BESS WG
Internet-Draft
Updates: 6514 6625 7524 (if approved)
Intended status: Standards Track
Expires: April 7, 2019

A. Dolganow
J. Kotalwar
Nokia
E. Rosen, Ed.
Z. Zhang
Juniper Networks, Inc.

October 4, 2018

Explicit Tracking with Wild Card Routes in Multicast VPN
draft-ietf-bess-mvpn-expl-track-11

Abstract

The MVPN specifications provide procedures to allow a multicast ingress node to invoke "explicit tracking" for a multicast flow or set of flows, thus learning the egress nodes for that flow or set of flows. However, the specifications are not completely clear about how the explicit tracking procedures work in certain scenarios. This document provides the necessary clarifications. It also specifies a new, optimized explicit tracking procedure. This new procedure allows an ingress node, by sending a single message, to request explicit tracking of each of a set of flows, where the set of flows is specified using a wildcard mechanism. This document updates RFCs 6514, 6625, and 7524.

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 April 7, 2019.

Copyright Notice

Table of Contents

1. Introduction 2
2. The Explicit Tracking Flags 5
3. Match for Tracking vs. Match for Reception 6
4. Ingress Node Initiation of Tracking 8
5. Egress Node Response to the Match for Tracking 10
5.1. General Egress Node Procedures 10
5.2. Responding to the LIR-pF Flag 11
5.3. When the Egress Node is an ABR or ASBR 14
6. Ingress Node Handling of Received Leaf A-D Routes with LIR-pF Set 15
7. Acknowledgments 16
8. IANA Considerations 16
9. Security Considerations 16
10. References 16
10.1. Normative References 16
10.2. Informative References 17
Authors' Addresses 17

1. Introduction

[RFC6513] and [RFC6514] define the "Selective Provider Multicast Service Interface Auto-Discovery route" (S-PMSI A-D route).

Per those RFCs, the S-PMSI A-D route contains a Network Layer Reachability Information (NLRI) field that identifies a particular multicast flow. In the terminology of those RFCs, each flow is denoted by (C-S,C-G), where C-S is an IP source address and C-G is an IP multicast address, both in the address space of a VPN customer. The (C-S,C-G) of the multicast flow is encoded into the NLRI field.

An S-PMSI A-D route also carries a PMSI Tunnel attribute (PTA). Typically, the PTA is used to identify a tunnel through the provider backbone network (a "P-tunnel").

By originating an S-PMSI A-D route identifying a particular multicast flow and a particular P-tunnel, a node is advertising the following:

if the node has to transmit packets of the identified flow over the backbone, it will transmit them through the identified tunnel.

[RFC6513] and [RFC6514] also define a procedure that allows an ingress node of particular multicast flow to determine the set of egress nodes that have requested to receive that flow from that ingress node. The ability of an ingress node to identify the egress nodes for a particular flow is known as "explicit tracking". An ingress node requests explicit tracking by setting a flag (the "Leaf Information Required" flag, or LIR) in the PTA. When an egress node receives an S-PMSI A-D route with LIR set, the egress node originates a Leaf A-D route whose NLRI field contains the NLRI from the corresponding S-PMSI A-D route. In this way, the egress node advertises that it has requested to receive the particular flow identified in the NLRI of that S-PMSI A-D route.

[RFC6513] and [RFC6514] also allow an ingress node to originate an S-PMSI A-D route whose PTA has LIR set, but which does not identify any P-tunnel. This mechanism can be used when it is desired to do explicit tracking of a flow without at the same time binding that flow to a particular P-tunnel.

[RFC6625] (and other RFCs that update it) extends the specification of S-PMSI A-D routes, and allows an S-PMSI A-D route to encode a wildcard in its NLRI. Either the C-S or the C-G or both can be replaced by wildcards. These routes are known as (C-*,C-S) S-PMSI A-D routes, or as (C-S,C-*) S-PMSI A-D routes, or as (C-*,C-*) S-PMSI A-D routes, depending on whether the C-S or C-G or both have been replaced by wildcards. These routes are known jointly as "wildcard S-PMSI A-D routes".

One purpose of this document is to clarify the way that the explicit tracking procedures of [RFC6513] and [RFC6514] are applied when wildcard S-PMSI A-D routes are used.

In addition, this document addresses the following scenario, which is not addressed in [RFC6513], [RFC6514], or [RFC6625]. Suppose an ingress node originates an S-PMSI A-D route whose NLRI specifies, for example, (C-*,C-*) (i.e., both C-S and C-G are replaced by wildcards), and whose PTA identifies a particular P-tunnel. Now suppose that the ingress node wants explicit tracking for each individual flow that it transmits (following the procedures of [RFC6625]) on that P-tunnel.

In this example, if the ingress node sets LIR in the PTA of the wildcard S-PMSI A-D route, each egress node that needs to receive a flow from the ingress node will respond with a Leaf A-D route whose NLRI specifies contains the (C-*,C-*) wildcard. This allows the

ingress node to determine the set of egress nodes that are interested in receiving flows from the ingress node. However, it does not allow the ingress node to determine exactly which flows are of interest to which egress nodes.

If the ingress node needs to determine which egress nodes are interested in receiving which flows, it needs to originate an S-PMSI A-D route for each individual (C-S,C-G) flow that it is transmitting, and it needs to set LIR in the PTA of each such route. However, since all the flows are being sent through the tunnel identified in the (C-*,C-*) S-PMSI A-D route, there is no need to identify a tunnel in the PTA of each (C-S,C-G) S-PMSI A-D route. Per Section 5 of [RFC6514], the PTA of the (C-S,C-G) S-PMSI A-D routes can specify "no tunnel information present". This procedure allows explicit tracking of individual flows, even though all those flows are assigned to tunnels in wildcard S-PMSI A-D routes.

However, this procedure requires several clarifications:

- The procedures of [RFC6625] do not address the case of an S-PMSI A-D route whose NLRI contains wild cards, but whose PTA specifies "no tunnel information present".

- If it is desired to send a set of flows through the same tunnel (where that tunnel is advertised in a wildcard S-PMSI A-D route), but it is also desired to explicitly track each individual flow transmitted over that tunnel, one has to send an S-PMSI A-D route (with LIR set in the PTA) for each individual flow. It would be more optimal if the ingress node could just send a single wildcard S-PMSI A-D route binding the set of flows to a particular tunnel, and have the egress nodes respond with Leaf A-D routes for each individual flow.

- [RFC6513] and [RFC6514] support the notion of "segmented P-tunnels", where "segmentation" occurs at Autonomous System Border Routers (ASBRs); [RFC7524] extends the notion of segmented P-tunnels so that segmentation can occur at Area Border Routers (ABRs). One can think of a segmented P-tunnel as passing through a number of "segmentation domains". In each segmentation domain, a given P-tunnel has an ingress node and a set of egress nodes. The explicit tracking procedures allow an ingress node of a particular segmentation domain to determine, for a particular flow or set of flows, the egress nodes of that segmentation domain. This has given rise to two further problems:

  * The explicit tracking procedures do not allow an ingress node to "see" past the boundaries of the segmentation domain.

* The prior specifications do not make it very clear whether a segmented tunnel egress node, upon receiving an S-PMSI A-D route whose PTA specifies "no tunnel information present", is expected to forward the S-PMSI A-D route, with the same PTA, to the next segmentation domain.

These problems are addressed in Section 5.3.

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2. The Explicit Tracking Flags

[RFC6514] defines one flag in the PTA, the "Leaf Information Required" (LIR) flag, that is used for explicit tracking.

This document defines a new flag in the flags field of the PMSI Tunnel attribute. This new flag is known as the "Leaf Information Required per Flow" bit (LIR-pF). This flag may be set in the PTA of a (C-*,C-*), (C-*,C-G), or (C-S,C-*) S-PMSI A-D route. The conditions under which it should be set by the originator of the route are discussed in Section 4. The significance of the flag in a received S-PMSI A-D route is discussed in Sections 5 and 5.2.

If the LIR-pF flag is set in the PTA of an S-PMSI A-D route, the LIR flag of that PTA MUST also be set.

Note that support for the LIR-pF flag is OPTIONAL. This flag SHOULD NOT be set unless it is known that all the PEs that are to receive the route carrying the PTA support the flag. How this is known is outside the scope of this document.

The LIR-pF flag may also be set in the PTA of a Leaf A-D route. The conditions under which it should be set by the originator of the route are discussed in Section 5.2. The significance of the flag in a received Leaf A-D route is discussed in Section 6.

Use of this flag in the PTA carried by other route types is outside the scope of this document. Use of this flag in the PTA carried by an S-PMSI A-D routes whose NLRI does not contain a wildcard is outside the scope of this document.

It is worth noting what will happen if the LIR-pF flag is set in the PTA of, for example, a (C-*,C-*) S-PMSI A-D route originated by an

ingress node, but one or more of the egress nodes do not support the LIR-pF flag:

1. The ingress node will not be able to determine the complete set of egress node that are expecting a particular multicast flow from that ingress node.

2. Depending upon the tunnel type, the ingress node may send a particular multicast flow only to the egress nodes that do support the LIR-pF flag. From the perspective of egress nodes that do not support LIR-pF, certain flows may appear to be "blackholed".

It is also worth noting that it is possible for an ingress node that sets the LIR-pF flag in an S-PMSI A-D route to detect the presence of egress nodes that do not support this flag.

Since an ingress node that sets the LIR-pF flag is also REQUIRED to set the LIR flag, egress nodes that do not support the LIR-pF flag will respond, as specified in [RFC6514], to the ingress node's (C-*,C-*) S-PMSI A-D route with a Leaf A-D route operator.

As will be discussed in Section 5.2, any Leaf A-D route originated in response to an S-PMSI A-D route that has LIR-pF set will carry a PTA whose LIR-pF flag is set. If an ingress node receives a Leaf A-D route whose "route key" field corresponds to the NLRI of an S-PMSI A-D route whose PTA has LIR-pF set, but the Leaf A-D route lacks a PTA or has a PTA where LIR-pF is clear, the ingress node can conclude that the egress node originating that Leaf A-D route does not support the LIR-pF flag.

The software at the ingress node MUST detect this, and MUST have a way of alerting the operator that the deployment is not properly configured.

3. Match for Tracking vs. Match for Reception

Section 3.2 of [RFC6625] specifies a set of rules for finding the S-PMSI A-D route that is the "match for data reception" (or more simply, the "match for reception") of a given (C-S,C-G) or (C-*,C-G) state. These rules do not take into account the fact that some S-PMSI A-D routes may not be carrying PTAs at all, or may be carrying PTAs that do not identify any P-tunnel. (A PTA that does not identify any P-tunnel is one whose "tunnel type" field has been set to "no tunnel information present", as specified in Section 5 of [RFC6514].)

The rules for finding a "match for reception" in [RFC6625] are hereby modified as follows:

> When applying the rules of Section 3.2.1 or 3.2.2 of [RFC6625], it is REQUIRED to ignore any S-PMSI A-D route that has no PTA, or whose PTA specifies "no tunnel information present".

There are other RFCs that update [RFC6625] and that modify the rules for finding a "match for reception". See, e.g., [RFC7582] and [RFC7900]. When applying those modified rules, it is REQUIRED to ignore any S-PMSI A-D route that has no PTA, or whose PTA specifies "no tunnel information present".

We also introduce a new notion, the "match for tracking":

> For a given C-flow ((C-S,C-G) or (C-*,C-G)) the "match for tracking" is chosen as follows. Ignore any S-PMSI A-D route that has no PTA. Also ignore any S-PMSI A-D route whose PTA specifies "no tunnel information present", but does not have either LIR or LIR-pF set. (That is, DO NOT ignore an S-PMSI A-D route that has a PTA specifying "no tunnel information present" unless its LIR and LIR-pF bits are both clear). Then apply the rules (from [RFC6625] and other documents that update [RFC6625]) for finding the "match for reception". The result (if any) is the match for tracking".
>
> Note that the procedure for finding the match for tracking takes into account S-PMSI A-D routes whose PTAs specify "no tunnel information present" and that have either LIR or LIR-pf set. The procedure for finding the match for reception ignores such routes.

We will clarify this with a few examples. In these examples, we assume that there is only one segmentation domain. In this case, the ingress and egress nodes are Provider Edge (PE) routers.

Suppose a given PE router, PE1, has chosen PE2 as the "upstream PE" ([RFC6513]) for a given flow (C-S1,C-G1). And suppose PE1 has installed the following two routes that were originated by PE2:

- Route1: A (C-*,C-*) S-PMSI A-D route, whose PTA specifies a tunnel.

- Route2: A (C-S1,C-G1) S-PMSI A-D route, whose PTA specifies "no tunnel information present" and has LIR set.

Route1 is (C-S1,C-G1)'s match for reception, and Route2 is (C-S1,C-G1)'s match for tracking.

Continuing this example, suppose:

- PE1 has chosen PE2 as the upstream PE for a different flow, (C-S2,C-G2).

- PE2 has not originated an S-PMSI A-D route for (C-S2,C-G2).

In this case, PE1 would consider Route1 to be (C-S2,C-G2)'s match for tracking as well as its match for reception.

Also note that if a match for tracking does not have the LIR flag or the LIR-pF flag set, no explicit tracking information will be generated. See Section 5.

As another example, suppose PE1 has installed the following two routes that were originated by PE2:

- Route1: A (C-*,C-*) S-PMSI A-D route (irrespective of whether the PTA specifies a tunnel)

- Route2: A (C-S1,C-G1) S-PMSI A-D route whose PTA specifies a tunnel.

Then Route2 is both the "match for reception" and the "match for tracking" for (C-S1,C-G1).

Note that for a particular C-flow, PE1's match for reception might be the same route as its match for tracking, or its match for reception might be a "less specific" route than its match for tracking. But its match for reception can never be a "more specific" route than its match for tracking.

4. Ingress Node Initiation of Tracking

An ingress node that needs to initiate explicit tracking for a particular flow or set of flows can do so by performing one of the following procedures:

1. An ingress node can initiate explicit tracking for (C-S1,C-G1) by originating an S-PMSI A-D route that identifies (C-S1,C-G1) in its NLRI, including a PTA in that route, and setting the LIR flag in that PTA. The PTA may specify a particular tunnel, or may specify "no tunnel information present".

   However, the PTA of the (C-S1,C-G1) S-PMSI A-D route SHOULD NOT specify "no tunnel information present" unless the ingress node also originates an A-D route carrying a PTA that specifies the tunnel to be used for carrying (C-S1,C-G1) traffic. Such a route

could be an I-PMSI A-D route, a (C-*,C-G1) S-PMSI A-D route, a (C-S1,C-*) S-PMSI A-D route, or a (C-*,C-*) S-PMSI A-D route. (There is no point in requesting explicit tracking for a given flow if there is no tunnel on which the flow is being carried.)

Note that if the ingress node originates a wildcard S-PMSI A-D route carrying a PTA specifying the tunnel to be used for carrying (C-S1,C-G1) traffic, and if that PTA has the LIR-pF bit set, then explicit tracking for (C-S1,C-G1) is requested by that S-PMSI A-D route. In that case, the ingress node SHOULD NOT originate a (C-S1,C-G1) S-PMSI A-D route whose PTA specifies "no tunnel information present"; such a route would not provide any additional functionality.

To terminate explicit tracking that has been initiated by an S-PMSI A-D route whose PTA specifies "no tunnel information present", the ingress node withdraws the route.

To terminate explicit tracking that has been initiated by an S-PMSI A-D route whose PTA specifies a tunnel, the ingress node re-originates the route without the LIR flag set.

2. The following procedure can be used if and only if it is known that the egress nodes support the optional LIR-pF flag. If the ingress node originates a wildcard S-PMSI A-D route, it can initiate explicit tracking for the individual flows that match the wildcard route by setting the LIR-pF flag in the PTA of the wildcard route. If an egress node needs to receive one or more flows for which that wildcard route is a match for tracking, the egress node will originate a Leaf A-D route for each such flow, as specified in Section 5.2).

   When following this procedure, the PTA of the S-PMSI A-D route may specify a tunnel, or may specify "no tunnel information present". The choice between these two options is determined by considerations that are outside the scope of this document.

   To terminate explicit tracking that has been initiated by an S-PMSI A-D route whose PTA specifies "no tunnel information present", the ingress node withdraws the route.

   To terminate explicit tracking that has been initiated by an S-PMSI A-D route whose PTA specifies a tunnel, the ingress node re-originates the route without either the LIR or LIR-pF flags set.

   Note that this procedure (procedure 2 of Section 4) may not yield the expected results if there are egress nodes that do not

support the LIR-pF flag, and hence SHOULD NOT be used in that case.

5. Egress Node Response to the Match for Tracking

5.1. General Egress Node Procedures

There are four cases to consider:

1. With regard to a particular (C-S,C-G) or (C-*,C-G) multicast state, the egress node's match for tracking is same as its match for reception, and neither LIR nor LIR-pF flags are on.

   In this case, the egress node does not originate a Leaf A-D route in response to the match for reception/tracking, and there is no explicit tracking of the flow. This document specifies no new procedures for this case.

2. With regard to a particular (C-S,C-G) or (C-*,C-G) multicast state, the egress node's match for tracking is the same as its match for reception, LIR is set, but LIR-pF is not set.

   In this case, a Leaf A-D route is originated by the egress node, corresponding to the S-PMSI A-D route that is the match for reception/tracking. Construction of the Leaf A-D route is as specified in [RFC6514]; this document specifies no new procedures for this case.

3. With regard to a particular (C-S,C-G) or (C-*,C-G) multicast state, the egress node's match for tracking is the same as its match for reception, and LIR-pF is set. The egress node MUST follow whatever procedures are required by other specifications, based on the match for reception. If the egress node supports the LIR-pF flag, it MUST also follow the procedures of Section 5.2.

4. With regard to a particular (C-S,C-G) or (C-*,C-G) multicast state, the egress node's match for tracking is not the same as its match for reception. This can only happen if the match for tracking has a PTA specifying "no tunnel information present", with either LIR or LIR-pF set. In this case, the egress node MUST respond, separately, BOTH to the match for tracking and to the match for reception.

   When responding to the match for reception, the egress node MUST ignore the LIR-pF flag. However, the LIR flag is processed normally per the procedures for the match for reception.

If the match for tracking has LIR set and if either (a) the egress node does not support LIR-pF, or (b) LIR-pF is not set, then the behavior of the egress node is not affected by the procedures of this document.

If the match for tracking has LIR-pF set, and the egress node supports the LIR-pF flag, the egress node must originate one or more Leaf A-D routes, as specified in Section 5.2.

Note that if LIR is set in the PTA of the match for reception, the egress node may need to originate one or more Leaf A-D routes corresponding to the match for tracking, as well as originating a Leaf A-D route corresponding to the match for reception.

5.2. Responding to the LIR-pF Flag

To respond to a match for tracking that has LIR-pF set, an egress node originates one or more Leaf A-D routes.

Suppose the egress node has multicast state for a (C-S,C-G) or a (C-*,C-G) flow, and has determined a particular S-PMSI A-D route, which has the LIR-pF flag set, to be the match for tracking for that flow. Then if the egress node supports the LIR-pF flag, it MUST originate a Leaf A-D route whose NLRI identifies that particular flow. Note that if a single S-PMSI A-D route (with wild cards) is the match for tracking for multiple flows, the egress node may need to originate multiple Leaf A-D routes, one for each such flow. We say that, from the perspective of a given egress node, a given S-PMSI A-D route tracks the set of flows for which it is the match for tracking. Each of the Leaf A-D routes originated in response to that S-PMSI A-D route tracks a single such flow.

The NLRI of each the Leaf A-D route that tracks a particular flow is constructed as follows. The "route key" field of the NLRI will have the following format:

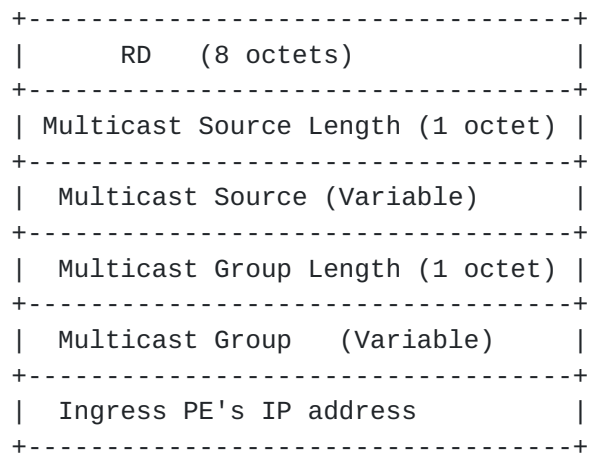

Figure 1: NLRI of S-PMSI A-D Route

- The "ingress PE" address is taken from the "originating router" field of the NLRI of the S-PMSI A-D route that is the match for tracking.

- The multicast source and group fields specify the S and G of one of the flow being tracked by this Leaf A-D route. If a (C-*,C-G) is being tracked by this Leaf A-D route, the source field is omitted, and its length is set to 0.

- The Route Distinguisher (RD) field is set to the value of the RD field from the NLRI of the S-PMSI A-D route.

The encoding of these Leaf A-D routes is similar to the encoding of the Leaf A-D routes described in section 6.2.2 of [RFC7524], which were designed for the support of "global table multicast". However, [RFC7524] sets the RD to either 0 or -1; following the procedures of the present document, the RD will never be 0 or -1. Therefore Leaf A-D routes constructed according to the procedures of this section can always be distinguished from the Leaf A-D routes constructed according to the procedures of section 6.2.2 of [RFC7524]. Also, Leaf A-D routes constructed according to the procedures of this section are VPN-specific routes, and will always carry an IP-address-specific Route Target, as specified in [RFC6514].

If a Leaf A-D route is originated as a response to a match for tracking whose PTA specifies "no tunnel information present", the Leaf A-D route MUST carry a PTA that specifies "no tunnel information present". The LIR-pF flag in this PTA MUST be set.

In the case where the match for tracking and the match for reception are the same, the PTA of the match may have both the LIR and the

LIR-pF flags set. This may cause the egress node to originate one Leaf A-D route in response to the LIR bit, and one or more Leaf A-D routes in response to the LIR-pF bit. Each such Leaf A-D route MUST have a PTA, and the LIR-pF flag of that PTA MUST be set. Note that when the match for tracking is the same as the match for reception, the PTA of the match for tracking/reception will have specified a tunnel type. The following rules specify how the PTA of the Leaf A-D route is to be constructed:

- o If the tunnel type of the PTA attached to the match for tracking/reception is Ingress Replication, the Leaf A-D route's PTA MAY specify Ingress Replication. In this case, the MPLS Label field of the PTA MAY be a non-zero value. If so, this label value will be used by the ingress PE when it transmits, to the egress PE, packets of the flow identified in the Leaf A-D route's NLRI.

  Alternatively, the egress PE MAY specify an MPLS label value of zero, or it MAY specify a tunnel type of "no tunnel information present". In either of these cases, when the ingress PE transmits packets of the identified flow to the egress PE, it will use the label that the egress PE specified in the PTA of the Leaf A-D route that it originated in response to the LIR bit of the match for reception.

- o If the tunnel type of the PTA attached to the match for tracking/reception is any of the other tunnel types listed in [RFC6514] Section 5, the PTA attached to the Leaf A-D route MUST specify a tunnel type of "no tunnel information present".

- o When additional tunnel types are defined, the specification for how MVPN is to use those tunnel types must also specify how to construct the PTA of a Leaf A-D route that is originated in response to the LIR-pF flag. As an example, see [BIER-MVPN].

Of course, an egress node that originates such Leaf A-D routes needs to remember which S-PMSI A-D route caused these Leaf A-D routes to be originated; if that S-PMSI A-D route is withdrawn, those Leaf A-D routes MUST be withdrawn.

Similarly, a Leaf A-D route needs to be withdrawn (either implicitly or explicitly) if the egress node changes its Upstream Multicast Hop (UMH) ([RFC6513]) for the flow that is identified in the Leaf A-D route's NLRI, or if the egress node that originated the route no longer needs to receive the flow identified in the NLRI of the route.

Note that an egress node may acquire (C-S,C-G) state or (C-*,C-G) state after it has already received the S-PMSI A-D that is the match for tracking for that state. In this case, a Leaf A-D route needs to

be originated at that time, and the egress node must remember that the new Leaf A-D route corresponds to that match for tracking.

Note that if a particular S-PMSI A-D route is a match for tracking but not a match for reception, the LIR bit in its PTA is ignored if the LIR-pF bit is set.

5.3. When the Egress Node is an ABR or ASBR

When segmented P-tunnels are used, the ingress and egress nodes may be ABRs or ASBRs. An egress ABR/ASBR that receives and installs an S-PMSI A-D route also forwards that route. If the PTA of an installed S-PMSI A-D route specifies a tunnel, the egress ABR/ASBR MAY change the PTA to specify a different tunnel type (as discussed in [RFC6514] and/or [RFC7524]). The egress ABR/ASBR may also need to originate a Leaf A-D route, as specified in [RFC6514] and/or [RFC7524].

Suppose the forwarded S-PMSI A-D route has a PTA specifying a tunnel, and also has LIR-pF set. The egress ABR/ASBR originates a corresponding Leaf A-D route for a given (C-S,C-G) only if it knows that it needs to receive that flow. It will know this by virtue of receiving a corresponding Leaf A-D route from downstream. (In the case where the PTA specifies a tunnel but LIR-pF is not set, this document does not specify any new procedures.)

The procedures in the remainder of this section apply only when an egress ABR/ASBR has installed an S-PMSI A-D route whose PTA specifies "no tunnel information present" but has LIR or LIR-pF set.

If the PTA of the installed S-PMSI A-D route specifies "no tunnel information present", the egress ABR/ASBR MUST pass the PTA along unchanged when it forwards the S-PMSI A-D route. (That is, a PTA specifying "no tunnel information present" MUST NOT be changed into a PTA specifying a tunnel.) Furthermore, if the PTA specifies "no tunnel information present", the LIR and LIR-pF flags in the PTA MUST be passed along unchanged.

As a result of propagating such an S-PMSI A-D route, the egress ABR/ASBR may receive one or more Leaf A-D routes that correspond to that S-PMSI A-D route. These routes will be received carrying an IP-address-specific Route Target (RT) Extended Community that specifies the address of the egress ABR/ASBR. The egress ABR/ASBR will propagate these Leaf A-D routes, after changing the RT as follows. The "global administrator" field of the modified RT will be set to the IP address taken either from the S-PMSI A-D route's next hop field ([RFC6514]), or from its Segmented P2MP Next Hop Extended Community ([RFC7524]).

This procedure enables the ingress PE to explicitly track the egress PEs for a given flow, even if segmented tunnels are being used. However, cross-domain explicit tracking utilizes S-PMSI A-D routes that do not specify tunnel information; therefore it can only be done when the S-PMSI A-D route which is a flow's match for tracking is different than the S-PMSI A-D route which is that flow's match for reception.

6. Ingress Node Handling of Received Leaf A-D Routes with LIR-pF Set

Consider the following situation:

- An ingress node, call it N, receives a Leaf A-D route, call it L.
- L carries an IP-address-specific RT identifying N.
- The route key field of L's NLRI is not identical to the NLRI of any current I-PMSI or S-PMSI A-D route originated by N.

Per the procedures of [RFC6514] and [RFC7524], such a Leaf A-D route does not cause any MVPN-specific action to be taken by N.

This document modifies those procedures in the case where there is a current S-PMSI A-D route with a wildcard NLRI, originated by N, to which L is a valid response according to the procedures of Section 5.2. In this case, L MUST be processed by N.

Suppose that L's PTA specifies a tunnel type of Ingress Replication, and that it also specifies a non-zero MPLS label. Then if N needs to send to L a packet belonging to the multicast flow or flows identified in L's NLRI, N MUST use the specified label.

If L's PTA meets any of the following conditions:

- It specifies a tunnel type of "no tunnel information present", or
- It specifies a tunnel type of Ingress Replication, but specifies an MPLS label of zero, or
- It specifies another of the tunnel types listed in Section 5 of [RFC6514],

then the action taken by N when it receives L is a local matter. In this case, the Leaf A-D route L provides N with explicit tracking information for the flow identified by L's NLRI. However, that information is for management/monitoring purposes and does not have an effect on the flow of multicast traffic.

If L's PTA specifies a tunnel type not mentioned above, the specification for how MVPN uses that tunnel type must specify the actions that N is to take upon receiving L. As an example, see [BIER-MVPN].

7. Acknowledgments

The authors wish to thank Robert Kebler for his ideas and comments. We also thank Stephane Litkowski for his thorough review and useful suggestions.

8. IANA Considerations

The LIR-pF flag needs to be added to the "P-Multicast Service Interface Tunnel (PMSI Tunnel) Attribute Flags" in the "Border Gateway Protocol (BGP) Parameters" registry. This registry is defined in [RFC7902]. The requested value is Bit Position 2. This document should be the reference.

9. Security Considerations

The Security Considerations of [RFC6513] and [RFC6514] apply.

By setting the LIR-pF flag in a single wildcard S-PMSI A-D route, a large number of Leaf A-D routes can be elicited. If this flag is set when not desired (through either error or malfeasance), a significant increase in control plane overhead can result. The specification of counter-measures for this problem is outside the scope of this document.

10. References

10.1. Normative References

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

[RFC6513] Rosen, E., Ed. and R. Aggarwal, Ed., "Multicast in MPLS/ BGP IP VPNs", RFC 6513, DOI 10.17487/RFC6513, February 2012, <https://www.rfc-editor.org/info/rfc6513>.

[RFC6514] Aggarwal, R., Rosen, E., Morin, T., and Y. Rekhter, "BGP Encodings and Procedures for Multicast in MPLS/BGP IP VPNs", RFC 6514, DOI 10.17487/RFC6514, February 2012, <https://www.rfc-editor.org/info/rfc6514>.

[RFC6625] Rosen, E., Ed., Rekhter, Y., Ed., Hendrickx, W., and R. Qiu, "Wildcards in Multicast VPN Auto-Discovery Routes", RFC 6625, DOI 10.17487/RFC6625, May 2012, <https://www.rfc-editor.org/info/rfc6625>.

[RFC7524] Rekhter, Y., Rosen, E., Aggarwal, R., Morin, T., Grosclaude, I., Leymann, N., and S. Saad, "Inter-Area Point-to-Multipoint (P2MP) Segmented Label Switched Paths (LSPs)", RFC 7524, DOI 10.17487/RFC7524, May 2015, <https://www.rfc-editor.org/info/rfc7524>.

[RFC7902] Rosen, E. and T. Morin, "Registry and Extensions for P-Multicast Service Interface Tunnel Attribute Flags", RFC 7902, DOI 10.17487/RFC7902, June 2016, <https://www.rfc-editor.org/info/rfc7902>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

10.2. Informative References

[BIER-MVPN] Rosen, E., Sivakumar, M., Aldrin, S., Dolganow, A., and T. Przygienda, "Multicast VPN Using BIER", internet-draft draft-ietf-bier-mvpn-11, March 2018.

[RFC7582] Rosen, E., Wijnands, IJ., Cai, Y., and A. Boers, "Multicast Virtual Private Network (MVPN): Using Bidirectional P-Tunnels", RFC 7582, DOI 10.17487/RFC7582, July 2015, <https://www.rfc-editor.org/info/rfc7582>.

[RFC7900] Rekhter, Y., Ed., Rosen, E., Ed., Aggarwal, R., Cai, Y., and T. Morin, "Extranet Multicast in BGP/IP MPLS VPNs", RFC 7900, DOI 10.17487/RFC7900, June 2016, <https://www.rfc-editor.org/info/rfc7900>.

Authors' Addresses

Andrew Dolganow
Nokia
438B Alexandra Rd #08-07/10
Alexandra Technopark
Singapore 119968
Singapore

Email: andrew.dolganow@nokia.com

Jayant Kotalwar
Nokia
701 East Middlefield Rd
Mountain View, California 94043
United States of America

Email: jayant.kotalwar@nokia.com

Eric C. Rosen (editor)
Juniper Networks, Inc.
10 Technology Park Drive
Westford, Massachusetts 01886
United States of America

Email: erosen@juniper.net

Zhaohui Zhang
Juniper Networks, Inc.
10 Technology Park Drive
Westford, Massachusetts 01886
United States of America

Email: zzhang@juniper.net