Abstract

This document describes a mechanism for a root node to discover the leaf nodes of an mLDP based P2MP/MP2MP LSP. Such kind of function could be used for multiplexing/aggregating root initiated and leaf initiated application which will use mLDP based P2MP/MP2MP LSP. Examples of root initiated applications are P2MP PW [I-D.ietf-pwe3-p2mp-pw], VPLS multicast [I-D.ietf-l2vpn-vpls-mcast], L3VPN multicast [RFC6513]. And examples of leaf initiated applications are statically configured mLDP based P2MP/MP2MP LSP, mLDP in-band signaling [I-D.ietf-mpls-mldp-in-band-signaling].

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on November 7, 2012.

Copyright Notice

Copyright (c) 2012 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.
Table of Contents

1. Introduction ................................................. 3
2. Motivation and problem statement ........................... 3
3. Terminology .................................................. 4
4. Leaf discovery mechanism ...................................... 4
   4.1. Leaf discovery mechanism based on T-LDP .................. 5
       4.1.1. Node operation .................................... 5
   4.2. Leaf discovery mechanism based on MP-BGP ................ 6
       4.2.1. mLDP leaf NLRI .................................... 6
       4.2.2. Node operation .................................... 7
5. Scalability .................................................... 7
6. Security Considerations ....................................... 7
7. IANA Considerations ........................................... 8
   7.1. MP-BGP .................................................. 8
8. Acknowledgement ............................................... 8
9. References ..................................................... 8
   9.1. Normative references ..................................... 8
   9.2. Informative References ................................... 8
Authors’ Addresses ................................................. 9
1. Introduction

This document describes a mechanism for a root node to discover the leaf nodes of an mLDP based P2MP/MP2MP LSP. Such kind of function could be used for multiplexing/aggregating root initiated and leaf initiated application which will use mLDP based P2MP/MP2MP LSP. Examples of root initiated applications are P2MP PW [I-D.ietf-pwe3-p2mp-pw], VPLS multicast [I-D.ietf-l2vpn-vpls-mcast], L3VPN multicast [RFC6513]. And examples of leaf initiated applications are statically configured mLDP based P2MP/MP2MP LSP, mLDP in-band signaling [I-D.ietf-mpls-mldp-in-band-signaling].

This draft provides a discovery mechanism based on a signaling session between each leaf and root node. Each leaf node would signal the leaf node information to root node through this session. There are two signaling protocols to be used for root initiated application, targeted LDP [RFC5036] or BGP auto-discovery using BGP Multiprotocol Extensions [RFC4760]. In order to reuse the signaling protocol of root initiated application, this document introduces both signaling protocols for mLDP leaf discovery.

2. Motivation and problem statement

The leaf initiated application mLDP in-band signaled P2MP LSP will trigger the leaf node to join from leaf node, which means none of the members belonging to a P2MP/MP2MP LSP topology knows all the other members of the P2MP/MP2MP LSP. This means that the root node cannot get the whole P2MP/MP2MP LSP membership information. This problem may cause some limitation for multiplexing/aggregation root initiated applications using mLDP LSPs.

Multicast VPLS [I-D.ietf-l2vpn-vpls-mcast] is a root initiated application. When setting up a inclusive P-Multicast tunnel, BGP A-D is used to do the VPLS membership auto-discovery. The mLDP based P2MP/MP2MP LSP will be set up when receiving auto-discovery routes through BGP A-D. The root node will only know the mLDP LSP leaf node information which is triggered by the specific BGP A-D mechanism. Let’s assume that a mLDP in-band signaling P2MP/MP2MP LSP_a (setup by leaf initiated application) already exist on the root node, and that LSP_a has the same leaf nodes as the P2MP LSP that VPLS multicast BGP A-D tries to set up. The root node does not know LSP_a leaf node information, and will set up mLDP based LSP_b triggered by BGP A-D with same root and leaf nodes.

This causes mLDP based LSP resources waste in the network as it may not be necessary to setup two mLDP LSPs with the same root and leaves in the same network.
The introduction of a leaf discovery mechanism for mLDP based P2MP/MP2MP LSP will enable leaf initiated applications to share one P2MP/MP2MP LSP with root initiated application of P2MP/MP2MP LSP by multiplexing/aggregating mechanism.

3. Terminology

mLDP: Multicast LDP.

T-LDP: Target LDP.

P2MP LSP: An LSP that has one Ingress LSR and one or more Egress LSRs.

MP2MP LSP: An LSP that connects a set of nodes, such that traffic sent by any node in the LSP is delivered to all others.

Bud LSR: An LSR that is an egress but also has one or more directly connected downstream LSRs.

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

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

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

Leaf node: A Leaf node can be either an Egress or Bud LSR when referred in the context of a P2MP LSP. In the context of a MP2MP LSP, an LSR is both Ingress and Egress for the same MP2MP LSP and can also be a Bud LSR.

P2MP FEC: The P2MP FEC Element consists of the address of the root of the P2MP LSP and an opaque value.

MP2MP FEC: MP2MP FEC consists of MP2MP downstream FEC and upstream FEC Element.

MP FEC: Includes both P2MP FEC and MP2MP FEC.

4. Leaf discovery mechanism

It would be beneficial if the mLDP leaf discovery mechanism can reuse the same signaling session as the root initiated application, without requiring additional session overload. This document defines two
leaf discovery mechanisms, one is based on T-LDP, the other is based on MP-BGP. Generally, the root initiated application with LDP as the main signaling mechanism, e.g, P2MP PW [I-D.ietf-pwe3-p2mp-pw], would use leaf discovery mechanism based on T-LDP, while application with MP-BGP as main signaling mechanism, e.g, VPLS Multicast[I-D.ietf-l2vpn-vpls-mcast], L3VPN Multicast [RFC6513] may use leaf discovery mechanism based on MP-BGP.

4.1. Leaf discovery mechanism based on T-LDP

This section will introduce the discovery mechanism based on T-LDP session. Each leaf node will report the leaf node information to root through this T-LDP session. It is required that there is a T-LDP session existed between each leaf node and root node. mLDP leaf discovery function will share the same mLDP P2MP capability described in section 2.1 of [RFC6388].

A LDP Label mapping message on the T-LDP session to the root with the MP FEC Element is used to convey the addition of the leaf membership to the root. The implicit NULL label is used to indicate that the mapping is from a leaf node. The Label Withdraw message is used to convey the deletion of the leaf membership to the root.

4.1.1. Node operation

The mLDP based P2MP/MP2MP LSP leaf discovery mechanism can be operated as follows.

For every leaf node, there will be a T-LDP session to be setup between root and leaf node. This T-LDP session can be setup automatically or manually, which depends on specific implementation.

When the leaf node is triggered to join one P2MP/MP2MP LSP, by various applications, the leaf node sends label mapping message to its upstream node (root or transit node). At the same time, the leaf node sends LDP label map message with MP FEC to its root node. When the root node receives the LDP label map message over T-LDP session with MP FEC, it will store the leaf node information associated with the specified P2MP/MP2MP LSP locally.

When the leaf node is triggered to leave one P2MP/MP2MP LSP, by various applications, the leaf node sends label withdraw message to its upstream node (root or transit node). At the same time, the leaf node sends LDP label withdraw message with MP FEC to its root node. When the root node receives the LDP label withdraw message over T-LDP with MP FEC, it will delete the leaf node information associated with the specified P2MP/MP2MP LSP locally.
4.2. Leaf discovery mechanism based on MP-BGP

This section will introduce the discovery mechanism based on MP-BGP [RFC4760]. Each leaf node will report the leaf node information to root through this BGP session.

4.2.1. mLDP leaf NLRI

This document defines a new BGP NLRI, called mLDP leaf NLRI. Following is the format of the mLDP leaf NLRI:

```
+-------------------------------------------+
|   mLDP MP FEC Element Length (1 octet)   |
+-------------------------------------------+
| mLDP MP FEC Element (Variable)           |
+-------------------------------------------+
| Leaf Node Address Length (1 octet)       |
+-------------------------------------------+
| Leaf Node Address (Variable)             |
+-------------------------------------------+
```

Figure 1

mLDP MP FEC Element may either contain P2MP FEC or MP2MP FEC element. Leaf Node Address field contains the leaf node IP address, and the value of length is 32 if it is IPv4 address, or 128 if it is IPv6 address. The NLRI field in the MP_REACH_NLRI and MP_UNREACH_NLRI is a mLDP MP FEC Element attached with Leaf Node Address. The mLDP leaf NLRI is advertised in BGP UPDATE messages using the MP_REACH_NLRI and MP_UNREACH_NLRI attributes [RFC4760]. The [AFI, SAFI] value pair used to identify this NLRI is (AFI=26(AFI for MPLS Multicast, pending, IANA allocation), SAFI=8(SAFI for mLDP leaf discovery, pending IANA allocation)).

In order for two BGP speakers to exchange mLDP leaf NLRI, they must use BGP Capabilities Advertisement to ensure that they both are capable of properly processing such NLRI. This is done as specified in [RFC4760], by using capability code 1 (multiprotocol BGP) with an AFI of 26 and an SAFI of mLDP leaf discovery.

The Next Hop field of MP_REACH_NLRI attribute shall be interpreted as an IPv4 address, whenever the length of the NextHop address is 4 octets, and as a IPv6 address, whenever the length of the NextHop address is 16 octets.
4.2.2. Node operation

The mLDP based P2MP/MP2MP LSP leaf discovery mechanism can be operated as follows.

When the leaf node is triggered to join one P2MP/MP2MP LSP, by various applications, the leaf node sends label mapping message to its upstream node (root or transit node). At the same time, the leaf node sends BGP UPDATE messages with MP_REACH_NLRI to its root node. The mLDP leaf NLRI will set Leaf Node Address to leaf node IP address, and next hop field to leaf node identifier. When the root node receives BGP UPDATE messages with MP_REACH_NLRI, it will store the leaf node information associated with the specified P2MP/MP2MP LSP locally.

When the leaf node is triggered to leave one P2MP/MP2MP LSP, by various applications, the leaf node sends label withdraw message to its upstream node (root or transit node). At the same time, the leaf node sends BGP UPDATE messages with MP_UNREACH_NLRI to its root node. The mLDP leaf NLRI will set Leaf Node Address to leaf node IP address, and next hop field to leaf node identifier. When the root node receives BGP UPDATE messages with MP_UNREACH_NLRI, it will delete the leaf node information associated with the specified P2MP/MP2MP LSP locally.

To constrain distribution of the mLDP leaf NLRI to the AS of the advertising PE the BGP Update message originated by the advertising PE SHOULD carry the NO_EXPORT Community [RFC1997].

5. Scalability

As recommended in section 4, leaf discovery will reuse the same signaling session as application, and will not setup additional sessions. For the application that uses T-LDP to do leaf discovery, all the leaf nodes have to setup T-LDP session to root node. There may be too many T-LDP sessions that have to be setup on the root node in the network, which will cause some scalability problem. This problem is caused by the application and out of scope of this draft.

6. Security Considerations

The same security considerations apply as for the multicast LDP specification, as described in [I-D. draft-ietf-mpls-mlps-ldp-p2mp], and MP-BGP, as described in [RFC4760].
7. IANA Considerations

7.1. MP-BGP

This document requires allocation of a new BGP AFI and SAFI.

A new AFI is allocated for MPLS Multicast function, the requested value has been pre-allocated as 26.

A new BGP SAFI for "Network Layer Reachability Information used for mLDP leaf discovery" from the IANA "Subsequence Address Family Identifiers (SAFI)" registry. The requested value has been pre-allocated as 8.

8. Acknowledgement

The author would like to thank Rahul Aggarwal, Dimitri Papadimitriou, IJsbrand Wijnands, Sandeep Bishnoi, Frederic Jounay and Simon DeLord for their valuable comments.

9. References

9.1. Normative references


9.2. Informative References


[I-D.ietf-pwe3-p2mp-pw]
Sivabalan, S., Boutros, S., and L. Martini, "Signaling
Root-Initiated Point-to-Multipoint Pseudowire using LDP",
draft-ietf-pwe3-p2mp-pw-04 (work in progress), March 2012.

[I-D.ietf-pwe3-p2mp-pw-requirements]
Bocci, M., Heron, G., and Y. Kamite, "Requirements and
Framework for Point-to-Multipoint Pseudowires over MPLS
PSNs", draft-ietf-pwe3-p2mp-pw-requirements-05 (work in
progress), September 2011.

[RFC6388] Wijnands, IJ., Minei, I., Kompella, K., and B. Thomas,
"Label Distribution Protocol Extensions for Point-to-
Multipoint and Multipoint-to-Multipoint Label Switched
Paths", RFC 6388, November 2011.

VPNs", RFC 6513, February 2012.

Authors’ Addresses

Lizhong Jin
ZTE Corporation
889, Bibo Road
Shanghai, 201203, China

Email: lizhong.jin@zte.com.cn

Kebo Liu
Nokia Siemens Networks
1122 North Qinzhou Road
Shanghai, 200233, China

Email: kebo.liu@nsn.com

Sriganesh Kini
Ericsson
300 Holger Way
San Jose, CA 95134

Email: sriganesh.kini@ericsson.com