Abstract

This document describes the use of Segment Routing (SR) control plane to setup LSP(s) for resilient MPLS ring networks. It specifies how clockwise and anti-clockwise ring LSP(s) are signaled using SR IGP protocol extensions, and how such ring LSP(s) are combined to achieve ring protection.

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 https://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 8, 2020.

Copyright Notice

Copyright (c) 2019 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 (https://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
1. Introduction

Ring topologies are very common in transport networks, and are ubiquitous in access and aggregation networks.

This draft introduces the necessary extensions to Segment Routing (SR) IGP signaling protocol(s) that are needed to establish Label Switched Paths (LSPs) for Resilient MPLS Rings (RMR). An RMR LSP is a multipoint to point (MP2P) LSP that is signaled using SR control plane extensions to IGPs (e.g. IS-IS or OSPF).

SR as defined in [RFC8402] allows for a flexible definition of end-to-end paths within IGP topologies by encoding the SR path as a sequence of one or more topological sub-paths, or "segments". Such
.segments can be advertised by link-state routing protocols such as IS-IS or OSPF.

Rings are auto-discovered using the mechanisms described in the [I-D.ietf-mpls-rmr]. Signaling extensions for IS-IS and OSPF are introduced in [I-D.kompella-isis-ospf-rmr] to enable the auto-discovery of ring topologies. After the ring topology is discovered, each node in the ring determines its Clockwise (CW) and Anti-clockwise (AC) ring neighbors and associated ring links.

[I-D.ietf-teas-rsvp-rmr-extension] describes the necessary RSVP-TE signaling protocol extensions that are needed to setup Resilient MPLS Ring (RMR) LSPs. [I-D.ietf-mpls-ldp-rmr-extensions] describes extensions to LDP [RFC5036] signaling protocol that are needed for LDP setup RMR LSPs.

This document describes how Segment Routing (SR) IGP control plane can be extended to allow for the setup of RMR LSPs and how a pair of CW and AC SR MPLS LSPs can provide the required protection for ring topologies.

The first revisions of this document will focus on the simple most common ring topologies in access networks that do not include any "express links". Future revisions of this document will expand on express link(s) for the more general rings.

2. Terminology

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 BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2.1. Ring Taxonomy

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 rings:

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, each identified by its unique RID.
Ring Node:
A member of a ring. Note that a node may belong to several rings.

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

Ring Master:
The ring master initiates the ring identification process. Mastership is indicated in the IGP by a two-bit field.

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

Ring Size:
The ring size for a given instantiation is N. This can change as nodes are added or removed, or go up or down.

Ring Links:
Links that connect ring neighbors.

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

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

Ring Identification:
The process of discovering ring nodes, ring links, link directions, and express links.

Ring SID:
Each ring node advertises 2 unique Ring Prefix Segment Identifiers (Ring SIDs). CW ring SID is advertised by ring node to receive traffic in clockwise direction. AC ring SID is advertised by ring node to receive traffic in anti-clockwise direction.

3. Protocol extensions

This section describes the necessary IGP protocol extensions to SR to allow signaling and establishing RMR LSP(s) to each node in a ring.

3.1. SR Ring Capability

A new sub-TLV is defined for SR RMR Ring Capability to advertise node support for signaling SR RMR LSP(s) using extensions in IS-IS and OSPF protocol(s).
If the SR RMR Ring Capability is not advertised by a node, such node is considered not capable of establishing SR RMR LSP(s) using SR control plane extensions.

The SR RMR Ring Capability sub-TLV has the following format:

```
+------------+-----------+-----------+
|    Type    |     Length|    Flags  |
+------------+-----------+-----------+
```

Type: TBD-1

Length: length in octets, 1.

Flags: 1 octet. Currently no flags are defined.

```
+------------+----------+
|   Reserved |
+------------+
```

3.1.1. IS-IS SR RMR Ring Capabilities Sub-TLV

In IS-IS, the Router Capability TLV TLV-242 defined in [RFC7981] is used to carry the SR RMR Ring Capability sub-TLV.

3.1.2. OSPF SR RMR Ring Capabilities Sub-TLV

In OSPF, a top-level TLV of the Router Information Opaque LSA (defined in [RFC7770]) is used to carry the SR RMR Ring Capability sub-TLV.

3.2. SR Ring Prefix Segment Identifier (Ring-SID)

In order to setup an SR RMR LSP(s) in CW and AC directions towards each node of the ring, this document defines a new Ring SR Prefix Segment Identifier (Ring-SID) sub-TLV.

The Ring-SID sub-TLV carries the IGP Segment Routing Ring-SID that is associated with the advertised prefix by the node. The Ring-SID is unique within a specific ring in an IGP domain.

For IS-IS protocol, the Ring-SID MAY be present in any of the following TLVs:

TLV-135 (Extended IPv4 reachability) defined in [RFC5305].
TLV-235 (Multitopology IPv4 Reachability) defined in [RFC5120].

TLV-236 (IPv6 IP Reachability) defined in [RFC5308].

TLV-237 (Multitopology IPv6 IP Reachability) defined in [RFC5120].

For OSPF protocol, the Ring-SID sub-TLV is carried as part of the OSPF Extended Prefix TLV defined in [RFC7684].

The Ring-SID sub-TLV has the following format:

```
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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Type        |     Length    |     Flags     |   Algorithm   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Ring ID (RID)                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                   Ring-SID Index/Label (variable)             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

Type:
TBD-2

Length:
9 or 10 depending on the size of the SID (index vs. absolute label)

Flags:
1 octet field of following flags: TBD.

Length:
length in octets, 1.

```
0 1 2 3 4 5 6 7
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|D |NP|M |E |V | | |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

D-Flag: identifies the direction towards the downstream neighbors. The possible values are:
0: CW next-hop(s) neighbor(s) derived after the completion of ring identification phase.

1: AC next-hop(s) neighbor(s) derived after the completion of ring identification phase.

NP-Flag: No-PHP flag. If set, then the penultimate hop MUST NOT pop the Prefix-SID before delivering packets to the node that advertised the Ring-SID.

M-Flag: Mapping Server Flag. If set, the SID was advertised by a Segment Routing Mapping Server as described in [I-D.ietf-spring-segment-routing-ldp-interop].

E-Flag: Explicit-Null Flag. If set, any upstream neighbor of the Prefix-SID originator MUST replace the Prefix-SID with the Explicit-NULL label (0 for IPv4) before forwarding the packet.

V-Flag: Value/Index Flag. If set, then the Prefix-SID carries an absolute value. If not set, then the Ring-SID carries an index.

Other bits: Reserved. These MUST be zero when sent and are ignored when received.

Algorithm:
Single octet identifying the algorithm to use to derive the downstream member ring next-hop(s) to reach a specific ring node.
The following values are defined by this document:

0: Include all next-hop(s) derived from the completion of ring identification process. The D-Flag indicates whether CW or AC are to be considered.

Other user-defined algorithms identifiers from the range 128-255 can be defined and used as described in [I-D.ietf-lsr-flex-algo].

Ring-ID:
The Ring ID that the Ring-SID is advertised for.

Ring-SID:
The 'Ring SID' MUST be unique within a given IGP domain (when the L-flag is not set).

SID/Index/Label:
as defined in [I-D.ietf-isis-segment-routing-extensions].
3.3. Ring Segment Identifier (Ring-SID)

An alternative means of propagating the CW and AC Ring SIDs is as a sub-TLV of the RMR TLV. This sub-TLV has the following format in IS-IS:

```
  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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Type (TBD)  |  Length (8)   | CW Ring SID ...               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              ... CW Ring SID  | AC Ring SID ...               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|              ... AC Ring SID  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where:

Type:
- is the type of the sub-TLV (TBD)

Length:
- 8

Value:
- The CW SID followed by the AC SID.

3.4. Ring-SID Propagation

The rules for the propagation of a Ring-SID follow those dictated in [I-D.ietf-isis-segment-routing-extensions] and [I-D.ietf-ospf-segment-routing-extensions].

4. Ring SR Signaling Procedures

4.1. Ring SID Assignment

As described in [I-D.ietaf-mpls-rmr], a ring of RID r can be either configured on all nodes participating in ring r, or be configured on select master member ring node(s) while running other nodes in promiscuous mode. All ring node(s) participating in a ring use the IGP extensions defined in [I-D.kompella-isis-ospf-rmr] to advertise their ring membership and to complete ring discovery and identification phase.

A unique CW and AC Ring-SID(s) are assigned to a prefix on residing on each ring member node participating in the ring.
4.1.1. Static Ring-SID Assignment

An operator can choose to statically assign and configure the unique CW and AC Ring-SID(s) associated with the prefixes residing on each member node of the ring. In such case, it is expected that Ring-SID assignment and management becomes the responsibility of the network operator in order to ensure global uniqueness.

When static provisioning of Ring-SID(s) on ring node(s) is implemented, the Ring-SID(s) sub-TLV(s) are explicitly advertised along with the prefix reachability advertisement. Examples of such explicit advertisement(s) for prefix-SID(s) are given in [I-D.ietf-isis-segment-routing-extensions], [I-D.ietf-ospf-segment-routing-extensions], and [I-D.ietf-ospf-ospfv3-segment-routing-extensions].

4.1.2. Ring-SID Assignment Using SRMS

It is possible to leverage the Segment Routing Mapping Server (SRMS) functionality as defined in [I-D.ietf-spring-segment-routing-ldp-interop] to assign, advertise, and manage Ring-SID(s) on behalf of all ring nodes in the network. This simplifies the burden on the operator to provision rings and Ring-SID(s) on network ring nodes, by restricting configuration to only select nodes in the ring (e.g. master node(s)).

The SRMS functionality consists of two functional blocks: a Mapping Server (MS) and a Mapping Client (MC). The SR MS functionality supports the advertisement of prefix-SID(s) to a prefix without the need to explicitly advertise such assignment within a prefix reachability advertisement. This document extends the MS functionality to allow it to advertise the Ring-SID to prefix mapping in a similar fashion.

The SR MC is any node that receives and uses the MS mapping advertisements. The MC interprets the SR mapping Ring-SID advertisement as an assignment of a Ring-SID to a prefix. Note that the SRMS node can serve as both an MS and an MC.

To implement the SRMS for assigning and managing Ring-SID(s), the network operator reserves a block of SR Ring SID indices and delegates it to the SRMS.

When a ring of RID r is configured/enabled on a ring master node, the SRMS learns of ring nodes participating in ring RID r using the ring node TLV defined in [I-D.kompella-isis-ospf-rmr]. Whenever the SRMS discovers a new ring node, it automatically assigns two unique Ring-
SID(s) - for CW and AC Ring-SID LSP(s) - from the available SID(s) within the Ring SID block available on the SRMS.

After CW and AC Ring-SID(s) are assigned to a ring node prefix, the SRMS can advertise the Ring-SID sub-TLV in a SID/label binding TLV as described in [I-D.ietf-isis-segment-routing-extensions] and [I-D.ietf-ospf-segment-routing-extensions].

4.1.3. Ring-SID Assignment Using DHCP

The Dynamic Host Configuration Protocol is uniquely well-suited for handling node and ring SID assignments. When ring directions have been established for all links in the ring, each node can request, as a DHCP client, a pair of ring SIDs. The DHCP server responds with two unique values from the SID block(s) for Ring SIDs with which it has been configured. The DHCP server SHOULD be configured with very long leases for such assignments, as well as "sticky" assignments; that is, should a lease expire, the pair of values assigned should not be offered to another client unless the server has run out of ring SID values; also, should the same client re-request ring SIDs, the server should return the same SIDs if at all possible.

Further details are provided in [I-D.kompella-spring-dhcp].

4.2. Ring SID LSP Setup

Any ring node that receives a Ring-SID advertisement either part of an explicit prefix TLV advertisement or part of a SID/label binding TLV advertised by the SRMS will perform the following depending on ring node role.

4.2.1. Egress Ring Node

An node that is member of a ring RID r can advertise a Ring-SID sub-TLV associated with local prefix. If the node learns of a Ring-SID sub-TLV associated to local prefix using the SID/label binding TLV advertised by the SRMS, the node first verify it is member of ring indicated in the Ring-SID sub-TLV. If not, the node does not process the Ring-SID sub-TLV any further.

Otherwise, when the node is member of the ring indicated in the Ring-SID sub-TLV it ensures that the corresponding local MPLS label from its SRGB is assigned and bound to the specific local prefix and RID. If no penultimate hop popping is desired, an egress Incoming Label Map (ILM) entry corresponding for the corresponding local label is installed in the forwarding table as usual.
4.2.2. Ingress and Transit Ring Nodes

An ingress or transit node that receives a Ring-SID sub-TLV advertisement for a remote prefix through an explicit prefix TLV advertisement, or through a SID/label binding TLV for a remote prefix will first verify that it is member of the ring indicated in the Ring-SID sub-TLV advertisement before processing it any further.

If the ingress or transit node are is not member of the Ring-SID’s ring, the node MUST not process the Ring-SID any further. Otherwise, if the ingress or transit node is member of the ring indicated in the Ring-SID, the ingress or transit node ensures that the corresponding local MPLS label in its SRGB is assigned and bound to the specific remote prefix and for the specific ring.

As described in Section 3.2, the Ring-SID sub-TLV carries an Algorithm identifier that indicates the procedure to derive the set of next-hop(s) to reach a CW or AC neighbor. The default algorithm (Algorithm 0) is to include all next-hop(s)/links between itself and the respective downstream neighbor. The D-Flag in the Flags field within the Ring-SID sub-TLV indicates whether the CW or AC downstream neighbors are intended.

An operator may make use of other user-defined Algorithms to allow an ingress or transit ring node to select specific link(s)/neighbors that connect to the downstream CW or AC neighbor.

Once the CW and AC next-hop(s) are determined, the ingress or transit router(s) of the RMR ring LSP add the corresponding ILM entries for the CW and AC labels and map each to the set of CW and AC Next-hop Label Forwarding Entries (NHLFE)’s.

4.2.3. Protection and Fastreroute

The ingress or a transit node that receives a Ring-SID advertisement from a remote ring node, in addition to assigning and programming the primary CW or AC next-hop(s) as described in the previous section, assigns the opposite direction AC or CW next-hop and its associated AC or CW Ring-SID as a backup path.

Upon failure, the ingress or a transit router that detects a local link failure of the Ring-SID’s primary next-hop, immediately diverts traffic on to the pre-programmed backup next-hop of the Ring-SID. Traffic originally flowing in a CW or AC direction will be diverted to flow on to flow in the AC or CW direction after the failure.

In case of a failure at a transit ring node along the path of a Ring-SID, any upstream router that learns of the downstream link failure
via IGP link updates, can locally reroute(s) the traffic towards the backup next-hop. This optimizes the repair path and avoids packets from being unnecessarily forwarded to the ring node where local fault occurred, and the be looped back on the opposite Ring-SID due to the fault.

5. IANA Considerations
   TODO.

6. Security Considerations
   It is not anticipated the extensions to IGP SR protocols described in this document may introduce additional security risk(s). Future revisions of this document will update this section with details about those risks.

7. Acknowledgement
   The authors would like to thank XX for reviewing and providing valuable feedback on this document.

8. Contributors
   Raveendra Torvi
   Juniper Networks
   Email: rtorvi@juniper.net

   Vishnu Pavan Beeram
   Juniper Networks
   Email: vbeeram@juniper.net

9. References
9.1. Normative References
   [I-D.ietf-isis-segment-routing-extensions]
[I-D.ietf-lsr-flex-algo]

[I-D.ietf-mpls-ldp-rmr-extensions]

[I-D.ietf-mpls-rmr]

[I-D.ietf-ospf-ospfv3-segment-routing-extensions]

[I-D.ietf-ospf-segment-routing-extensions]

[I-D.ietf-spring-segment-routing-ldp-interop]

[I-D.kompella-isis-ospf-rmr]
Kompella, K., "IGP Extensions for Resilient MPLS Rings", draft-kompella-isis-ospf-rmr-00 (work in progress), October 2016.

[I-D.kompella-spring-dhcp]
Kompella, K. and R. Bonica, "Using DHCP to Manage Node and Ring SID Assignment", draft-kompella-spring-dhcp-00 (work in progress), July 2019.

9.2.  Informative References

[I-D.ietf-teas-rsvp-rmr-extension]
  Deshmukh, A. and K. Kompella, "RSVP Extensions for RMR",
draft-ietf-teas-rsvp-rmr-extension-01 (work in progress),
  June 2018.

Authors’ Addresses

Kireeti Kompella
Juniper Networks
Email: kireeti.kompella@gmail.com

Tarek Saad
Juniper Networks
Email: tsaad@juniper.net

Abhishek Deshmukh
Juniper Networks
Email: adeshmukh@juniper.net