             Encoding mLDP FECs in the NLRI of BGP MCAST-VPN Routes

                   draft-ietf-l3vpn-mvpn-mldp-nlri-03.txt

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 Multicast Extensions to Label Distribution
   Protocol (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.

Status of this Memo

The list of current Internet-Drafts can be accessed at
http://www.ietf.org/ietf/1id-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at
http://www.ietf.org/shadow.html.

Table of Contents

[1](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](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" value is "MCAST-VPN" (5). The NLRI ("Network Layer
Reachability Information") field of these 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
("Multicast Extensions to Label Distribution Protocol") [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 Equivalance 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 [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](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 control protocol") states:

> Whenever a PE receives from one of its CEs 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 may
   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.

In addition, neither [MVPN-BGP] nor [MVPN-WILDCARDS] addresses the
use of "wild cards" when the MCAST-VPN NLRI encodes an MLDP FEC.
This document specifies a way to encode mLDP FEC Element wild cards
in the NLRI of the relevant BGP MCAST-VPN routes.

[3](3). **Encoding an mLDP FEC in the MCAST-VPN NLRI**

This section specifies the way to encode an mLDP FEC element in the
NLRI of the following three MCAST-VPN route types defined in [MVPN-
BGP]:

   - C-multicast Source Tree Join,

   - 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.

Per [MVPN-BGP], the first octet of the NLRI of an MCAST-VPN route is
a "route type". Only values 1-7 are defined. The high order 5 bits

of that octet are thus always zero.

This document updates [MVPN-BGP] by specifying a use for the high
order 2 bits of the "route type" octet. The following two values are
defined:

  - If the two high order bits are both zero, the NLRI is as
    specified in [MVPN-BGP] and/or [MVPN-WILDCARDS].

  - If the two high order bits have the value 01, the NLRI encoding
    is modified as follows: 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.

The other two possible values (11 and 10) for the two high order bits
may be used at a later time to identify other multicast control
protocols.

As a result, the NLRI of an S-PMSI A-D route with an mLDP FEC in its
NLRI 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 with an mLDP FEC in
its NLRI will consist of a Route Distinguisher, followed by the
Source AS, followed by the mLDP FEC.

In a Leaf A-D route that has been derived from an S-PMSI A-D route,
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 that has not been derived from an S-PMSI A-D
route, the "route key" field is as specified in [SEGMENTED-MVPN],
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.

An mLDP FEC element contains an "address family" field from IANA's
"Address Family Numbers" registry. This 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 an a BGP update whose
SAFI is MCAST-VPN, the address family of the root node (as indicated
in the FEC element) MUST "correspond to" the address family that is
identified in the AFI field of that BGP update. These two "address
family" fields are considered to "correspond" under the following

conditions:

   - they contain identical values, or

   - 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](LDP-MT)] and [mLDP-MT].

[4](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](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 should create a registry called "MCAST-VPN Route
   Types", referencing this document and [MVPN-BGP]. The allocation
   policy should be "Standards Action with Early Allocation", and the
   assignable values are in the range 0-0xFF. The following values
   should be assigned:

   - 0x00: Reserved

   - 0x01: Intra-AS I-PMSI A-D route (reference: [MVPN-BGP])

   - 0x02: Inter-AS I-PMSI A-D route (reference: [MVPN-BGP])

   - 0x03: S-PMSI A-D route for PIM as the C-multicast control
     protocol (reference: [MVPN-BGP])

    - 0x43: S-PMSI A-D route for mLDP as the C-multicast control
      protocol (reference: this document)

    - 0x04: Leaf A-D route for PIM as the C-multicast control protocol
      (reference: [MVPN-BGP])

    - 0x44: Leaf A-D route for mLDP as the C-multicast control protocol
      (reference: this document)

    - 0x05: Source Active A-D route for PIM as the C-multicast control
      protocol (reference: [MVPN-BGP])

    - 0x06: Shared Tree Join route for PIM as the C-multicast control
      protocol (reference: [MVPN-BGP])

    - 0x07: Source Tree Join route for PIM as the C-multicast control
      protocol (reference: [MVPN-BGP])

    - 0x47: Source Tree Join route for mLDP as the C-multicast control
      protocol (reference: this document)

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. Acknowledgments

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

8. Authors' Addresses

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

Eric C. Rosen
Cisco Systems, Inc.
1414 Massachusetts Avenue
Boxborough, MA, 01719
E-mail: erosen@cisco.com

Uwe Joorde
Deutsche Telekom
Hammer Str. 216-226
D-48153 Muenster, Germany
E-mail: Uwe.Joorde@telekom.de

[9](9). Normative References

[mLDP] "Label Distribution Protocol Extensions for Point-to-
Multipoint and Multipoint-to-Multipoint Label Switched Paths",
Wijnands, Minei, Kompella, Thomas, [RFC 6388](RFC 6388), November 2011

[MVPN-BGP] "BGP Encodings and Procedures for Multicast in MPLS/BGP IP
VPNs", Aggarwal, Rosen, Morin, Rekhter, [RFC 6514](RFC 6514), February 2012

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

[10](10). Informative References

[GTM] "Global Table Multicast with BGP-MVPN Procedures", Zhang,
Giuliano, Rosen, Subramanian, Pacella, Schiller, draft-zzhang-l3vpn-mvpn-global-table-mcast-01.txt, October 2013

[IGMP] "Internet Group Management Protocol, Version 3", Cain,
Deering, Kouvelas, Fenner, Thyagarajan, [RFC 3376](RFC 3376), October 2002

[LDP-MT] "LDP Extensions for Multi-Topology Routing", Zhao, et. al.,
draft-ietf-mpls-ldp-multi-topology-09.txt, October 2013

[mLDP-MT] "mLDP Extensions for Multi Topology Routing", Wijnands,
Raza, draft-iwijnand-mpls-mldp-multi-topology-03.txt, June 2013

[MVPN-WILDCARDS], "Wildcards in Multicast VPN Auto-Discovery Routes",
Rosen, Rekhter, Hendrickx, Qiu, [RFC 6625](RFC 6625), May 2012

[PIM] "Protocol Independent Multicast - Sparse Mode (PIM-SM)",
Fenner, Handley, Holbrook, Kouvelas, August 2006, [RFC 4601](RFC 4601)

   [SEAMLESS-MCAST] "Inter-Area P2MP Segmented LSPs", Rekhter, Aggarwal,
   Morin, Grosclaude, Leymann, Saad, draft-ietf-mpls-seamless-
   mcast-08.txt, November 2013