

INTERNET-DRAFT
Intended Status: Proposed Standard
Expires: May 4, 2017

Santosh Esale
Kireeti Kompella
Juniper Networks
October 31, 2016

LDP Extensions for RMR
draft-esale-ldp-rmr-extensions-00

Abstract

This document describes LDP extensions to signal Resilient MPLS Ring (RMR) Label Switched Paths (LSPs). An RMR LSP is a multipoint to point LSP signaled using LDP (Label Distribution Protocol). RMR Architecture document - draft-ietf-mpls-rmr-02 - describes why and how MPLS should be used in ring topologies.

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at
<http://www.ietf.org/lid-abstracts.html>

The list of Internet-Draft Shadow Directories can be accessed at
<http://www.ietf.org/shadow.html>

Copyright and License Notice

Copyright (c) 2016 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.	Terminology	3
3.	Protocol extensions	4
3.1.	Ring LSP Capability	4
3.2.	Ring FEC Element	4
4.	Ring Procedures	6
4.1	Upstream LSR	6
4.2	Ring Label Mapping Procedures	7
4.2.1	Definitions	7
4.2.2	Preliminary	7
4.2.3	Egress LSR	7
4.2.4	Ingress and Transit LSR	8
4.3	Equal Cost Multipath (ECMP)	8
4.4	Protection	9
5.	LSP Hierarchy	9
6.	Ring LSPs	10
7.	Security Considerations	11
8.	IANA Considerations	11
9.	Acknowledgments	11
10.	Contributors	11
11.	References	12
11.1	Normative References	12
11.2	Informative References	12
	Authors' Addresses	13

1 Introduction

This document describes LDP extensions to signal resilient MPLS ring (RMR) label switched paths (LSPs). An RMR LSP is a multipoint to point LSP signaled using LDP (Label Distribution Protocol). Architecture document of RMR - draft-ietf-mpls-rmr-02 - describes why and how MPLS should be used in ring topologies.

The ring is either auto-discovered or configured using IGP protocol such as OSPF or ISIS. IGP extensions for RMR will be described in a companion documents. After the ring discovery, each ring node acting as egress constructs and signals a clockwise (CW) and anti-clockwise (AC) ring FEC towards AC and CW direction respectively. Each transit node that receives the RMR FEC signals this LSP further in same direction using RMR link state database. In addition, it also adds a transit and ingress route for this LSP. Once the signaling is complete, every node in a ring should have two counter rotating LSPs in CW and AC direction to reach every other node on the ring.

2. Terminology

A ring consists of a subset of n nodes $\{R_i, 0 \leq i < n\}$. The direction from node R_i to R_{i+1} is defined as "clockwise" (CW) and the reverse direction is defined as "anti-clockwise" (AC). As there may be several rings in a graph, each ring is numbered with a distinct ring ID (RID).

The following terminology is used for ring LSPs:

Ring ID (RID): A non-zero number that identifies a ring; this is unique in some scope of a Service Provider's network. A node may belong to multiple rings.

Ring node: A member of a ring. Note that a device may belong to several rings.

Node index: A logical numbering of nodes in a ring, from zero upto one less than the ring size. Used purely for exposition in this document.

Ring neighbors: Nodes whose indices differ by one (modulo ring size).

Ring links: Links that connect ring neighbors.

Express links: Links that connect non-neighboring ring nodes.

MP2P LSP: Each LSP in the ring is a multipoint to point LSP such that LSP can have multiple ingress nodes and one egress node.

3. Protocol extensions

This section describes LDP extensions to signal RMR LSP in a ring.

3.1. Ring LSP Capability

RMR LSPs support for a LSR is advertised using LDP capabilities as defined in [RFC5561]. An implementation that supports the RMR procedures specified in this document MUST add the procedures pertaining to Capability Parameters for Initialization messages.

A new optional capability parameter TLV, RMR Capability, is defined. Following is the format of the RMR Capability Parameter:

```

      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
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|U|F| RMR Capability (TBD) | Length (= 1) |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|S| Reserved |
+---+---+---+---+---+

```

As described in [RFC5561]

U: set to 1. Ignore, if not known.

F: Set to 0. Do not forward.

S: MUST be set to 1 to advertise the RMR Capability TLV.

The RMR Capability TLV MUST be advertised in the LDP Initialization message. If the peer has not advertised the RMR capability, then label messages pertaining to RMR FEC Element MUST not be sent to the peer.

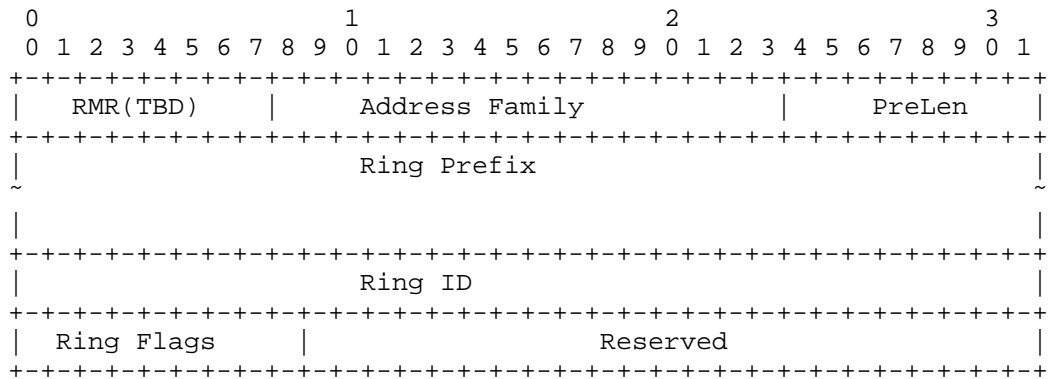
3.2. Ring FEC Element

In order to setup RMR LSP in clockwise and anti-clockwise direction for every ring node, this document defines new protocol entity, the RMR FEC Element, to be used as a FEC Element in the FEC TLV.

The RMR FEC Element consists of the ring address, ring identifier and ring flags which depicts ring direction. The combination of ring address, ring identifier and ring flags uniquely identifies a ring

LSP within the MPLS network.

The RMR FEC Element value encoding is as follows:



FEC Type

One octet quantity containing a value from FEC Type Name Space that encodes the fec type for a RMR LDP LSP.

Address Family

Two octet quantity containing a value from ADDRESS FAMILY NUMBERS in [ASSIGNED_AF] that encodes the address family for the address prefix in the Prefix field.

PreLen

One octet unsigned integer containing the length in bits of the address prefix that follows. A length of zero indicates a prefix that matches all addresses (the default destination); in this case, the Prefix itself is zero octets).

Prefix

An address prefix encoded according to the Address Family field, whose length, in bits, was specified in the PreLen field, padded to a byte boundary.

Ring ID (RID)

A four-octet non-zero number that identifies a ring; this is unique in some scope of a Service Provider's network.

Ring Flags

0	1	2	3	4	5	6	7

```

|RF | Reserved |
+-----+
The value of ring flags (RF) is defined as follows:
1: Clockwise (CW) FEC
2: Anti-clockwise (AC) FEC

```

4. Ring Procedures

Resilient MPLS ring architecture needs interaction between MPLS protocols such as LDP and RSVP and IGP to signal a RMR LSP.

4.1 Upstream LSR

This section describes how to select a upstream LSR for RMR LSP. Consider MPLS rings as follows:

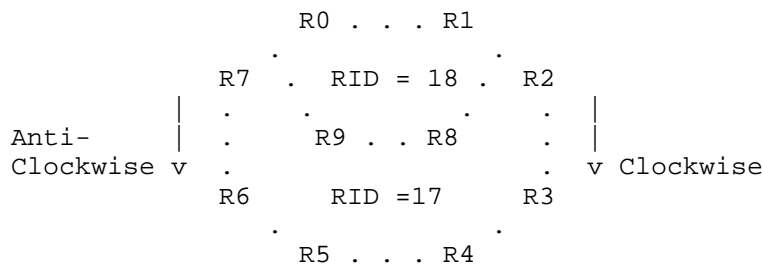


Figure 1: Two Rings with 10 nodes

During the discovery of a MPLS ring, IGP populates its link state database with ring information. After the discovery, there are just two paths - one in clockwise direction and other in anti-clockwise direction - for every ring neighbor on a specific ring. For instance, the following table shows router R0's path for ring 17 and 18 depicted in figure 1.

RID	CW neighbor	AC neighbor
17	R1	R7
18	R1	R9

Figure 2: R0's RMR upstream signaling table

IGP informs LDP that a new MPLS ring, RID 17, is discovered. A LDP transit LSR uses this information to establish RMR LSPs. For

instance, suppose R5 receives a FEC with prefix R0, RID 17 and ring flags AC. R5 knows that its clockwise path is R6 and anti-clockwise path is R4 to reach R0 and that the label map arrived from router R4 for a anti-clockwise LSP. Therefore, R5 selects the upstream session for this LSP as R6.

4.2 Ring Label Mapping Procedures

The procedures in the subsequent sections are organized by the role that a node plays to establish a ring LSP. Each node is egress for its own prefixes and transit for every prefix received with a Label Mapping message. Every transit node is also a ingress for that LSP.

4.2.1 Definitions

This section defines the notations for initiation, decoding, processing and propagation of RMR FEC Element.

1. RMR FEC Element $\langle P, R, C \rangle$ or $\langle P, R, A \rangle$: a FEC Element with egress prefix P, RID R and clockwise direction C or anti-clockwise direction A.
2. RMR Label Mapping $\langle P, R, C, L \rangle$ or $\langle P, R, A, L \rangle$: a Label Mapping message with a FEC TLV with a single RMR FEC Element $\langle P, R, C \rangle$ or $\langle P, R, A \rangle$ and Label TLV with label L. Label L MUST be allocated from the per-platform label space of the LSR sending the Label Mapping message. The use of the interface label space is outside the scope of this document.
3. RMR Label Withdraw $\langle P, R, C, L \rangle$ or $\langle P, R, A, L \rangle$: a Label Withdraw message with a FEC TLV with a single RMR FEC Element $\langle P, R, C \rangle$ or $\langle P, R, A \rangle$ and Label TLV with label L.
4. RMR LSP $\langle P, R, C \rangle$ or $\langle P, R, A \rangle$: A RMR LSP with egress prefix P, Ring ID R and clockwise direction C or anti-clockwise direction A.

4.2.2 Preliminary

A node X wishing to participate in LDP RMR signaling SHOULD negotiate the RMR capability with all its neighbors. When the IGP informs X of its RMR neighbors A and C for RID R, it MUST check that A and C have also negotiated the RMR capability with X. If these conditions are not satisfied, X cannot participate in signaling for ring R. This applies for all roles that X may play: ingress, transit and egress.

4.2.3 Egress LSR

Every ring node initiates two counter-rotating LSPs that egress on that node. After the IGP discovers the ring, LDP constructs the clockwise RMR FEC $\langle P, R, C \rangle$ and sends it in a Label Mapping message

to anti-clockwise neighbor. Similarly, LDP constructs a anti-clockwise RMR FEC <P, R, A> and sends it in a Label Mapping message to clockwise neighbor. This SHOULD establish a clockwise and anti-clockwise LSP - in terms of data traffic - in the clockwise and anti-clockwise direction respectively.

Furthermore, if a label other than implicit or explicit null is advertised for a LSP, LDP SHOULD add a pop route for this label in the Incoming Label Map (ILM) MPLS table.

When the node is no longer part of the ring, it SHOULD tear down its egress LSPs - CW and AC - by sending a label withdraw message.

4.2.4 Ingress and Transit LSR

When a transit LSR R5 depicted in figure 1 receives a label map message with RMR FEC Element <R0, 17, A, L1> from a downstream LDP session to R4, it SHOULD verify that R4 is indeed its anticlockwise neighbor for ring 17. If not, it SHOULD stop decoding the FEC TLV, abort processing the message containing the TLV, send an "Unknown FEC" Notification message to its LDP peer R4 signaling an error and close the session.

If the LSR encounters no other error while parsing the RMR FEC element, it allocates a Label L2 and determines a upstream LDP session as R6 using the algorithm described in section 'Upstream LSR'. It also programs a MPLS ILM table with label route L2 swapped to L1 and Ingress tunnel table with prefix R0 with label push L1 on all the LDP interfaces to R4, and sends the RMR FEC Element <R0, 17, A, L2> to R6.

If a session to the anti-clockwise neighbor for RID 17 depicted in Figure 2, namely R6, does not exist, the RMR FEC Element <R0, 17, A, L2> SHOULD not be propagated further. Similarly, when the upstream session changes because of ring topology change, transit LSR should send a label withdraw for RMR FEC Element <R0, 17, A, L2> to older upstream session R6 before sending Label Mapping message with RMR FEC Element <R0, 17, A, L2> to a new upstream session.

4.3 Equal Cost Multipath (ECMP)

A transit and ingress LSR of RMR LSP uses all the links between itself and downstream LSR to program transit and ingress route. Thus, ECMP works automatically for a LDP RMR LSP. A vendor could provide exception when necessary to this behavior by disabling certain ring links for RMR LSPs.

4.4 Protection

RMR uses the two counter-rotating LSPs to protect the other. Say that R5 wants to protect the LSP to R0 for RID 17. R5 receives RMR FEC Element <R0, 17, A, L1> from R4 and sends RMR FEC Element <R0, 17, A, L2> to R6. Then the primary path for the AC LSP is to swap L1 with L2 with next hop R4. Also, R5 receives RMR FEC Element <R0, 17, C, L3> from R6 and sends RMR FEC Element <R0, 17, C, L4> to R4. The primary path for the CW LSP is to swap L3 with L4. The protection path for the AC LSP is to swap L1 with L4 with next hop R6, thus sending the traffic back where it came from, but with a different label. The protection path for the CW LSP is to swap L3 with L2 with next hop R4.

5. LSP Hierarchy

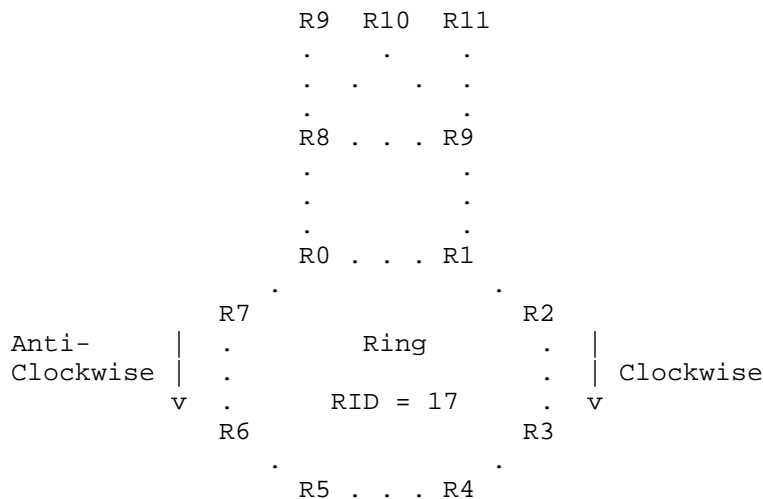


Figure 3: Ring 17 with rest of the Network

Suppose R5 needs to reach R10. Only RMR LSPs are setup inside the ring 17. Additionally, whenever services on R5 need to reach R10, R5 dynamically establishes a tLDP session to ring 17 master node R0 and R1. Further, suppose it only learns IPv4 and IPv6 FECs only over this session using [draft-ietf-mpls-app-aware-tldp-05]. Thus, in order to reach R10, R5 uses top label as RMR LSP label to R0 or R1 and bottom label as R10's FEC label received over tLDP session of R0 or R1 respectively.

6. Ring LSPs

An RMR LSP consists of two counter-rotating ring LSPs that start and end at the same node, say R1. As such, this appears to cause a loop, something that is normally to be avoided by LDP [RSVP-TE]. There are some benefits to this. Having a ring LSP allows the anchor node R1 to ping itself and thus verify the end-to-end operation of the LSP. This, in conjunction with link-level OAM, offers a good indication of the operational state of the LSP. [Also, having R1 be the ingress means that R1 can initiate the Path messages for the two ring LSPs. This avoids R1 having to coordinate with its neighbors to signal the LSPs, and simplifies the case where a ring update changes R1's ring neighbors.] The cost of this is a little more signaling and a couple more label entries in the LFIB. However, we will let implementation guide us to the wisdom of this approach.

7. Security Considerations

The Capability and RMR FEC procedures described in this document will not introduce any change to LDP Security Considerations section described in [RFC5036].

8. IANA Considerations

This document requires the assignment of a new code point for a Capability Parameter TLVs from the IANA managed LDP registry "TLV Type Name Space", corresponding to the advertisement of the RMR capability. IANA is requested to assign the lowest available value.

Value	Description	Reference
-----	-----	-----
TBD1	RMR capability	[this document]

This document requires the assignment of a new code point for a FEC type from the IANA managed LDP registry "Forwarding Equivalence Class (FEC) Type Name Space". IANA is requested to assign the lowest available value.

Value	Description	Reference
-----	-----	-----
TBD1	RMR FEC type	[this document]

9. Acknowledgments

TODO.

10. Contributors

Raveendra Torvi
Juniper Networks
10 Technology Park Dr
Westford, MA 01886
USA
Email: rtorvi@juniper.net

Ravi Singh
Juniper Networks
1133 Innovation Way
Sunnyvale, CA 94089
USA

Email: ravis@juniper.net

Abhishek Deshmukh
Juniper Networks
10 Technology Park Dr
Westford, MA 01886
USA
Email: adeshmukh@juniper.net

11. References

11.1 Normative References

- [I-D.ietf-mpls-rmr] Kompella, K. and L. Contreras, "Resilient MPLS Rings", draft-ietf-mpls-rmr-02 (work in progress), July 2016.
- [RFC5036] Andersson, L., Ed., Minei, I., Ed., and B. Thomas, Ed., "LDP Specification", RFC 5036, October 2007, <<http://www.rfc-editor.org/info/rfc5036>>.
- [RFC5561] Thomas, B., Raza, K., Aggarwal, S., Aggarwal, R., and JL. Le Roux, "LDP Capabilities", RFC 5561, July 2009, <<http://www.rfc-editor.org/info/rfc5561>>.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<http://www.rfc-editor.org/info/rfc2119>>.

11.2 Informative References

- [RFC6388] IJ. Wijnands, I. Minei, K. Kompella, B. Thomas, "Label Distribution Protocol Extensions for Point-to-Multipoint and Multipoint-to-Multipoint Label Switched Paths", RFC 6388, November 2011, <<http://www.rfc-editor.org/info/rfc6388>>
- [I-D.draft-deshmukh-rsvp-rmr-extension] A. Deshmukh, K. Kompella, "RSVP Extensions for RMR", draft-deshmukh-rsvp-rmr-extension-00 (work in progress), July 2016.
- [I-D.draft-ietf-mpls-app-aware-tldp] Santosh Esale, Raveendra Torvi, Luay Jalil, U. Chunduri, Kamran Raza, "Application-aware

Targeted LDP", draft-ietf-mpls-app-aware-tldp-05 (work in progress), June 2016.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997, <<http://www.rfc-editor.org/info/rfc2119>>.

Authors' Addresses

Santosh Esale
Juniper Networks
1133 Innovation Way
Sunnyvale, CA 94089
USA
Email: sesale@juniper.net

Kireeti Kompella
Juniper Networks
1133 Innovation Way
Sunnyvale, CA 94089
USA
Email: kireeti@juniper.net