Internet Engineering Task Force (IETF)
Request for Comments: 7441
Updates: 6514
Category: Standards Track
ISSN: 2070-1721

IJ. Wijnands
Cisco Systems, Inc.
E. Rosen
Juniper Networks, Inc.
U. Joorde
Deutsche Telekom

January 2015

Encoding Multipoint LDP (mLDP) Forwarding Equivalence Classes (FECs) in the NLRI of BGP MCAST-VPN Routes

Abstract

Many service providers offer "BGP/MPLS IP VPN" service to their customers. Existing IETF standards specify the procedures and protocols that a service provider uses in order to offer this service to customers who have IP unicast and IP multicast traffic in their VPNs. It is also desirable to be able to support customers who have MPLS multicast traffic in their VPNs. This document specifies the procedures and protocol extensions that are needed to support customers who use the Multipoint LDP (mLDP) as the control protocol for their MPLS multicast traffic. Existing standards do provide some support for customers who use mLDP, but only under a restrictive set of circumstances. This document generalizes the existing support to include all cases where the customer uses mLDP, without any restrictions. This document updates RFC 6514.

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/rfc7441.

Table of Contents

1. Introduction

Many service providers (SPs) offer BGP/MPLS IP VPN service to their customers. When a customer has IP multicast traffic in its VPN, the service provider needs to signal the customer multicast states across the backbone. A customer with IP multicast traffic is typically using PIM (Protocol Independent Multicast) [PIM] and/or IGMP (Internet Group Management Protocol) [IGMP] as the multicast control protocol in its VPN. The IP multicast states of these protocols are commonly denoted as "(S,G)" and/or "(*,G)" states, where "S" is a multicast source address and "G" is a multicast group address. [MVPN-BGP] specifies the way an SP may use BGP to signal a customer's IP multicast states across the SP backbone. This is done by using Multiprotocol BGP Updates whose Subsequent Address Family Identifier (SAFI) values contain the codepoint for MCAST-VPN (as defined in [MVPN-BGP]). The NLRI (Network Layer Reachability Information) field of these BGP Updates includes a customer Multicast Source field and a customer Multicast Group field, thus enabling the customer's (S,G) or (*,G) states to be encoded in the NLRI.

It is also desirable for the BGP/MPLS IP VPN service to be able to support customers who are using MPLS multicast, either instead of or in addition to IP multicast. This document specifies the procedures and protocol extensions needed to support customers who use mLDP [mLDP] to create and maintain Point-to-Multipoint (P2MP) and/or Multipoint-to-Multipoint (MP2MP) Label Switched Paths (LSPs). While mLDP is not the only protocol that can be used to create and maintain multipoint LSPs, consideration of other MPLS multicast control protocols is outside the scope of this document.

When a customer is using mLDP in its VPN, the customer multicast states associated with mLDP are denoted by an mLDP FEC Element (Forwarding Equivalence Class Element; see [mLDP]) instead of by an (S,G) or (*,G). Thus, it is necessary to have a way to encode a customer's mLDP FEC Elements in the NLRI field of the BGP MCAST-VPN routes.

While [MVPN-BGP] does specify a way of encoding an mLDP FEC Element in the MCAST-VPN NLRI field, the encoding specified therein makes a variety of restrictive assumptions about the customer's use of mLDP. (These assumptions are described in Section 2 of this document.) The purpose of this document is to update RFC 6514 [MVPN-BGP] so that customers using mLDP in their VPNs can be supported even when those assumptions do not hold.

Some SPs use the MVPN procedures to provide "global table multicast" service (i.e., multicast service that is not in the context of a VPN) to customers. Methods for doing this are specified in [GTM] and in [SEAMLESS-MCAST]. The procedures described in this document can be used along with the procedures of [GTM] or [SEAMLESS-MCAST] to provide global table multicast service to customers that use MPLS multicast in a global table context.

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

2. Why This Document Is Needed

An mLDP FEC Element consists of a FEC Type, a Root Node, and an Opaque Value. mLDP uses several FEC Types and, in particular, uses the FEC Type to distinguish between P2MP LSPs and MP2MP LSPs.

Section 11.1.2 of [MVPN-BGP] ("Originating Routes: mLDP as the C-Multicast Protocol") states:

> Whenever a PE [Provider Edge router] receives, from one of its CEs [Customer Edge routers], a P2MP Label Map <X, Y, L> over interface I, where X is the Root Node Address, Y is the Opaque Value, and L is an MPLS label ... the PE constructs a Source Tree Join C-multicast route whose MCAST-VPN NLRI contains X as the Multicast Source field, and Y as the Multicast Group field.

In other words, the Root Node of the mLDP FEC Element appears in the Multicast Source field, and the Opaque Value of the mLDP FEC Element appears in the Multicast Group field.

This method of encoding an mLDP FEC in an MCAST-VPN NLRI can only be used if all of the following conditions hold:

1. A customer using mLDP is not also using PIM/IGMP.

   The encoding in [MVPN-BGP] does not specify any way in which one can determine, upon receiving a BGP Update, whether the Multicast Group field contains an IP address or whether it contains an mLDP FEC Element Opaque Value. Therefore, it might not uniquely identify a customer multicast state if the customer is using both PIM/IGMP and mLDP in its VPN.

2. A customer using mLDP is using only the mLDP P2MP FEC Element and is not using the mLDP MP2MP FEC Element.

   The encoding in [MVPN-BGP] does not specify any way to encode the type of the mLDP FEC Element; it just assumes it to be a P2MP FEC Element.

3. A customer using mLDP is using only an mLDP Opaque Value type for which the Opaque Value is exactly 32 bits or 128 bits long.

   The use of Multicast Group fields that have other lengths is declared by [MVPN-BGP] to be "out of scope" of that document (see, e.g., Section 4.3 of that document).

   This condition holds if the customer uses only the mLDP "Generic LSP Identifier" Opaque Value type (defined in [mLDP]). However, mLDP supports many other Opaque Value types whose length is not restricted to be 32 or 128 bits.

The purpose of this document is to update [MVPN-BGP] so that customers using mLDP can be supported, even when these conditions do not hold.

3. Encoding an mLDP FEC in the MCAST-VPN NLRI

When mLDP is used as the customer multicast control protocol, [MVPN-BGP] presupposes that an mLDP FEC Element will be encoded in the NLRI of the following three MCAST-VPN route types:

- C-multicast Source Tree Join route,

- S-PMSI A-D route, and

- Leaf A-D route.

The other four MCAST-VPN route types defined in [MVPN-BGP] do not ever need to carry mLDP FEC Elements. The C-multicast Shared Tree Join route and the Source Active A-D route are used to communicate state about unidirectional shared trees; since mLDP does not have unidirectional shared trees, these routes are not used to signal mLDP states. The Intra-AS I-PMSI A-D route and the Inter-AS I-PMSI A-D route do not identify specific customer multicast states and hence do not carry any information that is specific to the customer’s multicast control protocol.

This document defines three new route types:

- C-Multicast Source Tree Join route for C-multicast mLDP,

- S-PMSI A-D route for C-multicast mLDP, and

- Leaf A-D route for C-multicast mLDP.

The term "C-multicast mLDP" in the names of these route types is intended to indicate that the NLRI of these routes contains mLDP FEC Elements.

Each of these route types corresponds to a route type defined in [MVPN-BGP]. IANA has been requested to allocate codepoints for these three route types such that (a) the high-order two bits have the value 0x01, and (b) the low-order six bits have the same value as the codepoints for the corresponding route types from [MVPN-BGP].

In general, the procedures defined in other MVPN specifications for the C-Multicast Source Tree Join route, the S-PMSI A-D route, and the Leaf A-D route also apply to the C-Multicast Source Tree Join route for C-multicast mLDP, the S-PMSI A-D route for C-multicast mLDP, and the Leaf A-D route for C-multicast mLDP, respectively. However, the NLRI of these three new route types is constructed differently than the NLRI of the corresponding routes from [MVPN-BGP]: the Multicast Source Length, Multicast Source, Multicast Group Length, and

Multicast Group fields are omitted, and in their place is a single mLDP FEC Element, as defined in [mLDP]. (See Section 2.2 of [mLDP] for a diagram of the mLDP FEC Element.)

As a result, the NLRI of an S-PMSI A-D route for C-multicast mLDP will consist of a Route Distinguisher, followed by the mLDP FEC, followed by the Originating Router's IP Address field.

The NLRI of a C-multicast Source Tree Join route for C-multicast mLDP will consist of a Route Distinguisher, followed by the Source AS, followed by the mLDP FEC.

In a Leaf A-D route for C-multicast mLDP that has been derived from an S-PMSI A-D route for C-multicast mLDP, the Route Key field remains the NLRI of the S-PMSI A-D route from which it was derived.

In a Leaf A-D route for C-multicast mLDP that has not been derived from an S-PMSI A-D, the Route Key field is as specified in [SEAMLESS-MCAST], except that the Multicast Source Length, Multicast Source, Multicast Group Length, and Multicast Group fields are omitted, and in their place is a single mLDP FEC Element. Thus, the Route Key field consists of a Route Distinguisher, an mLDP FEC Element, and the IP address of the Ingress PE router.

Please note that [BGP-ERR], Section 5.4 ("Typed NLRI") is applicable if the Route Type field of the NLRI of a received MCAST-VPN route contains an unrecognized value. Any such routes will be discarded.

An mLDP FEC Element contains an address family field whose value is from IANA's "Address Family Numbers" registry. The value of the address family field identifies the address family of the Root Node Address field of the FEC Element. When an mLDP FEC Element is encoded into the NLRI of a BGP Update whose SAFI is MCAST-VPN, the address family of the Root Node Address (as indicated in the FEC Element) MUST correspond to the address family that is identified in the Address Family Identifier (AFI) field of that BGP Update. These two address family fields are considered to correspond to each other under the following conditions:

- they contain identical values,

- the BGP Update's AFI field identifies IPv4 as the address family, and the mLDP FEC Element identifies "Multi-Topology IPv4" as the address family of the Root Node, or

- the BGP Update's AFI field identifies IPv6 as the address family, and the mLDP FEC Element identifies "Multi-Topology IPv6" as the address family of the Root Node.

For more information about the "multi-topology" address families, see [LDP-MT] and [mLDP-MT].

4. Wildcards

[MVPN-WILDCARDS] specifies encodings and procedures that allow "wildcards" to be specified in the NLRI of S-PMSI A-D routes. A set of rules are given that specify when a customer multicast flow "matches" a given S-PMSI A-D route whose NLRI contains wildcards. However, the use of these wildcards is defined only for the case where the customer is using PIM as its multicast control protocol. The use of wildcards when the customer is using mLDP as its multicast control protocol is outside the scope of this document.

5. IANA Considerations

[MVPN-BGP] does not create a registry for the allocation of new MCAST-VPN Route Type values. In retrospect, it seems that it should have done so. IANA has created a new registry called "BGP MCAST-VPN Route Types", which references this document and [MVPN-BGP]. The registry has been created under the top-level registry: "Border Gateway Protocol (BGP) Parameters" <http://www.iana.org/assignments/bgp-parameters>. The allocation policy is "Standards Action". Values may be assigned from one of several ranges:

- Range 0x01-0x3f: Generic/PIM Range. Values are assigned from this range when the NLRI format associated with the route type presupposes that PIM or IGMP is the C-multicast control protocol or when the NLRI format associated with the route type is independent of the C-multicast control protocol.

- Range 0x43-0x7f: mLDP Range. Values are assigned from this range when the NLRI format associated with the route type presupposes that mLDP is the C-multicast control protocol.

- Range 0x80-0xff: This range is reserved; values should not be assigned from this range.

In general, whenever an assignment is requested from this registry, two codepoints should be requested at the same time: one from the Generic/PIM range and one from the mLDP range. The two codepoints should have the same low-order 6 bits. If one of the two codepoints is not actually needed, it should be registered anyway and marked as "Reserved".

The "BGP MCAST-VPN Route Types" contains the following initial assignments:

Value	Meaning	Reference
0x00	Reserved	This RFC
0x01	Intra-AS I-PMSI A-D route	This RFC, [RFC6514]
0x02	Inter-AS I-PMSI A-D route	This RFC, [RFC6514]
0x03	S-PMSI A-D route	This RFC, [RFC6514]
0x04	Leaf A-D route	This RFC, [RFC6514]
0x05	Source Active A-D route	This RFC, [RFC6514]
0x06	Shared Tree Join route	This RFC, [RFC6514]
0x07	Source Tree Join route	This RFC, [RFC6514]
0x08-0x3f	Unassigned (Generic/PIM range)	This RFC
0x40-0x42	Reserved	This RFC
0x43	S-PMSI A-D route for C-multicast mLDP	This RFC
0x44	Leaf A-D route for C-multicast mLDP	This RFC
0x45-0x46	Reserved	This RFC
0x47	Source Tree Join route for C-multicast mLDP	This RFC
0x48-0x7f	Unassigned (mLDP range)	This RFC
0x80-0xff	Reserved	This RFC

6. Security Considerations

This document specifies a method of encoding an mLDP FEC Element in the NLRI of some of the BGP Update messages that are specified in [MVPN-BGP]. The security considerations of [mLDP] and of [MVPN-BGP] are applicable, but no new security considerations are raised.

7. References

7.1. Normative References

[mLDP] 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, <http://www.rfc-editor.org/info/rfc6388>.

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

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

7.2. Informative References

[BGP-ERR] Chen, E., Ed, Scudder, J., Mohapatra, P., and K. Patel, "Revised Error Handling for BGP UPDATE Messages", Work in Progress, draft-ietf-idr-error-handling-18, December 2014.

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

[IGMP] Cain, B., Deering, S., Kouvelas, I., Fenner, B., and A. Thyagarajan, "Internet Group Management Protocol, Version 3", RFC 3376, October 2002, <http://www.rfc-editor.org/info/rfc3376>.

[LDP-MT] Zhao, Q., Raza, K., Zhou, C., Fang, L., Li, L., and D. King, "LDP Extensions for Multi-Topology", RFC 7307, July 2014, <http://www.rfc-editor.org/info/rfc7307>.

[mLDP-MT] Wijnands, IJ. and K. Raza, "mLDP Extensions for Multi Topology Routing", Work in Progress, draft-iwijnand-mpls-mldp-multi-topology-03, June 2013.

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

[PIM] Fenner, B., Handley, M., Holbrook, H., and I. Kouvelas, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", RFC 4601, August 2006, <http://www.rfc-editor.org/info/rfc4601>.

[SEAMLESS-MCAST] Rekhter, Y., Aggarwal, R., Morin, T., Grosclaude, I., Leymann, N., and S. Saad, "Inter-Area P2MP Segmented LSPs", Work in Progress, draft-ietf-mpls-seamless-mcast-14, June 2014.

Acknowledgments

The authors wish to thank Pradosh Mohapatra and Saquib Najam for their ideas and comments. We also thank Yakov Rekhter and Martin Vigoureux for their comments.

Authors' Addresses

IJsbrand Wijnands
Cisco Systems, Inc.
De kleetlaan 6a Diegem 1831
Belgium
EMail: ice@cisco.com

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

Uwe Joorde
Deutsche Telekom
Dahlweg 100
D-48153 Muenster
Germany
EMail: Uwe.Joorde@telekom.de