Internet Engineering Task Force
Internet-Draft
Intended status: Standards Track
Expires: December 30, 2013

Yimin. Shen
Juniper Networks
Yuji. Kamite
NTT Communications Corporation
Eric. Osborne
Cisco Systems
June 28, 2013

RSVP Setup Protection
draft-shen-mpls-rsvp-setup-protection-03

Abstract

RFC 4090 specifies an RSVP facility-backup fast reroute mechanism for protecting LSPs against link and node failures. This document extends the mechanism to provide so-called "setup protection" for LSPs during their initial Path message signaling time. In particular, it enables a router to reroute an LSP via an existing bypass LSP, when there is a failure of the immediate downstream link or node along the desired path. Therefore, it can be used to avoid LSP signaling failure and reduce setup time in such kind of situation, and allow an LSP to be established temporarily over a bypass LSP when an alternative path can only be resolved at a much later time.

Status of This Memo

Table of Contents

1. Introduction

In RSVP facility-backup fast reroute (FRR) [RFC 4090], the router at a point of local repair (PLR) of an LSP can redirect traffic via a bypass LSP upon a failure of the immediate downstream link or node. Such protection is normally established after the LSP has been set up. This is because the PLR must know the label and address of the next-hop router (in link protection) or those of the next-next-hop router (in node protection), before it can select or create a bypass LSP to protect the LSP. The information of the label and the address is carried in the Resv message of the LSP.

Imagine a scenario where a new LSP is being signaled, and its Path message carries an EXPLICIT_ROUTE object (ERO) with a strict path that is statically configured or computed offline based on a topology that assumes no failure of the network. If a link or node along the path happens to be in a failure condition, RSVP signaling will stop at the router upstream adjacent to the failure, as the next hop in the strict path no longer matches the current network topology. This will be the case even if there is an existing bypass LSP protecting the link or node for some existing LSPs. In other words, this new LSP is not protected during the setup time, i.e. the initial Path message signaling.

In this situation, the network would normally rely on IGP to update traffic engineering (TE) information throughout the network, and the router upstream adjacent to the failure to send a PathErr message to trigger the ingress router to compute and signal a new path. However, this approach may not always be possible or desirable in the following scenarios:

1. Static strict path. As described above, if the ERO carries an explicit path with a sequence of strict hops that are statically configured or computed offline based on a topology assuming no network failure, the LSP will not be established until the path is modified. This is a typical case where CSPF calculation is disabled at the LSP's ingress router due to the operational preference of service provider.

2. LSPs with a strict requirement for setup time. IGP TE information flooding, PathErr message propagation and path re-computation and re-signaling may introduce a significant delay to LSP establishment. This may impact on LSP setup time, and even become unacceptable for LSPs that have a strict requirement for it, such as on-demand transport LSPs for real-time data or TV broadcast. For these LSPs, a guaranteed establishment and setup time are considered as more important than path optimality.

3. Sibling P2MP sub-LSPs sharing a common link. In this case, the new LSP is a sub-LSP of a P2MP LSP, and its desired path is supposed to share the failed link with an existing sibling sub-LSP, i.e. another sub-LSP of the same P2MP LSP, which is being protected by a bypass LSP. If the new sub-LSP is rerouted via a different path, it will not be able to share the data flow over the bypass LSP with that sibling sub-LSP, creating unnecessary traffic flow in the network.

For networks where a failure, delay or resignaling during LSP setup is not desirable, this document extends the RSVP facility-backup fast reroute mechanism to provide a graceful solution, called "setup

protection". During the initial Path message signaling of an LSP, if there is a link or node failure along the desired path, and if there is a bypass LSP protecting the link or node, the LSP can be signaled through the bypass LSP without a delay. The LSP will be established as if it were originally set up along the desired path (i.e. primary path) and then failed over to the bypass LSP after the failure. Meanwhile, actions may be taken to resolve the failure or resignal the LSP via an alternative path, by following procedures or timing appropriate to the service provider. The setup protection is applicable to both P2P LSPs and P2MP LSPs, when such kind of temporary rerouting is not considered as a violation of desired path, as in the case of the normal fast reroute. It may be enabled by policy on a per LSP basis.

2. Specification of Requirements

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. Theory of Operation

When an LSP is being signaled by RSVP, a Path message is sent hop by hop from the ingress router to the egress router, following the path defined by an ERO. The setup protection mechanism in this document enables a router to reroute the LSP via a bypass LSP, if the router detects a failure of the immediate downstream link or node represented by the next hop in the ERO, called "next ERO hop". In this case, the current router is referred to as a PLR.

The mechanism is only relevant when the Path message carries the "local protection desired" flag in the SESSION_ATTRIBUTE object [RFC 4090] and a new "setup protection desired" flag defined in this document (Section 3.1). That is, setup protection is explicitly requested for the LSP.

In link protection, the mechanism is only applicable when the next ERO hop received by a PLR is a strict hop. In node protection, the mechanism is only applicable when both the next and the next-next ERO hops received by the PLR are strict hops. Otherwise, setup protection would be unnecessary, as the router may perform a loose hop expansion to reroute the LSP via any alternative path around the downstream failure. The strict ERO hops ensure that the PLR can unambiguously decide the intended downstream link or node and reliably detect its status. In link protection, the strict next ERO hop also indicates the merge point (MP), i.e. the destination of the bypass LSP to be used to reroute the LSP. In node protection, the strict next-next ERO hop indicates the MP.

When performing setup protection, the PLR signals a backup LSP by tunneling Path message through the bypass LSP. Like the Path message of a backup LSP in the normal facility-backup FRR ([RFC 4090]), this Path message carries an address of the PLR as the sender address in SENDER_TEMPLATE object. In addition, the Path message also carries the information of the protected LSP (Section 3.2). When the MP receives the Path message, it terminates the backup LSP, and re-creates the protected LSP. If the MP is the egress router of the protected LSP, it terminates the protected LSP as well. If the MP is a transit router of the protected LSP, it signals the LSP further downstream.

Eventually, the LSP will be established end to end, with the backup LSP tunneled through the bypass LSP from the PLR to the MP. The RSVP state on the PLR and the MP and the RSVP messages generated by these routers are no different than those in a post-failure situation of a normal facility-backup FRR.

Later, when the failure is resolved, the PLR MAY revert the LSP to the primary path, in the same manner as the local revertive mode specified in [RFC 4090].

The setup protection MAY be enabled and disabled on a router based on configuration. For an LSP to be setup-protected, the mode MUST be enabled on both PLR and MP. If it is enabled on the PLR but disabled on the MP, the MP SHOULD reject the Path message of the backup LSP and send a PathErr message, as described Section 3.4.

3.1. New RSVP Attribute Flag

In order for an LSP to explicitly request setup protection, this document defines a new "setup protection desired" flag for the Attribute Flags TLV of the LSP_ATTRIBUTES object [RFC5420]. The flag is set by the ingress router in the Path message of the LSP, i.e. the protected LSP. It MUST be supported by all routers that intend to serve as PLRs for setup protection.

3.2. New RSVP Attributes TLVs

This document defines the following two new RSVP Attributes TLVs [RFC 5420]. They are used by a PLR to convey to an MP the original sender address in SENDER_TEMPLATE object of the Path message of a protected LSP.

o Protected LSP Sender IPv4 Address TLV

o Protected LSP Sender IPv6 Address TLV

One of the TLVs SHOULD be carried by the LSP_REQUIRED_ATTRIBUTES object of the Path message of the backup LSP that the PLR sends to the MP. The information is used by the MP to build Path message for the protected LSP. The MP SHOULD NOT propagate the TLV downstream via that Path message.

3.2.1. Protected LSP Sender IPv4 Address TLV

The Protected LSP Sender IPv4 Address TLV is defined with type TBD. It is allowed in LSP_REQUIRED_ATTRIBUTES object, and not allowed in LSP_ATTRIBUTES object. The encoding is as 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
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |             Type (TBD)        |           Length (8)          |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                              Value                            |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1

Type

TBD

Length

8

Value

Original sender address in the IPv4 SENDER_TEMPLATE object of the protected LSP.

3.2.2. Protected LSP Sender IPv6 Address TLV

The Protected LSP Sender IPv6 Address TLV is defined with type TBD. It is allowed in LSP_REQUIRED_ATTRIBUTES object, and not allowed in LSP_ATTRIBUTES object. The encoding is as 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
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |             Type (TBD)        |           Length (20)         |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                                                               |
 //                             Value                           //
```

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

Figure 2

Type

TBD

Length

20

Value

Original sender address in the IPv6 SENDER_TEMPLATE object of the protected LSP.

3.3. PLR behavior

When a router has a Path message to send out, if the Path message carries the "local protection desired" flag in the SESSION_ATTRIBUTE object and the "setup protection desired" flag in the LSP_ATTRIBUTES object, and if the next ERO hop is a strict IPv4 or IPv6 prefix, the router SHOULD validate the reachability of the prefix against routing tables, traffic engineering (TE) database, or a database that reflects the current status of the network topology. If the prefix is reachable and is one hop away from the current router, the router should send out the Path message as it is. Otherwise, there is a possibility that the link or node corresponding to the prefix has failed.

The router SHOULD further search for an existing bypass LSP that is protecting the prefix. If the protected LSP desires link protection, the destination of the bypass LSP (i.e. MP) must be the router that owns the prefix. If the LSP desires node protection and the next-next ERO hop of the LSP is a strict prefix, the MP must be the router that owns this prefix.

If a bypass LSP is not found by the above criteria, the router MUST originate a PathErr with code = 24 (routing problem) and sub-code = 2 (bad strict node).

If a bypass LSP is found, the router MUST act as a PLR for setup protection, and reroute the protected LSP via the bypass LSP. If multiple satisfactory bypass LSPs exist, the PLR MAY select one based on bandwidth constraints or local policies. Specifically, if the protected LSP is a sub-LSP of a P2MP LSP, a bypass LSP that is

protecting an existing sibling sub-LSP MUST be preferred, to minimize traffic duplication in the network.

The PLR SHOULD NOT send the Path message of the protected LSP any further. Instead, it MUST create a backup LSP, and send a Path message of the backup LSP to the MP via the bypass LSP. The Path message is constructed by using the sender template specific method [RFC 4090]. In particular, it has the sender address in the SENDER_TEMPLATE object set to an address of the PLR. It MUST carry an LSP_REQUIRED_ATTRIBUTES object with a Protected LSP Sender IPv4 Address TLV or Protected LSP Sender IPv6 Address TLV.

Upon receiving a Resv message of the backup LSP from the MP, the PLR SHOULD bring up both of the backup LSP and the protected LSP. If the PLR is the ingress router of the protected LSP, the LSP has been set up successfully. If the PLR is a transit router, it MUST send a Resv message upstream for the protected LSP, with the "local protection available" and "local protection in use" set to 1, and if applicable, the "node protection" and "bandwidth protection" flags set to 1, in the RRO hop corresponding to the PLR. The PLR SHOULD also originate a PathErr message with code = 25 (notify error) and sub-code = 3 (tunnel locally repaired), as if the LSP had just fell over to the bypass LSP.

The PLR SHOULD also install a forwarding entry for the protected LSP. In the typical case, the forwarding entry should result in two outgoing labels for packets. The inner label is the backup LSP's label, and the outer label is the bypass LSP's label. However, the forwarding entry may result in one or no label, if either or both of the backup LSP and the bypass LSP have the Implicit NULL label.

If the PLR receives a PathErr message when signaling the backup LSP, the PLR MUST NOT bring up the backup LSP or the protected LSP. If the PLR is a transit router of the protected LSP, it MUST propagate the PathErr message upstream for the protected LSP. Likewise, if the PLR receives a PathErr message of the backup LSP after the backup LSP and the primary LSP have previously been brought up, and the PLR is a transit router of the protected LSP, it SHOULD also propagate the PathErr message upstream for the protected LSP.

When the PLR receives a ResvTear message of the backup LSP, the PLR MUST bring down both the backup LSP and the protected LSP. If the PLR is a transit router of the protected LSP, it MUST send a ResvTear message upstream for the protected LSP.

In any cases where the PLR needs to bring down the protected LSP due to a received PathTear message, an RSVP state time-out, a configuration change, an administrative command, etc, the PLR MUST

also bring down the backup LSP by sending a PathTear message through the bypass LSP.

3.4. MP behavior

When an MP receives the Path message of a backup LSP, it MUST realize the setup protection situation based on the presence of Protected LSP Sender IPv4 Address TLV or Protected LSP Sender IPv6 Address TLV in LSP_REQUIRED_ATTRIBUTES object.

If setup protection mode is disabled on the MP, it MUST reject the Path message, by sending a PathErr with code = 2 (policy control failure) to the PLR.

Otherwise, the MP MUST terminate the backup LSP and re-create the protected LSP. If the MP is the egress router of the protected LSP, it MUST also terminate the protected LSP. If the MP is a transit router of the LSP, it MUST send a Path message downstream for the protected LSP. The Path message has the sender address in SENDER_TEMPLATE object set to the original address of the ingress router, based on the above received Protected LSP Sender IPv4 Address TLV or Protected LSP Sender IPv6 Address TLV. The Path message MUST NOT carry any Protected LSP Sender IPv4 Address TLV or Protected LSP Sender IPv6 Address TLV in LSP_REQUIRED_ATTRIBUTES object.

The MP MUST allocate a label for the backup LSP, and distribute it to the PLR via Resv message of the backup LSP. If the protected LSP is a sub-LSP of a P2MP LSP and there is an existing sibling sub-LSP whose backup LSP is tunneled through the same bypass LSP, the MP MUST allocate the same label as the sibling sub-LSP, in order to avoid traffic duplication at the PLR.

When the MP receives a PathTear message for the backup LSP, it MUST bring down both the backup LSP and the protected LSP. If the MP is a transit router of the protected LSP, it MUST send a PathTear message downstream for the protected LSP.

In any cases where the MP receives or originates a PathErr or ResvTear message for the protected LSP, the MP MUST send the same type of message to the PLR for the backup LSP.

3.5. Local Revertive Mode

When the failed link or node is restored, the PLR MAY revert the protected LSP to its desired primary path, by following the procedure of local revertive mode described in [RFC 4090].

4. IANA Considerations

This document defines a new flag for the Attribute Flags TLV, which is carried in the LSP_ATTRIBUTES Object of Path message. This flag is used to communicate whether setup protection is desired for an LSP. The value of the new flag needs to be assigned by IANA.

Setup Protection Desired: TBD

This document defines two new RSVP Attributes TLVs, which are carried in the LSP_REQUIRED_ATTRIBUTES object of Path message. The values of the new types need to be assigned by IANA.

Protected LSP Sender IPv4 Address TLV

Protected LSP Sender IPv6 Address TLV

5. Security Considerations

The security considerations discussed in RFC 3209, RFC 4090 and RFC 4875 apply to this document.

6. Acknowledgements

Thanks to Rahul Aggarwal, Disha Chopra, and Nischal Sheth for their contribution.

7. References

7.1. 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.

[RFC5420] Farrel, A., Papadimitriou, D., Vasseur, JP., and A. Ayyangarps, "Encoding of Attributes for MPLS LSP Establishment Using Resource Reservation Protocol Traffic Engineering (RSVP-TE)", RFC 5420, February 2009.

[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.

[RFC3471] Berger, L., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", RFC 3471, January 2003.

[RFC3472] Ashwood-Smith, P. and L. Berger, "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Constraint-based Routed Label Distribution Protocol (CR-LDP) Extensions", RFC 3472, January 2003.

[RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, January 2001.

7.2. Informative References

[RFC5920] Fang, L., "Security Framework for MPLS and GMPLS Networks", RFC 5920, July 2010.

Authors' Addresses

Yimin Shen
Juniper Networks
10 Technology Park Drive
Westford, MA 01886
USA

Phone: +1 9785890722
Email: yshen@juniper.net

Yuji Kamite
NTT Communications Corporation
Granpark Tower 3-4-1 Shibaura, Minato-ku
Tokyo 108-8118
Japan

Email: y.kamite@ntt.com

Eric Osborne
Cisco Systems

Email: eosborne@cisco.com