MPLS Working Group
Internet-Draft
Intended status: Standards Track
Expires: July 7, 2021

C. Ramachandran
V. Beeram
Juniper Networks
H. Sitaraman
Individual

January 3, 2021

Node Protection for RSVP-TE tunnels on a shared MPLS forwarding plane
draft-chandra-mpls-rsvp-shared-labels-np-05

Abstract

Segment Routed RSVP-TE tunnels provide the ability to use a shared MPLS forwarding plane at every hop of the Label Switched Path (LSP). The shared forwarding plane is realized with the use of 'Traffic Engineering (TE) link labels' that get shared by LSPs traversing these TE links. This paradigm helps significantly reduce the forwarding plane state required to support a large number of LSPs on a Label Switching Router (LSR). These tunnels require the ingress Label Edge Router (LER) to impose a stack of labels. If the ingress LER cannot impose the full label stack, it can use the assistance of one or more delegation hops along the path of the LSP to impose parts of the label stack.

The procedures for a Point of Local Repair (PLR) to provide local protection against link failures using facility backup for Segment Routed RSVP-TE tunnels are well defined and do not require specific protocol extensions. This document defines the procedures for a PLR to provide local protection against transit node failures using facility backup for these tunnels. The procedures defined in this document include protection against delegation hop failures.

Requirements Language

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.

Status of This Memo

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

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

working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on July 7, 2021.

Table of Contents

1. Introduction

With the advent of Traffic Engineering (TE) link labels and Segment Routed RSVP-TE Tunnels [RFC8577], a shared MPLS forwarding plane can be realized by allowing the TE link label to be shared by MPLS RSVP-TE Label Switched Paths (LSPs) traversing the link. The shared forwarding plane behavior helps reduce the amount of forwarding plane state required to support a large number of LSPs on a Label Switching Router (LSR).

Segment Routed RSVP-TE tunnels request the use of a shared forwarding plane at every hop of the LSP. The TE link label used at each hop is recorded in the Record Route object (RRO) of the Resv message. The ingress Label Edge Router (LER) uses this recorded information to construct a stack of labels that can be imposed on the packets steered on to the tunnel. In the scenario where the ingress LER cannot impose the full label stack, it can use the assistance of one or more delegation hops along the path of the LSP to impose parts of the label stack.

Facility backup is a local repair method [RFC4090] in which a bypass tunnel is used to provide protection against link or node failures for MPLS RSVP-TE LSPs at the Point of Local Repair (PLR). The facility backup procedures that provide protection against link failures for Segment Routed RSVP-TE LSPs are defined in [RFC8577]. This document defines the facility backup procedures that provide protection against node failures for these LSPs. These procedures include protection against delegation hop failures. The document also discusses the procedures for handling backwards compatibility scenarios where a node along the path of the LSP does not support the procedures defined in this document.

The procedures discussed in this document do not cover protection against ingress/egress node failures. They also do not apply to Point to Multipoint (P2MP) RSVP-TE Tunnels.

2. Terminology

The reader is expected to be familiar with the terminology specified in [RFC3209], [RFC4090] and [RFC8577]. Unless otherwise stated, the term LSPs in this document refer to Segment Routed RSVP-TE LSPs. The following additional terms are used in this document:

Primary forwarding action: The outbound label forwarding action performed at a PLR for a protected LSP before the occurrence of local failure.

Backup forwarding action: The outbound label forwarding action performed at a PLR for a protected LSP after the occurrence of local failure.

3. Node Protection Specific Procedures

A set of Segment Routed RSVP-TE LSPs can share a TE link label on an LSR only if all the LSPs in the set share the same outbound label forwarding action. For protected LSPs, having the same outbound label forwarding action means having the same primary forwarding action and the same backup forwarding action. In the case of LSPs that do not request local protection or LSPs that request only link protection, they can use the same outbound label forwarding action if they reach a common next-hop LSR via a common outgoing TE link. However, in the case of LSPs that request node protection, they can use the same outbound label forwarding action only if they reach a common next-next-hop LSR via a common outgoing TE link and a common next-hop LSR.

3.1. Applicability of this Document

The label allocation and signaling procedures defined in [RFC8577] can sufficiently cater to the following scenarios on an LSR:

(a) Offer no protection to LSPs that do not request local protection

(b) Offer no protection or link protection to LSPs that request link protection

(c) Offer no protection or link protection to LSPs that request node protection

The label allocation and signaling procedures defined in this document are meant to enable LSRs to offer node protection to LSPs that request node protection.

3.2. PLR Procedures for Protecting Next-Hop Non-Delegation LSR

If the protected next-hop LSR signals a TE link label for the LSP but does not set the Delegation Label flag in the RRO Label Subobject carried in Resv message, then the PLR SHOULD allocate multiple shared labels for the same TE link such that a unique label is allocated for every unique next-next-hop LSR that is reachable via the protected next-hop LSR.

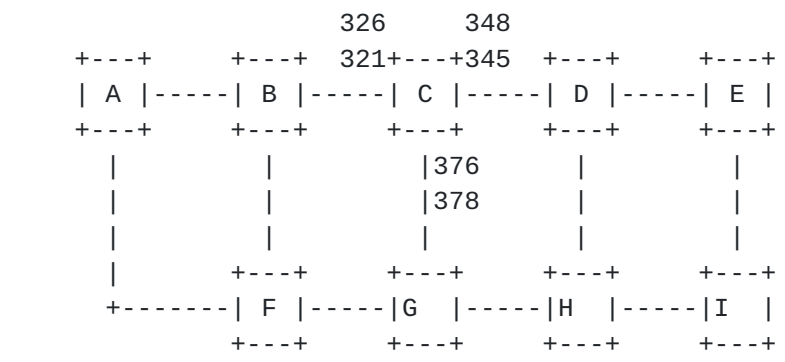


Figure 1: Per-nhop-nnhop label allocation

In the example shown in Figure 1, LSR C has allocated the following TE link labels:

321 for the TE link C-B to reach the next-next-hop LSR A
326 for the TE link C-B to reach the next-next-hop LSR F
345 for the TE link C-D to reach the next-next-hop LSR E
348 for the TE link C-D to reach the next-next-hop LSR H
376 for the TE link C-G to reach the next-next-hop LSR F
378 for the TE link C-G to reach the next-next-hop LSR H

If a LSP requesting node protection transits PLR C and if the protected next-hop LSR after C along the LSP path is not a delegation hop, then LSR C signals the respective TE link label depending on the next-next-hop LSR on the LSP path.

LSP path: A -> B -> C -> D -> E : Label = 345
LSP path: A -> B -> C -> D -> H : Label = 348
LSP path: A -> B -> C -> G -> H : Label = 378

In all LSP paths above, at PLR C, the protected next-hop LSRs D and G along the LSP paths signal TE link labels but are not delegation hops.

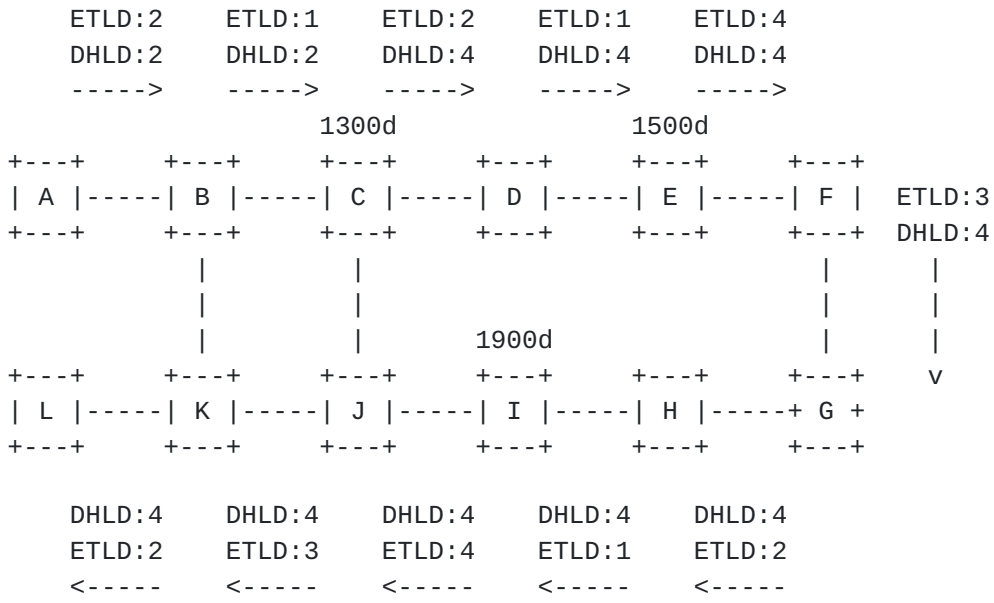
If the primary TE link is operational, LSR C will pop the TE link label and forward the packet to the corresponding next-hop LSR over that TE link. During local repair, LSR C will pop the TE link label and also the label beneath the top label, and forward the packet over the node protecting bypass tunnel to the appropriate next-next-hop LSR, which is the Merge Point (MP).

3.3. PLR Procedures for Protecting Next-Hop Delegation LSR

The outgoing backup label forwarding action corresponding to a label shared by LSPs requesting node protection MUST bypass the protected next-hop LSR. The PLR MUST push the label stack on behalf of the

next-hop delegation LSR. Hence, the number of labels that a delegation hop chooses to push also depends on the number of labels that the upstream hop (acting as PLR) along the primary LSP can push. This section extends the Effective Transport Label-Stack Depth (ETLD) signaling procedure specified in [RFC8577] for LSPs requesting node protection.

Considering Figure 2, assume LER A and LSR B can push a maximum of 3 labels on an MPLS packet while the remaining nodes can push a maximum of 5 labels. LER A originates a Path message with an ETLD of 2 after reserving space for the bypass tunnel label that should be pushed for backup forwarding action. In addition to setting the ETLD, LER A also sets the Delegation Helper Label Depth (DHLD) to 2 in the Path message. The DHLD is computed as the maximum number of labels that the node can push after reserving space for the NNHOP bypass tunnel label that should be pushed for backup forwarding action. The ETLD procedures dictate that each LSR add its own ETLD value before sending the Path message downstream. LSRs C, E and I are automatically selected as delegation hops by the time the Path message reaches the egress LER L. LSR C uses the DHLD signaled by the upstream LSR B as input when calculating the outgoing ETLD in the Path message.



Notation : <Label>d - delegation label

Figure 2: ETLD and DHLD signaling for node protection

As shown in Figure 2, delegation hop LSR C does not set outgoing ETLD to 4 that it would have normally set given that LSR C can push a maximum of 5 labels on an outgoing packet. Instead, LSR C sets the outgoing ETLD to the minimum of the ETLD that it computes and the DHLD value of its previous hop i.e. minimum(computed ETLD = 4, previous hop DHLD = 2).

The extension for signaling the DHLD in the Path message is defined in Section 4.1.

3.3.1. Label Allocation and Stacking

An LSR that decides to become a delegation hop for one or more LSPs requesting node protection MUST allocate a delegation label separate from delegation label assigned for LSPs that are offered no protection or link protection - even though the delegation segments share the same hops. In the example shown in Figure 2, the delegation hops LSRs C, E and I will set the Delegation Label flag in the Label sub-object that they add to the Resv message.

A PLR node that offers node protection to a delegation hop SHOULD be capable of helping the downstream delegation when the primary TE link to the delegation hop goes down. In the example shown in Figure 1, the LSRs B, D and H act as helpers for their respective downstream delegation hops. The PLR nodes that are delegation helpers along the path of LSPs requesting node protection SHOULD allocate a unique label for every delegation label signaled by the protected delegation node.

Before primary TE link failure, the PLR playing the role of a delegation helper pops the incoming label and forwards the packet on the primary TE link. During local repair, the delegation helper PLR pops the incoming label and also the label beneath it and pushes the label stack on behalf of the next hop delegation LSR and forwards the packet over the bypass tunnel.

Any LSR that creates label stack upstream of the delegation helper MUST include the label signaled by the delegation helper onto the outgoing label stack just as it uses the TE link label to construct outgoing label stack.

3.4. Backwards Compatibility

3.4.1. LSR does not Support Node Protection for Shared Labels

As defined in Section 3.1, any LSR along the path of an LSP requesting node protection may choose to instead offer no protection or link protection. Hence, it must be possible to build an LSP where

some LSRs along the path support the node protection extensions defined in this document whereas the rest of the LSRs support only [RFC8577]. Any PLR along the LSP path that does not support the procedures defined in this document MUST provide either no protection or link protection for the LSPs requesting node protection.

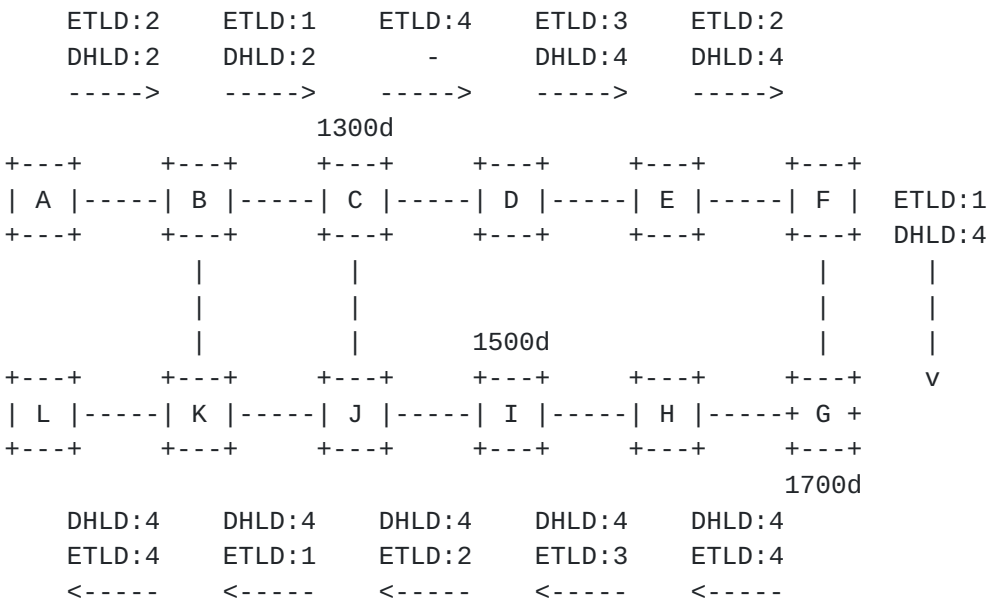


Figure 3: Delegation node does not support signaling DHLD

In Figure 3, assume LER A and LSR B can push a maximum of 3 labels to the MPLS packet while the remaining nodes can push a maximum of 5 labels. Also assume that LSR C supports the extensions defined in [RFC8577] but does not support the extensions defined in this document. Based on ETLD signaling procedure, LSR C will become a delegation hop. However, as LSR C cannot understand the DHLD signaled by the previous hop LSR B, LSR C will set outgoing ETLD to 4. If LSR C had supported the DHLD signaling, it would have set outgoing ETLD to 2 (see Section 3.3). When PLR B receives shared label from the protected next-hop LSR C in the Resv message, it must determine the number of labels it has to push in order to offer node protection from the RRO sub-object carried in Resv. As the label stack depth of the delegation hop C is greater than the number of labels LSR C can push, it must either not provide local protection or provide only link protection for the LSP.

3.4.2. Protected Hop does not Support Shared Labels

If the ingress LER has requested label stacking to reach delegation hop for the LSP requesting node protection, and if the next-hop LSR allocates a regular label for the LSP, then the LSR MUST also allocate a regular label for the LSP.

If the ingress LER has requested label stacking to reach the egress LER for the LSP requesting node protection, and if the next-hop LSR has allocated a regular label for the LSP, then the PLR MUST become a delegation hop and set the RRO Label Subobject delegation label flag in the RRO carried in Resv message. The PLR MUST set ETLD to 1 in its outgoing Path message.

3.4.3. PLR does not Support Shared Labels

If an LSR determines that its immediate upstream LSR (PLR) has not included an ETLD in the incoming Path message, then the LSR MUST become a delegation hop and set the ETLD to 1 in the outgoing Path message. The outgoing ETLD is set to 1 because the upstream LSR does not support shared labels and cannot push the label stack on behalf of this LSR.

4. Protocol Extensions

This section discusses the protocol extension required to support the procedures in Section 3.3

4.1. DHLD Encoding in ETLD Attributes TLV

Delegation Helper Label Depth (DHLD) is defined as the number of labels that an LSR has the capability to push while performing local repair protecting the next-hop delegation LSR. This document updates the ETLD Attributes TLV defined in [RFC8577]. The encoding of DHLD in the ETLD Attributes TLV is shown in Figure 4

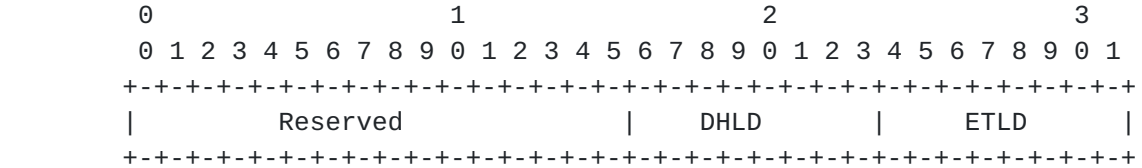


Figure 4: The ETLD Attributes TLV

The presence of ETLD Attributes TLV in the HOP_ATTRIBUTES sub-object [RFC7570] of the RRO object carried in Path message indicates that

the hop identified by the preceding IPv4 or IPv6 or Unnumbered Interface ID sub-object supports automatic delegation [RFC8577].

An implementation that supports this document MUST set the 8 bits from bit number 16 to bit number 23 with its DHLD value as indicated in Figure 4 when signaling Path message for an LSP for which node protection has been requested.

When processing the ETLD Attributes TLV of the previous hop LSR in the received Path message, the LSR checks whether it has to be the delegation hop based on the ETLD algorithm defined in [RFC8577].

If the LSR does not become a delegation hop along the LSP path, then no further action is required based on the DHLD value set by the previous hop.

If the LSR does become a delegation hop along the LSP path, then it MUST decode the 8 bit unsigned value from bit number 16 to bit number 23 as indicated in Figure 4. If the 8 bit value is zero, then the LSR MUST infer that the previous hop has not included DHLD in the ETLD Attributes TLV. If the 8 bit value is non-zero, then the LSR MUST consider that value as the DHLD value signaled by the previous hop LSR and use that DHLD value for computing its own outgoing ETLD.

5. Acknowledgements

The authors would like to thank Raveendra Torvi for his input from discussions.

6. IANA Considerations

This document includes no requests to IANA.

7. Security Considerations

This document does not introduce new security issues. The security considerations pertaining to the original RSVP protocol [RFC2205] and RSVP-TE [RFC3209] and those that are described in [RFC5920] remain relevant.

8. References

8.1. 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>.

[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>.

[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>.

[RFC7570] Margaria, C., Ed., Martinelli, G., Balls, S., and B. Wright, "Label Switched Path (LSP) Attribute in the Explicit Route Object (ERO)", RFC 7570, DOI 10.17487/RFC7570, July 2015, <https://www.rfc-editor.org/info/rfc7570>.

[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>.

[RFC8577] Sitaraman, H., Beeram, V., Parikh, T., and T. Saad, "Signaling RSVP-TE Tunnels on a Shared MPLS Forwarding Plane", RFC 8577, DOI 10.17487/RFC8577, April 2019, <https://www.rfc-editor.org/info/rfc8577>.

8.2. Informative References

[RFC5920] Fang, L., Ed., "Security Framework for MPLS and GMPLS Networks", RFC 5920, DOI 10.17487/RFC5920, July 2010, <https://www.rfc-editor.org/info/rfc5920>.

Authors' Addresses

Chandra Ramachandran
Juniper Networks

Email: csekar@juniper.net

Vishnu Pavan Beeram
Juniper Networks

Email: vbeeram@juniper.net

Harish Sitaraman
Individual

Email: harish.ietf@gmail.com