Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: December 20, 2015

H. Chen
Huawei Technologies
N. So
Tata Communications
A. Liu
Ericsson
T. Saad
Cisco Systems
F. Xu
Verizon

June 18, 2015

Extensions to RSVP-TE for LSP Egress Local Protection
draft-ietf-teas-rsvp-egress-protection-02.txt

Abstract

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

Status of this Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). 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 December 20, 2015.

Copyright Notice

Copyright (c) 2015 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

RFC 4090 describes two methods for protecting the transit nodes of a P2P LSP: one-to-one and facility protection. RFC 4875 specifies how to use them to protect the transit nodes of a P2MP LSP. However, they do not mention any local protection for an egress of an LSP.

To protect the egresses of an LSP (P2P or P2MP), an existing approach sets up a backup LSP from a backup ingress (or the ingress of the LSP) to the backup egresses, where each egress is paired with a backup egress and protected by the backup egress.

This approach uses more resources and provides slow fault recovery. This document specifies extensions to RSVP-TE for local protection of an egress of an LSP, which overcomes these disadvantages.

1.1. An Example of Egress Local Protection

Figure 1 shows an example of using backup LSPs to locally protect egresses of a primary P2MP LSP from ingress R1 to two egresses: L1 and L2. The primary LSP is represented by star(*) lines and backup LSPs by hyphen(-) lines.

La and Lb are the designated backup egresses for egresses L1 and L2 respectively. To distinguish an egress (e.g., L1) from a backup egress (e.g., La), an egress is called a primary egress if needed.

The backup LSP for protecting L1 is from its upstream node R3 to backup egress La. The one for protecting L2 is from R5 to Lb.

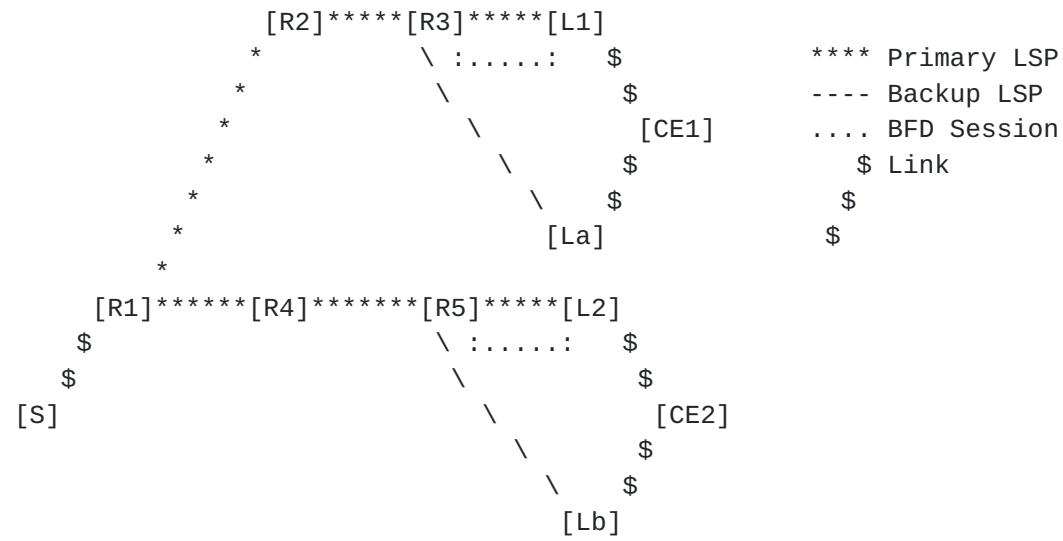


Figure 1: Backup LSP for Locally Protecting Egress

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 La, which delivers the traffic to CE1. The time for switching the traffic is within tens of milliseconds.

The failure of a primary egress (e.g., L1 in the figure) may be detected by its upstream node (e.g., R3 in the figure) through a BFD between the upstream node and the egress in MPLS networks. Exactly how the failure is detected is out of scope for this document.

1.2. Egress Local Protection with FRR

Using the egress local protection and the FRR, we can locally protect the egresses, the links and the transit nodes of an LSP. The traffic switchover time is within tens of milliseconds whenever an egress, any of the links and the transit nodes of the LSP fails.

The egress nodes of the LSP can be locally protected via the egress local protection. All the links and the transit nodes of the LSP can be locally protected through using the FRR.

The egress local protection may be generalized and used with the segment protection defined in RFC 4873. How it is generalized and used is out of scope for this document.

2. Conventions Used in This Document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119.

3. Terminology

This document uses terminologies defined in RFC 2205, RFC 3209, RFC 4090 and RFC 4875.

4. Protocol Extensions

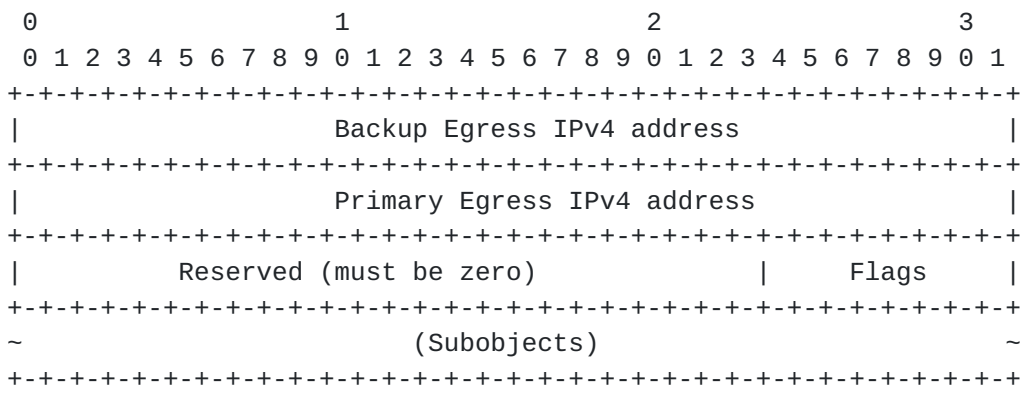
This section presents new RSVP objects.

4.1. EGRESS_BACKUP Object

A new object EGRESS_BACKUP is defined for egress local protection. It contains a backup egress for a primary egress.

4.1.1. EGRESS_BACKUP IPv4 Object

Class = EGRESS_BACKUP, EGRESS_BACKUP_IPv4 C-Type = 1



- o Backup Egress IPv4 address:
  IPv4 address of the backup egress node
- o Primary Egress IPv4 address:
  IPv4 address of the primary egress node
- o Flags
  0x01 S2L Sub LSP Backup Desired
  0x02 Other Sending UA Label

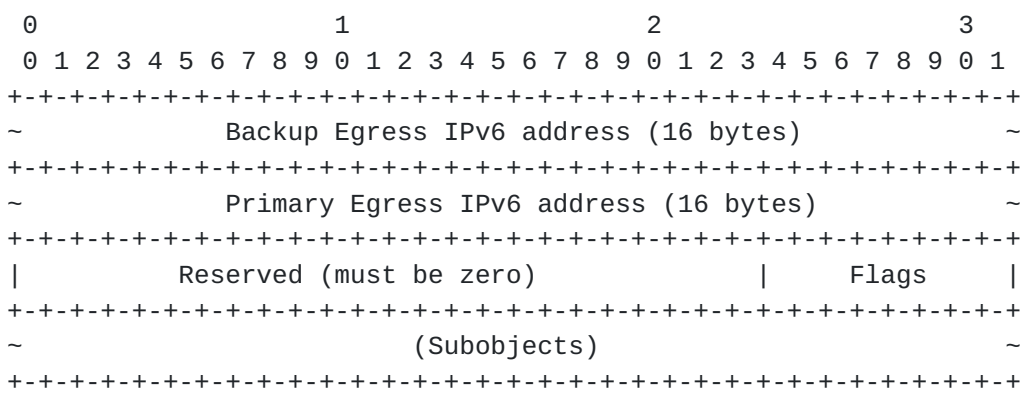
Flag "S2L Sub LSP Backup Desired" is used to indicate if S2L Sub LSP (ref to RFC 4875) is desired for protecting an egress of a P2MP LSP.

Flag "Other Sending UA Label" is used to indicate if another protocol is desired for sending a label as a UA label from a primary egress to a backup egress.

The Subobjects are TLVs. One of them is P2P LSP ID IPv4 subobject. Another is Label subobject.

4.1.2. EGRESS_BACKUP IPv6 Object

Class = EGRESS_BACKUP, EGRESS_BACKUP_IPv6 C-Type = 2

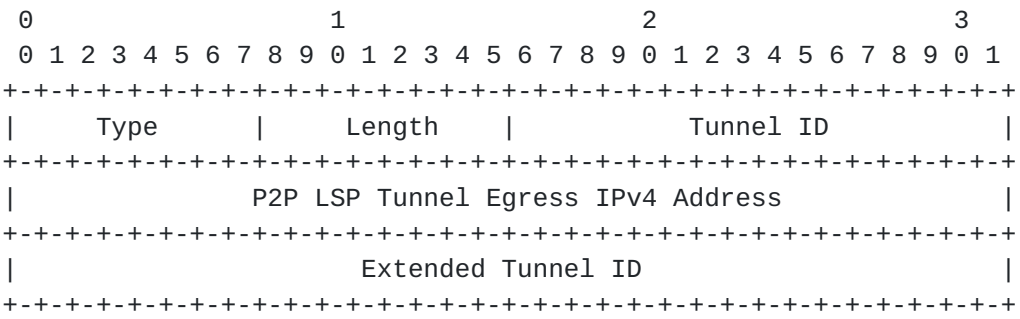


- o Backup Egress IPv6 address:
  IPv6 address of the backup egress node
- o Primary Egress IPv6 address:
  IPv6 address of the primary egress node
- o Flags
  0x01 S2L Sub LSP Backup Desired
  0x02 Other Sending UA Label

4.1.3. P2P LSP ID Subobject

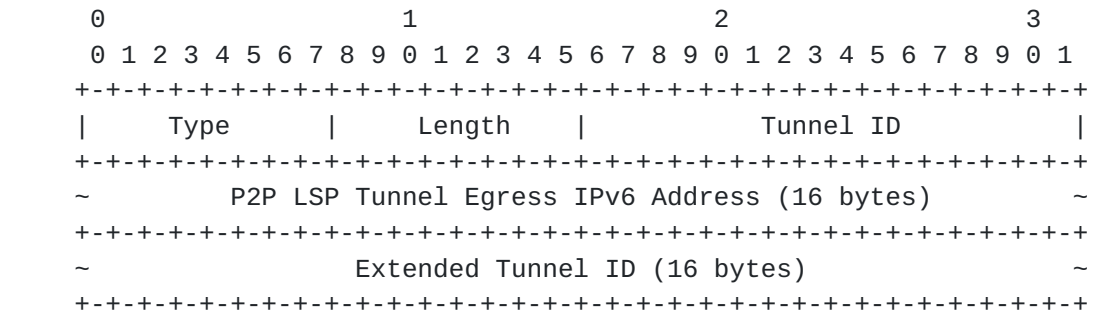
A P2P LSP ID subobject contains the information for identifying a backup LSP tunnel.

4.1.3.1. P2P LSP ID IPv4 Subobject



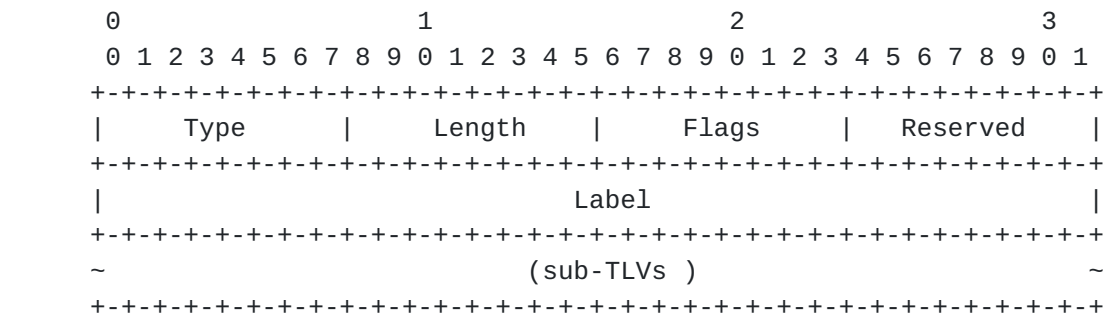
- o Type: 0x01 for P2P LSP ID IPv4
- o Length: The total length of the subobject in bytes, which is 12.
- o Tunnel ID:
  A 16-bit identifier that is constant over the life of the tunnel
- o P2P LSP Tunnel Egress IPv4 Address:
  IPv4 address of the egress of the tunnel
- o Extended Tunnel ID:
  A 4-byte identifier being constant over the life of the tunnel

4.1.3.2. P2P LSP ID IPv6 Subobject



o Type: 0x02 for P2P LSP ID IPv6
o Length: The total length of the subobject in bytes, which is 36.
o Tunnel ID:
A 16-bit identifier that is constant over the life of the tunnel
o P2P LSP Tunnel Egress IPv6 Address:
IPv6 address of the egress of the tunnel
o Extended Tunnel ID:
A 16-byte identifier being constant over the life of the tunnel

4.1.4. Label Subobject



o Type: 0x03 for Label
o Length: The total length of the subobject in bytes.
o Flags: 0x01 = Global Label
o Label: Value of Label

4.2. Path Message

A Path message is enhanced to carry the information about a backup egress for a primary egress of an LSP by including an egress backup descriptor list. The format of the message is illustrated below.

```
<Path Message> ::= <Common Header> [ <INTEGRITY> ]
                   [ [<MESSAGE_ID_ACK> | <MESSAGE_ID_NACK>] ...]
                   [ <MESSAGE_ID> ]<SESSION> <RSVP_HOP> <TIME_VALUES>
                   [ <EXPLICIT_ROUTE> ]
                   <LABEL_REQUEST> [ <PROTECTION> ] [ <LABEL_SET> ...]
                   [ <SESSION_ATTRIBUTE> ] [ <NOTIFY_REQUEST> ]
                   [ <ADMIN_STATUS> ] [ <POLICY_DATA> ... ]
                   <sender descriptor> [<S2L sub-LSP descriptor list>]
                   [<egress backup descriptor list>]
```

The egress backup descriptor list in the message is defined below. It is a sequence of EGRESS_BACKUP objects, each of which describes a pair of a primary egress and a backup egress.

```
<egress backup descriptor list> ::=
                    <egress backup descriptor>
                    [ <egress backup descriptor list> ]

<egress backup descriptor> ::= <EGRESS_BACKUP>
```

5. Egress Protection Behaviors

5.1. Ingress Behavior

To protect a primary egress of an LSP, the ingress 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 of an LSP, the ingress 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 of a P2MP LSP, the ingress MUST include an EGRESS_BACKUP object and set the "S2L Sub LSP Backup Desired" flag.

If another protocol is desired for sending a label as a upstream assigned label to a backup egress, the ingress MUST set the "Other Sending UA Label" flag.

A backup egress SHOULD be configured on the ingress of an LSP to protect a primary egress of the LSP.

The ingress MUST send a Path message for the LSP with the objects

above and an optional egress backup descriptor list for protecting egresses of the LSP. For each primary egress of the LSP to be protected, the ingress MUST add an EGRESS_BACKUP object into the list if the backup egress is given. The object MUST contain the primary egress and the backup egress for protecting the primary egress.

5.2. Transit Node and PLR Behavior

If a transit node of an LSP receives the Path message with an egress backup descriptor list and it is not an upstream node of any primary egress of the LSP, it MUST forward the list unchanged.

If the transit node is the upstream node of a primary egress to be protected, it determines the backup egress, obtains a path for the backup LSP and sets up the backup LSP along the path.

The PLR (upstream node of the primary egress) MUST extract the backup egress from the respective EGRESS_BACKUP object in the egress backup descriptor list. If no matching EGRESS_BACKUP object is found or the list is empty, the PLR applies a local policy to determine the backup egress and MUST add an EGRESS_BACKUP object with the backup egress and primary egress into a Path message to the primary egress.

After obtaining the backup egress, the PLR computes a backup path from itself to the backup egress. It excludes the primary egress to be protected when computing the path. Thus the PLR will not select any path via the primary egress.

The PLR MUST provide one-to-one backup protection for the primary egress 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 Sub-object for the primary egress in the Resv message according to the status of the primary egress and the backup LSP protecting the primary egress. For example, it sets the "local protection available" and the "node protection" flag indicating that the primary egress is protected when the backup LSP is up and ready for protecting the primary egress.

5.2.1. Signaling for One-to-One Protection

The behavior of the upstream node of a primary egress of an LSP as a PLR is the same as that of a PLR for one-to-one backup described in RFC 4090 except for that the upstream node as a PLR creates a backup LSP from itself to a backup egress.

If the LSP is a P2MP LSP and a primary egress of the LSP is also a

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

When the PLR detects the failure of the primary egress, it switches the packets from the primary LSP to the backup LSP to the backup egress. 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.

5.2.2. Signaling for Facility Protection

Except for backup LSP and downstream label, the behavior of the upstream node of the primary egress of a primary LSP as a PLR follows the PLR behavior for facility backup described in RFC 4090.

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

The PLR selects or creates a backup LSP from itself to the backup egress. 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 is created.

After getting the backup LSP, the PLR associates the backup LSP with a primary LSP for protecting its primary egress. The PLR records that the backup LSP is used to protect the primary LSP against its primary egress failure and MUST include an EGRESS_BACKUP object in the Path message to the primary egress. The object MUST contain the backup egress and the backup LSP ID. It indicates that the primary egress MUST send the backup egress the service label as UA label if there is a service carried by the LSP and the primary LSP label as UA label if the label is not implicit null.

A UA label MAY be sent via RSVP or another protocol (e.g., BGP). If "Other Sending UA Label" flag is one, the primary egress MUST send the UA labels to the backup egress through another protocol; otherwise, UA labels MUST be sent via RSVP.

After receiving the Path message with the EGRESS_BACKUP, the primary egress MUST include the information about the UA labels in the Resv message with an EGRESS_BACKUP object. When the PLR receives the Resv message with the information about the UA labels, it MUST include the information in the Path message for the backup LSP to the backup egress. Thus the UA labels are sent to the backup egress from the primary egress via RSVP.

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

5.2.3. Signaling for S2L Sub LSP Protection

The S2L Sub LSP Protection uses a S2L Sub LSP (ref to RFC 4875) as a backup LSP to protect a primary egress of a P2MP LSP. The PLR MUST determine to protect a primary egress of a P2MP LSP via S2L sub LSP protection when it receives a Path message with flag "S2L Sub LSP Backup Desired" set.

The PLR MUST set up the backup S2L sub LSP to the backup egress, create and maintain its state in the same way as of setting up a source to leaf (S2L) sub LSP defined in RFC 4875 from the signaling's point of view. It computes a path for the backup LSP from itself to the backup egress, constructs and sends a Path message along the path, 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, 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.

5.2.4. PLR Procedures during Local Repair

When the upstream node of a primary egress of an LSP as a PLR detects the failure of the primary egress, it follows the procedures defined in section 6.5 of RFC 4090. It SHOULD notify the ingress about the failure of the primary egress in the same way as a PLR notifies the ingress about the failure of a transit node.

Moreover, the PLR MUST let the upstream part of the primary LSP stay after the primary egress fails through sending Resv message to its upstream node along the primary LSP. The downstream part of the primary LSP from the PLR to the primary egress SHOULD be removed.

In the local revertive mode, the PLR SHOULD 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 SHOULD be switched back.

6. Considering Application Traffic

This section focuses on the application traffic carried by P2P LSPs. When a primary egress of a P2MP LSP fails, the application traffic carried by the P2MP LSP is delivered to the same destination by the backup egress since the inner label if any for the traffic is a upstream assigned label for every egress of the P2MP LSP.

6.1. A Typical Application

L3VPN is a typical application. An existing solution (refer to Figure 2) for protecting L3VPN traffic against egress failure includes: 1) A multi-hop BFD session between ingress R1 and egress L1 of primary LSP; 2) A backup LSP from ingress R1 to backup egress La; 3) La sends R1 VPN backup label and related information via BGP; 4) R1 has a VRF with two sets of routes: one uses primary LSP and L1 as next hop; the other uses backup LSP and La as next hop.

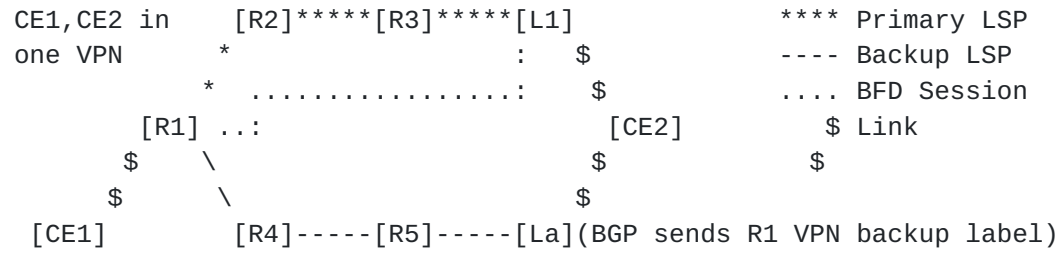


Figure 2: Protect Egress for L3VPN Traffic

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

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

A new solution (refer to Figure 3) with egress local protection for protecting L3VPN traffic includes: 1) A BFD session between R3 and egress L1 of primary LSP; 2) A backup LSP from R3 to backup egress La; 3) L1 sends La VPN label as UA label and related information; 4) L1 and La is virtualized as one. This can be achieved by configuring

a same local address on L1 and La, using the address as a destination of the LSP and BGP next hop for VPN traffic.

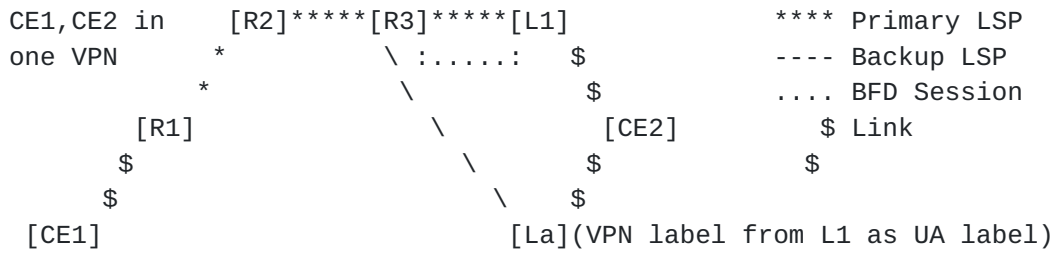


Figure 3: Locally Protect Egress for L3VPN Traffic

When R3 detects L1's failure, R3 sends the traffic from primary LSP via backup LSP to La, which delivers the traffic to CE2 using VPN label as UA label under the backup LSP label as a context label.

6.2. PLR Procedure for Applications

When the PLR gets a backup LSP from itself to a backup egress for protecting a primary egress of a primary LSP, it includes an EGRESS_BACKUP 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 sends the backup egress the application traffic label (e.g., VPN label) as UA label when needed.

6.3. Egress Procedures for Applications

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

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

7. Security Considerations

In principle this document does not introduce new security issues. The security considerations pertaining to RFC 4090, RFC 4875 and other RSVP protocols remain relevant.

Note that protecting a primary egress of a P2P LSP carrying service traffic through a backup egress requires that the backup egress trust the primary egress for the information received for a service label as UA label.

8. IANA Considerations

IANA is requested to administer the assignment of new values defined in this document and summarized in this section.

8.1. A New Class Number

IANA maintains a registry called "Class Names, Class Numbers, and Class Types" under "Resource Reservation Protocol-Traffic Engineering (RSVP-TE) Parameters". IANA is requested to assign a new Class Number for new object EGRESS_BACKUP as follows:

Class Names	Class Numbers	Class Types
EGRESS_BACKUP	TBD1 (>192)	1: EGRESS_BACKUP_IPv4
		2: EGRESS_BACKUP_IPv6

IANA is requested to assign Types for new TLVs in the new objects as follows:

Type	Name	Allowed in
1	P2P_LSP_ID_IPv4 TLV	EGRESS_BACKUP_IPv4
2	P2P_LSP_ID_IPv6 TLV	EGRESS_BACKUP_IPv6
3	Label TLV	EGRESS_BACKUP

9. Co-authors

Lu Huang, Mehmet Toy, Lei Liu, Zhenbin Li

10. Contributors

Boris Zhang
Telus Communications
200 Consilium Pl Floor 15
Toronto, ON M1H 3J3
Canada
Email: Boris.Zhang@telus.com

Nan Meng
Huawei Technologies
Huawei Bld., No.156 Beiqing Rd.
Beijing 100095
China
Email: mengnan@huawei.com

Vic Liu
China Mobile
No.32 Xuanwumen West Street, Xicheng District
Beijing, 100053
China
Email: liuzhiheng@chinamobile.com

11. Acknowledgement

The authors would like to thank Richard Li, Nobo Akiya, Lou Berger, Jeffrey Zhang, Lizhong Jin, Ravi Torvi, Eric Gray, Olufemi Komolafe, Michael Yue, Daniel King, Rob Rennison, Neil Harrison, Kannan Sampath, Yimin Shen, Ronhazli Adam and Quintin Zhao for their valuable comments and suggestions on this draft.

12. References

12.1. Normative References

[RFC2205] Braden, B., Zhang, L., Berson, S., Herzog, S., and S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", RFC 2205, September 1997.

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.

[RFC4090] Pan, P., Swallow, G., and A. Atlas, "Fast Reroute Extensions to RSVP-TE for LSP Tunnels", RFC 4090, May 2005.

[RFC4875] Aggarwal, R., Papadimitriou, D., and S. Yasukawa, "Extensions to Resource Reservation Protocol - Traffic Engineering (RSVP-TE) for Point-to-Multipoint TE Label Switched Paths (LSPs)", RFC 4875, May 2007.

[RFC5331] Aggarwal, R., Rekhter, Y., and E. Rosen, "MPLS Upstream Label Assignment and Context-Specific Label Space",

RFC 5331, August 2008.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

12.2. Informative References

[RFC4873] Berger, L., Bryskin, I., Papadimitriou, D., and A. Farrel, "GMPLS Segment Recovery", RFC 4873, May 2007.

Authors' Addresses

Huaimo Chen
Huawei Technologies
Boston, MA
USA

Email: huaimo.chen@huawei.com

Ning So
Tata Communications
2613 Fairbourne Cir.
Plano, TX 75082
USA

Email: ningso01@gmail.com

Autumn Liu
Ericsson
CA
USA

Email: autumn.liu@ericsson.com

Tarek Saad
Cisco Systems

Email: tsaad@cisco.com

Fengman Xu
Verizon
2400 N. Glenville Dr
Richardson, TX 75082
USA

Email: fengman.xu@verizon.com