

Internet Engineering Task Force (IETF)
Request for Comments: 8534
Updates: 6514, 6625, 7524, 7582, 7900
Category: Standards Track
ISSN: 2070-1721

A. Dolganow
J. Kotalwar
Nokia
E. Rosen, Ed.
Z. Zhang
Juniper Networks, Inc.
February 2019

Explicit Tracking with Wildcard Routes in Multicast VPN

Abstract

The base Multicast VPN (MVPN) specifications (RFCs 6513 and 6514) 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, 7524, 7582, and 7900.

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 7841.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at <https://www.rfc-editor.org/info/rfc8534>.

Copyright Notice

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

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

Table of Contents

| | |
|--|----|
| 1. Introduction | 2 |
| 1.1. Terminology | 5 |
| 2. The Explicit-Tracking Flags | 5 |
| 3. Match for Tracking versus Match for Reception | 7 |
| 4. Ingress Node Initiation of Tracking | 9 |
| 5. Egress Node Response to the Match for Tracking | 11 |
| 5.1. General Egress Node Procedures | 11 |
| 5.2. Responding to the LIR-pF Flag | 12 |
| 5.3. When the Egress Node Is an ABR or ASBR | 15 |
| 6. Ingress Node Handling of Received Leaf A-D Routes with LIR-pF Set | 16 |
| 7. IANA Considerations | 17 |
| 8. Security Considerations | 18 |
| 9. References | 19 |
| 9.1. Normative References | 19 |
| 9.2. Informative References | 20 |
| Acknowledgments | 20 |
| Authors' Addresses | 21 |

1. Introduction

The base Multicast VPN (MVPN) specifications, [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 a 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 flag) in the PTA. When an egress node receives an S-PMSI A-D route with the LIR flag 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 the LIR flag set but that does not identify any P-tunnel. This mechanism can be used when 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, (C-S,C-*) S-PMSI A-D routes, or (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 the LIR flag 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 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 the LIR flag 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 by wildcard S-PMSI A-D routes.

However, this procedure requires several clarifications:

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

This document clarifies the procedures for originating and receiving S-PMSI A-D routes and Leaf A-D routes. This document also adds new procedures to allow more efficient explicit tracking. The procedures being clarified and/or extended are discussed in multiple places in the documents being updated.

1.1. Terminology

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" flag (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 wildcard S-PMSI A-D route is discussed in Sections 5 and 5.2.

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.

Note that support for the LIR-pF flag is OPTIONAL. This flag SHOULD NOT be set in a route's PTA unless it is known that the flag is supported by all the Provider Edge (PE) routers that are to receive that route. Typically, this might mean that the ability to set this flag would be controlled by a configuration knob, and operators would not set this knob unless they know that all the relevant PEs support this feature. How this is known is outside the scope of this document.

This document only defines procedures for the LIR-pF flag when that flag is in the PTA of a wildcard S-PMSI A-D route or a Leaf A-D route. In all other cases, the flag SHOULD be clear, and its value SHOULD be ignored. Use of the flag in these other cases is outside the scope of this document.

Section 5 of [RFC6514] lists a number of tunnel types. We will refer to these as "6514-tunnel-types". Other tunnel types will be referred to as "non-6514-tunnel-types". This document specifies procedures for using the LIR-pF flag with 6514-tunnel-types. Procedures for using the LIR-pF flag with non-6514-tunnel-types are outside the scope of this document.

If it is desired to use a particular non-6514-tunnel-type in MVPN, there needs to be a specification for how that tunnel type is used in MVPN. If it is desired to use that tunnel type along with the LIR-pF flag, that specification (or a follow-on specification) will have to specify the rules for using the LIR-pF flag with that tunnel type. As an example, see [BIER-MVPN]. In the absence of such a specification (or in the absence of support for such a specification):

- o the originator of a route that carries a PTA SHOULD NOT set the LIR-pF flag in any PTA that specifies that tunnel type, and
- o the receiver of a route that carries a PTA specifying that tunnel type SHOULD treat the LIR-pF flag as if it were not set.

If the LIR-pF flag is set in the PTA of an S-PMSI A-D route, the originator of that route MUST also set the LIR flag. If the PTA of a received wildcard S-PMSI A-D route has the LIR-pF flag set but does not have the LIR flag set, the receiver MUST log the fact that the flags appear to have been improperly set. However, the route MUST then be processed normally (as if both flags were set), as specified in 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 nodes 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 the LIR-pF flag, 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.

As discussed in Section 5.2, any Leaf A-D route originated in response to an S-PMSI A-D route that has the LIR-pF flag 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 the LIR-pF flag set, but the Leaf A-D route lacks a PTA or has a PTA where the LIR-pF flag is clear, the ingress node can infer 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 versus 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 Sections 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 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" and has neither the LIR flag nor the LIR-pF flag 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 flags are both clear). Then apply the rules (from [RFC6625] and other documents that update it) 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 the LIR or LIR-pf flag 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 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:

- o Route1: A (C-*,C-*) S-PMSI A-D route whose PTA specifies a tunnel.
- o Route2: A (C-S1,C-G1) S-PMSI A-D route whose PTA specifies "no tunnel information present" and has the LIR flag set.

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

Continuing this example, suppose:

- o PE1 has chosen PE2 as the upstream PE for a different flow, (C-S2,C-G2).
- o PE2 has not originated an S-PMSI A-D route for (C-S2,C-G2).

In this case, PE1 would consider Route1 to be the match of (C-S2,C-G2) 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:

- o Route1: A (C-*,C-*) S-PMSI A-D route (irrespective of whether the PTA specifies a tunnel).
- o Route2: A (C-S1,C-G1) S-PMSI A-D route whose PTA specifies a tunnel.

In this case, 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 either a particular tunnel or "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 "Inclusive Provider Multicast Service Interface Auto-Discovery route" (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 flag 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 either a tunnel or "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; hence, it 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 the same as its match for reception, and neither the LIR nor the LIR-pF flags are set.

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, and the LIR flag is set, but the LIR-pF flag 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 follows whatever procedures are required by other specifications, based on the match for reception. However, any Leaf A-D route originated by the egress node as a result MUST have the LIR-pF flag set in its PTA. The egress node 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 the LIR flag or the LIR-pF flag set. In this case, the egress node MUST respond, separately, to *both* the match for tracking and the match for reception.

If a Leaf A-D route is originated in response to the match for reception, the LIR-pF flag in the Leaf A-D route's PTA MUST have the same value as the LIR-pF flag in the match for reception's PTA. In all other respects, the procedures for responding to the match for reception are not affected by this document.

If the match for tracking has the LIR flag set but the LIR-pF flag 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 the LIR-pF flag set, the egress node MUST follow the procedures of Section 5.2.

Note that if the LIR flag 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 the LIR-pF flag 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 wildcards) 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 Leaf A-D route that tracks a particular flow is constructed as follows. The Route Key field of the NLRI will have the format shown in Figure 1 (as defined in Sections 4.3 and 4.4 of [RFC6514]).

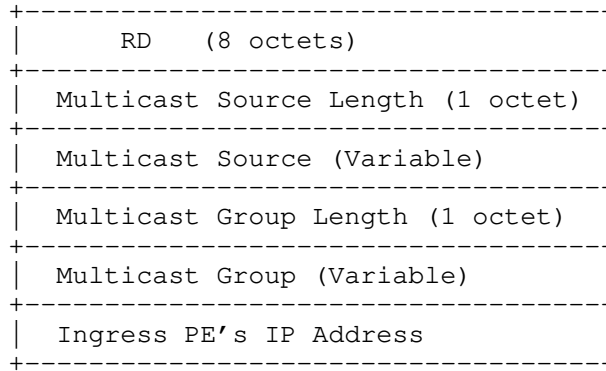


Figure 1: NLRI of S-PMSI A-D Route

- o 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. Section 2 of [RFC6515] explains how the receiver of a Leaf A-D route determines the length of this field and the address family of the PE's IP address.
- o The Multicast Source and Multicast Group fields respectively specify a source address (S) and a group address (G) that together identify the flow or flows being tracked by this Leaf A-D route. If a (C-*,C-G) is being tracked by this Leaf A-D route, the Multicast Source field is omitted, and the Multicast Source Length field is set to 0. In this case, the Leaf A-D route is known as a "wildcard Leaf A-D route".
- o 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, that document 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.

If an egress node originates multiple Leaf A-D routes in response to a single S-PMSI A-D route, and that S-PMSI A-D route is later withdrawn, then those Leaf A-D routes MUST also 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 that flow.

It is possible that an egress node will 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.

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

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. Some of the rules for constructing the PTA of the Leaf A-D route depend on the tunnel type, and some are independent of the tunnel type. No matter what the tunnel type is, the LIR-pF flag MUST be set.

If the match for tracking/reception is a wildcard S-PMSI A-D route, the egress node may originate a wildcard Leaf A-D route in response, as well as originating one or more non-wildcard Leaf A-D routes. Note that the LIR-pF flag MUST be set in the wildcard Leaf A-D route as well as in the non-wildcard Leaf A-D routes.

This document provides additional rules for constructing the PTA when the tunnel type is a 6514-tunnel-type (see Section 2).

As discussed in Section 2, if a non-6514-tunnel-type is being used, then presumably there is a specification for how that tunnel type is used in MVPN. If it is desired to use that tunnel type along with the LIR-pF flag, that specification (or a follow-on specification) will have to specify the additional rules for constructing the PTA. As an example, see [BIER-MVPN].

For 6514-tunnel-types, additional rules for constructing the PTA are as follows:

- 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 flag 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 6514-tunnel-types, the PTA attached to the Leaf A-D route MUST specify a tunnel type of "no tunnel information present".

It may happen that the tunnel type is a non-6514-tunnel type, but either (a) there is no specification for how to use that tunnel type with the LIR-pF flag or (b) there is such a specification, but the egress node does not support it. In that case, the egress node MUST treat the match for tracking/reception as if it had the LIR-pF flag clear.

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 received PTA of an installed S-PMSI A-D route specifies a tunnel, the egress ABR/ASBR MAY change the PTA before forwarding the route, in order 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 S-PMSI A-D route as received has a PTA specifying a tunnel and also has the LIR-pF flag 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 the LIR-pF flag 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 as received specifies "no tunnel information present" but has the LIR flag or the LIR-pF flag set.

If the received 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 its Segmented Point-to-Multipoint (P2MP) Next-Hop Extended Community [RFC7524]. The address from the Segmented P2MP Next-Hop Extended Community is used if that Extended Community is present; otherwise, the address from the Next-Hop field is used.

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 that is a flow's match for tracking is different from the S-PMSI A-D route that 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:

- o An ingress node, call it N, receives a Leaf A-D route, call it L.
- o L carries an IP-address-specific RT identifying N.
- o 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 wildcard S-PMSI A-D route, 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:

- o It specifies a tunnel type of "no tunnel information present", or
- o It specifies a tunnel type of Ingress Replication, but specifies an MPLS label of zero, or
- o It specifies any other 6514-tunnel-type,

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 any direct effect on the flow of multicast traffic.

If L's PTA specifies a non-6514-tunnel-type not mentioned above, presumably there is a specification for how MVPN uses that tunnel type. If the LIR-pF flag is to be used with that tunnel type, that specification must specify the actions that N is to take upon receiving L. As an example, see [BIER-MVPN]. In the absence of such a specification, the LIR-pF flag SHOULD BE ignored. See Section 2 for further discussion of non-6514-tunnel-types.

7. IANA Considerations

IANA has added the following entry to the "P-Multicast Service Interface (PMSI) Tunnel Attribute Flags" registry under the "Border Gateway Protocol (BGP) Parameters" registry. This registry is defined in [RFC7902]. The entry appears as:

- o Value: 2
- o Name: Leaf Information Required per-Flow (LIR-pF)
- o Reference: RFC 8534

8. 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. Properly protecting the control plane should prevent this kind of attack.

In the event such an attack occurs, mitigating it is unfortunately not very straightforward. The ingress node can take note of the fact that it is getting, in response to an S-PMSI A-D route that has LIR-pF clear, one or more Leaf A-D routes that have LIR-pF set. By default, the reception of such a route MUST be logged. However, it is possible for such log entries to be "false positives" that generate a lot of "noise" in the log; therefore, implementations SHOULD have a knob to disable this logging.

In theory, if one or more Leaf A-D routes with LIR-pF set arrive in response to an S-PMSI A-D route with LIR-pF clear, withdrawing the S-PMSI A-D route could put a stop to the attack. In practice, that is not likely to be a very good strategy, because:

- o Under normal operating conditions, there are some race conditions that may cause the ingress node to think it is being attacked, when in fact it is not.
- o If some egress nodes have a bug that causes them to set LIR-pF when it should be clear, withdrawing the S-PMSI A-D route will stop the flow of multicast data traffic to all the egress nodes, causing an unnecessary customer-visible disruption.
- o The same situation that caused the S-PMSI A-D route to be originated in the first place will still exist after the S-PMSI A-D route is withdrawn, so the route will just be re-originated.

In other words, any action that would ameliorate the effects of this sort of attack would likely have a negative effect during normal operation. Therefore, it is really better to rely on security mechanisms that protect the control plane generally than to have a mechanism that is focused on this one particular type of attack.

9. References

9.1. 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>>.
- [RFC6515] Aggarwal, R. and E. Rosen, "IPv4 and IPv6 Infrastructure Addresses in BGP Updates for Multicast VPN", RFC 6515, DOI 10.17487/RFC6515, February 2012, <<https://www.rfc-editor.org/info/rfc6515>>.
- [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>>.

9.2. Informative References

[BIER-MVPN]

Rosen, E., Sivakumar, M., Aldrin, S., Dolganow, A., and T. Przygienda, "Multicast VPN Using BIER", Work in Progress, 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>>.

Acknowledgments

The authors wish to thank Robert Kebler for his ideas and comments and Stephane Litkowski and Benjamin Kaduk for their thorough reviews and useful suggestions. We would also like to thank Mirja Kuhlewind for her attention to the Security Considerations section.

Authors' Addresses

Andrew Dolganow
Nokia
438B Alexandra Rd #08-07/10
Alexandra Technopark
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: erosen52@gmail.com

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

Email: zzhang@juniper.net

