Network Working Group   IJ. Wijnands, Ed.
Internet-Draft   Cisco Systems, Inc.
Intended status: Standards Track   P. Hitchen
Expires: April 9, 2012   BT
                                                         N. Leymann
                                                    Deutsche Telekom
                                                       W. Henderickx
                                                      Alcatel-Lucent
                                                     October 7, 2011

Multipoint Label Distribution Protocol
In-Band Signaling in a VPN Context
draft-wijnands-mpls-mldp-vpn-in-band-signaling-00

Abstract

   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 LSP
   (Label Switched Path) 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. [I-D.ietf-mpls-mldp-in-band-signaling] specifies the
   procedures for building such a hybrid MDT, using Protocol Independent
   Multicast (PIM) as the control protocol in the non-MPLS region of the
   network, and using Multipoint Extensions to Label Distribution
   Protocol (mLDP) in the MPLS region. This document extends those
   procedures so that they will work when the links between the MPLS and
   non-MPLS regions are [RFC4364] interfaces. While these procedures do
   not provide a good general multicast VPN solution, they are useful in
   certain specific situations.

This Internet-Draft will expire on April 9, 2012.

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. In such
   cases, it can be useful to begin building the MDT as a pure IP MDT,
   then convert it to an MPLS Multipoint LSP (Label Switched Path) 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.
   [I-D.ietf-mpls-mldp-in-band-signaling] specifies the procedures for
   building such a hybrid MDT, using Protocol Independent Multicast
   (PIM) [RFC4601] as the control protocol in the non-MPLS region of the
   network, and using Multipoint Extensions to Label Distribution
   Protocol (mLDP) [I-D.ietf-mpls-ldp-p2mp] in the MPLS region.

   For a given tree, one or more routers that are at the border between
   a non-MPLS-enabled region and an MPLS-enabled region receive PIM
   control messages from the non-MPLS-enabled region, and convert them
   to mLDP control messages to be sent into the MPLS-enabled region.
   Another router that is at the border between a non-MPLS-enabled
   region and an MPLS-enabled region receives mLDP control messages from
   the MPLS-enabled region and converts them to PIM control messages to
   be sent into a non-MPLS-enabled region.

   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
   [I-D.ietf-mpls-mldp-in-band-signaling] 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, those procedures 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 VPN as
   described in [RFC4364]. This specification extends the procedures of
   [I-D.ietf-mpls-mldp-in-band-signaling] so that they may be applied in
   the VPN context.

   As in [I-D.ietf-mpls-mldp-in-band-signaling], 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.

Each multicast tree in the non-MPLS enabled region is mapped 1-1 onto
a multipoint LSP in the MPLS-enabled region. For a variety of
reasons (discussed in [I-D.ietf-l3vpn-2547bis-mcast]), this is not a
suitable for a general purpose multicast VPN solution. But the
procedures described herein are much simpler than the general purpose
MVPN procedures, and are applicable when the 1-1 mapping is
acceptable, and when it is acceptable to use mLDP as the protocol for
setting up the multipoint LSPs. For example, some service providers
offer multicast content to their customers, but have chosen to use
VPNs to isolate the various customers and services. This is a very
different scenario than the general MVPN scenario, in which the
customers provide their own multicast content, out of the control of
the service provider.

Due to the 1-1 mapping and the multicast source and group information
being encoded in the mLDP FEC, there is deterministic mapping beween
the multicast tree and the mLDP LSP in the core network. This
improves and simplifies fault resolution.

In order to use the mLDP in-band signaling procedures for a
particular group address in a particular VPN, the Provider Edge (PE)
routers that attach to that VPN 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 ("Virtual Routing and
Forwarding table", defined in [RFC4364]). For those groups, and
those groups only, the procedures of this document are used instead
of the general purpose Multicast VPN procedures. This configuration
must be present in all PE routers that attach to sites containing
senders or receivers for the given set of group addresses.

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

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

   mLDP : Multicast LDP.

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

   P2MP LSP: An LSP that has one Ingress LSR and one or more Egress
      LSRs (see [I-D.ietf-mpls-ldp-p2mp]).

   MP2MP LSP: An LSP that connects a set of leaf nodes, acting
      indifferently as ingress or egress (see [I-D.ietf-mpls-ldp-p2mp]).

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

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

2. VPN In-band signaling for MP LSPs

   Suppose that a PE router, PE1, attaching to a particular VPN,
   receives a PIM Join(S,G) message over one of its VRF interfaces.
   Following the procedure of section 5.1 of
   [I-D.ietf-l3vpn-2547bis-mcast], PE1 determines the "upstream RD", the
   "upstream PE", and the "upstream multicast hop" (UMH) for the source
   address S. Please note that sections 5.1.1 and 5.1.2 of
   [I-D.ietf-l3vpn-2547bis-mcast] are applicable.

   In order to transport the multicast tree via a MP LSP using VPN 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
   [I-D.ietf-mpls-ldp-p2mp], 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 coding
   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
its path to the upstream PE. The "next hop" is usually the IGP next
hop, but see [I-D.napierala-mpls-targeted-mldp] for cases in which
the next hop is not the IGP next hop.

If the UMH and the upstream PE do not have the same IP address, the
procedures of section 2 of [I-D.ietf-mpls-mldp-recurs-fec] 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 types of MDT are
defined: 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
[I-D.ietf-mpls-mldp-in-band-signaling].

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: (to be assigned by IANA).

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: (to be assigned by IANA).

   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                               |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
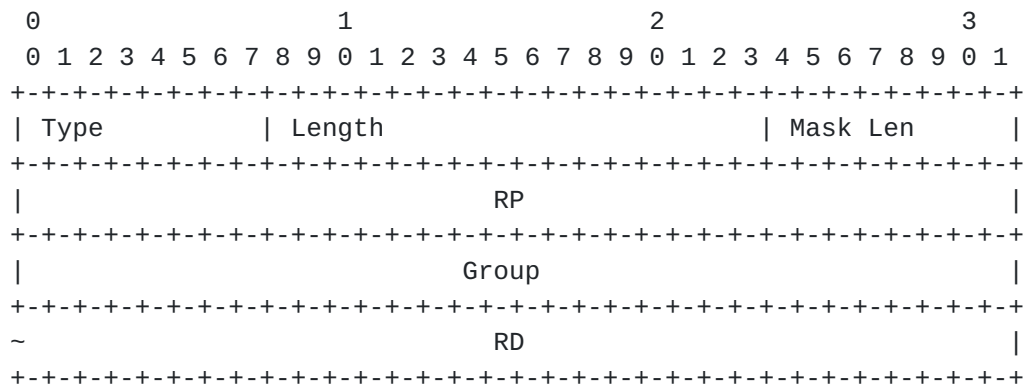

   Type: (to be assigned by IANA).

   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 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: (to be assigned by IANA).

   Length: 41

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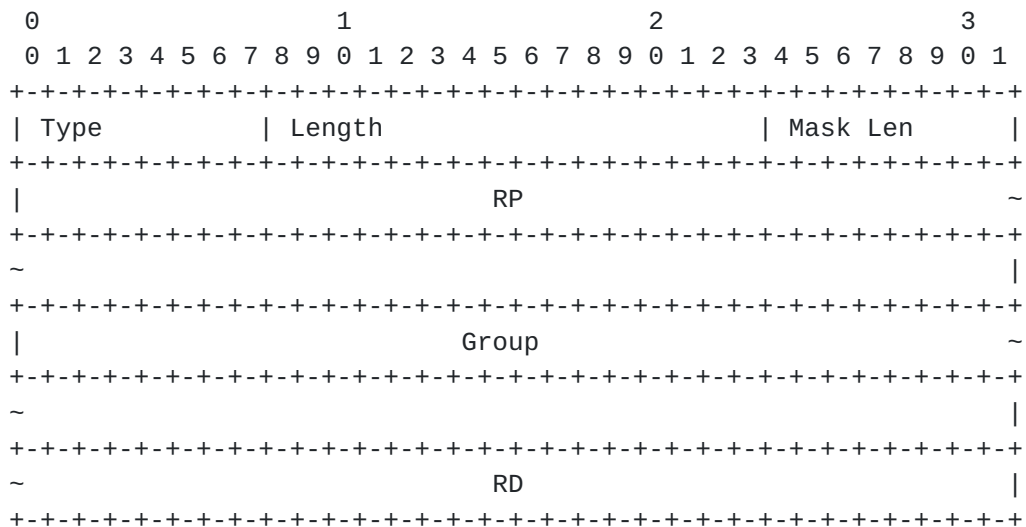   and used as a mask, 1 octet.

   RP: Rendezvous Point (RP) IPv6 address used for 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 [I-D.ietf-mpls-ldp-p2mp]

5. IANA considerations

[I-D.ietf-mpls-ldp-p2mp] defines a registry for "The LDP MP Opaque
Value Element Basic Type". This document requires the assignment of
four new code points in this registry:

Transit VPNv4 Source TLV type - requested 10

Transit VPNv6 Source TLV type - requested 11

Transit VPNv4 Bidir TLV type - requested 12

Transit VPNv6 Bidir TLV type - requested 13

6. Acknowledgments

Thanks to Eric Rosen and Andy Green for their comments on the draft.

7. References

7.1. Normative References

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

[RFC4601] Fenner, B., Handley, M., Holbrook, H., and I. Kouvelas,
           "Protocol Independent Multicast - Sparse Mode (PIM-SM):
           Protocol Specification (Revised)", RFC 4601, August 2006.

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

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

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

[I-D.ietf-mpls-ldp-p2mp]
            Minei, I., Kompella, K., Wijnands, I., and B. Thomas,
            "Label Distribution Protocol Extensions for Point-to-
            Multipoint and Multipoint-to-Multipoint Label Switched
            Paths", draft-ietf-mpls-ldp-p2mp-11 (work in progress),
            October 2010.

[I-D.ietf-mpls-mldp-in-band-signaling]
            Wijnands, I., Eckert, T., Leymann, N., and M. Napierala,
            "mLDP based in-band signaling for Point-to-Multipoint and
            Multipoint-to- Multipoint Label Switched Paths",
            draft-ietf-mpls-mldp-in-band-signaling-02 (work in
            progress), July 2010.

[I-D.ietf-mpls-mldp-recurs-fec]
            Wijnands, I., Rosen, E., Napierala, M., and N. Leymann,
            "Using mLDP through a Backbone where there is no Route to
            the Root", draft-ietf-mpls-mldp-recurs-fec-00 (work in
            progress), October 2010.

[I-D.ietf-l3vpn-2547bis-mcast]
            Aggarwal, R., Bandi, S., Cai, Y., Morin, T., Rekhter, Y.,
            Rosen, E., Wijnands, I., and S. Yasukawa, "Multicast in
            MPLS/BGP IP VPNs", draft-ietf-l3vpn-2547bis-mcast-10 (work
            in progress), January 2010.

7.2. Informative References

[I-D.napierala-mpls-targeted-mldp]
            Napierala, M., Rosen, E., and I. Wijnands, "Using LDP
            Multipoint Extensions on Targeted LDP Sessions",
            draft-napierala-mpls-targeted-mldp-00 (work in progress),
            January 2011.

Authors' Addresses

   IJsbrand Wijnands (editor)
   Cisco Systems, Inc.
   De kleetlaan 6a
   Diegem 1831
   Belgium

   Email: ice@cisco.com

   Paul Hitchen
   BT
   BT Adastral Park
   Ipswich IP53RE
   UK

   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