC, it MUST determine the upstream node to reach such a FEC. If the egress PE/ASBR and the ingress PE/ASBR are not in the same area, and the egress PE/ASBR is not in the backbone IGP area, then this upstream node would be an egress ABR. If the egress PE/ASBR is in the backbone area and the ingress PE/ASBR is not in the backbone area, then this upstream node would be an ingress ABR. If the egress PE/ASBR is in the same area as the ingress PE/ASBR, then this upstream node would be the ingress PE/ASBR.

6.1.1. Upstream Node for MVPN or VPLS

If the application is MVPN or VPLS, then the upstream node's IP address is the IP address determined from the Global Administrator field of the Inter-Area P2MP Segmented Next-Hop Extended Community. As described in Section 5, this Extended Community MUST be carried in the MVPN or VPLS A-D route from which the P2MP FEC of the inter-area P2MP segmented service LSP is determined.

6.1.2. Upstream Node for Global Table Multicast

If the application is global table multicast, then the unicast routes to multicast sources/RPs SHOULD carry the "VRF Route Import" Extended Community [RFC6514] where the IP address in the Global Administrator field is set to the IP address of the PE or ASBR advertising the unicast route. The Local Administrator field of this Extended Community MUST be set to 0 (note that this is in contrast to the case of MVPN, where the Local Administrator field carries a non-zero value that identifies a particular VRF on a PE that originates VPN-IP routes). If it is not desirable to advertise the VRF Route Import Extended Community in unicast routes, then unicast routes to multicast sources/RPs MUST be advertised using the multicast Subsequent Address Family Identifier (SAFI), i.e., SAFI 2, and such routes MUST carry the VRF Route Import Extended Community.

Further, if the application is global table multicast, then the BGP unicast routes that advertise the routes to the IP addresses of PEs/ASBRs/ABRs SHOULD carry the Inter-Area P2MP Segmented Next-Hop Extended Community. The IP address in the Global Administrator field of this Extended Community MUST be set to the IP address of the PE, ASBR, or ABR advertising the unicast route. The Local Administrator field of this Extended Community MUST be set to 0. If it is not desirable to advertise the Inter-Area P2MP Segmented Next-Hop Extended Community in BGP unicast routes, then the BGP unicast routes to the IP addresses of PEs/ASBRs/ABRs MUST be advertised using the multicast SAFI, i.e., SAFI 2, and such routes MUST carry the Inter-Area P2MP Segmented Next-Hop Extended Community. The procedures for handling the BGP Next Hop attribute of SAFI 2 routes are the same as those of handling regular unicast routes and MAY follow [SEAMLESS-MPLS].

If the application is global table multicast, then in order to determine the upstream node address, the egress PE first determines the ingress PE. In order to determine the ingress PE, the egress PE determines the best route to reach the S/RP. The ingress PE address is the IP address determined from the Global Administrator field of the VRF Route Import Extended Community that is carried in this route. Then, the egress PE finds the best unicast route to reach the ingress PE. The upstream node address is the IP address determined from the Global Administrator field of the Inter-Area P2MP Segmented Next-Hop Extended Community that is carried in this route.

6.2. Originating a Leaf A-D Route

If the P2MP FEC was derived from an MVPN or VPLS A-D route, and if the route carries a PMSI Tunnel attribute with the "Leaf Information Required" flag set, then the egress PE MUST originate a Leaf A-D route.

If the P2MP FEC was derived from a global table multicast (S/*,G), and the upstream node's address is not the same as the egress PE, then the egress PE MUST originate a Leaf A-D route.

6.2.1. Leaf A-D Route for MVPN and VPLS

If the P2MP FEC was derived from MVPN or VPLS A-D routes, then the Route Key field of the Leaf A-D route contains the NLRI of the A-D route from which the P2MP FEC was derived. This follows procedures for constructing Leaf A-D routes described in [RFC6514] [RFC7117].

6.2.2. Leaf A-D Route for Global Table Multicast

If the application is global table multicast, then the MCAST-VPN NLRI of the Leaf A-D route is constructed as follows.

The Route Key field of the MCAST-VPN NLRI has the following format:

```
+-----------------------------------+
|      RD   (8 octets)              |
+-----------------------------------+
| Multicast Source Length (1 octet) |
+-----------------------------------+
|  Multicast Source (Variable)      |
+-----------------------------------+
|  Multicast Group Length (1 octet) |
+-----------------------------------+
|  Multicast Group   (Variable)     |
+-----------------------------------+
|  Ingress PE's IP Address          |
+-----------------------------------+
```

RD is set to 0 for (S,G) state and all ones for (*,G) state, Multicast Source is set to S for (S,G) state or RP for (*,G) state, Multicast Group is set to G, and Multicast Source Length and Multicast Group Length are set to either 4 or 16 (depending on whether S/RP and G are IPv4 or IPv6 addresses).

The Ingress PE's IP address is determined as described in Section 6.1.

The Originating Router's IP Address field of the MCAST-VPN NLRI is set to the address of the local PE (the PE that originates the route).

Thus, the entire MCAST-VPN NLRI of the route has the following format:

```
+-----------------------------------+
|      Route Type = 4 (1 octet)     |
+-----------------------------------+
|         Length (1 octet)          |
+-----------------------------------+
|          RD   (8 octets)          |
+-----------------------------------+
| Multicast Source Length (1 octet) |
+-----------------------------------+
|  Multicast Source (Variable)      |
+-----------------------------------+
|  Multicast Group Length (1 octet) |
+-----------------------------------+
|  Multicast Group   (Variable)     |
+-----------------------------------+
|  Ingress PE's IP Address          |
+-----------------------------------+
|  Originating Router's IP Address  |
+-----------------------------------+
```

Note that the encoding of the MCAST-VPN NLRI for the Leaf A-D routes used for global table multicast is different from the encoding used by the Leaf A-D routes originated in response to S-PMSI or I-PMSI A-D routes. A router that receives a Leaf A-D route can distinguish between these two cases by examining the third octet of the MCAST-VPN NLRI of the route. If the value of this octet is 0x01, 0x02, or 0x03, then this Leaf A-D route was originated in response to an S-PMSI or I-PMSI A-D route. If the value of this octet is either 0x00 or 0xff, and octets 3 through 10 contain either all 0x00 or all 0xff, then this is a Leaf A-D route used for global table multicast.

When the PE deletes (S,G)/(*,G) state that was created as a result of receiving PIM or IGMP messages on one of its IP multicast interfaces, if the PE previously originated a Leaf A-D route for that state, the PE SHOULD withdraw that route.

An AS with an IPv4 network may provide global table multicast service for customers that use IPv6, and an AS with an IPv6 network may provide global table multicast service for customers that use IPv4. Therefore, the address family of the Ingress PE's IP Address field and the Originating Router's IP Address field in the Leaf A-D routes used for global table multicast MUST NOT be inferred from the AFI field of the associated MP_REACH_NLRI/MP_UNREACH_NLRI attribute of these routes. The address family is determined from the length of the address (a length of 4 octets for IPv4 addresses or a length of 16 octets for IPv6 addresses).

For example, if an AS with an IPv4 network is providing IPv6 multicast service to a customer, the Ingress PE's IP Address and Originating Router's IP Address in the Leaf A-D routes used for IPv6 global table multicast will be a 4-octet IPv4 address, even though the AFI of those routes will have the value 2.

Note that the Ingress PE's IP Address and the Originating Router's IP Address must be either both IPv4 or both IPv6 addresses; thus, they must be of the same length. Since the two variable-length fields (Multicast Source and Multicast Group) in the Leaf A-D routes used for global table multicast have their own Length field, from these two Length fields, and the Length field of the MCAST-VPN NLRI, one can compute the length of the Ingress PE's IP Address field and the Originating Router's IP Address field. If the computed length of these fields is neither 4 nor 16, the MP_REACH_NLRI attribute MUST be considered to be "incorrect", and MUST be handled as specified in Section 7 of [RFC4760].

6.2.3. Constructing the Rest of the Leaf A-D Route

The Next Hop field of the MP_REACH_NLRI attribute of the route SHOULD be set to the same IP address as the one carried in the Originating Router's IP Address field of the route.

When ingress replication is used to instantiate the egress area segment, the Leaf A-D route MUST carry a downstream-assigned label in the PMSI Tunnel attribute where the PMSI Tunnel Type is set to ingress replication. A PE MUST assign a distinct MPLS label for each Leaf A-D route originated by the PE.

To constrain distribution of this route, the originating PE constructs an IP-based Route Target Extended Community by placing the IP address of the upstream node in the Global Administrator field of the Extended Community, with the Local Administrator field of this community set to 0. The originating PE then adds this Route Target Extended Community to this Leaf A-D route. The upstream node's address is determined as specified in Section 6.1.

The PE then advertises this route to the upstream node.

6.3. PIM-SM in ASM Mode for Global Table Multicast

This specification allows two options for supporting global table multicast that are initiated using PIM-SM in ASM mode. The first option does not carry IP multicast shared trees over the MPLS network. The second option does carry shared trees over the MPLS network and provides support for switching from shared trees to source trees.

6.3.1. Option 1

This option does not carry IP multicast shared trees over the MPLS network. Therefore, when an (egress) PE creates (*,G) state (as a result of receiving PIM or IGMP messages on one of its IP multicast interfaces), the PE does not propagate this state using Leaf A-D routes.

6.3.1.1. Originating Source Active A-D Routes

Whenever an RP that is co-located with a PE discovers a new multicast source (as a result of receiving PIM Register or MSDP messages), the RP/PE SHOULD originate a BGP Source Active A-D route. Similarly, whenever, as a result of receiving MSDP messages, a PE that is not configured as an RP discovers a new multicast source, the PE SHOULD originate a BGP Source Active A-D route. The BGP Source Active A-D route carries a single MCAST-VPN NLRI constructed as follows:

+ The RD in this NLRI is set to 0.

+ The Multicast Source field MUST be set to S. The Multicast Source Length field is set appropriately to reflect this.

+ The Multicast Group field MUST be set to G. The Multicast Group Length field is set appropriately to reflect this.

The Route Target of this Source Active A-D route is an AS-specific Route Target Extended Community with the Global Administrator field set to the AS of the advertising RP/PE and the Local Administrator field set to 0.

To constrain distribution of the Source Active A-D route to the AS of the advertising RP, this route SHOULD carry the NO_EXPORT Community ([RFC1997]).

Using the normal BGP procedures, the Source Active A-D route is propagated to all other PEs within the AS.

Whenever the RP/PE discovers that the source is no longer active, the RP MUST withdraw the Source Active A-D route, if such a route was previously advertised by the RP.

6.3.1.2. Receiving BGP Source Active A-D Route by PE

As a result of receiving PIM or IGMP messages on one of its IP multicast interfaces, when an egress PE creates in its Tree Information Base (TIB) a new (*,G) entry with a non-empty outgoing interface list that contains one or more IP multicast interfaces, the PE MUST check if it has any Source Active A-D routes for that G. If there is such a route, S of that route is reachable via an MPLS interface, and the PE does not have (S,G) state in its TIB for (S,G) carried in the route, then the PE originates a Leaf A-D route carrying that (S,G), as specified in Section 6.2.2.

When an egress PE receives a new Source Active A-D route, the PE MUST check if its TIB contains an (*,G) entry with the same G as carried in the Source Active A-D route. If such an entry is found, S is reachable via an MPLS interface, and the PE does not have (S,G) state in its TIB for (S,G) carried in the route, then the PE originates a Leaf A-D route carrying that (S,G), as specified in Section 6.2.2.

6.3.1.3. Handling (S,G,rpt) State

Creation and deletion of (S,G,rpt) state on a PE that resulted from receiving PIM messages on one of its IP multicast interfaces do not result in any BGP actions by the PE.

6.3.2. Option 2

This option does carry IP multicast shared trees over the MPLS network. Therefore, when an egress PE creates (*,G) state (as a result of receiving PIM or IGMP messages on one of its IP multicast interfaces), the PE does propagate this state using Leaf A-D routes.

6.3.2.1. Originating Source Active A-D Routes

Whenever a PE creates an (S,G) state as a result of receiving Leaf A-D routes associated with the global table multicast service, if S is reachable via one of the IP multicast-capable interfaces, and the PE determines that G is in the PIM-SM in ASM mode range, the PE MUST originate a BGP Source Active A-D route. The route carries a single MCAST-VPN NLRI constructed as follows:

+ The RD in this NLRI is set to 0.

+ The Multicast Source field MUST be set to S. The Multicast Source Length field is set appropriately to reflect this.

+ The Multicast Group field MUST be set to G. The Multicast Group Length field is set appropriately to reflect this.

The Route Target of this Source Active A-D route is an AS-specific Route Target Extended Community with the Global Administrator field set to the AS of the advertising PE and the Local Administrator field set to 0.

To constrain distribution of the Source Active A-D route to the AS of the advertising PE, this route SHOULD carry the NO_EXPORT Community [RFC1997].

Using the normal BGP procedures, the Source Active A-D route is propagated to all other PEs within the AS.

Whenever the PE deletes the (S,G) state that was previously created as a result of receiving a Leaf A-D route for (S,G), the PE that deletes the state MUST also withdraw the Source Active A-D route, if such a route was advertised when the state was created.

6.3.2.2. Receiving BGP Source Active A-D Route

Procedures for receiving BGP Source Active A-D routes are the same as with Option 1.

6.3.2.3. Pruning Sources Off the Shared Tree

After receiving a new Source Active A-D route for (S,G), if a PE determines that (a) it has the (*,G) entry in its TIB, (b) the incoming interface (iif) list for that entry contains one of the IP interfaces, (c) an MPLS LSP is in the outgoing interface (oif) list for that entry, and (d) the PE does not originate a Leaf A-D route for (S,G), then the PE MUST transition the (S,G,rpt) downstream state to the Prune state. [Conceptually the PIM state machine on the PE will act "as if" it had received Prune(S,G,Rpt) from some other PE, without actually having received one.] Depending on the (S,G,rpt) state on the iifs, this may result in the PE using PIM procedures to prune S off the Shared (*,G) tree.

Transitioning the state machine to the Prune state SHOULD be done after a delay that is controlled by a timer. The value of the timer MUST be configurable. The purpose of this timer is to ensure that S is not pruned off the shared tree until all PEs have had time to receive the Source Active A-D route for (S,G).

The PE MUST keep the (S,G,rpt) downstream state machine in the Prune state for as long as (a) the outgoing interface list (oif) for (*,G) contains an MPLS LSP, (b) the PE has at least one Source Active A-D route for (S,G), and (c) the PE does not originate the Leaf A-D route for (S,G). Once any of these conditions become no longer valid, the PE MUST transition the (S,G,rpt) downstream state machine to the NoInfo state.

Note that, except for the scenario described in the first paragraph of this section, in all scenarios relying solely on PIM procedures on the PE is sufficient to ensure the correct behavior when pruning sources off the shared tree.

6.3.2.4. More on Handling (S,G,rpt) State

Creation and deletion of (S,G,rpt) state on a PE that resulted from receiving PIM messages on one of its IP multicast interfaces do not result in any BGP actions by the PE.

7. Egress ABR Procedures

This section describes the egress ABR procedures for constructing segmented inter-area P2MP LSPs.

7.1. Handling Leaf A-D Route on Egress ABR

When an egress ABR receives a Leaf A-D route and the Route Target Extended Community carried by the route contains the IP address of this ABR, the following procedures will be executed.

If the value of the third octet of the MCAST-VPN NLRI of the received Leaf A-D route is either 0x01, 0x02, or 0x03, this indicates that the Leaf A-D route was originated in response to an S-PMSI or I-PMSI A-D route (see Section 6.2.2). In this case, the egress ABR MUST find an S-PMSI or I-PMSI route whose NLRI has the same value as the Route Key field of the received Leaf A-D route. If such a matching route is found, then the Leaf A-D route MUST be accepted. If the Leaf A-D route is accepted and if it is the first Leaf A-D route update for the Route Key field in the route, or the withdrawal of the last Leaf A-D route for the Route Key field, then the following procedures will be executed.

If the RD of the received Leaf A-D route is set to all zeros or all ones, then the received Leaf A-D route is for the global table multicast service.

If the received Leaf A-D route is the first Leaf A-D route update for the Route Key field carried in the route, then the egress ABR originates a Leaf A-D route, whose MCAST-VPN NLRI is constructed as follows.

The Route Key field of the MCAST-VPN NLRI is the same as the Route Key field of the MCAST-VPN NLRI of the received Leaf A-D route. The Originating Router's IP Address field of the MCAST-VPN NLRI is set to the address of the local ABR (the ABR that originates the route).

The Next Hop field of the MP_REACH_NLRI attribute of the route SHOULD be set to the same IP address as the one carried in the Originating Router's IP Address field of the route.

To constrain distribution of this route, the originating egress ABR constructs an IP-based Route Target Extended Community by placing the IP address of the upstream node in the Global Administrator field of the Extended Community, with the Local Administrator field of this Extended Community set to 0, and sets the Extended Communities attribute of this Leaf A-D route to that Extended Community.

The upstream node's IP address is the IP address determined from the Global Administrator field of the Inter-Area P2MP Segmented Next-Hop Extended Community, where this Extended Community is obtained as follows. When the Leaf A-D route is for MVPN or VPLS, this Extended Community is the one carried in the I-PMSI/S-PMSI A-D route that matches the Leaf A-D route. When the Leaf A-D route is for global table multicast, this Extended Community is the one carried in the best unicast route to the Ingress PE. The Ingress PE address is determined from the received Leaf A-D route. The best unicast route MUST first be determined from multicast SAFI, i.e., SAFI 2 routes, if present.

The ABR then advertises this Leaf A-D route to the upstream node in the backbone area.

Mechanisms specified in [RFC4684] for constrained BGP route distribution can be used along with this specification to ensure that only the needed PE/ABR will have to process a said Leaf A-D route.

When ingress replication is used to instantiate the backbone area segment, the Leaf A-D route originated by the egress ABR MUST carry a downstream-assigned label in the PMSI Tunnel attribute where the Tunnel Type is set to ingress replication. The ABR MUST assign a distinct MPLS label for each Leaf A-D route that it originates.

In order to support global table multicast, an egress ABR MUST auto-configure an import AS-based Route Target Extended Community with the Global Administrator field set to the AS of the ABR and the Local Administrator field set to 0.

If the received Leaf A-D route is the withdrawal of the last Leaf A-D route for the Route Key carried in the route, then the egress ABR must withdraw the Leaf A-D route associated with that Route Key that has been previously advertised by the egress ABR in the backbone area.

7.2. P2MP LSP as the Intra-Area LSP in the Egress Area

This section describes procedures for using intra-area P2MP LSPs in the egress area. The procedures that are common to both P2MP RSVP-TE and P2MP LDP are described first, followed by procedures that are specific to the signaling protocol.

When P2MP LSPs are used as the intra-area LSPs, note that an existing intra-area P2MP LSP may be used solely for a particular inter-area P2MP service LSP or for other inter-area P2MP service LSPs as well.

The choice between the two options is purely local to the egress ABR. The first option provides one-to-one mapping between inter-area P2MP service LSPs and intra-area P2MP LSPs; the second option provides many-to-one mapping, thus allowing the aggregation of forwarding state.

7.2.1. Received Leaf A-D Route Is for MVPN or VPLS

If the value of the third octet of the MCAST-VPN NLRI of the received Leaf A-D route is either 0x01, 0x02, or 0x03, this indicates that the Leaf A-D route was originated in response to an MVPN or VPLS S-PMSI or I-PMSI A-D route (see Section 6.2.2). In this case, the ABR MUST re-advertise in the egress area the MVPN/VPLS A-D route that matches the Leaf A-D route to signal the binding of the intra-area P2MP LSP to the inter-area P2MP service LSP. This must be done if and only if (a) such a binding hasn't already been advertised or (b) the binding has changed. The re-advertised route MUST carry the Inter-area P2MP Segmented Next-Hop Extended Community.

The PMSI Tunnel attribute of the re-advertised route specifies either an intra-area P2MP RSVP-TE LSP or an intra-area P2MP LDP LSP rooted at the ABR and MUST also carry an upstream-assigned MPLS label. The upstream-assigned MPLS label MUST be set to Implicit NULL if the mapping between the inter-area P2MP service LSP and the intra-area P2MP LSP is one-to-one. If the mapping is many-to-one, the intra-area segment of the inter-area P2MP service LSP (referred to as the "inner" P2MP LSP) is constructed by nesting the inter-area P2MP service LSP in an intra-area P2MP LSP (referred to as the "outer" intra-area P2MP LSP), by using P2MP LSP hierarchy based on upstream-assigned MPLS labels [RFC5332].

If segments of multiple MVPN or VPLS S-PMSI service LSPs are carried over a given intra-area P2MP LSP, each of these segments MUST carry a distinct upstream-assigned label, even if all these service LSPs are for (C-S/*,C-G/*)s from the same MVPN/VPLS. Therefore, an ABR maintains a Label Forwarding Information Base (LFIB) state for each such S-PMSI traversing the ABR (that applies to both the ingress and the egress ABRs).

7.2.2. Received Leaf A-D Route Is for Global Table Multicast

When the RD of the received Leaf A-D route is set to all zeros or all ones, this is the case of inter-area P2MP service LSP being associated with the global table multicast service. The procedures for this are described below.

7.2.2.1. Global Table Multicast and S-PMSI A-D Routes

This section applies only if it is desired to send a particular (S,G) or (*,G) global table multicast flow to only those egress PEs that have receivers for that multicast flow.

If the egress ABR has not previously received (and re-advertised) an S-PMSI A-D route for (S,G) or (*,G) that has been originated by an ingress PE/ASBR (see Section 9.1), then the egress ABR MUST originate an S-PMSI A-D route. The PMSI Tunnel attribute of the route MUST contain the identity of the intra-area P2MP LSP and an upstream-assigned MPLS label (although this label may be an Implicit NULL -- see Section 3). The RD, Multicast Source Length, Multicast Source, Multicast Group Length (1 octet), and Multicast Group fields of the NLRI of this route are the same as those of the received Leaf A-D route. The Originating Router's IP Address field in the S-PMSI A-D route is the same as the Ingress PE's IP Address field in the received Leaf A-D route. The Route Target of this route is an AS-specific Route Target Extended Community with the Global Administrator field set to the AS of the advertising ABR and the Local Administrator field set to 0. The route MUST carry the Inter-Area P2MP Segmented Next-Hop Extended Community. This Extended Community is constructed following the procedures in Section 4.

The egress ABR MUST advertise this route into the egress area. PEs in the egress area that participate in the global table multicast will import this route based on the Route Target carried by the route.

A PE in the egress area that originated the Leaf A-D route SHOULD join the P2MP LSP advertised in the PMSI Tunnel attribute of the S-PMSI A-D route.

7.2.2.2. Global Table Multicast and Wildcard S-PMSI A-D Routes

It may be desirable for an ingress PE to carry multiple multicast flows associated with the global table multicast over the same inter-area P2MP service LSP. This can be achieved using wildcard, i.e., (*,*) S-PMSI A-D routes [RFC6625]. An ingress PE MAY advertise a wildcard S-PMSI A-D route as described in Section 9.

If the ingress PE originates a wildcard S-PMSI A-D route, and the egress ABR receives this route from the ingress ABR, then the egress ABR either (a) MUST re-advertise this route into the egress area with the PMSI Tunnel attribute containing the identifier of the intra-area P2MP LSP in the egress area and an upstream-assigned label (note that this label may be an Implicit NULL -- see Section 3) assigned to the inter-area wildcard S-PMSI or (b) MUST be able to disaggregate traffic carried over the wildcard S-PMSI onto the egress area (S,G) or (*,G) S-PMSIs. The procedures for such disaggregation require IP processing on the egress ABRs.

If the egress ABR advertises a wildcard S-PMSI A-D route into the egress area, this route MUST carry an AS-specific Route Target Extended Community with the Global Administrator field set to the AS of the advertising ABR and the Local Administrator field set to 0. PEs in the egress area that participate in the global table multicast will import this route.

A PE in the egress area SHOULD join the P2MP LSP advertised in the PMSI Tunnel attribute of the wildcard S-PMSI A-D route if (a) the Originating Router's IP Address field in the S-PMSI A-D route has the same value as the Ingress PE's IP Address in at least one of the Leaf A-D routes for global table multicast originated by the PE and (b) the upstream ABR for the Ingress PE's IP address in that Leaf A-D route is the egress ABR that advertises the wildcard S-PMSI A-D route.

7.2.3. Global Table Multicast and the Expected Upstream Node

If the mapping between the inter-area P2MP service LSP for global table multicast service and the intra-area P2MP LSP is many-to-one, then an egress PE must be able to determine whether a given multicast packet for a particular (S,G) is received from the "expected" upstream node. The expected node is the node towards which the Leaf A-D route is sent by the egress PE. Packets received from another upstream node for that (S,G) MUST be dropped. To allow the egress PE to determine the sender upstream node, the intra-area P2MP LSP MUST be signaled with no Penultimate Hop Popping (PHP), when the mapping between the inter-area P2MP service LSP for global table multicast service and the intra-area P2MP LSP is many-to-one.

Further, the egress ABR MUST first push onto the label stack the upstream-assigned label advertised in the S-PMSI A-D route, if the label is not the Implicit NULL.

7.2.4. P2MP LDP LSP as the Intra-Area P2MP LSP

The above procedures are sufficient if P2MP LDP LSPs are used as the intra-area P2MP LSP in the egress area.

7.2.5. P2MP RSVP-TE LSP as the Intra-Area P2MP LSP

If P2MP RSVP-TE LSP is used as the intra-area LSP in the egress area, then the egress ABR can either (a) graft the leaf (whose IP address is specified in the received Leaf A-D route) into an existing P2MP LSP rooted at the egress ABR, and use that LSP for carrying traffic for the inter-area segmented P2MP service LSP or (b) originate a new P2MP LSP to be used for carrying (S,G).

When the RD of the received Leaf A-D route is all zeros or all ones, the procedures are as described in Section 7.2.2.

Note also that the SESSION object that the egress ABR would use for the intra-area P2MP LSP need not encode the P2MP FEC from the received Leaf A-D route.

7.3. Ingress Replication in the Egress Area

When ingress replication is used to instantiate the egress area segment, the Leaf A-D route advertised by the egress PE MUST carry a downstream-assigned label in the PMSI Tunnel attribute where the Tunnel Type is set to ingress replication. We will call this label the egress PE downstream-assigned label.

The egress ABR MUST forward packets received from the backbone area intra-area segment, for a particular inter-area P2MP LSP, to all the egress PEs from which the egress ABR has imported a Leaf A-D route for the inter-area P2MP LSP. A packet to a particular egress PE is encapsulated, by the egress ABR, using an MPLS label stack the bottom label of which is the egress PE downstream-assigned label. The top label is the P2P RSVP-TE or the MP2P LDP label to reach the egress PE.

Note that these procedures ensure that an egress PE always receives packets only from the upstream node expected by the egress PE.

8. Ingress ABR Procedures

When an ingress ABR receives a Leaf A-D route and the Route Target Extended Community carried by the route contains the IP address of this ABR, the ingress ABR follows the same procedures as in Section 7, with egress ABR replaced by ingress ABR, backbone area replaced by ingress area, and backbone area segment replaced by ingress area segment.

In order to support global table multicast, the ingress ABR MUST be auto-configured with an import AS-based Route Target Extended Community whose Global Administrator field is set to the AS of the ABR and whose Local Administrator field is set to 0.

8.1. P2MP LSP as the Intra-Area LSP in the Backbone Area

The procedures for binding the backbone area segment of an inter-area P2MP LSP to the intra-area P2MP LSP in the backbone area are the same as in Sections 7 and 7.2, with egress PE being replaced by egress ABR, egress ABR being replaced by ingress ABR, and egress area being replaced by backbone area. This applies to the inter-area P2MP LSPs associated with either MVPN, VPLS, or global table multicast.

It is to be noted that, in the case of global table multicast, if the backbone area uses wildcard S-PMSI, then the egress area also SHOULD use wildcard S-PMSI for global table multicast, or the egress ABRs MUST be able to disaggregate traffic carried over the wildcard S-PMSI onto the egress area (S,G) or (*,G) S-PMSIs. The procedures for such disaggregation require IP processing on the egress ABRs.

8.2. Ingress Replication in the Backbone Area

When ingress replication is used to instantiate the backbone area segment, the Leaf A-D route advertised by the egress ABR MUST carry a downstream-assigned label in the PMSI Tunnel attribute where the Tunnel Type is set to ingress replication. We will call this the egress ABR downstream-assigned label. The egress ABR MUST assign a distinct MPLS label for each Leaf A-D route originated by the ABR.

The ingress ABR MUST forward packets received from the ingress area intra-area segment, for a particular inter-area P2MP LSP, to all the egress ABRs from which the ingress ABR has imported a Leaf A-D route for the inter-area P2MP LSP. A packet to a particular egress ABR is encapsulated, by the ingress ABR, using an MPLS label stack the bottom label of which is the egress ABR downstream-assigned label.

The top label is the P2P RSVP-TE or the MP2P LDP label to reach the egress ABR.

9. Ingress PE/ASBR Procedures

This section describes the ingress PE/ASBR procedures for constructing segmented inter-area P2MP LSPs.

When an ingress PE/ASBR receives a Leaf A-D route and the Route Target Extended Community carried by the route contains the IP address of this PE/ASBR, the following procedures will be executed.

If the value of the third octet of the MCAST-VPN NLRI of the received Leaf A-D route is either 0x01, 0x02, or 0x03, this indicates that the Leaf A-D route was originated in response to an S-PMSI or I-PMSI A-D route (see Section 6.2.2). In this case, the ingress PE/ASBR MUST find an S-PMSI or I-PMSI route whose NLRI has the same value as the Route Key field of the received Leaf A-D route. If such a matching route is found, then the Leaf A-D route MUST be accepted or else it MUST be discarded. If the Leaf A-D route is accepted, then it MUST be processed as per MVPN or VPLS procedures.

If the RD of the received A-D route is set to all zeros or all ones, then the received Leaf A-D route is for the global table multicast service. If this is the first Leaf A-D route for the Route Key carried in the route, the PIM implementation MUST set its upstream (S/RP,G) state machine to Joined state for the (S/RP,G) received via a Leaf A-D route update. Likewise, if this is the withdrawal of the last Leaf A-D route whose Route Key matches a particular (S/RP,G) state, the PIM implementation MUST set its upstream (S/RP,G) state machine to Prune state for the (S/RP,G).

9.1. P2MP LSP as the Intra-Area LSP in the Ingress Area

If the value of the third octet of the MCAST-VPN NLRI of the received Leaf A-D route is either 0x01, 0x02, or 0x03 (which indicates that the Leaf A-D route was originated in response to an S-PMSI or I-PMSI A-D route), and P2MP LSP is used as the intra-area LSP in the ingress area, then the procedures for binding the ingress area segment of the inter-area P2MP LSP to the intra-area P2MP LSP in the ingress area are the same as in Sections 7 and 7.2.

When the RD of the received Leaf A-D route is all zeros or all ones, as is the case where the inter-area service P2MP LSP is associated with the global table multicast service, the ingress PE MAY originate an S-PMSI A-D route with the RD, multicast source, and multicast group fields being the same as those in the received Leaf A-D route.

Further, in the case of global table multicast, an ingress PE MAY originate a wildcard S-PMSI A-D route as per the procedures in [RFC6625] with the RD set to 0. This route may be originated by the ingress PE based on configuration or based on the import of a Leaf A-D route with the RD set to 0. If an ingress PE originates such a route, then the ingress PE MAY decide not to originate (S,G) or (*,G) S-PMSI A-D routes.

The wildcard S-PMSI A-D route MUST carry the Inter-Area P2MP Segmented Next-Hop Extended Community. This Extended Community is constructed following the procedures in Section 4.

It is to be noted that, in the case of global table multicast, if the ingress area uses wildcard S-PMSI, then the backbone area also SHOULD use wildcard S-PMSI for global table multicast, or the ingress ABRs MUST be able to disaggregate traffic carried over the wildcard S-PMSI onto the backbone area (S,G) or (*,G) S-PMSIs. The procedures for such disaggregation require IP processing on the ingress ABRs.

9.2. Ingress Replication in the Ingress Area

When ingress replication is used to instantiate the ingress area segment, the Leaf A-D route advertised by the ingress ABR MUST carry a downstream-assigned label in the PMSI Tunnel attribute where the Tunnel Type is set to ingress replication. We will call this the ingress ABR downstream-assigned label. The ingress ABR MUST assign a distinct MPLS label for each Leaf A-D route originated by the ABR.

The ingress PE/ASBR MUST forward packets received from the CE, for a particular inter-area P2MP LSP, to all the ingress ABRs from which the ingress PE/ASBR has imported a Leaf A-D route for the inter-area P2MP LSP. A packet to a particular ingress ABR is encapsulated, by the ingress PE/ASBR, using an MPLS label stack the bottom label of which is the ingress ABR downstream-assigned label. The top label is the P2P RSVP-TE or the MP2P LDP label to reach the ingress ABR.

10. Common Tunnel Type in the Ingress and Egress Areas

For a given inter-area service P2MP LSP, the PE/ASBR that is the root of that LSP controls the type of the intra-area P-tunnel that carries the ingress area segment of that LSP. However, the type of the intra-area P-tunnel that carries the backbone area segment of that LSP may be different from the type of the intra-area P-tunnels that carry the ingress area segment and the egress area segment of that LSP. In that situation, if, for a given inter-area P2MP LSP, it is desirable/necessary to use the same type of tunnel for the intra-area P-tunnels that carry the ingress area segment and for the intra-area

P-tunnels that carry the egress area segment of that LSP, then the following procedures on the ingress ABR and egress ABR provide this functionality.

When an ingress ABR re-advertises into the backbone area a BGP MVPN I-PMSI, S-PMSI A-D route, or VPLS A-D route, the ingress ABR places the PMSI Tunnel attribute of this route into the ATTR_SET BGP attribute [RFC6368], adds this attribute to the re-advertised route, and then replaces the original PMSI Tunnel attribute with a new one (note that the Tunnel Type of the new attribute may be different from the Tunnel Type of the original attribute).

When an egress ABR re-advertises into the egress area a BGP MVPN I-PMSI or S-PMSI A-D route, or VPLS A-D route, if the route carries the ATTR_SET BGP attribute [RFC6368], the ABR sets the Tunnel Type of the PMSI Tunnel attribute in the re-advertised route to the Tunnel Type of the PMSI Tunnel attribute carried in the ATTR_SET BGP attribute, and removes the ATTR_SET from the route.

11. Placement of Ingress and Egress PEs

As described in the earlier sections, procedures in this document allow the placement of ingress and egress PEs in the backbone area. They also allow the placement of egress PEs in the ingress area or the placement of ingress PEs in the egress area.

For instance, suppose that in the ingress and egress areas, a global table multicast service is being provided, and the data is being sent over PIM-based IP/GRE P-tunnels. Suppose also that it is desired to carry that data over the backbone area through MPLS P-tunnels. This can be done if the ABR connecting the ingress area to the backbone follows the procedures that this document specifies for ingress PEs providing the global table multicast service, and if the ABR connecting the egress area to the backbone follows the procedures that this document specifies for egress PEs providing the global table multicast service.

If MVPN service is being provided in the ingress and egress areas, with the MVPN data carried in PIM-based IP/GRE P-tunnels, this same technique enables the MVPN data to be carried over the backbone in MPLS P-tunnels. The PIM-based IP/GRE P-tunnels in the ingress and egress areas are treated as global table multicasts, and the ABRs provide the ingress and egress PE functionality.

12. MVPN with Virtual Hub-and-Spoke

Procedures described in this document could be used in conjunction with the Virtual Hub-and-Spoke procedures [RFC7024].

This document does not place any restrictions on the placement of Virtual Hubs and Virtual Spokes.

In addition to I-PMSI/S-PMSI A-D routes, the procedures described in this document are applicable to Associated-V-spoke-only I-PMSI A-D routes.

In the scenario where a V-hub, as a result of importing an S-PMSI A-D route in its VRF, originates an S-PMSI A-D route targeted to its V-spokes (as specified in Section 7.8.2 of [RFC7024]), the V-hub SHOULD be able to control via configuration whether the Inter-Area P2MP Segmented Next-Hop Extended Community, if present in the received S-PMSI A-D route, should also be carried in the originated S-PMSI A-D route. By default, if the received S-PMSI A-D route carries the Inter-Area P2MP Segmented Next-Hop Extended Community, then the originated S-PMSI A-D route SHOULD also carry this Extended Community.

13. Data Plane

This section describes the data plane procedures on the ABRs, ingress PEs, egress PEs, and transit routers.

13.1. Data Plane Procedures on ABRs

When procedures in this document are followed to signal inter-area P2MP segmented LSPs, ABRs are required to perform only MPLS switching. When an ABR receives an MPLS packet from an "incoming" intra-area segment of the inter-area P2MP segmented LSP, it forwards the packet, based on MPLS switching, on to another "outgoing" intra-area segment of the inter-area P2MP segmented LSP.

If the outgoing intra-area segment is instantiated using a P2MP LSP, and if there is a one-to-one mapping between the outgoing intra-area segment and the P2MP LSP, then the ABR MUST pop the incoming segment's label stack and push the label stack of the outgoing P2MP LSP. If there is a many-to-one mapping between outgoing intra-area segments and the P2MP LSP, then the ABR MUST pop the incoming segment's label stack and first push the upstream-assigned label corresponding to the outgoing intra-area segment, if such a label has been assigned, and then push the label stack of the outgoing P2MP LSP.

If the outgoing intra-area segment is instantiated using ingress replication, then the ABR must pop the incoming segment's label stack and replicate the packet once to each leaf ABR or PE of the outgoing intra-area segment. The label stack of the packet sent to each such leaf MUST first include a downstream-assigned label assigned by the leaf to the segment, followed by the label stack of the P2P or MP2P LSP to the leaf.

13.2. Data Plane Procedures on Egress PEs

An egress PE must first identify the inter-area P2MP segmented LSP based on the incoming label stack. After this identification, the egress PE must forward the packet using the application that is bound to the inter-area P2MP segmented LSP.

Note that the application-specific forwarding for MVPN service may require the egress PE to determine whether the packets were received from the expected sender PE. When the application is MVPN, the FEC of an inter-area P2MP segmented LSP is at the granularity of the sender PE. Note that MVPN intra-AS I-PMSI A-D routes and S-PMSI A-D routes both carry the Originating Router's IP Address. Thus, an egress PE could associate the data arriving on P-tunnels advertised by these routes with the Originating Router's IP Address carried by these routes, which is the same as the ingress PE. Since a unique label stack is associated with each such FEC, the egress PE can determine the sender PE from the label stack.

Likewise for VPLS service, for the purposes of Media Access Control (MAC) learning the egress, the PE must be able to determine the "VE-ID" (VPLS Edge Device Identifier) from which the packets have been received. The FEC of the VPLS A-D routes carries the VE-ID. Thus, an egress PE could associate the data arriving on P-tunnels advertised by these routes with the VE-ID carried by these routes. Since a unique label stack is associated with each such FEC, the egress PE can perform MAC learning for packets received from a given VE-ID.

When the application is global table multicast, it is sufficient for the label stack to include identification of the sender upstream node. When P2MP LSPs are used, this requires that PHP MUST be turned off. When ingress replication is used, the egress PE knows the incoming downstream-assigned label to which it has bound a particular (S/*,G) and must accept packets with only that label for that (S/*,G).

13.3. Data Plane Procedures on Ingress PEs

An Ingress PE must perform application-specific forwarding procedures to identify the outgoing intra-area segment of an incoming packet.

If the outgoing intra-area segment is instantiated using a P2MP LSP, and if there is a one-to-one mapping between the outgoing intra-area segment and the P2MP LSP, then the ingress PE MUST encapsulate the packet in the label stack of the outgoing P2MP LSP. If there is a many-to-one mapping between outgoing intra-area segments and the P2MP LSP, then the PE MUST first push the upstream-assigned label corresponding to the outgoing intra-area segment, if such a label has been assigned, and then push the label stack of the outgoing P2MP LSP.

If the outgoing intra-area segment is instantiated using ingress replication, then the PE must replicate the packet once to each leaf ABR or PE of the outgoing intra-area segment. The label stack of the packet sent to each such leaf MUST first include a downstream-assigned label assigned by the leaf to the segment, followed by the label stack of the P2P or MP2P LSP to the leaf.

13.4. Data Plane Procedures on Transit Routers

When procedures in this document are followed to signal inter-area P2MP segmented LSPs, transit routers in each area perform only MPLS switching.

14. Support for Inter-Area Transport LSPs

This section describes OPTIONAL procedures that allow multiple (inter-area) P2MP LSPs to be aggregated into a single inter-area P2MP "transport LSP". The segmentation procedures, as specified in this document, are then applied to these inter-area P2MP transport LSPs, rather than being applied directly to the individual LSPs that are aggregated into the transport. In the following, the individual LSPs that are aggregated into a single transport LSP will be known as "service LSPs".

14.1. "Transport Tunnel" Tunnel Type

For the PMSI Tunnel attribute, we define a new Tunnel Type, called "Transport Tunnel", whose Tunnel Identifier is a tuple <Source PE Address, Local Number>. This Tunnel Type is assigned a value of 8. The Source PE Address is the address of the PE that originates the (service) A-D route that carries this attribute, and the Local Number is a number that is unique to the Source PE. The length of the Local Number part is the same as the length of the Source PE Address.

Thus, if the Source PE Address is an IPv4 address, then the Local Number part is 4 octets; if the Source PE Address is an IPv6 address, then the Local Number part is 16 octets.

14.2. Discovering Leaves of the Inter-Area P2MP Service LSP

When aggregating multiple P2MP LSPs using P2MP LSP hierarchy, determining the leaf nodes of the LSPs being aggregated is essential for being able to trade-off the overhead due to the P2MP LSPs versus suboptimal use of bandwidth due to the partial congruency of the LSPs being aggregated.

Therefore, if a PE that is a root of a given service P2MP LSP wants to aggregate this LSP with other (service) P2MP LSPs rooted at the same PE into an inter-area P2MP transport LSP, the PE should first determine all the leaf nodes of that service LSP, as well as those of the other service LSPs being aggregated.

To accomplish this, the PE sets the PMSI Tunnel attribute of the service A-D route (an I-PMSI or S-PMSI A-D route for MVPN or VPLS service) associated with that LSP as follows. The Tunnel Type is set to "No tunnel information present", and the "Leaf Information Required" flag is set to 1. The route MUST NOT carry the Inter-Area P2MP Segmented Next-Hop Extended Community. In contrast to the procedures for the MVPN and VPLS A-D routes described so far, the Route Reflectors that participate in the distribution of this route need not be ABRs.

14.3. Discovering P2MP FEC of P2MP Transport LSP

Once the ingress PE determines all the leaves of a given P2MP service LSP, the PE (using some local criteria) selects a particular inter-area transport P2MP LSP to be used for carrying the (inter-area) service P2MP LSP.

Once the PE selects the transport P2MP LSP, the PE (re-)originates the service A-D route. The PMSI Tunnel attribute of this route now carries the Tunnel Identifier of the selected transport LSP, with the Tunnel Type set to "Transport Tunnel". If the transport LSP carries multiple P2MP service LSPs, then the MPLS Label field in the attribute carries an upstream-assigned label assigned by the PE that is bound to the P2MP FEC of the inter-area P2MP service LSP. Otherwise, this field is set to Implicit NULL.

As described earlier, this service A-D route MUST NOT carry the Inter-Area P2MP Segmented Next-Hop Extended Community, and the Route Reflectors that participate in the distribution of this route need not be ABRs.

14.4. Egress PE Procedures for P2MP Transport LSP

When an egress PE receives and accepts an MVPN or VPLS service A-D route, if the "Leaf Information Required" flag in the PMSI Tunnel attribute of the received A-D route is set to 1, and the route does not carry the Inter-Area P2MP Segmented Next-Hop Extended Community, then the egress PE, following the "regular" MVPN or VPLS procedures associated with the received A-D route (as specified in [RFC6514] and [RFC7117]), originates a Leaf A-D route.

In addition, if the Tunnel Type in the PMSI Tunnel attribute of the received service A-D route is set to "Transport Tunnel", the egress PE originates yet another Leaf A-D route.

The format of the Route Key field of the MCAST-VPN NLRI of this additional Leaf A-D route is the same as defined in Section 6.2.2. The Route Key field of the MCAST-VPN NLRI of this route is constructed as follows:

RD (8 octets) - set to 0

Multicast Source Length, Multicast Source - constructed from the Source PE Address part of the Tunnel Identifier carried in the PMSI Tunnel attribute of the received service S-PMSI A-D route.

Multicast Group Length, Multicast Group - constructed from the Local Number part of the Tunnel Identifier carried in the PMSI Tunnel attribute of the received service S-PMSI A-D route.

Ingress PE IP Address - set to the Source PE Address part of the Tunnel Identifier carried in the PMSI Tunnel attribute of the received service S-PMSI A-D route.

The egress PE, when determining the upstream ABR, follows the procedures specified in Section 6.1 for global table multicast.

The egress PE constructs the rest of the Leaf A-D route following the procedures specified in Section 6.2.3.

From that point on we follow the procedures used for the Leaf A-D routes for global table multicast, as outlined below.

14.5. ABRs and Ingress PE Procedures for P2MP Transport LSP

In this section, we specify ingress and egress ABRs, as well as ingress PE procedures for P2MP transport LSPs.

When an egress ABR receives the Leaf A-D route, and P2MP LSP is used to instantiate the egress area segment of the inter-area transport LSP, the egress ABR will advertise into the egress area an S-PMSI A-D route. This route is constructed following the procedures in Section 7.2.2.1. The egress PE(s) will import this route.

The egress ABR will also propagate, with appropriate modifications, the received Leaf A-D route into the backbone area. This is irrespective of whether the egress area segment is instantiated using P2MP LSP or ingress replication.

If P2MP LSP is used to instantiate the backbone area segment of the inter-area transport LSP, then an ingress ABR will advertise into the backbone area an S-PMSI A-D route. This route is constructed following the procedures in Section 7.1.2.1. The egress ABR(s) will import this route.

The ingress ABR will also propagate, with appropriate modifications, the received Leaf A-D route into the ingress area towards the ingress/root PE. This is irrespective of whether the backbone area segment is instantiated using P2MP LSP or ingress replication.

Finally, if P2MP LSP is used to instantiate the ingress area segment, the ingress PE will advertise into the ingress area an S-PMSI A-D route with the RD, multicast source, and multicast group fields being the same as those in the received Leaf A-D route. The PMSI Tunnel attribute of this route contains the identity of the intra-area P2MP LSP used to instantiate the ingress area segment, and an upstream-assigned MPLS label. The ingress ABR(s) and other PE(s) in the ingress area, if any, will import this route. The ingress PE will use the (intra-area) P2MP LSP advertised in this route for carrying traffic associated with the original service A-D route advertised by the PE.

14.6. Discussion

Use of inter-area transport P2MP LSPs, as described in this section, creates a level of indirection between (inter-area) P2MP service LSPs, and intra-area transport LSPs that carry the service LSPs. Rather than segmenting (inter-area) service P2MP LSPs, and then aggregating (intra-area) segments of these service LSPs into intra-area transport LSPs, this approach first aggregates multiple (inter-area) service P2MP LSPs into a single inter-area transport P2MP LSP, then segments such inter-area transport P2MP LSPs, and then aggregates (intra-area) segments of these inter-area transport LSPs into intra-area transport LSPs.

On one hand, this approach could result in reducing the state (and the overhead associated with maintaining the state) on ABRs. This is because instead of requiring ABRs to maintain state for individual P2MP service LSPs, ABRs would need to maintain state only for the inter-area P2MP transport LSPs. Note, however, that this reduction is possible only if a single inter-area P2MP transport LSP aggregates more than one (inter-area) service LSP. In the absence of such aggregation, use of inter-area transport LSPs provides no benefits, yet results in extra overhead. And while such aggregation does allow reduced state on ABRs, it comes at a price, as described below.

As we mentioned before, aggregating multiple P2MP service LSPs into a single inter-area P2MP transport LSP requires the PE rooted at these LSPs to determine all the leaf nodes of the service LSPs to be aggregated. This means that the root PE has to track all the leaf PEs of these LSPs. In contrast, when one applies segmentation procedures directly to the P2MP service LSPs, the root PE has to track only the leaf PEs in its own IGP area, plus the ingress ABR(s). Likewise, an ingress ABR has to track only the egress ABRs. Finally, an egress ABR has to track only the leaf PEs in its own area. Therefore, while the total overhead of leaf tracking due to the P2MP service LSPs is about the same in both approaches, the distribution of this overhead is different. Specifically, when one uses inter-area P2MP transport LSPs, this overhead is concentrated on the ingress PE. When one applies segmentation procedures directly to the P2MP service LSPs, this overhead is distributed among the ingress PE and ABRs.

Moreover, when one uses inter-area P2MP transport LSPs, ABRs have to bear the overhead of leaf tracking for inter-area P2MP transport LSPs. In contrast, when one applies segmentation procedures directly to the P2MP service LSPs, there is no such overhead (as there are no inter-area P2MP transport LSPs).

Use of inter-area P2MP transport LSPs may also result in more bandwidth inefficiency, as compared to applying segmentation procedures directly to the P2MP service LSPs. This is because with inter-area P2MP transport LSPs the ABRs aggregate segments of inter-area P2MP transport LSPs, rather than segments of (inter-area) P2MP service LSPs. To illustrate this, consider the following example.

Assume PE1 in Area 1 is an ingress PE, with two multicast streams, (C-S1, C-G1) and (C-S2, C-G2), originated by an MVPN site connected to PE1. Assume that PE2 in Area 2 has an MVPN site with receivers for (C-S1, C-G1), PE3 and PE4 in Area 3 have an MVPN site with receivers for both (C-S1, C-G1) and (C-S2, C-G2). Finally, assume that PE5 in Area 4 has an MVPN site with receivers for (C-S2, C-G2).

When segmentation applies directly to the P2MP service LSPs, Area 2 would have just one intra-area transport LSP that would carry the egress area segment of the P2MP service LSP for (C-S1, C-G1); Area 3 would have just one intra-area transport LSP that would carry the egress area segments of both the P2MP service LSP for (C-S1, C-G1) and the P2MP service LSP for (C-S2, C-G2); Area 4 would have just one intra-area transport LSP that would carry the egress area segment of the P2MP service LSP for (C-S2, C-G2). Note that there is no bandwidth inefficiency in this case at all.

When using inter-area P2MP transport LSPs, to achieve the same state overhead on the routers within each of the egress areas (except for egress ABRs), PE1 would need to aggregate the P2MP service LSP for (C-S1, C-G1) and the P2MP service LSP for (C-S2, C-G2) into the same inter-area P2MP transport LSP. While such aggregation would reduce state on ABRs, it would also result in bandwidth inefficiency, as (C-S1, C-G1) will be delivered not just to PE2, PE3, and PE4, but also to PE5, which has no receivers for this stream. Likewise, (C-S2, C-G2) will be delivered not just to PE3, PE4, and PE5, but also to PE2, which has no receivers for this stream.

15. IANA Considerations

This document defines a new BGP Extended Community called "Inter-Area P2MP Segmented Next-Hop" (see Section 4). This may be either a Transitive IPv4-Address-Specific Extended Community or a Transitive IPv6-Address-Specific Extended Community. IANA has assigned the value 0x12 in the "Transitive IPv4-Address-Specific Extended Community Sub-Types" registry, and IANA has assigned the value 0x0012 in the "Transitive IPv6-Address-Specific Extended Community Types" registry. This document is the reference for both code points.

IANA has assigned the value 0x08 in the "P-Multicast Service Interface Tunnel (PMSI Tunnel) Tunnel Types" registry [RFC7385] as "Transport Tunnel" (see Section 14).

This document makes use of a Route Distinguisher whose value is all ones. The two-octet type field of this Route Distinguisher thus has the value 65535. IANA has assigned this value in the "Route Distinguisher Type Field" registry as "For Use Only in Certain Leaf A-D Routes", with this document as the reference.

16. Security Considerations

Procedures described in this document are subject to security threats similar to those experienced by any MPLS deployment. It is recommended that baseline security measures are considered as described in "Security Framework for MPLS and GMPLS Networks"

[RFC5920], in the mLDP specification [RFC6388], and in the P2MP RSVP-TE specification [RFC3209]. The security considerations of [RFC6513] are also applicable.

17. References

17.1. Normative References

[RFC1997] Chandra, R., Traina, P., and T. Li, "BGP Communities Attribute", RFC 1997, DOI 10.17487/RFC1997, August 1996, <http://www.rfc-editor.org/info/rfc1997>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <http://www.rfc-editor.org/info/rfc2119>.

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, DOI 10.17487/RFC3209, December 2001, <http://www.rfc-editor.org/info/rfc3209>.

[RFC4360] Sangli, S., Tappan, D., and Y. Rekhter, "BGP Extended Communities Attribute", RFC 4360, DOI 10.17487/RFC4360, February 2006, <http://www.rfc-editor.org/info/rfc4360>.

[RFC4456] Bates, T., Chen, E., and R. Chandra, "BGP Route Reflection: An Alternative to Full Mesh Internal BGP (IBGP)", RFC 4456, DOI 10.17487/RFC4456, April 2006, <http://www.rfc-editor.org/info/rfc4456>.

[RFC4684] Marques, P., Bonica, R., Fang, L., Martini, L., Raszuk, R., Patel, K., and J. Guichard, "Constrained Route Distribution for Border Gateway Protocol/MultiProtocol Label Switching (BGP/MPLS) Internet Protocol (IP) Virtual Private Networks (VPNs)", RFC 4684, DOI 10.17487/RFC4684, November 2006, <http://www.rfc-editor.org/info/rfc4684>.

[RFC4760] Bates, T., Chandra, R., Katz, D., and Y. Rekhter, "Multiprotocol Extensions for BGP-4", RFC 4760, DOI 10.17487/RFC4760, January 2007, <http://www.rfc-editor.org/info/rfc4760>.

[RFC4761] Kompella, K., Ed., and Y. Rekhter, Ed., "Virtual Private LAN Service (VPLS) Using BGP for Auto-Discovery and Signaling", RFC 4761, DOI 10.17487/RFC4761, January 2007, <http://www.rfc-editor.org/info/rfc4761>.

[RFC4875] Aggarwal, R., Ed., Papadimitriou, D., Ed., and S. Yasukawa, Ed., "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, <http://www.rfc-editor.org/info/rfc4875>.

[RFC5036] Andersson, L., Ed., Minei, I., Ed., and B. Thomas, Ed., "LDP Specification", RFC 5036, DOI 10.17487/RFC5036, October 2007, <http://www.rfc-editor.org/info/rfc5036>.

[RFC5331] Aggarwal, R., Rekhter, Y., and E. Rosen, "MPLS Upstream Label Assignment and Context-Specific Label Space", RFC 5331, DOI 10.17487/RFC5331, August 2008, <http://www.rfc-editor.org/info/rfc5331>.

[RFC5332] Eckert, T., Rosen, E., Ed., Aggarwal, R., and Y. Rekhter, "MPLS Multicast Encapsulations", RFC 5332, DOI 10.17487/RFC5332, August 2008, <http://www.rfc-editor.org/info/rfc5332>.

[RFC6074] Rosen, E., Davie, B., Radoaca, V., and W. Luo, "Provisioning, Auto-Discovery, and Signaling in Layer 2 Virtual Private Networks (L2VPNs)", RFC 6074, DOI 10.17487/RFC6074, January 2011, <http://www.rfc-editor.org/info/rfc6074>.

[RFC6368] Marques, P., Raszuk, R., Patel, K., Kumaki, K., and T. Yamagata, "Internal BGP as the Provider/Customer Edge Protocol for BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 6368, DOI 10.17487/RFC6368, September 2011, <http://www.rfc-editor.org/info/rfc6368>.

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

[RFC6513] Rosen, E., Ed., and R. Aggarwal, Ed., "Multicast in MPLS/BGP IP VPNs", RFC 6513, DOI 10.17487/RFC6513, February 2012, <http://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, <http://www.rfc-editor.org/info/rfc6514>.

[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, <http://www.rfc-editor.org/info/rfc6625>.

[RFC7117] Aggarwal, R., Ed., Kamite, Y., Fang, L., Rekhter, Y., and C. Kodeboniya, "Multicast in Virtual Private LAN Service (VPLS)", RFC 7117, DOI 10.17487/RFC7117, February 2014, <http://www.rfc-editor.org/info/rfc7117>.

[RFC7385] Andersson, L. and G. Swallow, "IANA Registry for P-Multicast Service Interface (PMSI) Tunnel Type Code Points", RFC 7385, DOI 10.17487/RFC7385, October 2014, <http://www.rfc-editor.org/info/rfc7385>.

17.2. Informative References

[GTM] Zhang, J, Giuliano, L, Rosen, E., Ed., Subramanian, K., Pacella, D., and J. Schiller, "Global Table Multicast with BGP-MVPN Procedures", Work in Progress, draft-ietf-bess-mvpn-global-table-mcast-00, November 2014.

[RFC5920] Fang, L., Ed., "Security Framework for MPLS and GMPLS Networks", RFC 5920, DOI 10.17487/RFC5920, July 2010, <http://www.rfc-editor.org/info/rfc5920>.

[RFC7024] Jeng, H., Uttaro, J., Jalil, L., Decraene, B., Rekhter, Y., and R. Aggarwal, "Virtual Hub-and-Spoke in BGP/MPLS VPNs", RFC 7024, DOI 10.17487/RFC7024, October 2013, <http://www.rfc-editor.org/info/rfc7024>.

[SEAMLESS-MPLS] Leymann, N., Ed., Decraene, B., Filsfils, C., Konstantynowicz, M., Ed., and D. Steinberg, "Seamless MPLS Architecture", Work in Progress, draft-ietf-mpls-seamless-mpls-07, June 2014.

Acknowledgements

We would like to thank Loa Andersson and Qin Wu for their review and comments.

Authors' Addresses

Yakov Rekhter
Juniper Networks
1194 North Mathilda Ave.
Sunnyvale, CA 94089
United States

Eric C Rosen
Juniper Networks
10 Technology Park Drive
Westford, MA 01886
United States
EMail: erosen@juniper.net

Rahul Aggarwal
EMail: raggarwa_1@yahoo.com

Thomas Morin
Orange
2, avenue Pierre-Marzin
22307 Lannion Cedex
France
EMail: thomas.morin@orange.com

Irene Grosclaude
Orange
2, avenue Pierre-Marzin
22307 Lannion Cedex
France
EMail: irene.grosclaude@orange.com

Nicolai Leymann
Deutsche Telekom AG
Winterfeldtstrasse 21
Berlin 10781
Germany
EMail: n.leymann@telekom.de

Samir Saad
AT&T
EMail: samirsaad1@outlook.com