              Multipoint LDP In-Band Signaling for
 Point-to-Multipoint and Multipoint-to-Multipoint Label Switched Paths

Abstract

   Consider an IP multicast tree, constructed by Protocol Independent
   Multicast (PIM), that needs to pass through an MPLS domain in which
   Multipoint LDP (mLDP) point-to-multipoint and/or multipoint-to-
   multipoint Labels Switched Paths (LSPs) can be created. The part of
   the IP multicast tree that traverses the MPLS domain can be
   instantiated as a multipoint LSP. When a PIM Join message is
   received at the border of the MPLS domain, information from that
   message is encoded into mLDP messages. When the mLDP messages reach
   the border of the next IP domain, the encoded information is used to
   generate PIM messages that can be sent through the IP domain. The
   result is an IP multicast tree consisting of a set of IP multicast
   sub-trees that are spliced together with a multipoint LSP. This
   document describes procedures regarding how IP multicast trees are
   spliced together with multipoint LSPs.

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/rfc6826.

Copyright Notice

Table of Contents

1. Introduction

   The mLDP (Multipoint LDP) [RFC6388] specification describes
   mechanisms for creating point-to-multipoint (P2MP) and multipoint-to-
   multipoint (MP2MP) LSPs (Label Switched Paths). These LSPs are
   typically used for transporting end-user multicast packets. However,
   the mLDP specification does not provide any rules for associating
   particular end-user multicast packets with any particular LSP. Other
   documents, like [RFC6513], describe applications in which out-of-band
   signaling protocols, such as PIM and BGP, are used to establish the
   mapping between an LSP and the multicast packets that need to be
   forwarded over the LSP.

   This document describes an application in which the information
   needed to establish the mapping between an LSP and the set of
   multicast packets to be forwarded over it is carried in the "opaque
   value" field of an mLDP FEC (Forwarding Equivalence Class) element.
   When an IP multicast tree (either a source-specific tree or a
   bidirectional tree) enters the MPLS network, the (S,G) or (*,G)
   information from the IP multicast control-plane state is carried in
   the opaque value field of the mLDP FEC message. As the tree leaves
   the MPLS network, this information is extracted from the FEC Element
   and used to build the IP multicast control plane. PIM messages can
   be sent outside the MPLS domain. Note that although the PIM control
   messages are sent periodically, the mLDP messages are not.

   Each IP multicast tree is mapped one-to-one to a P2MP or MP2MP LSP in
   the MPLS network. A network operator should expect to see as many
   LSPs in the MPLS network as there are IP multicast trees. A network
   operator should be aware how IP multicast state is created in the
   network to ensure that it does not exceed the scalability numbers of
   the protocol, either PIM or mLDP.

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

   ASM: PIM Any Source Multicast

   Egress LSR: One of potentially many destinations of an LSP; also
      referred to as leaf node in the case of P2MP and MP2MP LSPs.

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

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

IP multicast tree: An IP multicast distribution tree identified by
     an IP multicast Group address and, optionally, by a Source IP
     address, also referred to as (S,G) and (*,G).

LSR: Label Switching Router

LSP: Labels Switched Path

mLDP: Multipoint LDP

MP2MP LSP: An LSP that connects a set of leaf nodes that may each
     independently act as ingress or egress.

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

P2MP LSP: An LSP that has one Ingress Label Switching Router (LSR)
     and one or more Egress LSRs.

RP: PIM Rendezvous Point

SSM: PIM Source-Specific Multicast

Transit LSP: A P2MP or MP2MP LSP whose FEC Element contains the
     (S,G) or (*,G) identifying a particular IP multicast distribution
     tree.

Transit LSR: An LSR that has one or more directly connected
     downstream LSRs.

2. In-Band Signaling for MP LSPs

   Consider the following topology:

```
              |--- IPM ---|--- MPLS --|--- IPM ---|

         S/RP -- (A) - (U) - (C) - (D) -- (B) -- R

                         Figure 1
```

Nodes A and B are IP-multicast-capable routers and connect to a
Source/RP and a Receiver, respectively. Nodes U, C, and D are MPLS
Label Switched Routers (LSRs).

LSR D is attached to a network that is capable of MPLS multicast and
IP multicast (see figure 1), and D is required to create a IP
multicast tree due to a certain IP multicast event, like a PIM Join,
MSDP Source Announcement (SA) [RFC3618], BGP Source Active auto-
discovery route [SM-MLDP], or Rendezvous Point (RP) discovery.
Suppose that D can determine that the IP multicast tree needs to
travel through the MPLS network until it reaches LSR U. For
instance, when D looks up the route to the Source or RP [RFC4601] of
the IP multicast tree, it may discover that the route is a BGP route
with U as the BGP next hop. Then D may choose to set up a P2MP or an
MP2MP LSP, with U as root, and to make that LSP become part of the IP
multicast distribution tree. Note that other methods are possible to
determine that an IP multicast tree is to be transported across an
MPLS network using P2MP or MP2MP LSPs. However, these methods are
outside the scope of this document.

In order to establish a multicast tree via a P2MP or MP2MP LSP using
"in-band signaling", LSR D encodes a P2MP or MP2MP FEC Element, with
the IP address of LSR U as the "Root Node Address" and with the
source and the group encoded into the "opaque value" ([RFC6388],
Sections 2.2 and 3.2). Several different opaque value types are
defined in this document. LSR D MUST NOT use a particular opaque
value type unless it knows (through provisioning or through some
other means outside the scope of this document) that LSR U supports
the root node procedures for that opaque value type.

The particular type of FEC Element and opaque value used depends on
the IP address family being used, and on whether the multicast tree
being established is a source-specific or a bidirectional multicast
tree.

When an LSR receives a label mapping or withdraw whose FEC Element
contains one of the opaque value types defined in this document, and
that LSR is not the one identified by the "Root Node Address" field
of that FEC Element, the LSR follows the procedures provided in RFC
6388.

When an LSR receives a label mapping or withdraw whose FEC Element
contains one of the opaque value types defined in this document, and
that LSR is the one identified by the Root Node Address field of that
FEC Element, then the following procedure is executed. The multicast
source and group are extracted and passed to the multicast code. If
a label mapping is being processed, the multicast code will add the
downstream LDP neighbor to the olist of the corresponding (S,G) or

(*,G) state, creating such state if it does not already exist. If a
label withdraw is being processed, the multicast code will remove the
downstream LDP neighbor from the olist of the corresponding (S,G) or
(*,G) state. From this point on, normal PIM processing will occur.

Note that if the LSR identified by the Root Node Address field does
not recognize the opaque value type, the MP LSP will be established,
but the root node will not send any multicast data packets on it.

Source or RP addresses that are reachable in a VPN context are
outside the scope of this document.

Multicast groups that operate in PIM Dense-Mode are outside the scope
of this document.

2.1. Transiting Unidirectional IP Multicast Shared Trees

Nothing prevents PIM shared trees, used by PIM-SM in the ASM service
model, from being transported across an MPLS core. However, it is
not possible to prune individual sources from the shared tree without
the use of an additional out-of-band signaling protocol, like PIM or
BGP [SM-MLDP]. For this reason, transiting shared trees across a
transit LSP is outside the scope of this document.

2.2. Transiting IP Multicast Source Trees

IP multicast source trees can be created via PIM operating in either
SSM mode [RFC4607] or ASM mode [RFC4601]. When PIM-SM is used in ASM
mode, the usual means of discovering active sources is to join a
sparse-mode shared tree. However, this document does not provide any
method of establishing a sparse-mode shared tree across an MPLS
network. To apply the technique of this document to PIM-SM in ASM
mode, there must be some other means of discovering the active
sources. One possible means is the use of MSDP [RFC3618]. Another
possible means is to use BGP Source Active auto-discovery routes, as
documented in [SM-MLDP]. However, the method of discovering the
active sources is outside the scope of this document; as a result,
this document does not specify everything that is needed to support
the ASM service model using in-band signaling.

The source and group addresses are encoded into the a transit TLV as
specified in Sections 3.1 and 3.2.

2.3. Transiting IP Multicast Bidirectional Trees

   If a bidirectional IP multicast tree [RFC5015] has to be transported
   over an MPLS network using in-band signaling, as described in this
   document, it MUST be transported using an MP2MP LSPs. A
   bidirectional tree does not have a specific source address; the group
   address, subnet mask, and RP are relevant for multicast forwarding.
   This document does not provide procedures to discover RP-to-group
   mappings dynamically across an MPLS network and assumes the RP is
   statically defined. Support of dynamic RP mappings in combination
   with in-band signaling is outside the scope of this document.

   The RP for the group is used to select the ingress LSR and the root
   of the LSP. The group address is encoded according to the rules of
   Sections 3.3 or 3.4, depending on the IP version. The subnet mask
   associated with the bidirectional group is encoded in the Transit
   TLV. There are two types of bidirectional states in IP multicast,
   the group specific state and the RP state. The first type is
   typically created when a PIM Join has been received and has a subnet
   mask of 32 for IPv4 and 128 for IPv6. The RP state is typically
   created via the static RP mapping and has a variable subnet mask.
   The RP state is used to build a tree to the RP and is used for
   sender-only branches. Each state (group specific and RP state) will
   result in a separate MP2MP LSP. The merging of the two MP2MP LSPs
   will be done by PIM on the root LSR. No special procedures are
   necessary for PIM to merge the two LSPs. Each LSP is effectively
   treated as a PIM-enabled interface. Please see [RFC5015] for more
   details.

   For transporting the packets of a sender-only branch, we create a
   MP2MP LSP. Other sender-only branches will receive these packets and
   will not forward them because there are no receivers. These packets
   will be dropped. If that effect is undesirable, some other means of
   transport has to be established to forward packets to the root of the
   tree, for example, a multipoint-to-point LSP for example. A
   technique to unicast packets to the root of a P2MP or MP2MP LSP is
   documented in Section 3.2.2.1 of [MVPN-MSPMSI].

3. LSP Opaque Encodings

This section documents the different transit opaque encodings.

3.1. Transit IPv4 Source TLV

```
 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         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
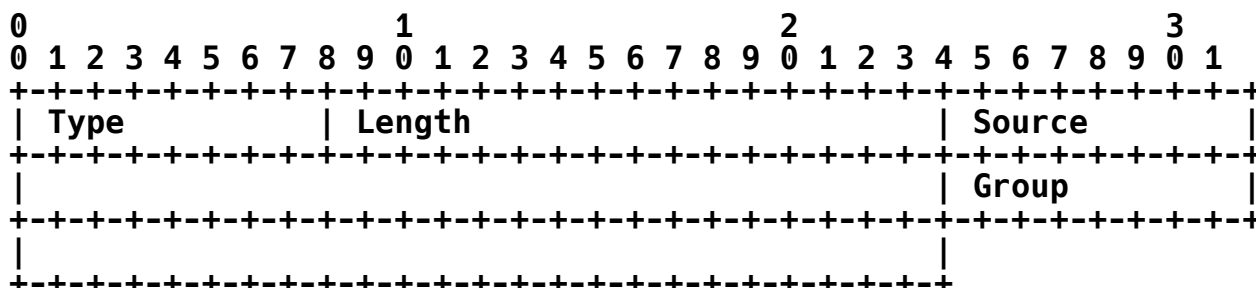
Type: 3

Length: 8 (octet size of Source and Group fields)

Source: IPv4 multicast source address, 4 octets

Group: IPv4 multicast group address, 4 octets

3.2. Transit IPv6 Source TLV

```
 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       ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
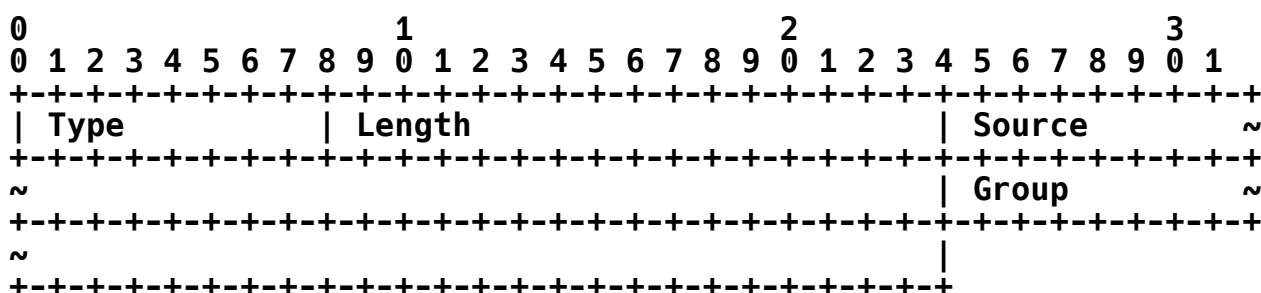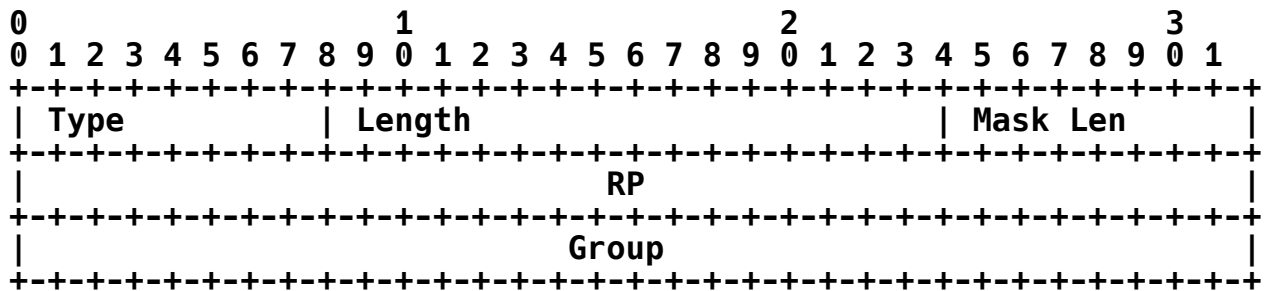
Type: 4

Length: 32 (octet size of Source and Group fields)

Source: IPv6 multicast source address, 16 octets

Group: IPv6 multicast group address, 16 octets.

3.3. Transit IPv4 Bidir TLV

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

Type: 5

Length: 9 (octet size of Mask Len, RP, and Group fields)

Mask Len: The number of contiguous one bits that are left-justified
   and used as a mask, 1 octet. Maximum value allowed is 32.

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

Group: IPv4 multicast group address, 4 octets.

3.4. Transit IPv6 Bidir TLV

```
 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                              ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
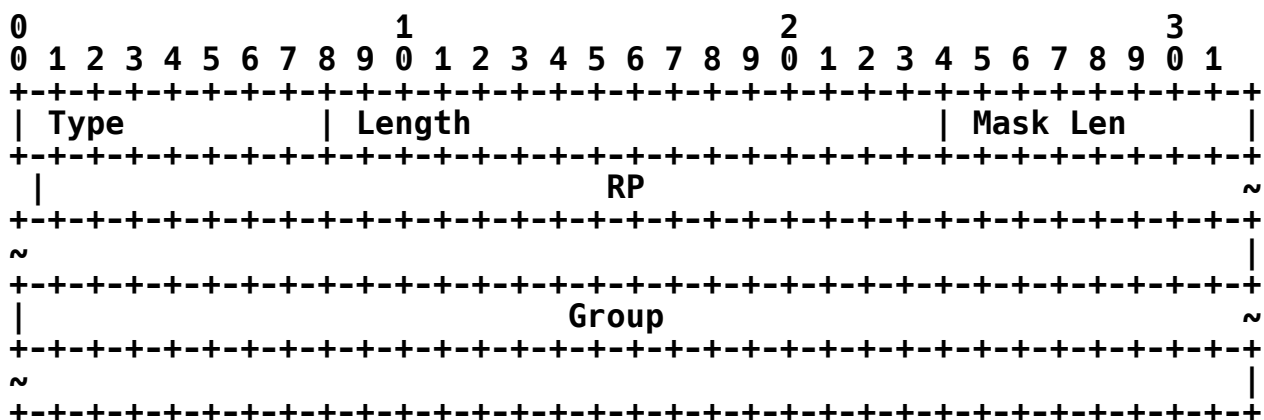
Type: 6

Length: 33 (octet size of Mask Len, RP and Group fields)

Mask Len: The number of contiguous one bits that are left-justified
   and used as a mask, 1 octet. Maximum value allowed is 128.

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

Group: IPv6 multicast group address, 16 octets.

4. Security Considerations

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

5. IANA Considerations

IANA has allocated the following values from the "LDP MP Opaque Value
Element basic type" registry: are:

Transit IPv4 Source TLV type - 3

Transit IPv6 Source TLV type - 4

Transit IPv4 Bidir TLV type - 5

Transit IPv6 Bidir TLV type - 6

6. References

6.1. Normative References

[RFC5036]   Andersson, L., Ed., Minei, I., Ed., and B. Thomas, Ed.,
            "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.

[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.

6.2. Informative References

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

[RFC4607]   Holbrook, H. and B. Cain, "Source-Specific Multicast for
            IP", RFC 4607, August 2006.

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

   [RFC3618]   Fenner, B., Ed., and D. Meyer, Ed., "Multicast Source
               Discovery Protocol (MSDP)", RFC 3618, October 2003.

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

   [SM-MLDP]   Rekhter, Y., Aggarwal, R., and N. Leymann, "Carrying PIM-
               SM in ASM mode Trees over P2MP mLDP LSPs", Work in
               Progress, August 2011.

   [MVPN-MSPMSI]
               Cai, Y., Rosen, E., Ed., Napierala, M., and A. Boers,
               MVPN: Optimized use of PIM via MS-PMSIs", February 2012.

7. Acknowledgments

Authors' Addresses

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

    EMail: ice@cisco.com

    Toerless Eckert
    Cisco Systems, Inc.
    170 Tasman Drive
    San Jose CA, 95134
    USA

    EMail: eckert@cisco.com

    Nicolai Leymann
    Deutsche Telekom
    Winterfeldtstrasse 21
    Berlin 10781
    Germany

    EMail: n.leymann@telekom.de

    Maria Napierala
    AT&T Labs
    200 Laurel Avenue
    Middletown NJ 07748
    USA

    EMail: mnapierala@att.com