Internet Engineering Task Force (IETF)
Request for Comments: 7246
Category: Standards Track
ISSN: 2070-1721
IJ. Wijnands, Ed.
Cisco Systems
P. Hitchen
BT
N. Leymann
Deutsche Telekom
W. Henderickx
Alcatel-Lucent
A. Gulko
Thomson Reuters
J. Tantsura
Ericsson
June 2014

Multipoint Label Distribution Protocol In-Band Signaling in a Virtual Routing and Forwarding (VRF) Table Context

Abstract

An IP Multicast Distribution Tree (MDT) may traverse both label switching (i.e., Multiprotocol Label Switching, or MPLS) and non-label switching regions of a network. Typically, the MDT begins and ends in non-MPLS regions, but travels through an MPLS region. In such cases, it can be useful to begin building the MDT as a pure IP MDT, then convert it to an MPLS Multipoint Label Switched Path (MP-LSP) when it enters an MPLS-enabled region, and then convert it back to a pure IP MDT when it enters a non-MPLS-enabled region. Other documents specify the procedures for building such a hybrid MDT, using Protocol Independent Multicast (PIM) in the non-MPLS region of the network, and using Multipoint Label Distribution Protocol (mLDP) in the MPLS region. This document extends those procedures to handle the case where the link connecting the two regions is a Virtual Routing and Forwarding (VRF) table link, as defined in the "BGP IP/MPLS VPN" specification. However, this document is primarily aimed at particular use cases where VRFs are used to support multicast applications other than multicast VPN.

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/rfc7246.

Copyright Notice

Table of Contents

1. Introduction

Sometimes an IP Multicast Distribution Tree (MDT) traverses both MPLS-enabled and non-MPLS-enabled regions of a network. Typically, the MDT begins and ends in non-MPLS regions, but travels through an MPLS region. In such cases, it can be useful to begin building the MDT as a pure IP MDT, then convert it to an MPLS Multipoint Label Switched Path (LSP) when it enters an MPLS-enabled region, and then convert it back to a pure IP MDT when it enters a non-MPLS-enabled region. Other documents specify the procedures for building such a hybrid MDT, using Protocol Independent Multicast (PIM) in the non-MPLS region of the network, and using Multipoint Label Distribution Protocol (mLDP) in the MPLS region. This document extends the procedures from [RFC6826] to handle the case where the link connecting the two regions is a Virtual Routing and Forwarding (VRF) table link, as defined in the "BGP IP/MPLS VPN" specification [RFC6513]. However, this document is primarily aimed at particular use cases where VRFs are used to support multicast applications other than multicast VPN.

In PIM, a tree is identified by a source address (or in the case of bidirectional trees [RFC5015], a rendezvous point address or "RPA") and a group address. The tree is built from the leaves up, by sending PIM control messages in the direction of the source address or the RPA. In mLDP, a tree is identified by a root address and an "opaque value", and is built by sending mLDP control messages in the direction of the root. The procedures of [RFC6826] explain how to convert a PIM <source address or RPA, group address> pair into an mLDP <root node, opaque value> pair and how to convert the mLDP <root node, opaque value> pair back into the original PIM <source address or RPA, group address> pair.

However, the procedures in [RFC6826] assume that the routers doing the PIM/mLDP conversion have routes to the source address or RPA in their global routing tables. Thus, the procedures cannot be applied exactly as specified when the interfaces connecting the non-MPLS-enabled region to the MPLS-enabled region are interfaces that belong to a VRF as described in [RFC4364]. This specification extends the procedures of [RFC6826] so that they may be applied in the VRF context.

As in [RFC6826], the scope of this document is limited to the case where the multicast content is distributed in the non-MPLS-enabled regions via PIM-created source-specific or bidirectional trees. Bidirectional trees are always mapped onto multipoint-to-multipoint LSPs, and source-specific trees are always mapped onto point-to-multipoint LSPs.

Note that the procedures described herein do not support non-bidirectional PIM Any-Source Multicast (ASM) groups, do not support the use of multicast trees other than mLDP multipoint LSPs in the core, and do not provide the capability to aggregate multiple PIM trees onto a single multipoint LSP. If any of those features are needed, they can be provided by the procedures of [RFC6513] and [RFC6514]. However, there are cases where multicast services are offered through interfaces associated with a VRF, and where mLDP is used in the core, but where aggregation is not desired. For example, some service providers offer multicast content to their customers, but have chosen to use VRFs to isolate the various customers and services. This is a simpler scenario than one in which the customers provide their own multicast content, out of the control of the service provider, and can be handled with a simpler solution. Also, when PIM trees are mapped one-to-one to multipoint LSPs, it is helpful for troubleshooting purposes to have the PIM source/group addresses encoded into the mLDP FEC (Forwarding Equivalence Class) element in what this document terms "mLDP in-band signaling".

In order to use the mLDP in-band signaling procedures for a particular group address in the context of a particular set of VRFs, those VRFs MUST be configured with a range of multicast group addresses for which mLDP in-band signaling is to be enabled. This configuration is per VRF defined in [RFC4364]). For those groups, and those groups only, the procedures of this document are used. For other groups, the general-purpose multicast VPN procedures MAY be used, although it is more likely this VRF is dedicated to mLDP in-band signaling procedures and all other groups are discarded. The configuration MUST be present in all PE routers that attach to sites containing senders or receivers for the given set of group addresses. Note that because the provider most likely owns the multicast content and the method of transportation across the network, which are both transparent to the end user, no coordination needs to happen between the end user and the provider.

1.1. Conventions Used in This Document

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

1.2. Terminology

In-band signaling: Using the opaque value of an mLDP FEC element to encode the (S,G) or (*,G) identifying a particular IP multicast tree.

Ingress LSR: Source of a P2MP LSP, also referred to as root node.

IP multicast tree: An IP multicast distribution tree identified by a source IP address and/or IP multicast destination address, also referred to as (S,G) and (*,G).

mLDP: Multipoint LDP.

MP LSP: A multipoint LSP, either a P2MP or an MP2MP LSP.

MP2MP LSP: An LSP that connects a set of leaf nodes, acting indifferently as ingress or egress (see [RFC6388]).

P2MP LSP: An LSP that has one Ingress LSR and one or more Egress LSRs (see [RFC6388]).

RPA: Rendezvous Point Address, the address that is used as the root of the distribution tree for a range of multicast groups.

RD: Route Distinguisher, an identifier that makes a route unique in the context of a VRF.

UMH: Upstream Multicast Hop, the upstream router in that is in the path to reach the source of the multicast flow.

VRF: Virtual Routing and Forwarding table.

2. VRF In-Band Signaling for MP LSPs

Suppose that a PE router, PE1, receives a PIM Join(S,G) message over one of its interfaces that is associated with a VRF. Following the procedure of Section 5.1 of [RFC6513], PE1 determines the "upstream RD", the "upstream PE", and the "upstream multicast hop" (UMH) for the source address S.

In order to transport the multicast tree via a multipoint (MP) LSP using VRF in-band signaling, an mLDP Label Mapping message is sent by PE1. This message will contain either a P2MP FEC or an MP2MP FEC (see [RFC6388]), depending upon whether the PIM tree being transported is a source-specific tree, or a bidirectional tree, respectively. The FEC contains a "root" and an "opaque value".

If the UMH and the upstream PE have the same IP address (i.e., the upstream PE is the UMH), then the root of the multipoint FEC is set to the IP address of the upstream PE. If, in the context of this VPN, (S,G) refers to a source-specific MDT, then the values of S, G, and the upstream RD are encoded into the opaque value. If, in the context of this VPN, G is a bidirectional group address, then S is replaced with the value of the RPA associated with G.

The encoding details are specified in Section 3. Conceptually, the multipoint FEC can be thought of as an ordered pair: {root=<Upstream-PE>; opaque_value=<S or RPA , G, Upstream-RD>}. The mLDP Label Mapping message is then sent by PE1 on its LDP session to the "next hop" on the message’s path to the upstream PE. The "next hop" is usually the directly connected next hop, but see [RFC7060] for cases in which the next hop is not directly connected.

If the UMH and the upstream PE do not have the same IP address, the procedures of Section 2 of [RFC6512] should be applied. The root node of the multipoint FEC is set to the UMH. The recursive opaque value is then set as follows: the root node is set to the upstream PE, and the opaque value is set to the multipoint FEC described in the previous paragraph. That is, the multipoint FEC can be thought of as the following recursive ordered pair: {root=<UMH>; opaque_value=<root=Upstream-PE, opaque_value=<S or RPA, G, Upstream-RD>>}.

The encoding of the multipoint FEC also specifies the "type" of PIM MDT being spliced onto the multipoint LSP. Four opaque encodings are defined in [RFC6826]: IPv4 source-specific tree, IPv6 source-specific tree, IPv4 bidirectional tree, and IPv6 bidirectional tree.

When a PE router receives an mLDP message with a P2MP or MP2MP FEC, where the PE router itself is the root node, and the opaque value is one of the types defined in Section 3, then it uses the RD encoded in the opaque value field to determine the VRF context. (This RD will be associated with one of the PEs VRFs.) Then, in the context of that VRF, the PE follows the procedure specified in section 2 of [RFC6826].

3. Encoding the Opaque Value of an LDP MP FEC

This section documents the different transit opaque encodings.

3.1. Transit VPNv4 Source TLV

This opaque value type is used when transporting a source-specific mode multicast tree whose source and group addresses are IPv4 addresses.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Type          | Length                        | Source
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                                                | Group
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                                                |               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                   RD                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Type: 250

Length: 16

Source: IPv4 multicast source address, 4 octets.

Group: IPv4 multicast group address, 4 octets.

RD: Route Distinguisher, 8 octets.

3.2. Transit VPNv6 Source TLV

This opaque value type is used when transporting a source-specific mode multicast tree whose source and group addresses are IPv6 addresses.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Type          | Length                        | Source        ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                                               | Group         ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                                               |               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                 RD                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Type: 251

Length: 40

Source: IPv6 multicast source address, 16 octets.

Group: IPv6 multicast group address, 16 octets.

RD: Route Distinguisher, 8 octets.

3.3. Transit VPNv4 Bidir TLV

This opaque value type is used when transporting a bidirectional multicast tree whose group address is an IPv4 address. The RP address is also an IPv4 address in this case.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Type          | Length                        | Mask Len      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              RP                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                             Group                             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                              RD                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Type: 9

Length: 17

Mask Len: The number of contiguous one bits that are left justified and used as a mask, 1 octet.

RP: Rendezvous Point (RP) IPv4 address used for the encoded Group, 4 octets.

Group: IPv4 multicast group address, 4 octets.

RD: Route Distinguisher, 8 octets.

3.4. Transit VPNv6 Bidir TLV

This opaque value type is used when transporting a bidirectional multicast tree whose group address is an IPv6 address. The RP address is also an IPv6 address in this case.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Type          | Length                        | Mask Len      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                              RP                               ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                            Group                              ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                              RD                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Type: 10

Length: 41

Mask Len: The number of contiguous one bits that are left justified and used as a mask, 1 octet.

RP: Rendezvous Point (RP) IPv6 address used for the encoded group, 16 octets.

Group: IPv6 multicast group address, 16 octets.

RD: Route Distinguisher, 8 octets.

4. Security Considerations

The same security considerations apply as for the base LDP specification, described in [RFC5036], and the base mLDP specification, described in [RFC6388].

Operators MUST configure packet filters to ensure that the mechanism described in this memo does not cause non-global-scoped IPv6 multicast packets to be tunneled outside of their intended scope.

5. IANA Considerations

[RFC6388] defines a registry for the "LDP MP Opaque Value Element basic type". IANA has assigned four new code points in this registry:

Transit VPNv4 Source TLV type - 250

Transit VPNv6 Source TLV type - 251

Transit VPNv4 Bidir TLV type - 9

Transit VPNv6 Bidir TLV type - 10

6. Acknowledgments

Thanks to Eric Rosen, Andy Green, Yakov Rekhter, and Eric Gray for their comments on the document.

7. References

7.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, February 2006.

[RFC5015] Handley, M., Kouvelas, I., Speakman, T., and L. Vicisano, "Bidirectional Protocol Independent Multicast (BIDIR-PIM)", RFC 5015, October 2007.

[RFC5036] Andersson, L., Ed., Minei, I., Ed., and B. Thomas, Ed., "LDP Specification", RFC 5036, October 2007.

[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, November 2011.

[RFC6512] Wijnands, IJ., Rosen, E., Napierala, M., and N. Leymann, "Using Multipoint LDP When the Backbone Has No Route to the Root", RFC 6512, February 2012.

[RFC6826] Wijnands, IJ., Ed., Eckert, T., Leymann, N., and M. Napierala, "Multipoint LDP In-Band Signaling for Point-to-Multipoint and Multipoint-to-Multipoint Label Switched Paths", RFC 6826, January 2013.

7.2. Informative References

[RFC6513] Rosen, E., Ed., and R. Aggarwal, Ed., "Multicast in MPLS/BGP IP VPNs", RFC 6513, February 2012.

[RFC6514] Aggarwal, R., Rosen, E., Morin, T., and Y. Rekhter, "BGP Encodings and Procedures for Multicast in MPLS/BGP IP VPNs", RFC 6514, February 2012.

[RFC7060] Napierala, M., Rosen, E., and IJ. Wijnands, "Using LDP Multipoint Extensions on Targeted LDP Sessions", RFC 7060, November 2013.

Authors' Addresses

IJsbrand Wijnands (editor)
Cisco Systems
De kleetlaan 6a
Diegem 1831
Belgium
EMail: ice@cisco.com

Paul Hitchen
BT
BT Adastral Park
Ipswich IP53RE
United Kingdom
EMail: paul.hitchen@bt.com

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

Wim Henderickx
Alcatel-Lucent
Copernicuslaan 50
Antwerp 2018
Belgium
EMail: wim.henderickx@alcatel-lucent.com

Arkadiy Gulko
Thomson Reuters
195 Broadway
New York, NY 10007
United States
EMail: arkadiy.gulko@thomsonreuters.com

Jeff Tantsura
Ericsson
300 Holger Way
San Jose, CA 95134
United States
EMail: jeff.tantsura@ericsson.com