Resilient MPLS Rings and Multicast
draft-zzhang-mpls-rmr-multicast-00

Abstract

This document discusses multicast with resilient mpls rings.

Requirements Language

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 RFC2119.

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 January 3, 2019.

Copyright Notice

Copyright (c) 2018 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.

Table of Contents

1. Introduction ................................................. 3
2. P2MP/MP2MP Tunnels on a Ring ............................. 3
   2.1. Tunnel Protection and FRR ............................... 4
3. End to End Tunnels with Rings ............................. 5
4. End to End Multicast with Rings ............................ 5
   4.1. End-to-end Native (x,g) Signaling/Forwarding in the
       Global Table ........................................... 5
   4.2. mLDP Inband Signaling in the Global Table or VRFs .... 6
   4.3. MVPN/GTM/EVPN/VPLS .................................. 6
       4.3.1. MVPN/GTM/EVPN/VPLS Tunnel Segmentation ........... 6
5. Summary ..................................................... 6
6. Security Considerations ...................................... 7
7. Acknowledgements ............................................ 7
8. References .................................................. 7
   8.1. Normative References ................................. 7
   8.2. Informative References ............................... 7
1. Introduction

[I-D.ietf-mpls-rmr] describes ...

This document discusses how multicast works with RMRs.

All existing multicast procedures and solutions can work as is. This include both mpls multicast tunnels and end-to-end multicast that makes use of multicast tunnels. Ring topology is just a special case of general topologies so all existing RSVP-TE P2MP and mLDP tunnel can be set up the same way. An Ingress Replication (IR) tunnel consists a bunch of p2p LSPs, and it does not matter whether a component LSP is a plain old LSP or a Ring LSP.

On the other hand, there are optimizations that could be done for RSVP-TE P2MP tunnel signaling (and perhaps FRR for both mLDP and RSVP-TE P2MP tunnels). This document describes that in high level (detailed protocol procedure will be specified in a separate draft), and also discusses end to end multicast when there are RMRs (even though the two are independent).

2. P2MP/MP2MP Tunnels on a Ring

Because mLDP label mapping messages are merged as they propagate from the leaves towards the root, ring topology does not lead to any further optimization.

For a conventionally signaled RSVP-TE P2MP tunnel, the ingress discovers the leaves and signals one sub-LSP for each leaf. Even though the forwarding state is merged at each hop (i.e, one incoming label mapping to multiple outgoing entries), the control plane maintains individual sub-LSP state. This leads to lots of redundant state on routers close to the ingress. With RMR, this can be optimized such that only a single LSP is signaled, with all the leaves listed in the PATH message. As the PATH message passes along the ring, the leaves send RESV messages, but only one RESV message reaches the tunnel ingress in each direction.

The ingress may send PATH messages in both directions, so that the tunnel is set up in such a way that minimum delay is incurred for traffic to reach all leaves. Alternatively, the ingress may send PATH message only in one direction for best bandwidth utilization. For example, a leaf D is three hops away from the ingress A in clockwise direction (A,B,C,D) and four hops away in the other direction (A,E,F,G,D), but G is also a leaf so it may be better to just send the PATH message in the anticlockwise direction.

Each router establishes forwarding state accordingly. Transit
routers switches traffic towards downstream. A transit router could also be a leaf router and in that case it does "drop and continue" - sends traffic off the ring and switches traffic downstream.

MP2MP RSVP-TE tunnel can also be easily achieved. The PATH message could carry a label for the downstream router to send traffic to its upstream. Then the ingress and each leaf can use the same tunnel to send traffic to each other.

The PATH message does not need to list all leaves. As long as a leaf somehow determines that it is a leaf, it can send RESV message when it receives the PATH message. This makes it a leaf-initiated, which may have advantages in certain situations.

2.1. Tunnel Protection and FRR

Each node on a ring signals two counter-rotating MP2P RSVP-TE LSPs to itself. As these LSPs are self-signaled after the discovery of the ring, they can be used to protect P2MP LSPs on ring. So neither mLDP nor RSVP-TE has to setup a separate P2P bypass LSPs for link and node protection. For instance, consider a ring with 8 nodes.

```
Root
R0 . . . R1
   .
R7       R2 Leaf
   |
| Anti-  | .  RID =17 .  |
| Clockwise v .  .  v Clockwise
Leaf R6       R3
   .
R5 . . . R4  Leaf
```

Figure 1: Ring with 8 nodes

Further, suppose a P2MP LSP is signaled with R0 as a root and R2, R4 and R6 as leafs. The P2MP LSP is formed as follows:
In the event of a link failure between R2 and R3, R2, the PLR, tunnels P2MP LSP traffic on an anti-clockwise ring LSP to R3. Once the traffic is out of a tunnel on R3, it uses the regular P2MP LSP to reach R4. Similarly in the event of a node failure R3, R2, the PLR, tunnels P2MP LSP traffic to R4, the MP, which is also the leaf. Thus, the P2MP LSP uses the existing RSVP-TE ring lsps for link and node protection.

3. End to End Tunnels with Rings

Consider a provider network that consists of one or more rings, optionally with a general topology connecting the rings. MVPN [RFC6514], EVPN [RFC7432], or GTM [RFC7716] services are provided where end-to-end multipoint tunnels are needed across the entire network.

If the end to end tunnels are RSVP-TE P2MP tunnels, there is not much optimization that can be done for RMR, unless overlay-assisted tunnel segmentation is used. That is described in Section 4.3.1.

If the end to end tunnels are mLDP tunnels and RSVP-TE signaling is desired on part of the network, mLDP Over Targeted Sessions [RFC7060] can be used without the help from the overlay, and that works even for genereral topologies. If RSVP-TE signaling is used on the rings, then the optimization for setting up the P2MP tunnel on the ring as described earlier can be used.

4. End to End Multicast with Rings

Consider a network that consists of one or more rings, optionally with a general topology connecting the rings. In this network, end-to-end multicast can take various forms described below.

4.1. End-to-end Native (x,g) Signaling/Forwarding in the Global Table

This is typically signaled by PIM end to end. Optionally IGMP/MLD proxy could be used. This works for any topology and RMR does not
make any difference.

4.2. mLDP Inband Signaling in the Global Table or VRFs

This is specified in [RFC6826, RFC7246, RFC7438]. An mLDP tunnel is created for each (x,g) multicast tree. RMR does not make any differences here.

4.3. MVPN/GTM/EVPN/VPLS

GTM here specifically refers to using BGP-MVPN procedures for multicast in the global table, as specified in [RFC7716]. The difference from the earlier described global table multicast is that the global table is treated as a VRF for signaling purpose. MVPN/GTM/EVPN/VPLS uses overlay multicast signaling to signal customer multicast state and tunnel binding. PE-PE multipoint underlay tunnels are used to distribute multicast packets among PEs. Any kind of tunnel can be used, whether the provider network has rings or not, with or without the RMR related optimizations.

4.3.1. MVPN/GTM/EVPN/VPLS Tunnel Segmentation

The MVPN/GTM/EVPN/VPLS PEs could span across ASes or areas. When the PE-PE multipoint tunnels cannot be signaled across AS/area boundaries, segmentation procedures can be used, as specified in [RFC6514, RFC7024] and [I-D.ietf-bess-evpn-bum-procedure-updates]. With the base MVPN/GTM/EVPN/VPLS procedures, PEs advertise I/S-PMSI A-D routes to signal traffic to tunnel binding, and the routes carry type and identification of multi-point tunnels used to carry corresponding traffic. With segmentation, the ASBRs/ABRs become segmentation points and they change the tunnel type/identification when they re-advertise the routes to the next AS/area. With this, each AS/area has its own tunnel of different type/identification, stitched together by the ASBRs/ABRs.

With segmentation, different RMRs could have their own tunnels, and RSVP-TE P2MP optimizations for RMRs could be applied. Notice that this is different from Section 3 in that overlay signaling is involved.

5. Summary

As described above, multicast in the presence of RMRs can work as is. RSVP-TE P2MP tunnel signaling can be optimized (to be specified separately). Tunnel protection/FRR can also be optimized for mLDP/RSPV-TE P2MP tunnels.
6. Security Considerations

To be added.

7. Acknowledgements

8. References

8.1. Normative References


8.2. Informative References


Authors’ Addresses

Zhaohui Zhang
Juniper Networks

EMail: zzhang@juniper.net

Santosh Esale
Juniper Networks, Inc.

EMail: sesale@juniper.net