Abstract

This document defines a mechanism for the reoptimization of loosely routed MPLS Traffic Engineering LSPs. A loosely routed LSP follows a path specified as a combination of strict and loose hop(s) that contains at least one loose hop and zero or more strict hop(s). The path calculation (which implies an ERO expansion) to reach a loose
hop is performed by the previous hop defined in the TE LSP path. This
document proposes a mechanism that allows:

- The TE LSP head-end LSR to trigger a new path re-evaluation on
every hop having a next hop defined as a loose hop,

- A mid-point LSR to signal to the head-end LSR that either a better
  path exists to reach a loose hop (compared to the current path in
  use) or that the TE LSP must be reoptimized because of some
  maintenance required on the TE LSP path. A better path is defined as
  a lower cost path, where the cost is determined by the metric used to
  compute the path.

The proposed mechanism applies to intra-domain and inter-domain (IGP
area or Autonomous System) packet and non-packet TE LSPs when the
path is defined as a list of loose hops or when a strict hop is a
non-specific abstract node (e.g. IGP area, Autonomous Systems).

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 [ii].

Table of contents

1. Introduction....................................................... 3
2. Establishment of a loosely routed TE LSP........................ 3
3. Reoptimization of a loosely routed TE LSP path................... 4
4. Signalling extensions............................................. 5
  4.1 Path re-evaluation request.................................... 5
  4.2 New error value sub-code..................................... 5
5. Mode of operation................................................ 6
  5.1 Head-end reoptimization control.............................. 6
  5.2 Reoptimization triggers....................................... 6
  5.3 Head-end request versus mid-point explicit notification modes
  ............................................................ 6
  5.3.1 Head-end request mode...................................... 7
  5.3.2 Mid-point explicit notification mode....................... 8
  5.3.3 ERO caching................................................ 9
6. Interoperability.................................................. 9
7. Security considerations.......................................... 9
8. Acknowledgments.................................................. 9
9. Intellectual property considerations............................ 9
  9.1 IPR Disclosure Acknowledgement................................ 9
10. References...................................................... 10
    Normative references........................................... 10
    Informative references......................................... 10
11. Author’s Addresses............................................. 11
    Full Copyright Statement...................................... 11

Vasseur, Ikejiri and Zhang [Page 2]
1. Introduction

The Traffic Engineering Work Group has specified a set of requirements for inter-area [INTER-AREA-TE-REQ] and inter-AS [INTER-AS-TE-REQ] MPLS Traffic Engineering. Both requirements documents specify the need for some mechanism providing an option for the head-end to control the reoptimization process, should a more optimal path exist in a downstream domain (IGP area or Autonomous System).

This document defines a solution to meet this requirement, in addition to a mechanism to notify a Head-end LSR of the existence of such a more optimal path or the need to reoptimize due to some maintenance required in a downstream domain.

2. Establishment of a loosely routed TE LSP

A loosely routed explicit path is a path specified as a combination of strict and loose hop(s) that contains at least one loose hop and a set of zero or more strict hop(s). Loose hops are listed in the ERO object of the RSVP Path message with the L flag of the Ipv4 or the IPv6 prefix sub-object set, as defined in [RSVP-TE]. In this case, each LSR along the path whose next hop is specified as a loose hop or a non-specific abstract node triggers a path computation (also referred to as an ERO expansion), before forwarding the RSVP Path message downstream. The path computation may either be performed by means of CSPF or any Path Computation Element (PCE) and can be partial (up to the next loose hop) or complete (up to the TE LSP destination).

Note that the examples in the rest of this document are provided in the context of MPLS inter-area TE but the proposed mechanism equally applies to loosely routed paths within a single routing domain and across multiple Autonomous Systems.

The examples below are provided with OSPF as the IGP but the described set of mechanisms similarly apply to IS-IS.

An example of an explicit loosely routed TE LSP signaling.

<---area 1--><-area 0--><-area 2->

R1---R2----R3----R6    R8---R10
 |     |    |    |   |
 |     |    |    / |  |
 |     |    | /   |   |
R4--------R5----R7----R9----R11

Assumptions
- R3, R5, R8 and R9 are ABRs
The path an inter-area TE LSP T1 from R1 (head-End LSR) to R11 (tail-end LSR) is defined on R1 as the following loosely routed path: R1-R3(loose)-R8(loose)-R11(loose). R3, R8 and R11 are defined as loose hops.

Step 1: R1 determines that the next hop (R3) is a loose hop (not directly connected to R1) and then performs an ERO expansion operation to reach the next loose hops R3 either by means of CSPF or any other PCE-based path computation method. The new ERO becomes: R2(S)-R3(S)-R8(L)-R11(L) where:

S: Strict hop (L=0)
L: Loose hop (L=1)

The R1-R2-R3 path obeys T1's set of constraints.

Step 2: the RSVP Path message is then forwarded by R1 following the ERO path and reaches R3 with the following content: R8(L)-R11(L)

Step 3: R3 determines that the next hop (R8) is a loose hop (not directly connected to R3) and then performs an ERO expansion operation to reach the next loose hops R8 either by means of CSPF or any other PCE-based path computation method. The new ERO becomes: R6(S)-R7(S)-R8(S)-R11(L)

Note: in this example, the assumption is made that the path is computed on a per loose hop basis, also referred to a partial route computation. Note that PCE-based mechanisms may also allow for full route computation (up to the final destination).

Step 4: the same procedure applies at R8 to reach T1's destination (R11).

3. Reoptimization of a loosely routed TE LSP path

Once a loosely routed explicit TE LSP is set up, it is maintained through normal RSVP procedures. During TE LSP life time, a more optimal path might appear between an LSR and its next loose hop (for the sake of illustration, suppose in the example above that a link between R6 and R8 is added or restored that provides a preferable path between R3 and R8 (R3-R6-R8) than the existing R3-R6-R7-R8 path). Since a preferable (e.g. shorter) path might not be visible from the head-end LSR by means of the IGP if it does not belong to the head-end IGP area, the head-end cannot make use of this shorter path (and reroute the LSP using a make before break) when appropriate. Hence, some mechanism is required to detect the existence of such a preferable path and to notify the head-end accordingly.

This document defines a mechanism that allows:
- A head-end LSR to trigger on every LSR whose next hop is a loose hop or an abstract node the re-evaluation of the current path in order to detect a potential more optimal path,

- A mid-point LSR whose next hop is a loose-hop or an abstract node to signal (using a new Error value sub-code carried in a Path Error message) to the head-end that a more preferable path exists (a path with a lower cost, where the cost definition is determined by some metric).

Then once the existence of such a preferable path is notified to the head-end LSR, the head-end LSR can decide (depending on the TE LSP characteristics) whether to perform a TE LSP graceful reoptimization.

There is another scenario whereby notifying the head-end of the existence of a better path is desirable: if the current path is about to fail due to some (link or node) required maintenance (see also [GR-SHUT]).

This allows the head-end to reoptimize a TE LSP making use of the non-disruptive make before break procedure if and only if a preferable path exists and if such a reoptimization is desired.

4. Signalling extensions

New ERO flags and Error value sub-codes are proposed in this document (to be assigned by IANA).

4.1 Path re-evaluation request

The following new flag of the SESSION_ATTRIBUTE object (C-Type 1 and 7) is defined (suggested value to be confirmed by IANA):

Path re-evaluation request: 0x20

This flag indicates that a path re-evaluation (of the current path in use) is requested. Note that this does not trigger any LSP Reroute but instead just signal the request to evaluate whether a preferable path exists.

Note: in case of link bundling for instance, although the resulting ERO might be identical, this might give the opportunity for a mid-point LSR to locally select another link within a bundle, although strictly speaking, the ERO has not changed.

4.2 New error value sub-code

As defined in [RSVP-TE], the ERROR-CODE 25 in ERROR SPEC object corresponds to a Notify Error.
This document adds three new error value sub-codes (suggested values to be confirmed by IANA):

6  Preferable path exists
7  Local link maintenance required
8  Local node maintenance required

The details about the local maintenance required modes are detailed in section 5.3.2

5. Mode of operation

5.1 Head-end reoptimization control

The notification process of a preferable path (shorter path or new path due to some maintenance required on the current path) is by nature de-correlated from the reoptimization operation. In other words, the location where a potentially preferable path is discovered does not have to be where the TE LSP is actually reoptimized. This document applies to the context of a head-end reoptimization.

5.2 Reoptimization triggers

There are three possible reoptimization triggers:

- Timer-based: a reoptimization is triggered (process evaluating whether a more optimal path can be found) when a configurable timer expires,
- Event-driven: a reoptimization is triggered when a particular network event occurs (such as a Link-UP event),
- Operator-driven: a reoptimization is manually triggered by the Operator.

It is RECOMMENDED for an implementation supporting the extensions proposed in this document to support the aforementioned modes as path re-evaluation triggers.

5.3 Head-end request versus mid-point explicit notification modes

This document defines two modes:

1) Head-end requesting mode: the request for a new path evaluation of a loosely routed TE LSP is requested by the head-end LSR.

2) Mid-point explicit notification: a mid-point LSR having determined that a preferable path (than the current path is use) exists or having the need to perform a link/node local maintenance explicitly notifies the head-end LSR which will in turn decide whether to perform a reoptimization.
5.3.1 Head-end request mode

In this mode, when a timer-based reoptimization is triggered on the head-end LSR or the operator manually requests a reoptimization, the head-end LSR immediately sends an RSVP Path message with the "Path re-evaluation request" bit of the SESSION-ATTRIBUTE object set. This bit is then cleared in subsequent RSVP path messages sent downstream.

Upon receiving a Path message with the "Path re-evaluation request" bit set, every LSR for which the next abstract node contained in the ERO is defined as a loose hop/abstract node, performs the following set of actions:

A path re-evaluation is triggered and the newly computed path is compared to the existing path:

- If a preferable path can be found, the LSR MUST immediately send a Path Error to the head-end LSR (Error code 25 (Notify), Error sub-code=6 (better path exists)). At this point, the LSR MAY decide to clear the "Path re-evaluation request" bit of the SESSION-ATTRIBUTE object in subsequent RSVP Path messages sent downstream: this mode is the RECOMMENDED mode for the reasons described below.

The sending of a Path Error Notify message "Preferable path exists" to the head-end LSR will notify the head-end LSR of the existence of a preferable path (e.g. in a downstream area/AS or in another location within a single domain). Hence, triggering additional path re-evaluations on downstream nodes is unnecessary. The only motivation to forward subsequent RSVP Path messages with the "Path re-evaluation request" bit of the SESSION-ATTRIBUTE object set would be to trigger path re-evaluation on downstream nodes that could in turn cache some potentially better paths downstream with the objective to reduce the signaling setup delay, should a reoptimization be performed by the head-end LSR.

- If no preferable path can be found, the recommended mode is for an LSR to relay the request (by setting the "Path re-evaluation" bit of the SESSION-ATTRIBUTE object in RSVP path message sent downstream).

By preferable path, we mean a path having a lower cost. By default, an LSR uses the TE metric to compute the shortest path that obeys a set of constraints. Note that the head-end LSR might use the METRIC-TYPE object (defined in [PATH-COMP]) in its path message to request the LSR having a next hop defined as a loose hop or an abstract node in the ERO to use another metric to determine a preferable path.
If the RSVP Path message with the Path re-evaluation request bit set is lost, then the next request will be sent when the next reoptimization trigger will occur on the head-end LSR. The solution to handle RSVP reliable messaging has been defined in [REFRESH-REDUCTION].

The network administrator may decide to establish some local policy specifying to ignore such request or to consider those requests not more frequently than a certain rate.

The proposed mechanism does not make any assumption of the path computation method performed by the ERO expansion process: it can either be local to each LSR in charge of computing the path to the next loose hop/abstract node or PCE based.

5.3.2 Mid-point explicit notification mode

In this mode, a mid-point LSR whose next hop is a loose hop or an abstract node can locally trigger a path re-evaluation when a configurable timer expires, some specific events occur (e.g. link-up event for example) or the user explicitly requests it. If a preferable path is found compared to the existing one, the LSR sends a Path Error to the head-end LSR (Error code 25 (Notify), Error sub-code=6 (preferable path exists)).

There are other circumstances whereby a mid-point LSR MAY send an RSVP PathError message with the objective for the TE LSP to be rerouted by its head-end LSR: when a link or a node will go down for local maintenance reasons. In this case, the mid-point LSR where the local maintenance must be performed is responsible for sending an RSVP PathError message with Error code 25 and Error sub-code=7 or 8 depending on the affected network element (link or node). Then the first upstream node having performed the ERO expansion MUST perform the following set of actions:

- The link (sub-code=7) or the node (sub-code=8) MUST be locally registered for further reference (the TE database must be updated)

- The RSVP Path Error message MUST be immediately forwarded upstream to the head-end LSR. Note that in the case of TE LSP spanning multiple administrative domains, it may be desirable for the boundary LSR to modify the RSVP PathError message and insert its own address for confidentiality reason.

Upon receiving a PathError message with Error code 25 and Error sub-code 7 or 8, the Head-end LSR MUST perform a TE LSP reoptimization.

Note that those modes are not exclusive: both the timer and event-driven reoptimization triggers can be implemented on the head-end...
and/or any mid-point LSR with potentially different timer values for the timer driven reoptimization case.

A head-end LSR MAY decide upon receiving an explicit mid-point notification to delay its next path re-evaluation request.

5.3.3 ERO caching

Once a mid-point LSR has determined that a preferable path exists (after a reoptimization request has been received by the head-end LSR or the reoptimization timer on the mid-point has fired), the more optimal path MAY be cached on the mid-point LSR for a limited amount of time to avoid having to recompute a path once the head-LSR performs a make before break. This mode is optional.

6. Interoperability

An LSR not supporting the Path re-evaluation request bit of the SESSION-ATTRIBUTE object SHALL forward it unmodified.

Any head-end LSR not supporting a PathError Error code 25 message with Error sub-code = 6, 7 or 8 MUST just silently ignore such Path Error messages.

7. Security considerations

This document defines a mechanism for a mid-point LSR to notify the head-end LSR of this existence of a preferable path or the need to reroute the TE LSP for maintenance purposes. Hence, in case of a TE LSP spanning multiple administrative domains, it may be desirable for a boundary LSR to modify the PathError message (Code 25, Error sub-code=6 or 7) so as to preserve confidentiality across domains.

8. Acknowledgments

The authors would like to thank Carol Iturralde, Miya Kohno, Francois Le Faucheur, Philip Matthews, Jim Gibson, Raymond Zhang, Jean-Louis Le Roux, Kenji Kumaki, Anca Zafir for their useful comments. A special thank to Adrian Farrel for his very valuable inputs.

9. Intellectual property considerations

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information
on the procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at http://www.ietf.org/ipr.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

9.1 IPR Disclosure Acknowledgement

By submitting this Internet-Draft, I certify that any applicable patent or other IPR claims of which I am aware have been disclosed, and any of which I become aware will be disclosed, in accordance with RFC 3668.

10. References

Normative references


Informative references


11. Author’s Addresses

Jean-Philippe Vasseur  
CISCO Systems, Inc.  
300 Beaver Brook  
Boxborough, MA 01719  
USA  
Email: jpv@cisco.com

Yuichi Ikejiri  
NTT Communications Corporation  
1-1-6, Uchisaiwai-cho, Chiyoda-ku  
Tokyo 100-8019  
JAPAN  
Email: y.ikejiri@ntt.com

Raymond Zhang  
Infonet Services Corporation  
2160 E. Grand Ave.  
El Segundo, CA 90025  
USA  
Email: raymond_zhang@infonet.com

Full Copyright Statement

"Copyright (C) The Internet Society (year). This document is subject to the rights, licenses and restrictions contained in BCP 78, and except as set forth therein, the authors retain all their rights."

"This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE."