Abstract

This document presents a mechanism aimed at providing segment routing forwarding option for MRT-FRR. It defines five new MRT Profiles, each using SR-LSP, SR-tunnel, etc., forwarding option respectively. These MRT Profiles will be very useful in the case of segment routing network without LDP deployed. On the other hand, MRT-FRR is also a competitive solution comparing with others for segment routing network.

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 March 2, 2017.

Copyright 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. Five New MRT Forwarding Options ....................... 3
   2.1. Option 1: MRT SR-tunnel Option ..................... 4
       2.1.1. Forwarding SR Label Unicast Traffic over MRT Paths .. 5
       2.1.2. Forwarding IP Unicast Traffic over MRT Paths .... 6
       2.1.3. Inter-area Forwarding Behavior .................. 6
       2.1.4. Prefixes Multiply Attached to the MRT Island .... 6
       2.1.5. FIB examples .................................. 6
   2.2. Option 2: MRT SR-LSP Tunneling with multi-prefix-sid Option .................. 7
       2.2.1. Forwarding SR Label Unicast Traffic over MRT Paths .. 8
       2.2.2. Forwarding IP Unicast Traffic over MRT Paths .... 8
       2.2.3. Inter-area Forwarding Behavior .................. 9
       2.2.4. Prefixes Multiply Attached to the MRT Island .... 9
       2.2.5. FIB examples .................................. 9
   2.3. Option 3: MRT SR-LSP Tunneling with multi-SRGB Option .. 10
       2.3.1. Forwarding SR Label Unicast Traffic over MRT Paths .. 11
       2.3.2. Forwarding IP Unicast Traffic over MRT Paths .... 12
       2.3.3. Inter-area Forwarding Behavior .................. 12
       2.3.4. Prefixes Multiply Attached to the MRT Island .... 12
       2.3.5. FIB examples .................................. 12
   2.4. Option 4: MRT SR-LSP Non-tunneling with multi-SRGB Option 13
       2.4.1. Forwarding SR Label Unicast Traffic over MRT Paths .. 14
       2.4.2. Forwarding IP Unicast Traffic over MRT Paths .... 15
       2.4.3. Inter-area Forwarding Behavior .................. 15
       2.4.4. Prefixes Multiply Attached to the MRT Island .... 15
       2.4.5. FIB examples .................................. 16
   2.5. Option 5: MRT SR-LSP Non-tunneling with multi-prefix-sid Option .................. 18
3. Interoperation ........................................... 18
   3.1. Multiple Profiles Coexistence ....................... 19
   3.2. Multiple Profiles Interworking ...................... 20
4. Support protecting segment list ......................... 21
   4.1. Received segment list is (segment F or S->F, vpn label) or (segment F or S->F) ................. 22
   4.2. Received segment list is (segment F or S->F, segment D1 or F->D1, segment D2, ..., segment Dn, vpn label) ... 23
   4.3. Received segment list is (segment D1, segment D2, ..., segment Dn, vpn label) .................. 24
5. Security Considerations ................................. 25
6. IANA Considerations ..................................... 25
7. References ............................................. 26
   7.1. Normative References ............................... 26
1. Introduction

MRT-FRR [RFC7812] creates two alternate forwarding trees that are distinct from the primary next-hop forwarding used during stable operation. These two trees are maximally diverse from each other, providing link and node protection for 100% of paths and failures as long as the failure does not cut the network into multiple pieces. MRT-FRR has defined some forwarding mechanism options, including LDP Label Option 1A and 1B, IP Tunnel Option 2A and 2B. Among these options, LDP Label Option 1A is a more appropriate choice, which is selected as the forwarding mechanism in MRT default profile.

Segment Routing [I-D.ietf-spring-segment-routing] allows to enforce a flow through any path and service chain while maintaining per-flow state only at the ingress node of the SR domain. Segment Routing can be directly applied to the MPLS architecture [RFC3031] with no change on the forwarding plane. A segment is encoded as an MPLS label. An ordered list of segments is encoded as a stack of labels. For each IGP-prefix segment in the segment list, it will forward packet according to shortest path to destination corresponding to the prefix, that has similar characteristics as an LDP LSP, so we can term it as an SR LSP. For the whole segment list, it will specify a forwarding path, other than the normal shortest path, so we can term it as an SR tunnel.

It is necessary to introduce MRT-FRR function as a new choice to segment routing network for reliability. First, MRT-FRR will bring a more competitive FRR solution to SR. Second, SR forwarding mechanism will greatly simplify MRT-FRR function.

2. Five New MRT Forwarding Options

The following five new options for MRT forwarding mechanisms are introduced.

1. MRT SR-tunnel Option
2. MRT SR-LSP Tunneling with multi-prefix-sid Option
3. MRT SR-LSP Tunneling with multi-SRGB Option
4. MRT SR-LSP Non-tunneling with multi-SRGB Option
5. MRT SR-LSP Non-tunneling with multi-prefix-sid Option
Respectively, we can define 5 new MRT profiles. Each profile is almost same as the default MRT profile, except the forwarding mechanisms would be set to the above value respectively, and allocate different MRT MT-IDs. For example, the MRT profile which support "MRT SR-tunnel Option" forwarding mechanism would be as the following:

MRT Algorithm: MRT Lowpoint algorithm defined in [RFC7811].

MRT-Red MT-ID: To be allocated.

MRT-Blue MT-ID: To be allocated.

GADAG Root Selection Policy: Among the routers in the MRT Island with the lowest numerical value advertised for GADAG Root Selection Priority, an implementation MUST pick the router with the highest Router ID to be the GADAG root. Note that a lower numerical value for GADAG Root Selection Priority indicates a higher preference for selection.

Forwarding Mechanisms: MRT SR-tunnel Option

Recalculation: Recalculation of MRTs SHOULD occur. This allows the MRT forwarding topologies to support IP/LDP fast-reroute traffic.

Area/Level Border Behavior: ABRs/LBRs SHOULD ensure that traffic leaving the area also exits the MRT-Red or MRT-Blue forwarding topology.

2.1. Option 1: MRT SR-tunnel Option

MRT SR-tunnel Option will treat the whole MRT-red or MRT-blue path as a segment list. Only FEC-label binding for the default topology-scoped FEC is needed. The specified segment list will direct the packet along the expected MRT path, but totally in the default topology. In order to forward packet along the expected MRT path strictly, the specified segment list suggests to be an adjacency segment list.
In Figure 1 above, supposed that the MRT-Red path is S-N1-N2-N3, which is used to protect the default primary nexthop F. Packets could be pushed to the SR-tunnel with segment list \{N1,N2,N3\} when the primary nexthop F is failed.

We can optimize a long adjacency segment list to a short node segment list. However, packets forwarding along a node segment list will not ensure to be strictly forwarded along the expected MRT path, as each node segment will always forward packets along the shortest path according to the newest topology that maybe different with the topology when doing optimization, also as each transit node can do local repair again for further failures, that could cause packet-forwarding micro-loop among two or more routers. Note that node segment list allows to try best to protect traffic for multiple simultaneous failures. Using adjacency segment list or node segment list is an implementation choice. Optimization from a adjacency segment list to a node segment list is out of the scope of this document.

Note that we can support tunneling traffic along an MRT to a tunnel endpoint that is not the destination of the traffic.

2.1.1. Forwarding SR Label Unicast Traffic over MRT Paths

When a PLR receives an SR label packet that needs to be forwarded on the MRT-Red (for example), it first does a label swap operation, replacing the incoming SR label assigned by the PLR for the FEC with the SR label assigned by the tunnel endpoint for that FEC, then push the whole label stack corresponding to the segment list of the MRT-Red path. The packet will forward to the tunnel endpoint (also MRT egress), then leave the MRT path to the shortest path to the destination.
2.1.2.  Forwarding IP Unicast Traffic over MRT Paths

When a PLR receives an IP packet that needs to be forwarded on the MRT-Red to a particular tunnel endpoint, it does a label stack push operation. The label stack pushed is the segment list of the MRT-Red path. The packet will forward to the tunnel endpoint (also MRT egress), then leave the MRT path to the shortest path to the destination.

2.1.3.  Inter-area Forwarding Behavior

As [RFC7812] described, it is desirable for traffic leaving the area to also exit MRT-Red or MRT-Blue and return to shortest path forwarding. However, this option always terminates the SR-tunnel corresponding to an MRT path at the MRT egress, the packets will leave the MRT path and continue to be forwarded according to the default topology-scoped SR label or IP header. There is no MRT specific prefix-sid in this option to create MRT specific SR LIB entry. Other considerations see "section 3. interoperation".

2.1.4.  Prefixes Multiply Attached to the MRT Island

This option will use tunnel endpoint selection approach to protection for both multihomed prefix (outside the area) and destinations outside the MRT Island (but inside the area). The forwarding behavior for these prefixes and ones that inside the MRT Island are totally same.

2.1.5.  FIB examples

A node S in the MRT Island will maintain the following FTN entry for a destination prefix: (suppose that the primary nexthop is F, and the MRT-red path is N1-N2-N3 that is used to protect the primary nexthop.)

example 1

KEY: MT-ID 0, prefix
Primary NHLFE: F, SRGB_F[prefix-sid]
According to adjacency segment list, the backup NHLFE would be:
Backup NHLFE: N1, adj-sid-N1toN2 ;outermost label
    adj-sid-N2toN3
    SRGB_N3[prefix-sid]
According to node segment list, the backup NHLFE would be:
Backup NHLFE: N1, SRGB_N1[N1-sid] ;outermost label
    SRGB_N1[N2-sid]
    SRGB_N2[N3-sid]
    SRGB_N3[prefix-sid]
Also, a LIB entry for FEC (MT-ID 0, prefix) is as the following:

```
example 2

KEY: SRGB_S[prefix-sid]
Primary NHLFE: F, SRGB_F[prefix-sid]
According to adjacency segment list, the backup NHLFE would be:
Backup NHLFE: N1, adj-sid-N1toN2 ;outermost label
        adj-sid-N2toN3
        SRGB_N3[prefix-sid]
According to node segment list, the backup NHLFE would be:
Backup NHLFE: N1, SRGB_N1[N1-sid] ;outermost label
        SRGB_N1[N2-sid]
        SRGB_N2[N3-sid]
        SRGB_N3[prefix-sid]
```

### 2.2. Option 2: MRT SR-LSP Tunneling with multi-prefix-sid Option

This Option will allocate a different FEC-label binding for each of the three topology-scoped FECs, i.e. default-FEC, red-FEC, blue-FEC, originated by routers inside the MRT Island. For example, When a packet is received with an SR label corresponding to red-FEC, an MRT transit router will determine the next hop for the MRT-Red forwarding topology for that FEC, swap the incoming SR label with the outgoing SR label corresponding to the MRT-Red next-hop router, and forward the packet.

The above FEC-label binding is calculated by prefix-sid for each of the three topology-scoped prefixes depending on default topology-scoped SRGB. Each router inside the MRT Island supporting this option will allocate three node-sids, i.e. default-node-sid, red-node-sid, blue-node-sid. In fact, We only need allocate node-sid for each of the three topology-scoped Node-flag prefixes for each router inside the MRT Island. Other prefixes inside the MRT Island could also be allocated the corresponding MRT topology-scoped prefix-sids besides the default topology-scoped prefix-sid, but it is not necessary for this option. And, prefixes outside the MRT Island, especially outside the area/level, need never allocate MRT topology-scoped prefix-sids, as a simple reason is that routers outsid the MRT Island have no idea about the MRT Island. Packets to the later two class prefixes could be tunneled in the SR-LSP to the tunnel endpoint which is inside the MRT Island. The MRT specific prefix-sid MUST only be advertised intra the area which the MRT Island belongs to.

Note that we can support tunneling traffic along an MRT to a tunnel endpoint that is not the destination of the traffic.
In Figure 2 above, supposed that the MRT-Red path is S-N1-N2-N3, which is used to protect the default primary nexthop F to destination D. Packets could be pushed to the SR LSP destined to N3 with label computed by MT-red node-sid 31 when the primary nexthop F is failed. In MT-red topology there is an LSP build from ingress S to egress N3.

2.2.1. Forwarding SR Label Unicast Traffic over MRT Paths

When a PLR receives an SR label packet that matches a LIB entry corresponding to a Node-flag prefix inside the MRT Island, and needs to be forwarded on the MRT-Red (for example), it does a label swap operation, replacing the incoming SR label for the FEC with the MRT-Red SR label for that FEC received from the next-hop router in the MRT-Red computed by the PLR. MRT transit router will determine the next hop for the MRT-Red forwarding topology for that FEC, swap the incoming MRT SR label with the outgoing MRT SR label learned from the MRT-Red next-hop router, and forward the packet.

When a PLR receives an SR label packet that matches a LIB entry corresponding to a prefix without Node-flag inside the MRT Island, or a prefix outside the MRT Island, and needs to be forwarded on the MRT-Red (for example) path to a tunnel endpoint inside the MRT Island, it first does a label swap operation, replacing the incoming SR label assigned by the PLR for the FEC with the SR label assigned by the tunnel endpoint for that FEC, then pushes the MRT-Red SR label to the tunnel endpoint. The packet will forward to the tunnel endpoint (also MRT egress), then leave the MRT path to the shortest path to the destination.

2.2.2. Forwarding IP Unicast Traffic over MRT Paths

When a PLR receives an IP packet that matches an FTN entry corresponding to a Node-flag prefix inside the MRT Island, and needs to be forwarded on the MRT-Red (for example), it does a label push operation, pushing the MRT-Red SR label for that FEC received from the next-hop router in the MRT-Red computed by the PLR. MRT transit
router will determine the next hop for the MRT-Red forwarding topology for that FEC, swap the incoming MRT SR label with the outgoing MRT SR label learned from the MRT-Red next-hop router, and forward the packet.

When a PLR receives an IP label packet that matches an FTN entry corresponding to a prefix without Node-flag inside the MRT Island, or a prefix outside the MRT Island, and needs to be forwarded on the MRT-Red (for example) path to a tunnel endpoint inside the MRT Island, it first pushes the SR label assigned by the tunnel endpoint for that FEC, then pushes the MRT-Red SR label to the tunnel endpoint. The packet will forward to the tunnel endpoint (also MRT egress), then leave the MRT path to the shortest path to the destination.

2.2.3. Inter-area Forwarding Behavior

This option always terminates the SR-LSP corresponding to an MRT path at the MRT egress, the packets will leave the MRT path and continue to be forwarded according to the default topology-scoped SR label or IP header. Although ABR will compute the MRT specific SR LIB for an MRT specific prefix-sid corresponding to a destination prefix with Node-flag inside the MRT Island that the ABR belongs to, the MRT specific prefix-sid will never leak to another area. It is impossible for the ABR to receive packets, from one area, containing an MRT specific SR label to the destination prefix which belongs to another area. Other considerations see "section 3. interoperation".

2.2.4. Prefixes Multiply Attached to the MRT Island

This option will use tunnel endpoint selection approach to protection for both multihomed prefix (outside the area) and destinations outside the MRT Island (but inside the area). The forwarding behavior for these prefixes and ones without Node-flag that inside the MRT Island are totally same.

2.2.5. FIB examples

A node S in the MRT Island will maintain the following FTN entry for a destination Node-flag prefix inside the MRT Island: (suppose that the primary nexthop is F, and the MRT-red path is N1-N2-N3 that is used to protect the primary nexthop.)

example 1

```
KEY: MT-ID 0, prefix
Primary NHLFE: F, SRGB_F[default-prefix-sid]
Backup NHLFE: N1, SRGB_N1[red-prefix-sid]
```
Also, a LIB entry for FEC (MT-ID 0, prefix) is as the following:

example 2

KEY: SRGB_S[default-prefix-sid]
Primary NHLFE: F, SRGB_F[default-prefix-sid]
Backup NHLFE: N1, SRGB_N1[red-prefix-sid]

Nodes along the above MRT-red path will maintain the red SR LIB entry. For example, a LIB entry for FEC (MT-red, prefix) in N1 node is as the following:

example 3

KEY: SRGB_N1[red-prefix-sid]
NHLFE: N2, SRGB_N2[red-prefix-sid]

Node S can also maintain the following FTN entry for a destination prefix without Node-flag inside the MRT Island or a destination prefix outside the MRT Island: (suppose that the primary nexthop is F, and the MRT-red path is N1-N2-N3 that is used to protect the primary nexthop.)

example 4

KEY: MT-ID 0, prefix
Primary NHLFE: F, SRGB_F[default-prefix-sid]
Backup NHLFE: N1, SRGB_N1[red-N3-sid] ;outermost label
SRGB_N3[default-prefix-sid]

Also, a LIB entry for FEC (MT-ID 0, prefix) is as the following:

example 5

KEY: SRGB_S[default-prefix-sid]
Primary NHLFE: F, SRGB_F[default-prefix-sid]
Backup NHLFE: N1, SRGB_N1[red-N3-sid] ;outermost label
SRGB_N3[default-prefix-sid]

2.3. Option 3: MRT SR-LSP Tunneling with multi-SRGB Option

This Option will allocate a different FEC-label binding for each of the three topology-scoped FECs, i.e. default-FEC, red-FEC, blue-FEC, originated by routers inside the MRT Island. For example, when a packet is received with an SR label corresponding to red-FEC, an MRT transit router will determine the next hop for the MRT-Red forwarding topology for that FEC, swap the incoming SR label with the outgoing
SR label corresponding to the MRT-Red next-hop router, and forward the packet.

The above FEC-label binding is calculated by prefix-sid for the default topology-scoped prefix depending on each topology-scoped SRGBs. Each router inside the MRT Island supporting this option will assign three SRGBs, i.e. default-SRGB, red-SRGB, blue-SRGB. In fact, we only need calculate MRT SR labels for Node-flag prefixes for each router inside the MRT Island. Other prefixes inside the MRT Island and all prefixes outside the MRT Island could also calculate the MRT SR labels depending on the MRT topology-scoped SRGBs besides the default SR label depending on the default topology-scoped SRGB, but it is not necessary for this option. Packets to the later two class prefixes could be tunneled in the SR-LSP to the tunnel endpoint which is inside the MRT Island.

The details of advertising SRGB per multi-topology will be described in the related IGP extension document. The MRT specific SRGB MUST only be advertised intra the area which the MRT Island belongs to.

Note that we can support tunneling traffic along an MRT to a tunnel endpoint that is not the destination of the traffic.

```
[F]---{___}................[D]
    |                           .
    |                           .
    |                           . for node N3:
[S]---[N1]---[N2]---[N3]...... MT-default SRGB:[100~200]
                           ================>
                           MT-red SRGB:[201~300]
                           MT-blue SRGB:[301~400]
MRT-Red path
```

Figure 3: MRT SR-LSP Tunneling with multi-SRGB Forwarding Mechanism

In Figure 3 above, supposed that the MRT-Red path is S-N1-N2-N3, which is used to protect the default primary nexthop F to destination D. Packets could be pushed to the SR LSP destined to N3 with label computed by MT-red SRGB when the primary nexthop F is failed. In MT-red topology there is an LSP build from ingress S to egress N3.

2.3.1. Forwarding SR Label Unicast Traffic over MRT Paths

Same as section 2.2.1.
2.3.2. Forwarding IP Unicast Traffic over MRT Paths

Same as section 2.2.2.

2.3.3. Inter-area Forwarding Behavior

This option always terminates the SR-LSP corresponding to an MRT path at the MRT egress, the packets will leave the MRT path and continue to be forwarded according to the default topology-scoped SR label or IP header. Suppose that area1 connects area2 by an ABR, MRT Island 1 is created inside area1 and MRT Island 2 inside area2. Although the ABR will compute the MRT specific SR LIB, according to specific MRT SRGB, for a destination prefix with Node-flag inside the MRT Island 2 that the ABR belongs to, a PLR also supporting this option inside the MRT Island 1 will never compute the MRT specific SR LIB for this prefix. It is impossible for the ABR to receive packets, from one area, containing an MRT specific SR label to the destination prefix which belongs to another area, if both MRT Islands support the same option. Other considerations see "section 3. interoperation".

2.3.4. Prefixes Multiply Attached to the MRT Island

This option will use tunnel endpoint selection approach to protection for both multihomed prefix (outside the area) and destinations outside the MRT Island (but inside the area). The forwarding behavior for these prefixes and ones without Node-flag that inside the MRT Island are totally same.

2.3.5. FIB examples

A node S in the MRT Island will maintain the following FTN entry for a destination Node-flag prefix inside the MRT Island: (suppose that the primary nexthop is F, and the MRT-red path is N1-N2-N3 that is used to protect the primary nexthop.)

example 1

KEY: MT-ID 0, prefix
Primary NHLFE: F, default_SRGB_F[prefix-sid]
Backup NHLFE: N1, red_SRGB_N1[prefix-sid]

Also, a LIB entry for FEC (MT-ID 0, prefix) is as the following:

example 2

KEY: default_SRGB_S[prefix-sid]
Primary NHLFE: F, default_SRGB_F[prefix-sid]
Backup NHLFE: N1, red_SRGB_N1[prefix-sid]
Nodes along the above MRT-red path will maintain the red SR LIB entry. For example, a LIB entry for FEC (MT-red, prefix) in N1 node is as the following:

example 3

KEY: red_SRGB_N1[prefix-sid]
NHLFE: N2, red_SRGB_N2[prefix-sid]

Node S can also maintain the following FTN entry for a destination prefix without Node-flag inside the MRT Island or a destination prefix outside the MRT Island: (suppose that the primary nexthop is F, and the MRT-red path is N1-N2-N3 that is used to protect the primary nexthop.)

example 4

KEY: MT-ID 0, prefix
Primary NHLFE: F, default_SRGB_F[prefix-sid]
Backup NHLFE: N1, red_SRGB_N1[N3-sid] ;outermost label
default_SRGB_N3[prefix-sid]

Also, a LIB entry for FEC (MT-ID 0, prefix) is as the following:

example 5

KEY: default_SRGB_S[prefix-sid]
Primary NHLFE: F, default_SRGB_F[prefix-sid]
Backup NHLFE: N1, red_SRGB_N1[N3-sid] ;outermost label
default_SRGB_N3[prefix-sid]

2.4. Option 4: MRT SR-LSP Non-tunneling with multi-SRGB Option

This Option will allocate a different FEC-label binding for each of the three topology-scoped FECs, i.e. default-FEC, red-FEC, blue-FEC, originated by any routers in the SR domain. For example, When a packet is received with an SR label corresponding to red-FEC, an MRT transit router will determine the next hop for the MRT-Red forwarding topology for that FEC, swap the incoming SR label with the outgoing SR label corresponding to the MRT-Red next-hop router, and forward the packet.

The above FEC-label binding is caculated by prefix-sid for the default topology-scoped prefix depending on each topology-scoped SRGBs. Each router inside the MRT Island supporting this option will assign three SRGBs, i.e. default-SRGB, red-SRGB, blue-SRGB. Difference between this option and the option described in section
2.3 is that we create MRT SR-LSP for every destination prefix, not only for Node-flag prefix inside the MRT Island.

Note that we only need to allocate and advertise default topology-scoped prefix-sid in the SR domain, there are no red and blue topology-scoped prefix-sids. But we need to advertise topology-scoped SRGBs, i.e. default-SRGB, red-SRGB, blue-SRGB, respectively intra the area.

The details of advertising SRGB per multi-topology will be described in the related IGP extension document. The MRT specific SRGB MUST only be advertised intra the area which the MRT Island belongs to.

```
[F]--(____)----------[D]
    |               |
    |               |
[S]--[N1]--[N2]--[N3].......
  ================>
     MT-default SRGB:[100~200]
     MT-red SRGB:[201~300]
     MT-blue SRGB:[301~400]
```

Figure 4: MRT SR-LSP Tunneling with multi-SRGB
Forwarding Mechanism

In Figure 4 above, supposed that the MRT-Red path is S-N1-N2-N3, which is used to protect the default primary nexthop F to destination D. Packets could be pushed to the SR LSP destined to D with label computed by MT-red SRGB when the primary nexthop F is failed. In MT-red topology there is an LSP build from ingress S to egress D, at node N3, the incoming label is computed by its own MT-red SRGB, the outgoing label is computed by next-hop’s MT-default SRGB.

2.4.1. Forwarding SR Label Unicast Traffic over MRT Paths

When a PLR receives an SR label packet that needs to be forwarded on the MRT-Red (for example), it does a label swap operation, replacing the incoming SR label for the FEC with the MRT-Red SR label for that FEC received from the next-hop router in the MRT-Red computed by the PLR. MRT transit router will determine the next hop for the MRT-Red forwarding topology for that FEC, swap the incoming MRT SR label with the outgoing MRT SR label learned from the MRT-Red next-hop router, and forward the packet.
2.4.2.  Forwarding IP Unicast Traffic over MRT Paths

When a PLR receives an IP packet that needs to be forwarded on the MRT-Red (for example), it does a label push operation, pushing the MRT-Red SR label for the FEC received from the next-hop router in the MRT-Red computed by the PLR. MRT transit router will determine the next hop for the MRT-Red forwarding topology for that FEC, swap the incoming MRT SR label with the outgoing MRT SR label learned from the MRT-Red next-hop router, and forward the packet.

2.4.3.  Inter-area Forwarding Behavior

For OSPF, the ABR is responsible for advertising the proper prefix-sid to each neighbor. Assume that an ABR has allocated a default topology-scoped prefix-sid for a particular destination. To those routers in the same area as the best route to the destination, the ABR advertises the prefix-sid as normal. However, to routers in other areas, the ABR advertises the prefix-sid with Rainbow-Flag. Prefix-sid with Rainbow-Flag causes the receiving routers to use default-SRGB of the next-hop to calculate SR out-label for the MRT-Blue MT-ID and for the MRT-Red MT-ID. The details of advertising prefix-sid with Rainbow-Flag will be described in another document.

The ABR installs all next hops for the best area: primary next hops for default-SRGB, MRT-Blue next hops for blue-SRGB, and MRT-Red next hops for red-SRGB. Because the ABR advertised prefix-sid with Rainbow-Flag to neighbors not in the best area, packets from those neighbors will arrive at the ABR with a label calculated by default-SRGB of the ABR and will be forwarded into the best area along the default topology. Other considerations see "section 3. interoperation".

The same essential behavior also applies to IS-IS if one substitutes IS-IS level for OSPF area. More details see "Appendix A" of [RFC7812].

2.4.4.  Prefixes Multiply Attached to the MRT Island

This option will use named proxy-node approach to protection for both multihomed prefix (outside the area) and destinations outside the MRT Island (but inside the area).

According to [RFC7812], a proxy-node attachment router has a special forwarding role. A proxy-node attachment router is similar with other routers in the MRT Island to compute and install three ILM entries for each (MT-ID 0, prefix) outside the MRT Island based on three different FEC-label bindings. In particular, it will compute SR out-label, for MRT SR ILM, based on the default-SRGB of the LFIN
whose cost was used in the selection or the remote endpoint of the interface that caused the router to advertise the prefix. However, it will always compute SR out-label based on the default-SRGB of the shortest path nexthop for default SR ILM.

Note that if the proxy-node attachment router is an ABR, and it also belongs to another MRT Island supporting this option in the area which the destination prefix belongs to. It will compute SR out-label, for MRT SR ILM, based on the red-SRGB or blue-SRGB of the MRT-red or MRT blue nexthops, in despite of whether or not advertising the destination prefix to another area. In this case, we can refer to section "2.4.3. Inter-area Forwarding Behavior".

When a packet is received destined to (MRT-Red, prefix) or (MRT-Blue, prefix), if the proxy-node attachment router is an IBR, it MUST swap to the shortest path forwarding topology (e.g., swap to the label calculated by default-SRGB of the IN) and forward the packet to the IN whose cost was used in the selection. If the proxy-node attachment router is not an IBR, then the packet MUST be removed from the MRT forwarding topology and sent along the interface(s) that caused the router to advertise the prefix; this interface might be out of the area/level/AS.

2.4.5. FIB examples

A node S in the MRT Island will maintain the following FTN entry for a destination prefix: (suppose that the primary nexthop is F, and the MRT-red path is N1-N2-N3 that is used to protect the primary nexthop.)

example 1

- KEY: MT-ID 0, prefix
  - Primary NHLFE: F, default_SRGB_F[prefix-sid]
  - Backup NHLFE: N1, red_SRGB_N1[prefix-sid]

Also, a LIB entry for FEC (MT-ID 0, prefix) is as the following:

example 2

- KEY: default_SRGB_S[prefix-sid]
  - Primary NHLFE: F, default_SRGB_F[prefix-sid]
  - Backup NHLFE: N1, red_SRGB_N1[prefix-sid]

Nodes along the above MRT-red path will maintain the red SR LIB entry. For example, a LIB entry for FEC (MT-red, prefix) in N1 node is as the following:
example 3

KEY: red_SRGB_N1[prefix-sid]
NHLFE: N2, red_SRGB_N2[prefix-sid]

In particular, if S is a proxy-node attachment router, and suppose that the LFIN whose cost was used in the selection or the remote endpoint of the interface that caused S to advertise the prefix is M, the FTN entry maintained by S would be like the following:

example 4

KEY: MT-ID 0, prefix
Primary NHLFE: F, default_SRGB_F[prefix-sid]
Backup NHLFE: M, default_SRGB_M[prefix-sid]

Also, the LIB entry for FEC (MT-ID 0, prefix) maintained by S would be like the following:

example 5

KEY: default_SRGB_S[prefix-sid]
Primary NHLFE: F, default_SRGB_F[prefix-sid]
Backup NHLFE: M, default_SRGB_M[prefix-sid]

Also, the LIB entry for FEC (MT-red, prefix) maintained by S would be like the following:

example 6

KEY: red_SRGB_S[prefix-sid]
NHLFE: M, default_SRGB_M[prefix-sid]

As notice in section 2.4.4, when S is a proxy-node attachment router and an ABR, and it also belongs to another MRT Island supporting this option in the area which the destination prefix belongs to. S will maintain FTN/LIB entries just like example 1,2,3, in despite of whether or not advertising the destination prefix to another area.

When S has received the the prefix-sid advertisement corresponding to (MT-ID 0, prefix) with a Rainbow-Flag from its MRT-red nexthop (an ABR), it will maintain the FTN entry as the following:

example 7

KEY: MT-ID 0, prefix
Primary NHLFE: F, default_SRGB_F[prefix-sid]
Backup NHLFE: ABR, default_SRGB_ABR[prefix-sid]
Also, a LIB entry for FEC (MT-ID 0, prefix) is as the following:

default_SRGB_S[prefix-sid]
Primary NHLFE: F, default_SRGB_F[prefix-sid]
Backup NHLFE: ABR, default_SRGB_ABR[prefix-sid]

Also, a LIB entry for FEC (MT-red, prefix) is as the following:

red_SRGB_S[prefix-sid]
NHLFE: ABR, default_SRGB_ABR[prefix-sid]

2.5. Option 5: MRT SR-LSP Non-tunneling with multi-prefix-sid Option

This Option will allocate a different FEC-label binding for each of the three topology-scoped FECs, i.e. default-FEC, red-FEC, blue-FEC, originated by any routers in the SR domain. For example, When a packet is received with an red SR label corresponding to red-FEC, an MRT transit router will determine the next hop for the MRT-Red forwarding topology for that FEC, swap the incoming red SR label with the outgoing red SR label corresponding to the MRT-Red next-hop router, and forward the packet.

Difference between this option and the option described in section 2.2 is that we create MRT SR-LSP for every destination prefix, not only for prefix with Node-flag inside the MRT Island. However, this option can not address the destination prefixes which are outside the MRT Island, especially outside the area/level, because routers outside the MRT Island have no idea about the MRT Island. It is impossible for these routers to allocate the red or blue topology-scoped prefix-sids, besides the default topology-scoped prefix-sid. Another reason is that an MRT specific prefix-sid can never leak between areas.

This option is waiting for future study.

3. Interoperation

It is possible that a set of routers inside an area support two or more options at the same time. It is also possible that one set of routers, supporting one option, interconnect another set of routers that support another option, both sets could be in the same area or not. More detailed descriptions of these sceneria are given below.
In Figure 5 above, the shortest path nexthop from S to D is F. S, D, N1, N2, and N3 support the above all profiles. N4, N5, and N6 support profile with "MRT SR-LSP Tunneling with multi-prefix-sid Option" forwarding mechanism only. N7, N8, and N9 support profile with "MRT SR-LSP Tunneling with multi-SRGB Option" forwarding mechanism only. N10, N11, and N12 support profile with "MRT SR-LSP Non-tunneling with multi-SRGB Option" forwarding mechanism only. So, there would be four MRT Islands in Figure 1. For each MRT Island, suppose that the MRT-red nexthop from S to D is same as the shortest path nexthop, i.e., F, the MRT-blue nexthop from S to D is Ni. Specifically, the MRT-blue nexthop is N1 for MRT Island with "MRT SR-tunnel Option" forwarding mechanism, the MRT-blue nexthop is ECMP {N1,N4} for MRT Island with "MRT SR-LSP Tunneling with multi-prefix-sid Option" forwarding mechanism, the MRT-blue nexthop is ECMP {N1,N7} for MRT Island with "MRT SR-LSP Tunneling with multi-SRGB Option" forwarding mechanism, the MRT-blue nexthop is ECMP {N1,N10} for MRT Island with "MRT SR-LSP Non-tunneling with multi-SRGB Option" forwarding mechanism.

From the perspective of node S, the FTN entry to D corresponding to each MRT Island is different with each other, e.g., some have single MRT-blue nexthop, some have ECMP MRT-blue nexthops. These FTN entries MUST be distinguished by different MT-IDs.

From the perspective of node N1 (also other transit node), the ILM entry to D corresponding to each MRT Island is also different with each other. These ILM entries MUST be distinguished by different SR in-label. Take a look at profile with "MRT SR-LSP Tunneling with multi-SRGB Option" forwarding mechanism and profile with "MRT SR-LSP Non-tunneling with multi-SRGB Option" forwarding mechanism, they both use MRT specific SRGB to calculate MRT specific SR label for a default topology-scoped prefix-sid. The MRT specific SRGB MUST be different.
between these two profiles. For the same reason, the MRT specific prefix-sid MUST be different between the profile with "MRT SR-LSP Tunneling with multi-prefix-sid Option" forwarding mechanism and the profile with "MRT SR-LSP Non-tunneling with multi-prefix-sid Option" forwarding mechanism.

### 3.2. Multiple Profiles Interworking

![Diagram](image)

In Figure 6 above, MRT Island 1 contains S, A, B, C, X. MRT Island 2 contains X, P, Q, R, D. If MRT Island 1 is supporting the profile with "MRT SR-tunnel Option" forwarding mechanism, MRT Island 2 could be any other profiles. The packets to from S to destination D will terminate the SR-tunnel corresponding to an MRT path at node X, then return to shortest path to D.

If MRT Island 1 is supporting the profile with "MRT SR-LSP Tunneling with multi-prefix-sid Option" or "MRT SR-LSP Tunneling with multi-SRGB Option" forwarding mechanism, MRT Island 2 could also be any other profiles. The packets to from S to destination D will terminate the MRT specific SR-LSP corresponding to an MRT path at node X, then return to shortest path to D. Note that each node in MRT Island 1 never compute MRT specific SR LIB for destination D which is outside MRT Island 1, although X in MRT Island 2 maybe compute it. X will never receive a packet, from MRT Island 1, which includes an MRT specific SR label directly to destination D.

If MRT Island 1 is supporting the profile with "MRT SR-LSP Non-tunneling with multi-SRGB Option" forwarding mechanism, each node in MRT Island 1 will compute MRT specific SR LIB for destination D directly. Especially, in node X, i.e., the proxy-node attachment router, the MRT specific SR LIB computed for destination D will forward packets to default topology. Hence, if MRT Island 2 is also supporting the profile with "MRT SR-LSP Non-tunneling with multi-SRGB Option" forwarding mechanism, X will also compute MRT specific SR LIB for destination D which will forward packets along MRT path, overwrite the above default topology forwarding information. In this case, X MUST advertise prefix-sid with Rainbow flag corresponding to (MT-ID 0, prefix-D) to neighbors in MRT Island 1, in order to avoid
receiving packets including an MRT specific SR label for destination D. If MRT Island 2 is supporting "MRT SR-LSP Tunneling with multi-SRGB Option" forwarding mechanism, X will also compute MRT specific SR LIB for destination D which will forward packets along MRT path, but in-label is different because of different MRT specific SRGB.

4. Support protecting segment list

As RFC7811 said, in some cases the PLR S can not find MRT alternates to destination D for local repair to protect the primary nexthop F. For example, if F is D, only link protection is possible. And if both MRT-Red and MRT-Blue use the primary next hop, then the primary next hop must be a cut-link, so either MRT could be pruned to avoid the failed primary next-hop interface.

Thus, if S received a packet which contains a segment list only including node segment F, we just comply with the MRT computing result for destination F, try best to forward packet along the MRT path which has not use the primary next hop. But if S received a packet which contains a segment list not only including node segment F as the first segment, but also other subsequent segments, S can temporarily change the destination to another node D’ rather than F, and send packet along the MRT path to destination D’. A simple condition which decided to set another destination D’ is that the original destination is a direct neighbor of S. We can easily check this condition during computing SPF routes, and set the related flag in the corresponding FTN/ILM entries. In dataplane, if the top label of the label stack of the packet S received matches an SR-ILM entry which has set the above flag to indicate the FEC is from a direct neighbor and the primary next hop of the ILM entry is failed, in despite of alternate next hops (e.g, MRT alternates) existing, S MAY check if the next label of the label stack is also an SR-based label, if so, deduce the next node-segment and directly forward the packet to the next node-segment according to the FTN/ILM entry corresponding to the next node-segment after twice label POP operations on the original label stack. However, S can also simply choose to forward packets along the alternate next hops of ILM entry corresponding to the first segment to avoid the aboved complexities. It is a local choice.

Similarly, if the top label of the label stack of the packet S received matches an SR-ILM entry corresponding to a local adjacency segment, S MAY also do the above check and process if the adjacency is failed, and can also simply choose to forward packets along the backup adjacencies. It is also a local choice.

It’s a bit complicated for the dataplane to deduce the next node-segment based on the label stack of the received packet. First, It
can deduce the first node-segment (i.e., F) based on the SR-ILM entry matched by the top label, the node information can get from the prefix FEC (F) or from the remote node-id of the adjacency FEC (S->F). Then, we can get the SRGB of the first node-segment (i.e., SRGB_F), and check if the next label is in SRGB_F, if so, we know that the next segment is a prefix segment and deduce the prefix-sid, which is then used to compute a label based on SRGB_S to match the ILM entry correspond to the next node-segment. If the next label is not in SRGB_F, S can check all F’s local adjacencies to see if there is a label equal to the next label, if so, we know that the next segment is an adjacency segment and deduce the next node-segment from the remote node-id of the matched adjacency, and an FTN entry could be match by the next node-segment. Note that S can maintain ILM entries for F’s all local adjacencies for the above purpose, although these ILM entries are not used to forward packets directly. The FTN or ILM entry of the next segment could be used to directly forward the packet to the next node-segment after twice label POP operations on the original label stack. If the next label is neither in SRGB_F nor matching F’s any local adjacencies, it must be a non-SR label, in this case, packets must be simply forwarded along the alternate next hops of ILM entry corresponding to the first segment, if there are no alternate next hops, packets would be dropped.

All possible cases are systematically described in the following sections.

4.1. Received segment list is {segment F or S->F, vpn label} or {segment F or S->F}

We can determine that only one SR related label existed in label stack, and the segment is a direct neighbor along the shortest path. In this case, only link protection is possible.

For MRT forwarding mechanism option 1, the MRT-FRR behavior is:

POP segment F or S->F
set D’ = F, goto FTN entry of D’, it will:
PUSH SRGB_E[SID_F] ;E is MRT Egress
PUSH the label stack of the SR-tunnel(from S to MRT Egress)

For MRT forwarding mechanism option 2, the MRT-FRR behavior is:

POP segment F or S->F
set D’ = F, goto FTN entry of D’, it will:
when F is inside MRT Island:
PUSH SRGB_N[mrt_SID_F] ;N is MRT Nexthop
when F is outside MRT Island:
no MRT FRR provided
For MRT forwarding mechanism option 3, the MRT-FRR behavior is:

POP segment F or S->F
set D' = F, goto FTN entry of D', it will:
when F is inside MRT Island:
   PUSH mrt_SRGB_N[SID_F]
when F is outside MRT Island:
   no MRT FRR provided

For MRT forwarding mechanism option 4, the MRT-FRR behavior is:

POP segment F or S->F
set D' = F, goto FTN entry of D', it will:
when S is not the proxy-node attachment router:
   no_rainbow:PUSH mrt_SRGB_N[SID_F]
   rainbow:PUSH default_SRGB_N[SID_F]
when S is the proxy-node attachment router:
   no MRT FRR provided

4.2. Received segment list is {segment F or S->F, segment D1 or F->D1, segment D2, ..., segment Dn, vpn label}

The first segment is a direct neighbor along the shortest path, followed by other segments.

For MRT forwarding mechanism option 1, the MRT-FRR behavior is:

POP segment F or S->F
POP segment D1 or F->D1
change D' from F to D1, goto FTN or ILM entry of D', it will:
when primary next hop for destination D' is valid:
   PUSH SRGB_H[SID_D1] ;H is primary next hop
when primary next hop for destination D' is not valid:
   PUSH SRGB_E[SID_D1], ;E is MRT Egress
   PUSH label stack of the SR-tunnel(from S to MRT Egress)

For MRT forwarding mechanism option 2, the MRT-FRR behavior is:
POP segment F or S->F
POP segment D1 or F->D1
change D' from F to D1, goto FTN or ILM entry of D', it will:
when primary next hop for destination D' is valid:
PUSH SRGB_H[default_SID_D1]
when primary next hop for destination D' is not valid:
when D1 is inside MRT Island:
PUSH SRGB_N[mrt_SID_D1] ;N is MRT Nexthop
when D1 is outside MRT Island:
PUSH SRGB_E[default_SID_D1]
PUSH SRGB_N[mrt_SID_E]

For MRT forwarding mechanism option 3, the MRT-FRR behavior is:

POP segment F or S->F
POP segment D1 or F->D1
change D' from F to D1, goto FTN or ILM entry of D', it will:
when primary next hop for destination D' is valid:
PUSH default_SRGB_H[SID_D1]
when primary next hop for destination D' is not valid:
when D1 is inside MRT Island:
PUSH mrt_SRGB_N[SID_D1]
when D1 is outside MRT Island:
PUSH default_SRGB_E[SID_D1]
PUSH mrt_SRGB_N[SID_E]

For MRT forwarding mechanism option 4, the MRT-FRR behavior is:

POP segment F or S->F
POP segment D1 or F->D1
change D' from F to D1, goto FTN or ILM entry of D', it will:
when primary next hop for destination D' is valid:
PUSH default_SRGB_H[SID_D1]
when primary next hop for destination D' is not valid:
when S is not the proxy-node attachment router:
  no_rainbow:PUSH mrt_SRGB_N[SID_D1]
  rainbow:PUSH default_SRGB_N[SID_D1]
when S is the proxy-node attachment router:
PUSH default_SRGB_M[SID_D1] ;M is nexthop outside Island

4.3. Received segment list is {segment D1, segment D2, ..., segment Dn, vpn label}

The first segment is not a direct neighbor along the shortest path, followed by other segments.

For MRT forwarding mechanism option 1, the MRT-FRR behavior is:
set D' = D1, goto ILM entry of D', it will:
    swap segment D1, i.e., SRGB_S[SID_D1] with SRGB_E[SID_D1]
    ;E is MRT Egress
    PUSH label stack of the SR-tunnel (from S to MRT Egress)

For MRT forwarding mechanism option 2, the MRT-FRR behavior is:

    set D' = D1, goto ILM entry of D', it will:
    when D1 is inside MRT Island:
        swap segment D1, i.e., SRGB_S[default_SID_D1] with
        SRGB_N[mrt_SID_D1]
        ;N is MRT Nexthop
    when D1 is outside MRT Island:
        swap segment D1, i.e., SRGB_S[default_SID_D1] with
        SRGB_E[default_SID_D1]
        PUSH SRGB_N[mrt_SID_E]

For MRT forwarding mechanism option 3, the MRT-FRR behavior is:

    set D' = D1, goto ILM entry of D', it will:
    when D1 is inside MRT Island:
        swap segment D1, i.e., default_SRGB_S[SID_D1] with
        mrt_SRGB_N[SID_D1]
    when D1 is outside MRT Island:
        swap segment D1, i.e., default_SRGB_S[SID_D1] with
        default_SRGB_E[SID_D1]
        PUSH mrt_SRGB_N[SID_E]

For MRT forwarding mechanism option 4, the MRT-FRR behavior is:

    set D' = D1, goto ILM entry of D', it will:
    when S is not the proxy-node attachment router:
        no_rainbow: swap segment D1, i.e., default_SRGB_S[SID_D1]
        with mrt_SRGB_N[SID_D1]
        rainbow: swap segment D1, i.e., default_SRGB_S[SID_D1] with
        default_SRGB_N[SID_D1]
    when S is the proxy-node attachment router:
        swap segment D1, i.e., default_SRGB_S[SID_D1] with
        default_SRGB_M[SID_D1]
        ;M is nexthop outside Island

5. Security Considerations
   TBD.

6. IANA Considerations
   TBD
7. References

7.1. Normative References

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

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

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

[I-D.ietf-spring-segment-routing-mpls]


7.2. Informative References

Authors’ Addresses

Shaofu Peng
ZTE Corporation
No.68 Zijinghua Road, Yuhuatai District
Nanjing  210012
China

Email: peng.shaofu@zte.com.cn

Ran Chen
ZTE Corporation
No.50 Software Avenue, Yuhuatai District
Nanjing  210012
China

Email: chen.ran@zte.com.cn