Internet Engineering Task Force (IETF)
Request for Comments: 5710
Category: Standards Track
ISSN: 2070-1721

L. Berger
LabN
D. Papadimitriou
Alcatel Lucent
JP. Vasseur
Cisco

January 2010

PathErr Message Triggered MPLS and GMPLS LSP Reroutes

Abstract

This document describes how Resource ReserVation Protocol (RSVP) PathErr messages may be used to trigger rerouting of Multi-Protocol Label Switching (MPLS) and Generalized MPLS (GMPLS) point-to-point Traffic Engineering (TE) Label Switched Paths (LSPs) without first removing LSP state or resources. Such LSP rerouting may be desirable in a number of cases, including, for example, soft-preemption and graceful shutdown. This document describes the usage of existing Standards Track mechanisms to support LSP rerouting. In this case, it relies on mechanisms already defined as part of RSVP-TE and simply describes a sequence of actions to be executed. While existing protocol definitions can be used to support reroute applications, this document also defines a new reroute-specific error code to allow for the future definition of reroute-application-specific error values.

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 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc5710.

Table of Contents

1. Introduction

The Resource ReserVation Protocol (RSVP), see [RFC2205], has been extended to support the control of Traffic Engineering (TE) Label Switched Paths (LSPs) for both Multi-Protocol Label Switching (MPLS) and Generalized MPLS (GMPLS) in, respectively, [RFC3209] and [RFC3473]. In all cases, a PathErr message is used to report errors to nodes upstream of the error-detecting node. As defined in [RFC2205] and left unmodified by [RFC3209], PathErr messages "do not change path state in the nodes through which they pass". Notwithstanding this definition, PathErr messages are most commonly used to report errors during LSP establishment, i.e., the RSVP-TE processing that occurs prior to the ingress receiving a Resv message. (See [RFC5711] for a broader discussion on PathErr message handling.) Support for such usage was enhanced via the introduction of the Path_State_Removed flag in [RFC3473], which enables a processing node to free related LSP state and resources. The usage of PathErr messages during LSP establishment was further covered in [RFC4920], which describes in detail how a node may indicate that it or one of its associated resources should be avoided, i.e., routed around, during LSP establishment.

PathErr messages can also be used to support a number of other cases that can occur after an LSP is established. This document focuses on the cases where PathErr messages can be used for a node to indicate that it desires an upstream node to reroute an LSP around the indicating node or resources associated with the indicating node. Some examples of such cases are soft-preemption and graceful shutdown (see [RFC5712] and [GRACEFUL]).

This document uses the terminology "reroute request" to refer to the indication by a node that an upstream reroute should take place. This document describes how a node can initiate a reroute request without disrupting LSP data traffic or, when so desired, with the disruption of data traffic and removal of LSP-associated state and resources. The applicability of this document is limited to point-to-point LSPs. Support for point-to-multipoint LSPs are for further study.

The mechanisms used to indicate reroute requests are derived from the mechanisms described in [RFC4920] and the error codes defined in [RFC4736]. This document describes (1) how a non-disruptive reroute request may be issued and, (2) based on an optional "timeout" period, how rerouting may be forced by removing LSP state and associated resources and signaling such removal. While this document describes how existing protocol definitions can be used to support rerouting, it also defines a new reroute-specific error code to allow for the future definition of reroute-application-specific error values.

1.1. 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 [RFC2119].

2. Reroute Requests

This section describes how a downstream node can indicate that it desires a node upstream (along the LSP path) to initiate the rerouting of an LSP, and how the upstream nodes can respond to such a request. Initiating nodes, transit nodes, and ingress nodes are described separately.

2.1. Processing at Requesting Node

When a transit or egress node desires to request the rerouting of an established LSP, it first determines if it can act on the reroute request locally. Such a check MUST be performed on the condition that the Explicit Route Object (ERO), see [RFC3209], received in the LSP's incoming Path message does not preclude LSP rerouting. Examples of requests that may trigger reroutes are avoiding an outgoing interface, a component, label resource, or a next hop not explicitly listed in the ERO. In all cases, the actual repair action SHOULD be performed after verification that the local policy allows local repair for that LSP/state. That is, any traffic-rerouting action (associated to this state) must be initiated and completed only as allowed by local node policy.

When the node cannot act locally, it MUST issue a PathErr message indicating its inability to perform local rerouting. The PathErr message MUST contain an ERROR_SPEC object of the format defined in [RFC2205] or [RFC3473]. Such a message MUST include one of the following combinations of error codes and error values:

1. "Notify/Local node maintenance required" to support backwards compatibility and to reroute around the local node.

2. "Notify/Local link maintenance required" to support backwards compatibility and to reroute around a local interface.

3. "Reroute/<any Reroute error value>" for future compatibility and when backwards compatibility is not a concern.

The rest of the ERROR_SPEC object is constructed based on the local rerouting decision and the resource that is to be avoided by an upstream node. It is important to note that the address and TLVs carried by the ERROR_SPEC object identify the resource to be avoided and not the error code and value.

When the reroute decision redirects traffic around the local node, the local node MUST be indicated in the ERROR_SPEC object. Otherwise, i.e., when the reroute decision does not redirect traffic around the local node, the impacted interface MUST be indicated in the ERROR_SPEC object and the IF_ID [RFC3473] ERROR_SPEC object formats SHOULD be used to indicate the impacted interface.

The IF_ID [RFC3473] ERROR_SPEC object format MUST be used to indicate a reroute request that is more specific than an interface. The TLVs defined in [RFC3471], as updated by [RFC3477], [RFC4201], and [RFC4920] MAY be used to provide specific, additional reroute request information, e.g., reroute around a specific label. The principles related to ERROR_SPEC object construction, defined in Section 6.3.1 of [RFC4920], SHOULD be followed.

2.1.1. Reroute Request Timeouts

Reroute request timeouts are used to remove an LSP when there is no response to a reroute request. A reroute request timeout is used when an LSP is to be removed at the expiration of the reroute request timeout period. When such LSP removal is desired, and after initiating a reroute request, the initiating node MUST initiate a timeout during which it expects to receive a response to the reroute request. Valid responses are a PathTear message or a trigger Path message with an ERO, avoiding the resource that was indicated in the reroute request. If either type of message is received, the timeout period MUST be canceled and no further action is needed. Note, normal refresh processing is not modified by the introduction of reroute request timeouts. Such processing may result in Path state being removed during the timeout period, in which case the timeout period MUST also be canceled.

If the reroute request timeout is reached, the initiating node MUST remove the LSP and its associated state and resources. Removal of LSP state is indicated downstream via a corresponding PathTear message. Removal is indicated upstream via a PathErr message with the error code of "Service preempted". The Path_State_Removed flag MUST be set if supported. When the Path_State_Removed flag is not supported, a corresponding ResvTear MUST also be sent.

2.2. Processing at Upstream Node

When a transit node's policy permits it to support reroute request processing and local repair, the node MUST examine incoming PathErr messages to see it the node can perform a requested reroute. A reroute request is indicated in a received PathErr message, which carries one of the error code and value combinations listed above in Section 2.1. Note that a conformant implementation MUST check for any of the three combinations listed in Section 2.1.

A transit node MAY act on a reroute request locally when the ERO received in the LSP's incoming Path message does not preclude the reroute. As before, examples include loosely routed LSP next hops. When the reroute request can be processed locally, standard, local repair processing MUST be followed. The node SHOULD limit the number of local repair attempts. Again, the expected norm is for local repair, and thereby this case, to be precluded due to policy.

When the transit node supports [RFC4920] and is a boundary node, and Boundary rerouting is allowed, it SHOULD use a route request as a trigger to reroute the LSP. (Per [RFC4920], the Flags field of the LSP_ATTRIBUTES object of the initial Path message indicates "Boundary rerouting".) In the case the node triggers rerouting, it first MUST identify an alternate path within the domain. When such a path is available, the node MUST terminate the PathErr message and issue a Path message reflecting the identified alternate path. Processing then continues per [RFC4920]. When an alternate path is not available, the node cannot act on the reroute request.

When a transit node cannot act on a reroute request locally, per standard processing, it MUST propagate the received PathErr message to the previous hop.

2.3. Processing at Ingress

When reroute processing is supported, an ingress node MUST check received PathErr messages to identify them as indicating reroute requests. A reroute request is indicated in a received PathErr message, which carries one of the error code and value combinations listed above in Section 2.1. Note that a conformant implementation MUST check for any of the three combinations listed in Section 2.1.

Upon receiving a reroute request, the ingress MUST attempt to identify an alternate path, avoiding the node, interface, resource, etc. identified within the ERROR_SPEC object. When an alternate path cannot be identified, the reroute request MUST be discarded. When an

alternate path is identified, a corresponding make-before-break LSP SHOULD be initiated and standard make-before-break procedures MUST be followed.

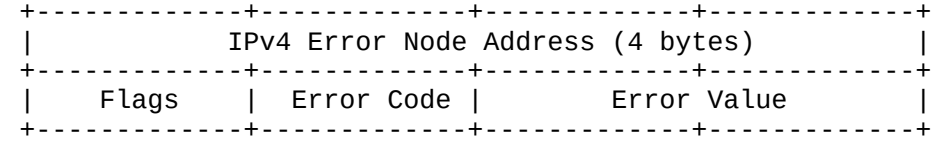
3. Example Reroute Requests

This section provides example reroute requests. This section is informative rather than prescriptive. Reroute requests are always sent via PathErr messages. As described above, a PathErr message may contain either an [RFC2205] format ERROR_SPEC object, or an IF_ID [RFC3473] format ERROR_SPEC object; it is the address and TLVs carried by the ERROR_SPEC object, and not the error value, that indicates the resource that is to be avoided by the reroute.

3.1. Node Reroute Request

To indicate that the node should be avoided by an upstream node, the node originating the reroute may format the ERROR_SPEC per [RFC2205], for example:

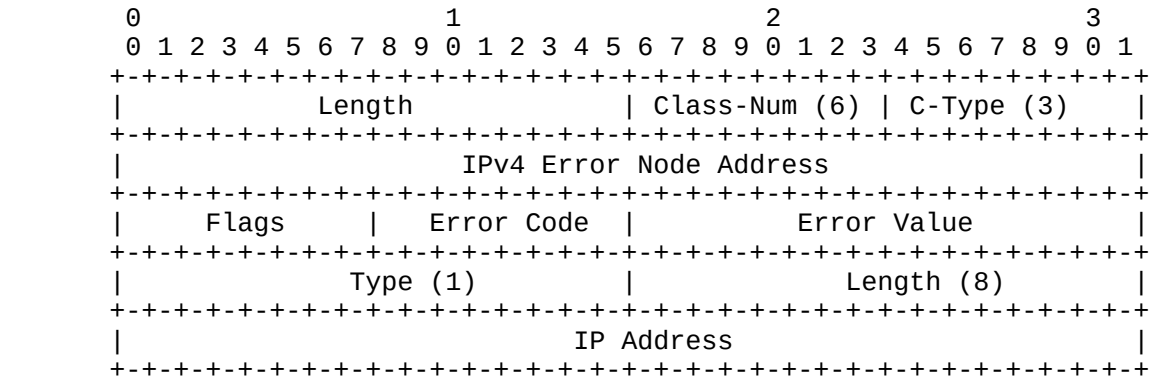
o IPv4 ERROR_SPEC object: Class = 6, C-Type = 1



The node address is set to the local node’s TE router address. Error code is set to either "Notify/Local node maintenance required" or "Reroute/<any Reroute error value>".

3.2. Interface Reroute Request

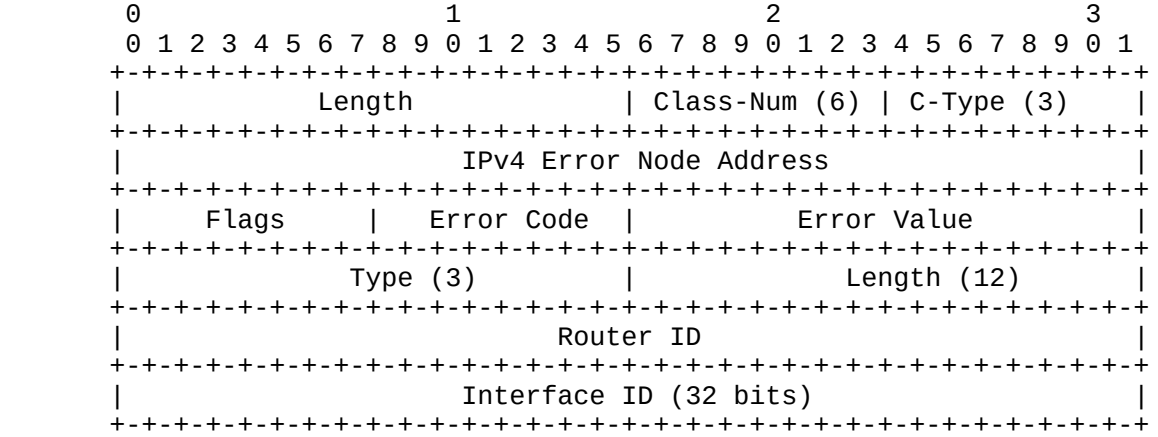
To indicate that a numbered interface should be avoided by an upstream node, the node originating the reroute may format the ERROR_SPEC per [RFC3473], for example:



The node address is set to the local node's TE router address. Error code is set to either "Notify/Local link maintenance required" or "Reroute/<any Reroute error value>". IP address is set to the TE address of the interface to be avoided.

3.3. Component Reroute Request

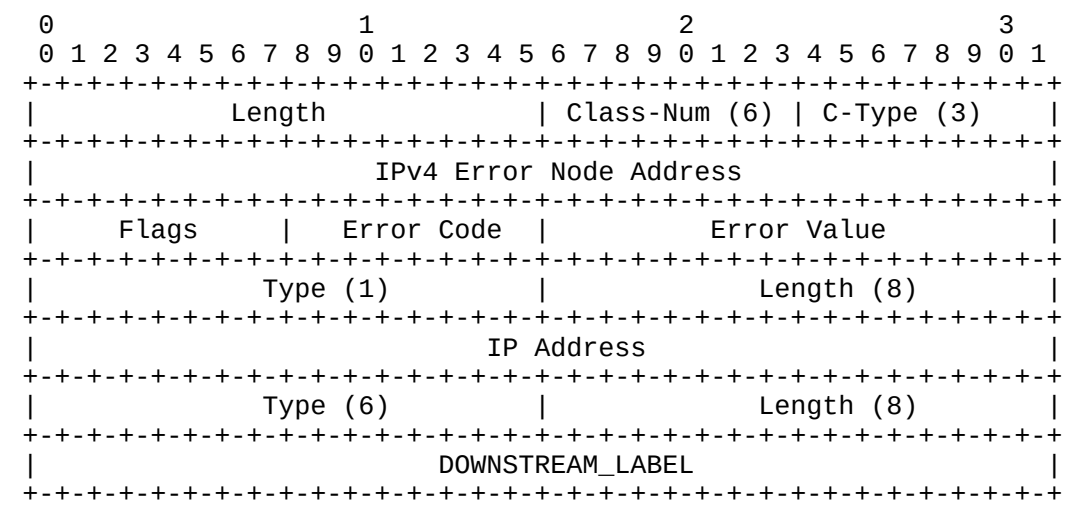
To indicate that an unnumbered component should be avoided by an upstream node, the node originating the reroute formats the ERROR_SPEC per [RFC4201], for example:



The node address is set to the local TE address used in the advertisement of the bundle associated with the component. Error code is set to either "Notify/Local link maintenance required" or "Reroute/<any Reroute error value>". Router ID is set to the local router ID, and Interface ID is the identifier assigned to the component link by the local node.

3.4. Label Reroute Request

To indicate that a label should be avoided by an upstream node, the node originating the reroute may format the ERROR_SPEC per [RFC4920], for example:



The node address is set to the local node's TE router address. Error code is set to either "Notify/Local link maintenance required" or "Reroute/<any Reroute error value>". IP address is set to the TE address of the interface that supports the label to be avoided. DOWNSTREAM_LABEL indicates the label to be avoided.

4. IANA Considerations

IANA assigned values for namespaces defined in this document and reviewed in this section.

IANA made the assignment in the "Error Codes and Globally-Defined Error Value Sub-Codes" section of the "RSVP Parameters" registry:

34 Reroute [RFC5710]

This error code has the following defined Error Value sub-code:

0 = Generic LSP reroute request

Reroute error values should be allocated based on the following allocation policy as defined in [RFC5226].

Range	Registration Procedures
0-32767	IETF Consensus
32768-65535	Private Use

5. Security Considerations

Sections 9 of [RFC4920] and [RFC4736] should be used as the starting point for reviewing the security considerations related to the formats and mechanisms discussed in this document. This document introduces a new error code, but this code is functionally equivalent to existing semantics, in particular, the error code/error value combinations of "Notify/Local node maintenance required" and "Notify/Local link maintenance required". As such, this document introduces no new security considerations beyond what already applies to these existing formats and mechanisms. Future documents may define new error values; any considerations specific to those values should be discussed in the document defining them.

6. References

6.1. Normative References

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

[RFC2205] Braden, R., Ed., 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.

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

[RFC3473] Berger, L., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, January 2003.

[RFC3477] Kompella, K. and Y. Rekhter, "Signalling Unnumbered Links in Resource ReSerVation Protocol - Traffic Engineering (RSVP-TE)", RFC 3477, January 2003.

[RFC4201] Kompella, K., Rekhter, Y., and L. Berger, "Link Bundling in MPLS Traffic Engineering (TE)", RFC 4201, October 2005.

[RFC4920] Farrel, A., Ed., Satyanarayana, A., Iwata, A., Fujita, N., and G. Ash, "Crankback Signaling Extensions for MPLS and GMPLS RSVP-TE", RFC 4920, July 2007.

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, May 2008.

6.2. Informative References

[RFC4736] Vasseur, JP., Ed., Ikejiri, Y., and R. Zhang, "Reoptimization of Multiprotocol Label Switching (MPLS) Traffic Engineering (TE) Loosely Routed Label Switched Path (LSP)", RFC 4736, November 2006.

[GRACEFUL] Ali, Z., Vasseur, JP., Zamfir, A., and J. Newton, "Graceful Shutdown in MPLS and Generalized MPLS Traffic Engineering Networks", Work in Progress, September 2009.

[RFC5711] Vasseur, JP., Ed., Swallow, G., and I. Minei, "Node Behavior upon Originating and Receiving Resource Reservation Protocol (RSVP) Path Error Messages", RFC 5711, January 2010.

[RFC5712] Meyer, M., Ed. and JP. Vasseur, Ed., "MPLS Traffic Engineering Soft Preemption", RFC 5712, January 2010.

7. Acknowledgments

This document was conceived along with Matthew Meyer. George Swallow provided valuable feedback. The General Area Review Team (Gen-ART) review was performed by Francis Dupont.

Authors' Addresses

Lou Berger
LabN Consulting, L.L.C.
Phone: +1-301-468-9228
EMail: lberger@labn.net

Dimitri Papadimitriou
Alcatel Lucent
Francis Wellesplein 1,
B-2018 Antwerpen, Belgium
Phone: +32 3 240-8491
EMail: Dimitri.Papadimitriou@alcatel-lucent.be

JP Vasseur
Cisco Systems, Inc
11, Rue Camille Desmoulins
L'Atlantis
92782 Issy Les Moulineaux
France
EMail: jpv@cisco.com