Multipoint LDP (mLDP) In-Band Signaling with Wildcards

Abstract

There are scenarios in which an IP multicast tree traverses an MPLS domain. In these scenarios, it can be desirable to convert the IP multicast tree "seamlessly" into an MPLS Multipoint Label Switched Path (MP-LSP) when it enters the MPLS domain, and then to convert it back to an IP multicast tree when it exits the MPLS domain. Previous documents specify procedures that allow certain kinds of IP multicast trees (either Source-Specific Multicast trees or Bidirectional Multicast trees) to be attached to an MPLS Multipoint Label Switched Path (MP-LSP). However, the previous documents do not specify procedures for attaching IP Any-Source Multicast trees to MP-LSPs, nor do they specify procedures for aggregating multiple IP multicast trees onto a single MP-LSP. This document specifies the procedures to support these functions. It does so by defining "wildcard" encodings that make it possible to specify, when setting up an MP-LSP, that a set of IP multicast trees, or a shared IP multicast tree, should be attached to that MP-LSP. Support for non-bidirectional IP Any-Source Multicast trees is subject to certain applicability restrictions that are discussed in this document. This document updates RFCs 6826 and 7246.
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/rfc7438.

Copyright Notice

Copyright (c) 2015 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.
1.  Introduction

[RFC6826] and [RFC7246] specify procedures for mLDP (Multipoint LDP) that allow an IP multicast tree (either a Source-Specific Multicast tree or a Bidirectional Multicast tree) to be attached "seamlessly" to an MPLS Multipoint Label Switched Path (MP-LSP). This can be useful, for example, when there is multicast data that originates in a domain that supports IP multicast, which then has to be forwarded across a domain that supports MPLS multicast and then has to forwarded across another domain that supports IP multicast. By attaching an IP multicast tree to an MP-LSP, data that is traveling along the IP multicast tree can be moved seamlessly to the MP-LSP when it enters the MPLS multicast domain. The data then travels along the MP-LSP through the MPLS domain. When the data reaches the boundary of the MPLS domain, it can be moved seamlessly to an IP multicast tree. This ability to attach IP multicast trees to MPLS MP-LSPs can be useful in either VPN context or global context.

In mLDP, every MP-LSP is identified by the combination of a "root node" (or "Ingress Label Switching Router (LSR)") and an "Opaque Value" that, in the context of the root node, uniquely identifies the MP-LSP. These are encoded into an mLDP "Forwarding Equivalence Class (FEC) Element". To set up an MP-LSP, the Egress LSRs originate mLDP
control messages containing the FEC element. A given FEC Element value identifies a single MP-LSP and is passed upstream from the Egress LSRs, through the intermediate LSRs, to the Ingress LSR.

In IP multicast, a multicast tree is identified by the combination of an IP source address ("S") and an IP group address ("G"), usually written as ")(S,G)". A tree carrying traffic of multiple sources is identified by its group address, and the identifier is written as "http://www.example.com".

When an MP-LSP is being set up, the procedures of [RFC6826] and [RFC7246], known as "mLDP in-band signaling", allow the Egress LSRs of the MP-LSP to encode the identifier of an IP multicast tree in the "Opaque Value" field of the mLDP FEC Element that identifies the MP-LSP. Only the Egress and Ingress LSRs are aware that the mLDP FEC Elements contain encodings of the IP multicast tree identifier; intermediate nodes along the MP-LSP do not take any account of the internal structure of the FEC Element’s Opaque Value, and the internal structure of the Opaque Value does not affect the operation of mLDP. By using mLDP in-band signaling, the Egress LSRs of an MP-LSP inform the Ingress LSR that they expect traffic of the identified IP multicast tree (and only that traffic) to be carried on the MP-LSP. That is, mLDP in-band signaling not only sets up the MP-LSP, it also binds a given IP multicast tree to the MP-LSP.

If multicast is being done in a VPN context [RFC7246], then the mLDP FEC elements also contain a "Route Distinguisher" (RD) (see [RFC7246]), as the IP multicast trees are identified not merely by ")(S,G)" but by ")(RD,S,G)". The procedures of this document are also applicable in this case. Of course, when an Ingress LSR processes an in-band signaling Opaque Value that contains an RD, it does so in the context of the VPN associated with that RD.

If mLDP in-band signaling is not used, then some other protocol must be used to bind an IP multicast tree to the MP-LSP; this requires additional communication mechanisms between the Ingress LSR and the Egress LSRs of the MP-LSP. The purpose of mLDP in-band signaling is to eliminate the need for these other protocols.

When following the procedures of [RFC6826] and [RFC7246] for non-bidirectional trees, the Opaque Value has an IP source address (S) and an IP group address (G) encoded into it, thus enabling it to identify a particular IP multicast (S,G) tree. Only a single IP source-specific multicast tree (i.e., a single "(S,G)") can be identified in a given FEC element. As a result, a given MP-LSP can carry data from only a single IP source-specific multicast tree (i.e., a single "(S,G) tree")". However, there are scenarios in which it would be desirable to aggregate a number of (S,G) trees on a
single MP-LSP. Aggregation allows a given number of IP multicast trees to use a smaller number of MP-LSPs, thus saving state in the network.

In addition, [RFC6826] and [RFC7246] do not support the attachment of an Any-Source Multicast (ASM) shared tree to an MP-LSP, except in the case where the ASM shared tree is a bidirectional tree (i.e., a tree set up by BIDIR-PIM [RFC5015]). However, there are scenarios in which it would be desirable to attach a non-bidirectional ASM shared tree to an MP-LSP.

This document specifies a way to encode an mLDP "Opaque Value" in which either the "S" or the "G" or both are replaced by a "wildcard" (written as "*"). Procedures are described for using the wildcard encoding to map non-bidirectional ASM shared trees to MP-LSPs and for mapping multiple (S,G) trees (with a common value of S or a common value of G) to a single MP-LSP.

Some example scenarios where wildcard encoding is useful are

- PIM shared tree forwarding with "threshold infinity";
- IGMP/Multicast Listener Discovery (MLD) proxying; and
- Selective Source mapping.

These scenarios are discussed in Section 4. Note that this list of scenarios is not meant to be exhaustive.

This document specifies only the mLDP procedures that are specific to the use of wildcards. mLDP in-band signaling procedures that are not specific to the use of wildcards can be found in [RFC6826] and [RFC7246]. Unless otherwise specified in this document, those procedures still apply when wildcards are used.

2. Terminology and Definitions

Readers of this document are assumed to be familiar with the terminology and concepts of the documents listed as Normative References. For convenience, some of the more frequently used terms appear below.

IGMP:
   Internet Group Management Protocol.
In-band signaling:
Using the opaque value of a mLDP FEC element to carry the (S,G) or
(*,G) identifying a particular IP multicast tree. This document
also allows (S,*) to be encoded in the opaque value; see
Section 6.

Ingress LSR:
Root node of a MP-LSP. When mLDP in-band signaling is used, the
Ingress LSR receives mLDP messages about a particular MP-LSP from
downstream and emits IP multicast control messages upstream. The
set of IP multicast control messages that are emitted upstream
depends upon the contents of the LDP Opaque Value TLVs. The
Ingress LSR also receives IP multicast data messages from upstream
and sends them downstream as MPLS packets on an MP-LSP.

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

MLD:
Multicast Listener Discovery.

mLDP:
Multipoint LDP.

MP-LSP:
A Point-to-Multipoint (P2MP) or Multipoint-to-Multipoint (MP2MP)
LSP.

PIM:
Protocol Independent Multicast.

PIM-ASM:
PIM Any-Source Multicast.

PIM-SM:
PIM Sparse Mode.

PIM-SSM:
PIM Source-Specific Multicast.

RP:
The PIM Rendezvous Point.
Egress LSR:
The Egress LSRs of an MP-LSP are LSPs that receive MPLS multicast
data packets from upstream on that MP-LSP, and that forward that
data downstream as IP multicast data packets. The Egress LSRs
also receive IP multicast control messages from downstream and
send mLDP control messages upstream. When in-band signaling is
used, the Egress LSRs construct Opaque Value TLVs that contain IP
source and/or group addresses based on the contents of the IP
multicast control messages received from downstream.

Threshold Infinity:
A PIM-SM procedure where no source-specific multicast (S,G) trees
are created for multicast packets that are forwarded down the
shared tree (*,G).

TLV:
A protocol element consisting of a type field, followed by a
length field, followed by a value field. Note that the value
field of a TLV may be subdivided into a number of subfields.

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

3. Wildcards in mLDP Opaque Value TLVs

[RFC6826] and [RFC7246] define the following Opaque Value TLVs:
Transit IPv4 Source TLV, Transit IPv6 Source TLV, Transit VPNv4
Source TLV, and Transit VPNv6 Source TLV. The value field of each
such TLV is divided into a number of subfields, one of which contains
an IP source address, and one of which contains an IP group address.
Per those documents, these fields must contain valid IP addresses.

This document extends the definition of those TLVs by allowing either
the IP source address field or the IP group address field (or both)
to specify a "wildcard" rather than a valid IP address.

3.1. Encoding the Wildcards

A value of all zeroes in the IP source address subfield is used to
represent a wildcard source address. A value of all zeroes in the IP
group address subfield is used to represent the wildcard group
address. Note that the lengths of these subfields are as specified
in the previous documents.
3.2. Wildcard Semantics

If the IP source address subfield contains the wildcard, and the IP group address subfield contains an IP multicast group address that is NOT in the SSM address range (see Section 4.8 of [RFC4601]), then the TLV identifies a PIM-SM shared tree. Please see Section 3.4 for the applicability restrictions that apply to this case.

If the IP source address subfield contains the wildcard, and the IP group address subfield contains an IP multicast group address that is in the SSM address range, then the TLV identifies the collection of PIM trees with the given group address.

If the IP source address subfield contains a non-zero IP address, and the IP group address subfield contains the wildcard, the TLV identifies the collection of PIM-SSM trees that have the source address as their root.

Procedures for the use of the wildcards are discussed in Sections 4, 5, and 6. Please note that, as always, the structure of an Opaque Value TLV does not affect the operation of mLDP. The structure is meaningful only to the IP multicast modules at the Ingress and Egress LSRs.

Procedures for the use of a wildcard group in the following TLVs (defined in [RFC6826] or [RFC7246]) are outside the scope of the current document: Transit IPv4 Bidir TLV, Transit IPv6 Bidir TLV, Transit VPNv4 Bidir TLV, and Transit VPNv6 Bidir TLV.

Procedures for the use of both a wildcard source and a wildcard group in the same TLV are outside the scope of the current document.

Note that the Bidir TLVs do not have a source address subfield, and hence the notion of a wildcard source is not applicable to them.

3.3. Backwards Compatibility

The procedures of this document do not change the behavior described in [RFC6826] and [RFC7246].

A correctly operating Egress LSR that supports [RFC6826] and/or [RFC7246], but that does not support this document, will never generate mLDP FEC Element Opaque values that contain source or group wildcards.

Neither [RFC6826] nor [RFC7246] specifies the behavior of an Ingress LSR that receives mLDP FEC Element Opaque values that contain zeroes in the source address or group address subfields. However, if an
Ingress LSR supports [RFC6826] and/or [RFC7246], but does not support this document, then it has no choice but to treat any such received FEC elements as invalid; the procedures specified in [RFC6826] and [RFC7246] do not work when the Opaque values contain zeroes in the source address or group address subfields.

The procedures of this document thus presuppose that if an Egress LSR uses wildcard encodings when setting up an MP-LSP, then the Ingress LSR (i.e., the root of the multipoint LSP) supports the procedures of this document. An Egress LSR MUST NOT use wildcard encodings when setting up a particular multipoint LSP unless it is known a priori that the Ingress LSR supports the procedures of this document. How this is known is outside the scope of this document.

3.4. Applicability Restrictions with Regard to ASM

In general, support for non-bidirectional PIM-ASM trees requires (a) a procedure for determining the set of sources for a given ASM tree ("source discovery"), and (b) a procedure for pruning a particular source off a shared tree ("source pruning"). No such procedures are specified in this document. Therefore, the combination of a wildcard source with an ASM group address MUST NOT be used unless it is known a priori that neither source discovery nor source pruning are needed. How this is known is outside the scope of this document. Section 4 describes some use cases in which source discovery and source pruning are not needed.

There are, of course, use cases where source discovery and/or source pruning is needed. These can be handled with procedures such as those specified in [RFC6513], [RFC6514], and [GTM]. Use of mLDP in-band signaling is NOT RECOMMENDED for those cases.

4. Some Wildcard Use Cases

This section discusses a number of wildcard use cases. The set of use cases here is not meant to be exhaustive. In each of these use cases, the Egress LSRs construct mLDP Opaque Value TLVs that contain wildcards in the IP source address or IP group address subfields.

4.1. PIM Shared Tree Forwarding

PIM [RFC4601] has the concept of a "shared tree", identified as (*,G). This concept is only applicable when G is an IP multicast group address that is not in the SSM address range (i.e., is an ASM group address). Every ASM group is associated with a Rendezvous Point (RP), and the (*,G) tree is built towards the RP (i.e., its root is the RP). The RP for group G is responsible for forwarding
packets down the (*,G) tree. The packets forwarded down the (*,G)
tree may be from any multicast source, as long as they have an IP
destination address of G.

The RP learns about all the multicast sources for a given group and
then joins a source-specific tree for each such source. That is,
when the RP for G learns that S has multicast data to send to G, the
RP joins the (S,G) tree. When the RP receives multicast data from S
that is destined to G, the RP forwards the data down the (*,G) tree.
There are several different ways that the RP may learn about the
sources for a given group. The RP may learn of sources via PIM
Register messages [RFC4601], via Multicast Source Discovery Protocol
(MSDP) [RFC3618], or by observing packets from a source that is
directly connected to the RP.

In PIM, a PIM router that has receivers for a particular ASM
multicast group G (known as a "last hop" router for G) will first
join the (*,G) tree. As it receives multicast traffic on the (*,G)
tree, it learns (by examining the IP headers of the multicast data
packets) the sources that are transmitting to G. Typically, when a
last hop router for group G learns that source S is transmitting to
G, the last hop router joins the (S,G) tree and "prunes" S off the
(*,G) tree. This allows each last hop router to receive the
multicast data along the shortest path from the source to the last
hop router. (Full details of this behavior can be found in
[RFC4601].)

In some cases, however, a last hop router for group G may decide not
to join the source trees, but rather to keep receiving all the
traffic for G from the (*,G) tree. In this case, we say that the
last hop router has "threshold infinity" for group G. This is
optional behavior documented in [RFC4601]. "Threshold infinity" is
often used in deployments where the RP is between the multicast
sources and the multicast receivers for group G, i.e., in deployments
where it is known that the shortest path from any source to any
receiver of the group goes through the RP. In these deployments,
there is no advantage for a last hop router to join a source tree
since the data is already traveling along the shortest path. The
only effect of executing the complicated procedures for joining a
source tree and pruning the source off the shared tree would be to
increase the amount of multicast routing state that has to be
maintained in the network.

To efficiently use mLDP in-band signaling in this scenario, it is
necessary for the Egress LSRs to construct an Opaque Value TLV that
identifies a (*,G) tree. This is done by using the wildcard in the
IP source address subfield and setting the IP group address subfield
to G.
Note that these mLDP in-band signaling procedures do not support PIM-ASM in scenarios where "threshold infinity" is not used.

4.2. IGMP/MLD Proxying

There are scenarios where the multicast senders and receivers are directly connected to an MPLS routing domain, and where it is desired to use mLDP rather than PIM to set up "trees" through that domain.

In these scenarios, we can apply "IGMP/MLD proxying" and eliminate the use of PIM. The senders and receivers consider the MPLS domain to be single hop between each other. [RFC4605] documents procedures where a multicast routing protocol is not necessary to build a "simple tree". Within the MPLS domain, mLDP will be used to build an MP-LSP, but this is hidden from the senders and receivers. The procedures defined in [RFC4605] are applicable since the senders and receivers are considered to be one hop away from each other.

For mLDP to build the necessary MP-LSP, it needs to know the root of the tree. Following the procedures as defined in [RFC4605], we depend on manual configuration of the mLDP root for the ASM multicast group. Since the MP-LSP for a given ASM multicast group will carry traffic from all the sources for that group, the Opaque Value TLV used to construct the MP-LSP will contain a wildcard in the IP source address subfield.

4.3. Selective Source Mapping

In many IPTV deployments, the content servers are gathered into a small number of sites. Popular channels are often statically configured and forwarded over a core MPLS network to the Egress routers. Since these channels are statically defined, they MAY also be forwarded over a multipoint LSP with wildcard encoding. The sort of wildcard encoding that needs to be used (source and/or group) depends on the source/group allocation policy of the IPTV provider. Other options are to use MSDP [RFC3618] or BGP "Auto-Discovery" procedures [RFC6513] for source discovery by the Ingress LSR. Based on the received wildcard, the Ingress LSR can select from the set of IP multicast streams for which it has state.

5. Procedures for Wildcard Source Usage

The decision to use mLDP in-band signaling is made by the IP multicast component of an Egress LSR, based on provisioned policy. The decision to use (or not to use) a wildcard in the IP source address subfield of an mLDP Opaque Value TLV is also made by the IP multicast component, again based on provisioned policy. Following are some example policies that may be useful:
1. Suppose that PIM is enabled, an Egress LSR needs to join a non-bidirectional ASM group G, and the RP for G is reachable via a BGP route. The Egress LSR may choose the BGP Next Hop of the route to the RP to be the Ingress LSR (root node) of the MP-LSP corresponding to the (*,G) tree (see also Section 7). The Egress LSR may identify the (*,G) tree by using an mLDP Opaque Value TLV whose IP source address subfield contains a wildcard, and whose IP group address subfield contains G.

2. Suppose that PIM is not enabled for group G, and an IGMP/MLD group membership report for G has been received by an Egress LSR. The Egress LSR may determine the "proxy device" for G (see Section 4.2). It can then set up an MP-LSP for which the proxy device is the Ingress LSR. The Egress LSR needs to signal the Ingress LSR that the MP-LSP is to carry traffic belonging to group G; it does this by using an Opaque Value TLV whose IP source address subfield contains a wildcard, and whose IP group address subfield contains G.

As the policies needed in any one deployment may be very different than the policies needed in another, this document does not specify any particular set of policies as being mandatory to implement.

When the Ingress LSR receives an mLDP Opaque Value TLV that has been defined for in-band signaling, the information from the subfields of that TLV is passed to the IP multicast component of the Ingress LSR. If the IP source address subfield contains a wildcard, the IP multicast component must determine how to process it. The processing MUST follow the rules below:

1. If PIM is enabled and the group identified in the Opaque Value TLV is a non-bidirectional ASM group, the Ingress LSR acts as if it had received a (*,G) IGMP/MLD report from a downstream node, and the procedures defined in [RFC4601] are followed.

2. If PIM is enabled and the identified group is a PIM-SSM group, all multicast sources known for the group on the Ingress LSR are to be forwarded down the MP-LSP. In this scenario, it is assumed that the Ingress LSR is already receiving all the necessary traffic. How the Ingress LSR receives this traffic is outside the scope of this document.

3. If PIM is not enabled for the identified group, the Ingress LSR acts as if it had received a (*,G) IGMP/MLD report from a downstream node, and the procedures as defined in [RFC4605] are followed. The Ingress LSR should forward the (*,G) packets to the Egress LSR through the MP-LSP identified by the Opaque Value TLV. (See also Section 4.2.)
6. Procedures for Wildcard Group Usage

The decision to use mLDP in-band signaling is made by the IP multicast component of an Egress LSR based on provisioned policy. The decision to use (or not to use) a wildcard in the IP group address subfield of an mLDP Opaque Value TLV is also made by the IP multicast component of the Egress LSR, again based on provisioned policy. As the policies needed in any one deployment may be very different than the policies needed in another, this document does not specify any particular set of policies as being mandatory to implement.

When the Ingress LSR (i.e., the root node of the MP-LSP) receives an mLDP Opaque Value TLV that has been defined for in-band signaling, the information from the subfields of that TLV is passed to the IP multicast component of the Ingress LSR. If the IP group address subfield contains a wildcard, then the Ingress LSR examines its IP multicast routing table to find all the IP multicast streams whose IP source address is the address specified in the IP source address subfield of the TLV. All these streams SHOULD be forwarded down the MP-LSP identified by the Opaque Value TLV. Note that some of these streams may have SSM group addresses, while some may have ASM group addresses.

7. Determining the MP-LSP Root (Ingress LSR)

[RFC6826] and [RFC7246] describe procedures by which an Egress LSR may determine the MP-LSP root node address corresponding to a given (S,G) IP multicast stream. That determination is based upon the IP address of the source ("S") of the multicast stream. To follow the procedures of this document, it is necessary to determine the MP-LSP root node corresponding to a given (*,G) set of IP multicast streams. The only difference from the above mentioned procedures is that the Proxy device or RP address is used instead of the source to discover the mLDP root node address.

Other procedures for determining the root node are also allowed, as determined by policy.

8. Anycast RP

In the scenarios where mLDP in-band signaling is used, it is unlikely that the RP-to-group mappings are being dynamically distributed over the MPLS core. It is more likely that the RP address is statically configured at each multicast site. In these scenarios, it is advisable to configure an Anycast RP address at each site in order to provide redundancy. See [RFC3446] for more details.
9. Security Considerations

There are no security considerations other than ones already mentioned in [RFC5036], [RFC6826], and [RFC7246].

10. References

10.1. Normative References


10.2. Informative References


Acknowledgements

We would like to thank Loa Andersson, Pranjal Dutta, Lizhong Jin, and Curtis Villamizar for their review and comments.
Authors’ Addresses

IJsbrand Wijnands (editor)
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

Arkadiy Gulko
Thomson Reuters
195 Broadway
New York, NY 10007
United States

EMail: arkadiy.gulko@thomsonreuters.com

Uwe Joorde
Deutsche Telekom
Hammer Str. 216-226
Muenster D-48153
Germany

EMail: Uwe.Joorde@telekom.de

Jeff Tantsura
Ericsson
300 Holger Way
San Jose, CA 95134
United States

EMail: jeff.tantsura@ericsson.com