  Extensions to RSVP-TE for Label Switched Path (LSP) Egress Protection

Abstract

   This document describes extensions to Resource Reservation Protocol -
   Traffic Engineering (RSVP-TE) for locally protecting the egress
   node(s) of a Point-to-Point (P2P) or Point-to-Multipoint (P2MP)
   Traffic Engineered (TE) Label Switched Path (LSP).

Status of This Memo

   This is an Internet Standards Track document.

   This document is a product of the Internet Engineering Task Force
   (IETF). It represents the consensus of the IETF community. It has
   received public review and has been approved for publication by the
   Internet Engineering Steering Group (IESG). Further information on
   Internet Standards is available in Section 2 of RFC 7841.

   Information about the current status of this document, any errata,
   and how to provide feedback on it may be obtained at
   https://www.rfc-editor.org/info/rfc8400.

Table of Contents

1. Introduction

   [RFC4090] describes two methods for locally protecting the transit
   nodes of a P2P LSP: one-to-one and facility protection. [RFC4875]
   specifies how these methods can be used to protect the transit nodes
   of a P2MP LSP. These documents do not discuss the procedures for
   locally protecting the egress node(s) of an LSP.

   This document fills that void and specifies extensions to RSVP-TE for
   local protection of the egress node(s) of an LSP. "Egress node" and
   "egress" are used interchangeably.

Chen, et al.   Standards Track   [Page 2]

1.1. Local Protection of Egress Nodes

   In general, locally protecting an egress node of an LSP means that
   when the egress node fails, the traffic that the LSP carries will be
   delivered to its destination by the direct upstream node of the
   egress node to a backup egress node. Without protecting the egress
   node of the LSP, when the egress node fails, the traffic will be lost
   (i.e., the traffic will not be delivered to its destination).

   Figure 1 shows an example of using backup LSPs to locally protect
   egress nodes L1 and L2 of a primary P2MP LSP starting from ingress
   node R1. La and Lb are the designated backup egress nodes for
   primary egress nodes L1 and L2, respectively. The backup LSP for
   protecting L1 is from its upstream node R3 to backup egress node La,
   and the backup LSP for protecting L2 is from R5 to Lb.

```
                    *******  *******              S   Source
                 [R2]-----[R3]-----[L1]           CEx Customer Edge
                  */          &\        \         Rx  Non-Egress
                  */          &\         \        Lx  Egress
                 */           &\      [CE1]        *** Primary LSP
                */            &\       /           &&& Backup LSP
               */             &\      /
              */              [La]
             */
            */
           */
          */ ********  ********  *******
       [S]---[R1]------[R4]------[R5]-----[L2]
                          &\           \
                          &\            \
                          &\        [CE2]
                          &\         /
                           &\       /
                          [Lb]
```
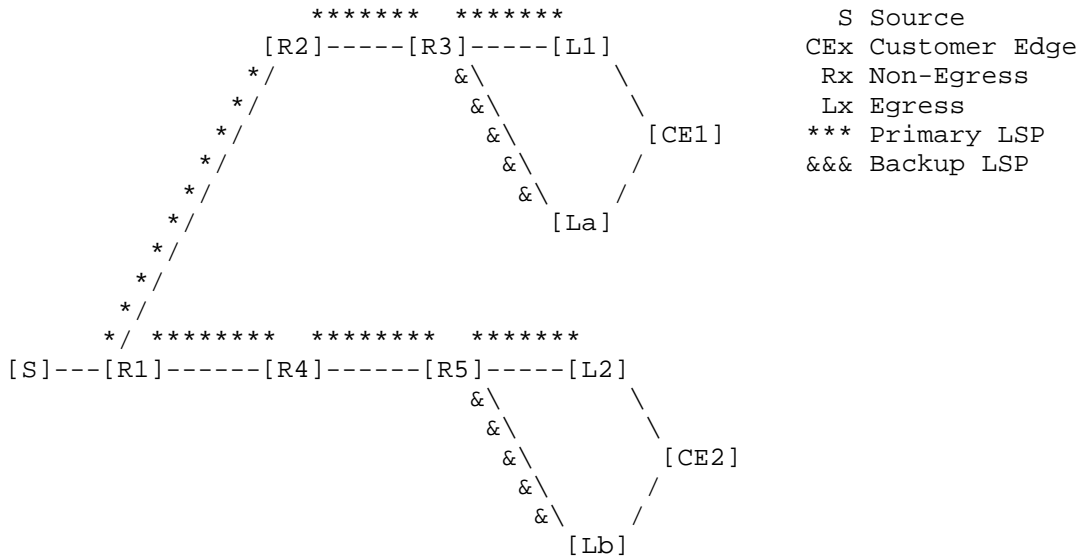
                Figure 1: Backup LSP for Locally Protecting Egress

   During normal operations, the traffic carried by the P2MP LSP is sent
   through R3 to L1, which delivers the traffic to its destination CE1.
   When R3 detects the failure of L1, R3 switches the traffic to the
   backup LSP to backup egress node La, which delivers the traffic to
   CE1. The time for switching the traffic is within tens of
   milliseconds.

   The exact mechanism by which the failure of the primary egress node
   is detected by the upstream node R3 is out of the scope of this
   document.

In the beginning, the primary P2MP LSP from ingress node R1 to
primary egress nodes L1 and L2 is configured. It may be used to
transport the traffic from source S, which is connected to R1, to
destinations CE1 and CE2, which are connected to L1 and L2,
respectively.

To protect the primary egress nodes L1 and L2, one configures on the
ingress node R1 a backup egress node for L1, another backup egress
node for L2, and other options. After the configuration, the ingress
node sends a Path message for the LSP with information such as the
Secondary Explicit Route Objects (SEROs), refer to Section 4.1,
containing the backup egress nodes for protecting the primary egress
nodes.

After receiving the Path message with the information, the upstream
node of a primary egress node sets up a backup LSP to the
corresponding backup egress node for protecting the primary egress
node.

2. Conventions Used in This Document

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.

3. Terminology

The following terminology is used in this document.

LSP: Label Switched Path

TE: Traffic Engineering

P2MP: Point-to-Multipoint

P2P: Point-to-Point

LSR: Label Switching Router

RSVP: Resource Reservation Protocol

S2L: Source-to-Leaf

SERO: Secondary Explicit Route Object

RRO: Record Route Object

   BFD: Bidirectional Forwarding Detection

   VPN: Virtual Private Network

   L3VPN: Layer 3 VPN

   VRF: Virtual Routing and Forwarding

   LFIB: Label Forwarding Information Base

   UA: Upstream Assigned

   PLR: Point of Local Repair

   BGP: Border Gateway Protocol

   CE: Customer Edge

   PE: Provider Edge

4. Protocol Extensions

4.1. Extensions to SERO

   The Secondary Explicit Route Object (SERO) is defined in [RFC4873].
   The format of the SERO is reused.

   The SERO used for protecting a primary egress node of a primary LSP
   may be added into the Path messages for the LSP and sent from the
   ingress node of the LSP to the upstream node of the egress node. It
   contains three subobjects.

   The first subobject (refer to Section 4.2 of [RFC4873]) indicates the
   branch node that is to originate the backup LSP (to a backup egress
   node). The branch node is typically the direct upstream node of the
   primary egress node of the primary LSP. If the direct upstream node
   does not support local protection against the failure of the primary
   egress node, the branch node can be any (upstream) node on the
   primary LSP. In this case, the backup LSP from the branch node to
   the backup egress node protects against failures on the segment of
   the primary LSP from the branch node to, and including, the primary
   egress node.

   The second subobject is an Egress Protection subobject, which is a
   PROTECTION object with a new C-Type (3). The format of the Egress
   Protection subobject is defined as follows:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|L|    Type       |     Length      |   Reserved   |  C-Type (3)  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Reserved                       |E-Flags|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Optional Subobjects                        |
~                                                               ~
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
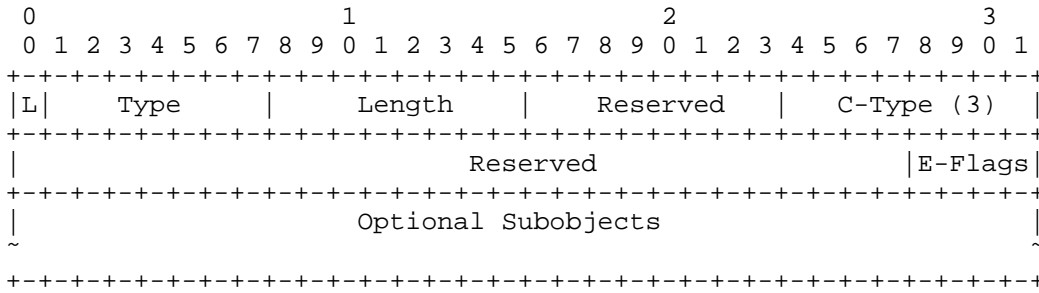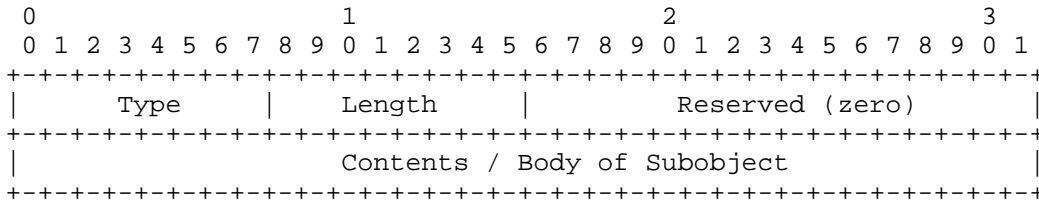
E-Flags are defined for local protection of egress nodes.

Bit 31 ("egress local protection" flag): It is the least significant
   bit of the 32-bit word and is set to 1, which indicates that local
   protection of egress nodes is desired.

Bit 30 ("S2L sub-LSP backup desired" flag): It is the second least
   significant bit of the 32-bit word and is set to 1, which
   indicates an S2L sub-LSP (refer to [RFC4875]) is desired for
   protecting an egress node of a P2MP LSP.

The Reserved parts MUST be set to zero on transmission and MUST be
ignored on receipt.

Four optional subobjects are defined: they are IPv4 and IPv6 primary
egress node subobjects as well as IPv4 and IPv6 P2P LSP ID
subobjects. IPv4 and IPv6 primary egress node subobjects indicate
the IPv4 and IPv6 address of the primary egress node, respectively.
IPv4 and IPv6 P2P LSP ID subobjects contain the information for
identifying IPv4 and IPv6 backup P2P LSP tunnels, respectively.
Their contents are described in Sections 4.1.1 through 4.1.2.2. They
have 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 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |     Length      |      Reserved (zero)        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                  Contents / Body of Subobject                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

where Type is the type of a subobject and Length is the total size of
the subobject in bytes, including Type, Length, and Contents fields.
The Reserved field MUST be set to zero on transmission and MUST be
ignored on receipt.

The third (final) subobject (refer to Section 4.2 of [RFC4873]) in
the SERO contains the egress node of the backup LSP, i.e., the
address of the backup egress node in the place of the merge node.

After the upstream node of the primary egress node (a.k.a. the branch
node) receives the SERO and determines a backup egress node for the
primary egress node, it computes a path from itself to the backup
egress node and sets up a backup LSP along the path for protecting
the primary egress node according to the information in the
FAST_REROUTE object in the Path message. For example, if facility
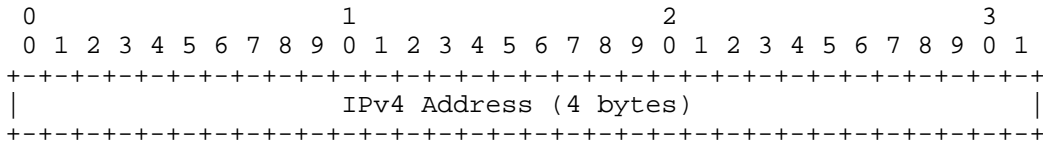protection is desired, it is provided for the primary egress node.

The upstream node constructs a new SERO based on the SERO received
and adds the new SERO into the Path message for the backup LSP. The
new SERO also contains three subobjects as the SERO for the primary
LSP. The first subobject in the new SERO indicates the upstream
node, which may be copied from the first subobject in the SERO
received. The second subobject in the new SERO includes a primary
egress node, which indicates the address of the primary egress node.
The third one contains the backup egress node.

The upstream node updates the SERO in the Path message for the
primary LSP. The Egress Protection subobject in the SERO contains a
subobject called a P2P LSP ID subobject, which contains the
information for identifying the backup LSP. The final subobject in
the SERO indicates the address of the backup egress node.

4.1.1. Primary Egress Subobject

There are two primary egress subobjects: the IPv4 primary egress
subobject and the IPv6 primary egress subobject.

The Type of an IPv4 primary egress subobject is 1, and the body of
the subobject is given below:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     IPv4 Address (4 bytes)                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

o IPv4 Address: The IPv4 address of the primary egress node.

   The Type of an IPv6 primary egress subobject is 2, and the body of
   the subobject is shown below:

```
    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
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                    IPv6 Address (16 bytes)                    |
   ~                                                               ~
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
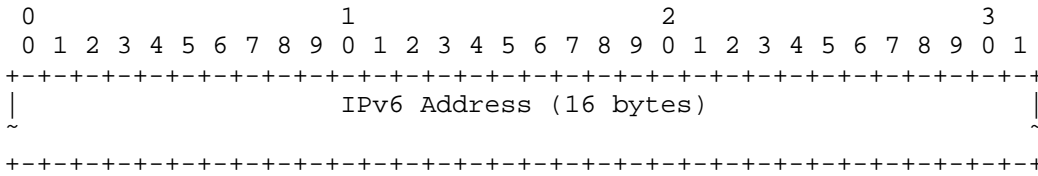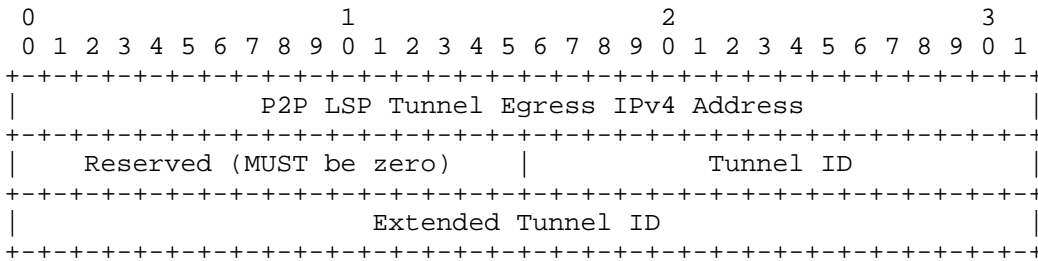
   o IPv6 Address: The IPv6 address of the primary egress node.

4.1.2. P2P LSP ID Subobject

   A P2P LSP ID subobject contains the information for identifying a
   backup P2P LSP tunnel.

4.1.2.1. IPv4 P2P LSP ID Subobject

   The Type of an IPv4 P2P LSP ID subobject is 3, and the body of the
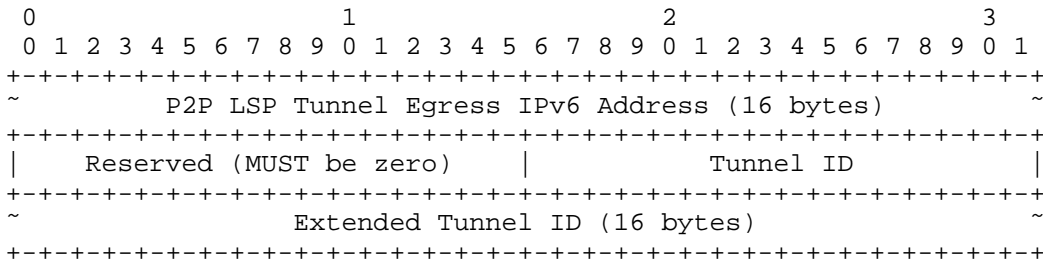   subobject is shown below:

```
    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
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                P2P LSP Tunnel Egress IPv4 Address             |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |      Reserved (MUST be zero)    |           Tunnel ID         |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                        Extended Tunnel ID                     |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

   o P2P LSP Tunnel Egress IPv4 Address: The IPv4 address of the egress
      node of the tunnel.

   o Tunnel ID (refer to [RFC4875] and [RFC3209]): A 16-bit identifier
      that remains constant over the life of the tunnel and occupies the
      least significant 16 bits of the 32-bit word.

   o Extended Tunnel ID (refer to [RFC4875] and [RFC3209]): A 4-byte
      identifier that remains constant over the life of the tunnel.

4.1.2.2. IPv6 P2P LSP ID Subobject

   The Type of an IPv6 P2P LSP ID subobject is 4, and the body of the
   subobject is illustrated below:

```
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~          P2P LSP Tunnel Egress IPv6 Address (16 bytes)       ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Reserved (MUST be zero)     |            Tunnel ID         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~                   Extended Tunnel ID (16 bytes)              ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

   o P2P LSP Tunnel Egress IPv6 Address: The IPv6 address of the egress
      node of the tunnel.

   o Tunnel ID (refer to [RFC4875] and [RFC3209]): A 16-bit identifier
      that remains constant over the life of the tunnel and occupies the
      least significant 16 bits of the 32-bit word.

   o Extended Tunnel ID (refer to [RFC4875] and [RFC3209]): A 16-byte
      identifier that remains constant over the life of the tunnel.

5. Egress Protection Behaviors

5.1. Ingress Behavior

   To protect a primary egress node of an LSP, the ingress node MUST set
   the "label recording desired" flag and the "node protection desired"
   flag in the SESSION_ATTRIBUTE object.

   If one-to-one backup or facility backup is desired to protect a
   primary egress node of an LSP, the ingress node MUST include a
   FAST_REROUTE object and set the "one-to-one backup desired" or
   "facility backup desired" flag, respectively.

   If S2L sub-LSP backup is desired to protect a primary egress node of
   a P2MP LSP, the ingress node MUST set the "S2L sub-LSP backup
   desired" flag in an SERO object.

   The decision to instantiate a backup egress node for protecting the
   primary egress node of an LSP can be initiated by either the ingress
   node or the primary egress node of that LSP, but not both.

   A backup egress node MUST be configured on the ingress node of an LSP
   to protect a primary egress node of the LSP if and only if the backup
   egress node is not configured on the primary egress node (refer to
   Section 5.2).

   The ingress node MUST send a Path message for the LSP with the
   objects above and the SEROs for protecting egress nodes of the LSP if
   protection of the egress nodes is desired. For each primary egress
   node of the LSP to be protected, the ingress node MUST add an SERO
   object into the Path message if the backup egress node, or some
   options, are given. If the backup egress node is given, then the
   final subobject in the SERO contains it; otherwise, the address in
   the final subobject is zero.

5.2. Primary Egress Behavior

   To protect a primary egress node of an LSP, a backup egress node MUST
   be configured on the primary egress node of the LSP to protect the
   primary egress node if and only if the backup egress node is not
   configured on the ingress node of the LSP (refer to Section 5.1).

   If the backup egress node is configured on the primary egress node of
   the LSP, the primary egress node MUST send its upstream node a Resv
   message for the LSP with an SERO for protecting the primary egress
   node. It sets the flags in the SERO in the same way as an ingress
   node.

   If the LSP carries the service traffic with a service label, the
   primary egress node sends its corresponding backup egress node the
   information about the service label as a UA label (refer to
   [RFC5331]) and the related forwarding.

5.3. Backup Egress Behavior

   When a backup egress node receives a Path message for an LSP, it
   determines whether the LSP is used for egress local protection by
   checking the SERO with an Egress Protection subobject in the message.
   If there is an Egress Protection subobject in the Path message for
   the LSP and the "egress local protection" flag in the object is set
   to 1, the LSP is the backup LSP for local protection of an egress
   node. The primary egress node to be protected is in the primary
   egress subobject in the SERO.

   When the backup egress node receives the information about a UA label
   and its related forwarding from the primary egress node, it uses the
   backup LSP label as a context label and creates a forwarding entry
   using the information about the UA label and the related forwarding.

This forwarding entry is in a forwarding table for the primary egress
node.

When the primary egress node fails, its upstream node switches the
traffic from the primary LSP to the backup LSP to the backup egress
node, which delivers the traffic to its receiver, such as a CE, using
the backup LSP label as a context label to get the forwarding table
for the primary egress node and using the service label as a UA label
to find the forwarding entry in the table to forward the traffic to
the receiver.

5.4. Transit Node and PLR Behavior

If a transit node of an LSP receives the Path message with the SEROs
and it is not an upstream node of any primary egress node of the LSP
as a branch node, it MUST forward them unchanged.

If the transit node is the upstream node of a primary egress node to
be protected as a branch node, it determines the backup egress node,
obtains a path for the backup LSP, and sets up the backup LSP along
the path. If the upstream node receives the Resv message with an
SERO object, it MUST send its upstream node the Resv message without
the object.

The PLR (which is the upstream node of the primary egress node a.k.a.
the branch node) MUST extract the backup egress node from the
respective SERO object in either a Path or a Resv message. If no
matching SERO object is found, the PLR tries to find the backup
egress node, which is not the primary egress node but has the same IP
address as the destination IP address of the LSP.

Note that if a backup egress node is not configured explicitly for
protecting a primary egress node, the primary egress node and the
backup egress node SHOULD have the same local address configured, and
the cost to the local address on the backup egress node SHOULD be
much bigger than the cost to the local address on the primary egress
node. Thus, the primary egress node and backup egress node are
considered as a "virtual node". Note that the backup egress node is
different from this local address (e.g., from the primary egress
node's point of view). In other words, it is identified by an
address different from this local address.

After obtaining the backup egress node, the PLR computes a backup
path from itself to the backup egress node and sets up a backup LSP
along the path. It excludes the segment including the primary egress
node to be protected when computing the path. The PLR sends the
primary egress node a Path message with an SERO for the primary LSP,

which indicates the backup egress node by the final subobject in the
SERO. The PLR puts an SERO into the Path messages for the backup
LSP, which indicates the primary egress node.

The PLR MUST provide one-to-one backup protection for the primary
egress node if the "one-to-one backup desired" flag is set in the
message; otherwise, it MUST provide facility backup protection if the
"facility backup desired" flag is set.

The PLR MUST set the protection flags in the RRO subobject for the
primary egress node in the Resv message according to the status of
the primary egress node and the backup LSP protecting the primary
egress node. For example, it sets the "local protection available"
flag and the "node protection" flag, which indicate that the primary
egress node is protected when the backup LSP is up and ready to
protect the primary egress node.

5.4.1. Signaling for One-to-One Protection

The behavior of the upstream node of a primary egress node of an LSP
(as a PLR) is the same as that of a PLR for one-to-one backup
described in [RFC4090], except that the upstream node (as a PLR)
creates a backup LSP from itself to a backup egress node in a session
different from the primary LSP.

If the LSP is a P2MP LSP and a primary egress node of the LSP is also
a transit node (i.e., bud node), the upstream node of the primary
egress node (as a PLR) creates a backup LSP from itself to each of
the next hops of the primary egress node.

When the PLR detects the failure of the primary egress node, it
switches the packets from the primary LSP to the backup LSP to the
backup egress node. For the failure of the bud node of a P2MP LSP,
the PLR also switches the packets to the backup LSPs to the bud
node's next hops, where the packets are merged into the primary LSP.

5.4.2. Signaling for Facility Protection

Except for backup LSP and downstream label, the behavior of the
upstream node of the primary egress node of a primary LSP (as a PLR)
follows the PLR behavior for facility backup, which is described in
[RFC4090].

For a number of primary P2P LSPs going through the same PLR to the
same primary egress node, the primary egress node of these LSPs MAY
be protected by one backup LSP from the PLR to the backup egress node
designated for protecting the primary egress node.

The PLR selects or creates a backup LSP from itself to the backup
egress node. If there is a backup LSP that satisfies the constraints
given in the Path message, then this one is selected; otherwise, a
new backup LSP to the backup egress node is created.

After getting the backup LSP, the PLR associates the backup LSP with
a primary LSP for protecting its primary egress node. The PLR
records that the backup LSP is used to protect the primary LSP
against its primary egress node failure and MUST include an SERO
object in the Path message for the primary LSP. The object MUST
contain the backup LSP ID. It indicates that the primary egress node
MUST send the backup egress node the service label as a UA label and
also send the information about forwarding the traffic to its
destination using the label if there is a service carried by the LSP
and the primary LSP label as a UA label (if the label is not implicit
null). How a UA label is sent is out of scope for this document
(refer to [FRAMEWK]).

When the PLR detects the failure of the primary egress node, it
redirects the packets from the primary LSP into the backup LSP to the
backup egress node and keeps the primary LSP label from the primary
egress node in the label stack if the label is not implicit null.
The backup egress node delivers the packets to the same destinations
as the primary egress node using the backup LSP label as a context
label and the labels under as UA labels.

5.4.3. Signaling for S2L Sub-LSP Protection

   The S2L sub-LSP protection uses an S2L sub-LSP (refer to [RFC4875])
   as a backup LSP to protect a primary egress node of a P2MP LSP. The
   PLR MUST determine to protect a primary egress node of a P2MP LSP via
   S2L sub-LSP protection when it receives a Path message with the "S2L
   sub-LSP backup desired" flag set.

   The PLR MUST set up the backup S2L sub-LSP to the backup egress node
   and create and maintain its state in the same way as if setting up a
   S2L sub-LSP defined in [RFC4875] from the signaling's point of view.
   It computes a path for the backup LSP from itself to the backup
   egress node, constructs and sends a Path message along the path, and
   receives and processes a Resv message responding to the Path message.

   After receiving the Resv message for the backup LSP, the PLR creates
   a forwarding entry with an inactive state or flag called "inactive
   forwarding entry". This inactive forwarding entry is not used to
   forward any data traffic during normal operations.

When the PLR detects the failure of the primary egress node, it
changes the forwarding entry for the backup LSP to "active". Thus,
the PLR forwards the traffic to the backup egress through the backup
LSP, which sends the traffic to its destination.

5.4.4. PLR Procedures during Local Repair

When the upstream node of a primary egress node of an LSP (as a PLR)
detects the failure of the primary egress node, it follows the
procedures defined in Section 6.5 of [RFC4090]. It SHOULD notify the
ingress node about the failure of the primary egress node in the same
way as a PLR notifies the ingress node about the failure of a transit
node.

Moreover, the PLR MUST let the upstream part of the primary LSP stay
alive after the primary egress node fails by sending the Resv message
to its upstream node along the primary LSP. The downstream part of
the primary LSP from the PLR to the primary egress node SHOULD be
removed. When a bypass LSP from the PLR to a backup egress node
protects the primary egress node, the PLR MUST NOT send any Path
message for the primary LSP through the bypass LSP to the backup
egress node.

In the local revertive mode, the PLR will re-signal each of the
primary LSPs that were routed over the restored resource once it
detects that the resource is restored. Every primary LSP
successfully re-signaled along the restored resource will be switched
back.

Note that the procedure for protecting the primary egress node is
triggered on the PLR if the primary egress node failure is
determined. If link (from PLR to primary egress node) failure and
primary egress node alive are determined, then the link protection
procedure is triggered on the PLR. How to determine these is out of
scope for this document.

6. Application Traffic Considerations

This section focuses on an example with application traffic carried
by P2P LSPs.

6.1. A Typical Application

L3VPN is a typical application. Figure 2 below shows a simple VPN
that consists of two CEs, CE1 and CE2, connected to two PEs, R1 and
L1, respectively. There is a P2P LSP from R1 to L1, which is
represented by stars (****). This LSP is called the primary LSP. R1
is the ingress node of the LSP and L1 is the (primary) egress node of

the LSP. R1 sends the VPN traffic received from CE1 through the P2P
LSP to L1, which delivers the traffic to CE2. R1 sends the VPN
traffic with an LSP label and a VPN label via the LSP. When the
traffic reaches the egress node L1 of the LSP, L1 pops the LSP label
and uses the VPN label to deliver the traffic to CE2.

In previous solutions based on ingress protection to protect the VPN
traffic against failure of the egress node L1 of the LSP, when the
egress node fails, the ingress node R1 of the LSP does the reroute
(refer to Figure 2). This solution entailed:

1. A multi-hop BFD session between ingress node R1 and egress node
    L1 of the primary LSP. The BFD session is represented by dots
    (....).

2. A backup LSP from ingress node R1 to backup egress node La, which
    is indicated by ampersands (&&&&).

3. La sends R1 a VPN backup label and related information via BGP.

4. R1 has a VRF with two sets of routes for CE2: one set uses the
    primary LSP and L1 as the next hop; the other uses the backup LSP
    and La as the next hop.

```
                    *****     *****
 CE1,CE2 in    [R2]-----[R3]-----[L1]              **** Primary LSP
 one VPN       */               :    \             &&&& Backup LSP
              */ .................:    \            .... BFD Session
  [CE1]--[R1] ..:                        [CE2]
           &\                          /
            &\                        /
             [R4]-----[R5]-----[La](BGP sends R1 VPN backup label)
                 &&&&&     &&&&&
```
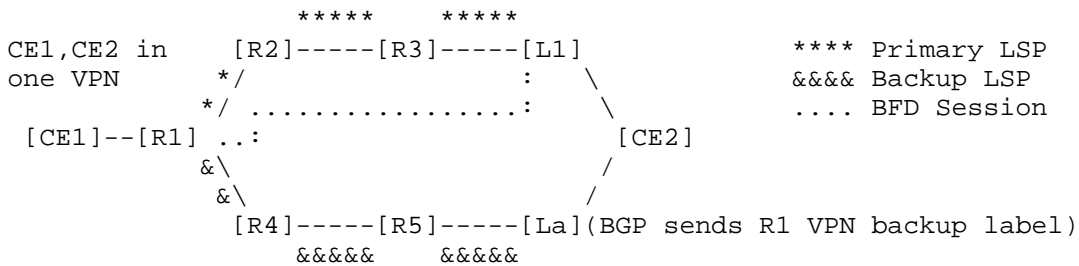
                Figure 2: Protect Egress for L3VPN Traffic

In normal operations, R1 sends the VPN traffic from CE1 through the
primary LSP with the VPN label received from L1 as the inner label to
L1, which delivers the traffic to CE2 using the VPN label.

When R1 detects the failure of L1, R1 sends the traffic from CE1 via
the backup LSP with the VPN backup label received from La as the
inner label to La, which delivers the traffic to CE2 using the VPN
backup label.

The solution defined in this document that uses egress local
protection for protecting L3VPN traffic entails (refer to Figure 3):

1. A BFD session between R3 (i.e., upstream node of L1) and egress
    node L1 of the primary LSP. This is different from the BFD
    session in Figure 2, which is a multi-hop between ingress node R1
    and egress node L1. The PLR R3 is closer to L1 than the ingress
    node R1. It may detect the failure of the egress node L1 faster
    and more reliably. Therefore, this solution can provide faster
    protection for failure of an egress node.

2. A backup LSP from R3 to backup egress node La. This is different
    from the backup LSP in Figure 2, which is an end-to-end LSP from
    ingress node R1 to backup egress node La.

3. Primary egress node L1 sends backup egress node La the VPN label
    as a UA label and also sends related information. The backup
    egress node La uses the backup LSP label as a context label and
    creates a forwarding entry using the VPN label in an LFIB for the
    primary egress node L1.

4. L1 and La are virtualized as one node (or address). R1 has a VRF
    with one set of routes for CE2, using the primary LSP from R1 to
    L1 and a virtualized node as the next hop. This can be achieved
    by configuring the same local address on L1 and La using the
    address as a destination of the LSP and BGP next hop for the VPN
    traffic. The cost to L1 is configured to be less than the cost
    to La.

```
                       *****     *****
   CE1,CE2 in     [R2]-----[R3]-----[L1]              **** Primary LSP
   one VPN       */          &\:.....:   \            &&&& Backup LSP
             */            &\         \          .... BFD Session
    [CE1]--[R1]              &\          [CE2]
                              &\     /
                               &\    /
                               [La](VPN label from L1 as a UA label)
```
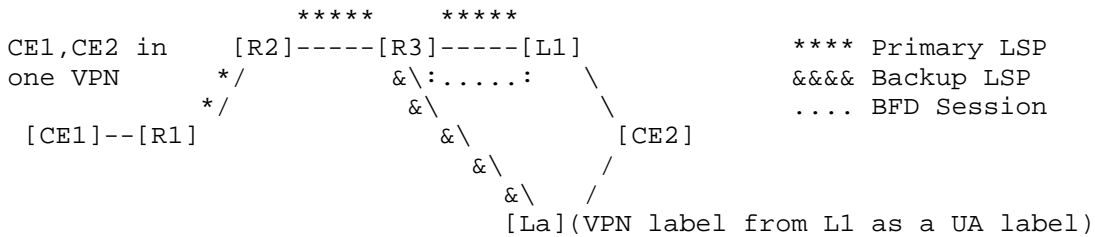
                Figure 3: Locally Protect Egress for L3VPN Traffic

In normal operations, R1 sends the VPN traffic from CE1 via the
primary LSP with the VPN label as an inner label to L1, which
delivers the traffic to CE2 using the VPN label.

When the primary egress node L1 fails, its upstream node R3 detects
it and switches the VPN traffic from the primary LSP to the backup
LSP to La, which delivers the traffic to CE2 using the backup LSP

label as a context label to get the LFIB for L1 and the VPN label as
a UA label to find the forwarding entry in the LFIB to forward the
traffic to CE2.

6.2. PLR Procedure for Applications

When the PLR gets a backup LSP from itself to a backup egress node
for protecting a primary egress node of a primary LSP, it includes an
SERO object in the Path message for the primary LSP. The object
contains the ID information of the backup LSP and indicates that the
primary egress node sends the backup egress node the application
traffic label (e.g., the VPN label) as a UA label when needed.

6.3. Egress Procedures for Applications

When a primary egress node of an LSP sends the ingress node of the
LSP a label for an application such as a VPN label, it sends the
label (as a UA label) to the backup egress node for protecting the
primary egress node. Exactly how the label is sent is out of scope
for this document.

When the backup egress node receives a UA label from the primary
egress node, it adds a forwarding entry with the label into the LFIB
for the primary egress node. When the backup egress node receives a
packet from the backup LSP, it uses the top label as a context label
to find the LFIB for the primary egress node and uses the inner label
to deliver the packet to the same destination as the primary egress
node according to the LFIB.

7. Security Considerations

This document builds upon existing work, specifically, the security
considerations of [RFC4090], [RFC4875], [RFC3209], and [RFC2205]
continue to apply. Additionally, protecting a primary egress node of
a P2P LSP carrying service traffic through a backup egress node
requires out-of-band communication between the primary egress node
and the backup egress node in order for the primary egress node to
convey a service label as a UA label and also convey its related
forwarding information to the backup egress node. It is important to
confirm that the identifiers used to identify the primary and backup
egress nodes in the LSP are verified to match with the identifiers
used in the out-of-band protocol (such as BGP).

8. IANA Considerations

   IANA maintains a registry called "Class Names, Class Numbers, and
   Class Types" under "Resource Reservation Protocol (RSVP) Parameters".
   IANA has assigned a new C-Type under the PROTECTION object class,
   Class Number 37:

      Value   Description   Definition
      -----   -----------   ----------
      3   Egress Protection   Section 4.1

   IANA has created and now maintains a registry under the PROTECTION
   object class (Class Number 37) and Egress Protection (C-Type 3).
   Initial values for the registry are given below. Future assignments
   are to be made through IETF Review [RFC8216].

      Value   Description   Definition
      -----   -----------   ----------
       0   Reserved
       1   IPv4_PRIMARY_EGRESS Section 4.1.1
       2   IPv6_PRIMARY_EGRESS Section 4.1.1
       3   IPv4_P2P_LSP_ID   Section 4.1.2
       4   IPv6_P2P_LSP_ID   Section 4.1.2
       5-127   Unassigned
       128-255   Reserved

9. References

9.1. Normative References

   [RFC2119]   Bradner, S., "Key words for use in RFCs to Indicate
               Requirement Levels", BCP 14, RFC 2119,
               DOI 10.17487/RFC2119, March 1997,
               <https://www.rfc-editor.org/info/rfc2119>.

   [RFC3209]   Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V.,
               and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP
               Tunnels", RFC 3209, DOI 10.17487/RFC3209, December 2001,
               <https://www.rfc-editor.org/info/rfc3209>.

   [RFC4090]   Pan, P., Ed., Swallow, G., Ed., and A. Atlas, Ed., "Fast
               Reroute Extensions to RSVP-TE for LSP Tunnels", RFC 4090,
               DOI 10.17487/RFC4090, May 2005,
               <https://www.rfc-editor.org/info/rfc4090>.

   [RFC4873]   Berger, L., Bryskin, I., Papadimitriou, D., and A. Farrel,
               "GMPLS Segment Recovery", RFC 4873, DOI 10.17487/RFC4873,
               May 2007, <https://www.rfc-editor.org/info/rfc4873>.

   [RFC4875] Aggarwal, R., Ed., Papadimitriou, D., Ed., and S.
              Yasukawa, Ed., "Extensions to Resource Reservation
              Protocol - Traffic Engineering (RSVP-TE) for Point-to-
              Multipoint TE Label Switched Paths (LSPs)", RFC 4875,
              DOI 10.17487/RFC4875, May 2007,
              <https://www.rfc-editor.org/info/rfc4875>.

   [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC
              2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174,
              May 2017, <https://www.rfc-editor.org/info/rfc8174>.

   [RFC8216] Pantos, R., Ed. and W. May, "HTTP Live Streaming",
              RFC 8216, DOI 10.17487/RFC8216, August 2017,
              <https://www.rfc-editor.org/info/rfc8216>.

9.2. Informative References

   [FRAMEWK] Shen, Y., Jeganathan, J., Decraene, B., Gredler, H.,
              Michel, C., Chen, H., and Y. Jiang, "MPLS Egress
              Protection Framework", Work in Progress, draft-ietf-mpls-
              egress-protection-framework-00, January 2018.

   [RFC2205] Braden, R., Ed., Zhang, L., Berson, S., Herzog, S., and S.
              Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1
              Functional Specification", RFC 2205, DOI 10.17487/RFC2205,
              September 1997, <https://www.rfc-editor.org/info/rfc2205>.

   [RFC5331] Aggarwal, R., Rekhter, Y., and E. Rosen, "MPLS Upstream
              Label Assignment and Context-Specific Label Space",
              RFC 5331, DOI 10.17487/RFC5331, August 2008,
              <https://www.rfc-editor.org/info/rfc5331>.

Contributors

   The following people contributed significantly to the content of this
   document and should be considered coauthors:

      Ning So
      Tata
      Email: ningso01@gmail.com

      Mehmet Toy
      Verizon
      Email: mehmet.toy@verizon.com

      Lei Liu
      Fujitsu
      Email: lliu@us.fujitsu.com

      Zhenbin Li
      Huawei Technologies
      Email: lizhenbin@huawei.com

   We also acknowledge the contributions of the following individuals:

      Boris Zhang
      Telus Communications
      Email: Boris.Zhang@telus.com

      Nan Meng
      Huawei Technologies
      Email: mengnan@huawei.com

      Prejeeth Kaladharan
      Huawei Technologies
      Email: prejeeth@gmail.com

      Vic Liu
      China Mobile
      Email: liu.cmri@gmail.com

Authors' Addresses

    Huaimo Chen
    Huawei Technologies
    Boston, MA
    United States of America

    Email: huaimo.chen@huawei.com

    Autumn Liu
    Ciena
    United States of America

    Email: hliu@ciena.com

    Tarek Saad
    Cisco Systems

    Email: tsaad@cisco.com

    Fengman Xu
    Verizon
    2400 N. Glenville Dr
    Richardson, TX 75082
    United States of America

    Email: fengman.xu@verizon.com

    Lu Huang
    China Mobile
    No.32 Xuanwumen West Street, Xicheng District
    Beijing 100053
    China

    Email: huanglu@chinamobile.com