Network Working Group
Internet Draft
Intended status: Standards Track
Expires: September 2014

Yakov Rekhter
Juniper Networks

Rahul Aggarwal
Arktan

Nicolai Leymann
Deutsche Telekom

Wim Henderickx
Alcatel-Lucent

Quintin Zhao
Huawei

Richard Li
Huawei

March 4 2014

Carrying PIM-SM in ASM mode Trees over P2MP mLDP LSPs

draft-ietf-mpls-pim-sm-over-mldp-00.txt

Status of this Memo

Copyright Notice

Abstract

When IP multicast trees created by PIM-SM in Any Source Multicast (ASM) mode need to pass through an MPLS domain, it may be desirable to map such trees to Point-to-Multipoint Label Switched Paths. This document describes how to accomplish this in the case where such Point-to-Multipoint Label Switched Paths are established using mLDP.

Table of Contents

1. Specification of Requirements

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].

2. Introduction

[RFC6826] describes how to map Point-to-Multipoint Label Switched Paths (P2MP LSPs) created by mLDP [mLDP] to multicast trees created by PIM-SM in SSM mode [RFC4607]. This document describes how to map mLDP P2MP trees to multicast trees created by PIM-SM in ASM mode. It describes two possible options for doing this.

An implementation MAY support Option 1, as described in Section 3 of this document. An implementation MUST support Option 2, as described in Section 4 of this document.

Note that from a deployment point of view these two options are mutually exclusive. That is on the same network one could either deploy Option 1, or Option 2, but not both.

The reader of this document is expected to be familiar with PIM-SM [RFC4601] and mLDP [mLDP].

This document relies on the procedures in [RFC6826] to support Source Trees. E.g., following these procedures an LSR may initiate a mLDP Label Map with the Transit IPv4/IPv6 Source TLV for (S, G) when receiving PIM (S,G) Join.

This document uses BGP Source Active auto-discovery routes, as defined in [MVPN-BGP].

In a deployment scenario where the service provider has provisioned the network in such a way that the RP for a particular ASM group G is always between the receivers and the sources. If the network is provisioned in this manner, the ingress PE for (S,G) is always the same as the ingress PE for the RP, and thus the Source Active A-D routes are never needed. If it is known a priori that the network is provisioned in this manner, mLDP in-band signaling can be supported using a different set of procedures, as specified in [draft-wijnands]. A service provider will provision the PE routers either to use [draft-wijnands] procedures or to use the procedures of this document.

Like [RFC6826], each IP multicast tree is mapped one-to-one to a P2MP LSP in the MPLS network. This type of service works well if the number of LSPs that are created is under control of the MPLS network operator, or if the number of LSPs for a particular service are known to be limited in number.

It is to be noted that the existing BGP MVPN [MVPN-BGP] procedures may be used to map Internet IP multicast trees to P2MP LSPs. These procedures would accomplish this for IP multicast trees created by PIM-SM in SSM mode as well as for IP multicast trees created by PIM-SM in ASM mode. Furthermore, these procedures would also support the ability to aggregate multiple IP multicast trees to one P2MP LSP in the MPLS network. The details of this particular approach are out of scope of this document.

This document assumes that a given LSR may have some of its interfaces IP multicast capable, while other interfaces being MPLS

capable.

3. Option 1 - Non-transitive mapping of IP multicast shared tree

This option does not transit IP multicast shared trees over the MPLS network. Therefore, when an LSR creates (*, G) state (as a result of receiving PIM messages on one of its IP multicast interfaces), the LSR does not propagate this state in mLDP.

3.1. Originating Source Active auto-discovery routes (Option 1)

Whenever (as a result of receiving either PIM Register or MSDP messages) a Rendezvous Point (RP) discovers a new multicast source, the RP SHOULD originate a BGP Source Active auto-discovery route. The route carries a single MCAST-VPN NLRI [MVPN-BGP] constructed as follows:

+ The Route Distinguisher (RD) in this NLRI is set to 0.

+ The Multicast Source field MUST be set to S. This could be either an IPv4 or an IPv6 address. The Multicast Source Length field is set appropriately to reflect this.

+ The Multicast Group field MUST be set to G. This could be either an IPv4 or an IPv6 address. The Multicast Group Length field is set appropriately to reflect this.

To constrain distribution of the Source Active auto-discovery route to the AS of the advertising RP this route SHOULD carry the NO_EXPORT Community ([RFC1997]).

Using the normal BGP procedures the Source Active auto-discovery route is propagated to all other LSRs within the AS.

Whenever the RP discovers that the source is no longer active, the RP MUST withdraw the Source Active auto-discovery route, if such a route was previously advertised by the RP.

3.2. Receiving BGP Source Active auto-discovery route by LSR

Consider an LSR that has some of its interfaces capable of IP multicast and some capable of MPLS multicast.

When as a result of receiving PIM messages on one of its IP multicast interfaces such LSR creates in its Tree Information Base (TIB) a new (*, G) entry with a non-empty outgoing interface list that contains one or more IP multicast interfaces, the LSR MUST check if it has any Source Active auto-discovery routes for that G. If there is such a route, S of that route is reachable via an MPLS interface, and the LSR does not have (S, G) state in its TIB for (S, G) carried in the route, then the LSR originates the mLDP Label Map with the Transit IPv4/IPv6 Source TLV carrying (S,G), as specified in [RFC6826].

When an LSR receives a new Source Active auto-discovery route, the LSR MUST check if its TIB contains an (*, G) entry with the same G as carried in the Source Active auto-discovery route. If such an entry is found, S is reachable via an MPLS interface, and the LSR does not have (S, G) state in its TIB for (S, G) carried in the route, then the LSR originates an mLDP Label Map with the Transit IPv4/IPv6 Source TLV carrying (S,G), as specified in [RFC6826].

3.3. Handling (S, G, RPT-bit) state

Creation and deletion of (S, G, RPT-bit) PIM state ([RFC4601]) on a LSR that resulted from receiving PIM messages on one of its IP multicast interfaces does not result in any mLDP and/or BGP actions by the LSR.

4. Option 2 - Transitive mapping of IP multicast shared tree

This option enables transit of IP multicast shared trees over the MPLS network. Therefore, when an LSR creates (*, G) state as a result of receiving PIM messages on one of its IP multicast interfaces, the LSR does propagate this state in mLDP, as described below.

Note that in the deployment scenarios where for a given G none of the PEs originate an (S, G) mLDP Label Map with the Transit IPv4/IPv6 Source TLV, no Source Active auto-discovery routes will be used. One other scenario where no Source Active auto-discovery routes will be used is described in section "Originating Source Active auto-discovery routes (Option 2)". In all these scenarios the only part of Option 2 that will be used is the in-band signaling for IP Multicast Shared Tree, as described in the next section.

4.1. In-band signaling for IP Multicast Shared Tree

To provide support for in-band mLDP signaling of IP multicast shared trees this document defines two new mLDP TLVs: Transit IPv4 Shared Tree TLV, and Transit IPv6 Shared Tree TLV.

These two TLVs have exactly the same encoding/format as the IPv4/IPv6 Source Tree TLVs defined in [RFC6826], except that instead of the Source field they have the RP field, and this field carries the address of the RP, as follows:

Transit IPv4 Shared Tree 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                        | RP
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                                                | Group
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                                                |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Type: TBD (to be assigned by IANA).

Length: 8

RP: IPv4 RP address, 4 octets.

Group: IPv4 multicast group address, 4 octets.

Transit IPv6 Shared Tree 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                        | RP            ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                                               | Group         ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Type: TBD (to be assigned by IANA).

Length: 32

RP: IPv6 RP address, 16 octets.

Group: IPv6 multicast group address, 16 octets.

Procedures for in-band signaling for IP multicast shared trees with mLDP follow the same procedures as for in-band signaling for IP multicast source trees specified in [RFC6826], except that while the latter signals (S,G) state using Transit IPv4/IPv6 Source TLVs, the former signals (*,G) state using Transit IPv4/IPv6 Shared Tree TLVs.

4.2. Originating Source Active auto-discovery routes (Option 2)

Consider an LSR that has some of its interfaces capable of IP multicast and some capable of MPLS multicast.

Whenever such LSR creates an (S, G) state as a result of receiving an mLDP Label Map with the Transit IPv4/IPv6 Source TLV for (S, G), if all of the following are true:

+ S is reachable via one of the IP multicast capable interfaces,

+ the LSR determines that G is in the PIM-SM in ASM mode range, and

+ the LSR does not have an (*, G) state with one of the IP multicast capable interfaces as an incoming interface (iif) for that state

the LSR MUST originate a BGP Source Active auto-discovery route.

The route carries a single MCAST-VPN NLRI constructed as follows:

+ The RD in this NLRI is set to 0.

+ The Multicast Source field MUST be set to S. The Multicast Source Length field is set appropriately to reflect this.

+ The Multicast Group field MUST be set to G. The Multicast Group Length field is set appropriately to reflect this.

To constrain distribution of the Source Active auto-discovery route to the AS of the advertising LSR this route SHOULD carry the

NO_EXPORT Community ([RFC1997]).

Using the normal BGP procedures the Source Active auto-discovery route is propagated to all other LSRs within the AS.

Whenever the LSR deletes the (S,G) state that was previously created as a result of receiving an mLDP Label Map with the Transit IPv4/IPv6 Source TLV for (S,G), the LSR that deletes the state MUST also withdraw the Source Active auto-discovery route, if such a route was advertised when the state was created.

Note that whenever an LSR creates an (S,G) state as a result of receiving an mLDP Label Map with the Transit IPv4/IPv6 Source TLV for (S,G) with S reachable via one of the IP multicast capable interfaces, and the LSR already has a (*,G) state with RP reachable via one of the IP multicast capable interfaces as a result of receiving an mLDP Label Map with the Transit IPv4/IPv6 Shared Tree TLV for (*,G), the LSR does not originate a Source Active auto-discovery route.

4.3. Receiving BGP Source Active auto-discovery route

Procedures for receiving BGP Source Active auto-discovery routes are the same as with Option 1.

4.4. Pruning Sources off the Shared Tree

If after receiving a new Source Active auto-discovery route for (S,G) the LSR determines that (a) it has the (*, G) entry in its TIB, (b) the incoming interface list (iif) for that entry contains one of the IP interfaces, (c) at least one of the MPLS interfaces is in the outgoing interface list (oif) for that entry, and (d) the LSR does not originate an mLDP Label Mapping message for (S,G) with the Transit IPv4/IPv6 Source TLV, then the LSR MUST transition the (S,G,RPT-bit) downstream state to the Prune state. [Conceptually the PIM state machine on the LSR will act "as if" it had received Prune(S,G,rpt) on one of its MPLS interfaces, without actually having received one.] Depending on the (S,G,RPT-bit) state on the iif, this may result in the LSR using PIM procedures to prune S off the Shared (*,G) tree.

The LSR MUST keep the (S,G,RPT-bit) downstream state machine in the Prune state for as long as (a) the outgoing interface list (oif) for (*, G) contains one of the MPLS interfaces, and (b) the LSR has at least one Source Active auto-discovery route for (S,G), and (c) the LSR does not originate the mLDP Label Mapping message for (S,G) with

the Transit IPv4/IPv6 Source TLV. Once either of these conditions become no longer valid, the LSR MUST transition the (S,G,RPT-bit) downstream state machine to the NoInfo state.

Note that except for the scenario described in the first paragraph of this section, in all other scenarios relying solely on PIM procedures on the LSR is sufficient to ensure the correct behavior when pruning sources off the shared tree.

4.5. More on handling (S,G,RPT-bit) state

Creation and deletion of (S,G,RPT-bit) state on a LSR that resulted from receiving PIM messages on one of its IP multicast interfaces does not result in any mLDP and/or BGP actions by the LSR.

5. IANA Considerations

This document requires allocation from the LDP MP Opaque Value Element type name space managed by IANA the following two new mLDP TLVs: Transit IPv4 Shared Tree TLV, and Transit IPv6 Shared Tree TLV.

6. Security Considerations

All the security considerations for mLDP ([mLDP]) apply here.

7. Acknowledgements

Use of Source Active auto-discovery routes was borrowed from [MVPN-BGP]. Some text in this document was borrowed from [MVPN-BGP].

Some of the text in this document was borrowed from [RFC6826].

We would like to acknowledge Arkadiy Gulko for his review and comments.

We would also like to thank Xuxiaohu, Gregory Mirsky, and Rajiv Asati for their review and comments.

8. Normative References

[mLDP] Minei, I., "Label Distribution Protocol Extensions for Point-to- Multipoint and Multipoint-to-Multipoint Label Switched Paths", RFC6388, November 2011.

[RFC6826] "In-band signaling for Point-to-Multipoint and Multipoint-to-Multipoint Label Switched Paths", I. Wijnands et al., RFC6826, January 2013

[MVPN-BGP] "BGP Encodings and Procedures for Multicast in MPLS/BGP IP VPNs", R. Aggarwal et al., RFC6514, February 2012

[RFC1997] R. Chandra, P. Traina, T. Li, "BGP Communities Attribute", RFC1997, August 1996.

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

9. 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.

[draft-wijnands] Wijnands IJ, et. al., "mLDP In-Band Signaling with Wildcards", draft-wijnands-mpls-mldp-in-band-wildcard-encoding, work in progress

10. Authors' Addresses

Yakov Rekhter
Juniper Networks, Inc.
e-mail: yakov@juniper.net

Rahul Aggarwal
e-mail: raggarwa_1@yahoo.com

Nicolai Leymann
Deutsche Telekom
Winterfeldtstrasse 21
Berlin 10781

Germany
e-mail: nicolai.leymann@t-systems.com

Wim Henderickx
Alcatel-Lucent
Email: wim.henderickx@alcatel-lucent.com

Richard Li
Huawei
Email: renwei.li@huawei.com

Quintin Zhao
Huawei
Email: quintin.zhao@huawei.com