Inter domain Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS) Traffic Engineering - RSVP-TE extensions

draft-ietf-ccamp-inter-domain-rsvp-te-07.txt

Status of this Memo

Abstract

This document describes procedures and protocol extensions for the
use of Resource ReserVation Protocol Traffic Engineering (RSVP-TE)
signaling in Multiprotocol Label Switching Traffic Engineering
(MPLS-TE) packet networks and Generalized MPLS (GMPLS) packet and
non-packet networks to support the establishment and maintenance of
Label Switched Paths that cross domain boundaries.

For the purpose of this document, a domain is considered to be any
collection of network elements within a common realm of address space
or path computation responsibility. Examples of such domains include
Autonomous Systems, IGP routing areas, and GMPLS overlay networks.

Table of Contents

1. Introduction

The requirements for inter-area and inter-AS Multiprotocol Label
Switching (MPLS) Traffic Engineering (TE) are stated in [RFC4105] and
[RFC4216] respectively. Many of these requirements also apply to
Generalized MPLS (GMPLS) networks. The framework for inter-domain
MPLS-TE is provided in [RFC4726].

This document presents procedures and extensions to Resource
Reservation Protocol Traffic Engineering (RSVP-TE) signaling for the
setup and maintenance of traffic engineered Label Switched Paths (TE
LSPs) that span multiple domains in MPLS-TE or GMPLS networks. The
signaling procedures described in this document are applicable to
MPLS-TE packet LSPs established using RSVP-TE ([RFC3209]) and all
LSPs (packet and non-packet) that use RSVP-TE GMPLS extensions as
described in [RFC3473].

Three different signaling methods for inter-domain RSVP-TE signaling
are identified in [RFC4726]. Contiguous LSPs are
achieved using the procedures of [RFC3209] and [RFC3473] to create a
single end-to-end LSP that spans all domains. Nested LSPs are
established using the techniques described in [RFC4206] to carry the
end-to-end LSP in a separate tunnel across each domain. Stitched LSPs
are established using the procedures of [LSP-STITCHING] to construct
an end-to-end LSP from the concatenation of separate LSPs each
spanning a domain.

This document defines the RSVP-TE protocol extensions necessary to
control and select which of the three signaling mechanisms is used
for any one end-to-end inter-domain TE LSP.

For the purpose of this document, a domain is considered to be any
collection of network elements within a common realm of address space
or path computation responsibility. Examples of such domains include
Autonomous Systems, IGP areas, and GMPLS overlay networks
([RFC4208]).

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 RFC 2119 [RFC2119].

1.2. Terminology

AS: Autonomous System.

ASBR: Autonmous System Border Router. A router used to connect
together ASes of a different or the same Service Provider via one or
more Inter-AS links.

Bypass Tunnel: An LSP that is used to protect a set of LSPs passing
over a common facility.

ERO: Explicit Route Object.

FA: Forwarding Adjacency.

LSR: Label Switching Router.

MP: Merge Point. The node where bypass tunnels meet the protected
LSP.

NHOP bypass tunnel: Next-Hop Bypass Tunnel. A backup tunnel, which
bypasses a single link of the protected LSP.

   NNHOP bypass tunnel: Next-Next-Hop Bypass Tunnel. A backup tunnel,
   which bypasses a single node of the protected LSP.

   PLR: Point of Local Repair. The ingress of a bypass tunnel.

   RRO: Record Route Object.

   TE link: Traffic Engineering link.

2. Signaling Overview

   The RSVP-TE signaling of a TE LSP within a single domain is described
   in [RFC3209] and [RFC3473]. Inter-domain TE LSPs can be supported by
   one of three options as described in [RFC4726] and set out in the
   next section:

   - contiguous LSPs
   - nested LSPs
   - stitched LSPs.

   In fact, as pointed out in [RFC4726], any combination of these three
   options may be used in the course of an end-to-end inter-domain LSP.
   That is, the options should be considered as per-domain transit
   options so that an end-to-end inter-domain LSP that starts in domain
   A, transits domains B, C and D, and ends in domain E might use an LSP
   that runs contiguously from the ingress in domain A, through domain B
   to the border with domain C. Domain C might be transited using the
   nested LSP option to reach the border with domain D, and domain D
   might be transited using the stitched LSP option to reach the border
   with domain E, from where a normal LSP runs to the egress.

   This document describes the RSVP-TE signaling extensions required to
   control which of the three signaling mechanisms is used and to select
   which of the three signaling mechanisms is used.

   The specific protocol extensions required to signal each LSP type are
   described in other documents and are out of scope for this document.
   Similarly, the routing extensions and path computation techniques
   necessary for the establishment of inter-domain LSPs are out of
   scope. An implementation of a transit LSR is unaware of the options
   for inter-domain TE LSPs since it sees only TE LSPs. An
   implementation of a domain border LSR has to decide what mechanisms
   of inter-domain TE LSP support to include, but must in any case
   support contiguous inter-domain TE LSPs since this is the default
   mode of operation for RSVP-TE. Failure to support either or both of
   nested LSPs or stitched LSPs, restricts the operators options, but
   does not prevent the establishment of inter-domain TE LSPs.

2.1. Signaling Options

There are three ways in which an RSVP-TE LSP could be signaled across multiple domains:

Contiguous
   A contiguous TE LSP is a single TE LSP that is set up across
   multiple domains using RSVP-TE signaling procedures described in
   [RFC3209] and [RFC3473]. No additional TE LSPs are required to
   create a contiguous TE LSP and the same RSVP-TE information for
   the TE LSP is maintained along the entire LSP path. In particular,
   the TE LSP has the same RSVP-TE session and LSP ID at every LSR
   along its path.

Nesting
   One or more TE LSPs may be nested within another TE LSP as
   described in [RFC4206]. This technique can be used to nest one
   or more inter-domain TE LSPs into an intra-domain hierarchical LSP
   (H-LSP). The label stacking construct is used to achieve nesting
   in packet networks. In the rest of this document, the term H-LSP
   is used to refer to an LSP that allows other LSPs to be nested
   within it. An H-LSP may be advertised as a TE link within the same
   instance of the routing protocol as was used to advertise the TE
   links from which it was created, in which case it is a Forwarding
   Adjacency (FA) [RFC4206].

Stitching
   The concept of LSP stitching as well as the required signaling
   procedures is described in [LSP-STITCHING]. This technique can be
   used to stitch together shorter LSPs (LSP segments) to create a
   single, longer LSP. The LSP segments of an inter-domain LSP may be
   intra-domain LSPs or inter-domain LSPs.

   The process of stitching in the data plane results in a single,
   end-to-end contiguous LSP. But in the control plane each segment is
   signaled as a separate LSP (with distinct RSVP sessions) and the
   end-to-end LSP is signaled as yet another LSP with its own RSVP
   session. Thus the control plane operation for LSP stitching is very
   similar to that for nesting.

An end-to-end inter-domain TE LSP may be achieved using one or more
of the signaling techniques described. The choice is a matter of
policy for the node requesting LSP setup (the ingress) and policy for
each successive domain border node. On receipt of an LSP setup
request (RSVP-TE Path message) for an inter-domain TE LSP, the
decision of whether to signal the LSP contiguously or whether to nest
or stitch it to another TE LSP, depends on the parameters signaled
from the ingress node and on the configuration of the local node.

The stitching segment LSP or H-LSP used to cross a domain may be
pre-established or signaled dynamically based on the demand caused by
the arrival of the inter-domain TE LSP setup request.

3. Procedures on the Domain Border Node

Whether an inter-domain TE LSP is contiguous, nested, or stitched is
limited by the signaling methods supported by or configured on the
intermediate nodes, and it is usually the domain border nodes where
this restriction applies since other transit nodes are oblivious to
the mechanism in use. The ingress of the LSP may further restrict the
choice by setting parameters in the Path message when it is signaled.

When a domain border node receives the RSVP Path message for an
inter-domain TE LSP setup, it MUST carry out the following procedures
before it can forward the Path message to the next node along the
path:

  1. Apply policies for the domain and the domain border node. These
     policies may restrict the establishment of inter-domain TE LSPs.
     In case of a policy failure, the node SHOULD fail the setup and
     send a PathErr message with error code "Policy control failure"/
     "Inter-domain policy failure".

  2. Determine the signaling method to be used to cross the domain.
     If the ingress node of the inter-domain TE LSP has specified
     restrictions on the methods to be used, these MUST be adhered
     to. Within the freedom allowed by the ingress node, the domain
     border node MAY choose any method according to local
     configuration and policies. If no resultant signaling method is
     available or allowed, the domain border node MUST send a PathErr
     message with an error code as described in Section 4.1.

     Thus, for example, an ingress may request a contiguous LSP
     because it wishes to exert maximal control over the LSP's path
     and to control when re-optimization takes place. But the
     operator of a transit domain may decide (for example) that only
     LSP stitching is allowed for exactly the reason that it gives
     the operator the chance to reoptimize their own domain under
     their own control. In this case, the policy applied at the entry
     to the transit domain will result in the return of a PathErr
     message and the ingress has a choice to:
     - find another path avoiding the transit domain
     - relax his requirements
     - fail to provide the service.

  3. Carry out ERO procedures as described in Section 3 in addition
     to the procedures in [RFC3209] and [RFC3473].

   4. Perform any path computations as required to determine the path
      across the domain and potentially to select the exit point from
      the domain.

      The path computation procedure is outside the scope of this
      document. A path computation option is specified in
      [INTER-DOMAIN-PD-PATH-COMP], and another option is to use a
      Path Computation Element (PCE) [RFC4655].

      4a. In the case of nesting or stitching, either find an existing
          intra-domain TE LSP to carry the inter-domain TE LSP or
          signal a new one, depending on local policy.

      In event of a path computation failure, a PathErr message SHOULD
      be sent with error code "Routing Problem" using an error value
      selected according to the reason for computation failure. A
      domain border node MAY opt so silently discard the Path message
      in this case as described in Section 8.

   In the event of the receipt of a PathErr message reporting
   signaling failure from within the domain or reported from a
   downstream domain, the domain border node MAY apply crankback
   procedures as described in Section 3.2. If crankback is not
   applied, or is exhausted, the border node MUST continue with
   PathErr processing as described in [RFC3209] and [RFC3473].

   In the event of successful processing of a Path or Resv message,
   the domain border node MUST carry out RRO procedures as described
   in Section 3.3.

3.1. Rules on ERO Processing

   The ERO that a domain border node receives in the Path message was
   supplied by the ingress node of the TE LSP and may have been updated
   by other nodes (for example, other domain border nodes) as the Path
   message was propagated. The content of the ERO depends on several
   factors including:

   - the path computation techniques used
   - the degree of TE visibility available to the nodes performing
     path computation
   - policy at the nodes creating/modifying the ERO.

   In general, H-LSPs and LSP segments are used between domain border
   nodes, but there is no restriction on the use of such LSPs to span
   multiple hops entirely within a domain. Therefore, the discussion
   that follows may be equally applied to any node within a domain
   although the term 'domain border node' continues to be used for
   clarity.

When a Path message reaches the domain border node, the following
rules apply for ERO processing and for further signaling.

1. If there are any policies related to ERO processing for the
   LSP, they MUST be applied and corresponding actions taken. For
   example, there might be a policy to reject EROs that identify
   nodes within the domain. In case of inter-domain LSP setup
   failures due to policy failures related to ERO processing, the
   node SHOULD issue a PathErr with error code "Policy control
   failure"/"Inter-domain explicit route rejected", but MAY be
   configured to silently discard the Path message or to return a
   different error code for security reasons.

2. Section 8.2 of [RFC4206] describes how a node at the edge of a
   region processes the ERO in the incoming Path message and uses
   this ERO, to either find an existing H-LSP or signal a new H-LSP
   using the ERO hops. This process includes adjusting the ERO
   before sending the Path message to the next hop. These
   procedures MUST be followed for nesting or stitching of
   inter-domain TE LSPs.

3. If an ERO subobject identifies a TE link formed by the
   advertisement of an H-LSP or LSP segment (whether numbered or
   unnumbered), contiguous signaling MUST NOT be used. The node MUST
   either use nesting or stitching according to the capabilities of
   the LSP that forms the TE link, the parameters signaled in the
   Path message, and local policy. If there is a conflict between the
   capabilities of the LSP that forms the TE link indicated in the
   ERO and the parameters on the Path message, the domain border node
   SHOULD send a PathErr with error code "Routing Problem"/"ERO
   conflicts with inter-domain signaling method", but MAY be
   configured to silently discard the Path message or to return a
   different error code for security reasons.

4. An ERO in a Path message received by a domain border node may
   have a loose hop as the next hop. This may be an IP address or
   an AS number. In such cases, the ERO MUST be expanded to
   determine the path to the next hop using some form of path
   computation that may, itself, generate loose hops.

5. In the absence of any ERO subobjects beyond the local domain
   border node, the LSP egress (the destination encoded in the RSVP
   Session object) MUST be considered as the next loose hop and
   rule 4 applied.

6. In the event of any other failures processing the ERO, a PathErr
   message SHOULD be sent as described in [RFC3209] or [RFC3473], but
   a domain border router MAY be configured to silently discard the
   Path message or to return a different error code for security

reasons.

3.2. LSP Setup Failure and Crankback

   When an error occurs during LSP setup a PathErr message is sent back
   towards the LSP ingress node to report the problem. If the LSP
   traverses multiple domains, this PathErr will be seen successively by
   each domain border node.

   Domain border nodes MAY apply local policies to restrict the
   propagation of information about the contents of the domain. For
   example, a domain border node MAY replace the information in a
   PathErr message that indicates a specific failure at a specific node
   with information that reports a more general error with the entire
   domain. These procedures are similar to those described for the
   borders of overlay networks in [RFC4208].

   However:

   - A domain border node MUST NOT suppress the propagation of a PathErr
     message except to attempt rerouting as described below.

   - Nodes other than domain border nodes SHOULD NOT modify the contents
     of a PathErr message.

   - Domain border nodes SHOULD NOT modify the contents of a PathErr
     message unless domain confidentiality is a specific requirement.

   Domain border nodes provide an opportunity for crankback rerouting
   [RFC4920]. On receipt of a PathErr message generated because of an
   LSP setup failure, a domain border node MAY hold the PathErr and make
   further attempts to establish the LSP if allowed by local policy and
   by the parameters signaled on the Path message for the LSP. Such
   attempts might involve the computation of alternate routes through
   the domain, or the selection of different downstream domains. If a
   subsequent attempt is successful, the domain border router MUST
   discard the held PathErr message, but if all subsequent attempts are
   unsuccessful, the domain border router MUST send the PathErr upstream
   toward the ingress node. In this latter case, the domain border
   router MAY change the information in the PathErr message to provide
   further crankback details and information aggregation as described
   in [RFC4920].

   Crankback rerouting MAY also be used to handle the failure of LSPs
   after they have been established [RFC4920].

3.3. RRO Processing Across Domains

[RFC3209] defines the RRO as an optional used for loop detection and
for providing information about the hops traversed by LSPs.

As described for overlay networks in [RFC4208], a domain border node
MAY filter or modify the information provided in an RRO for
confidentiality reasons according to local policy. For example, a
series of identifiers of hops within a domain MAY be replaced with
the domain identifier (such as the AS number) or be removed entirely
leaving just the domain border nodes.

Note that a domain border router MUST NOT mask its own presence, and
MUST include itself in the RRO.

Such filtering of RRO information does not hamper the working of the
signaling protocol, but the subsequent information loss may render
management diagnostic procedures inoperable or at least make them
more complicated requiring the coordination of administrators of
multiple domains.

Similarly protocol procedures that depend on the presence of RRO
information may become inefficient. For example, the fast reroute
procedures defined in [RFC4090] use information in the RRO to
determine the labels to use and the downstream MP.

3.4. Notify Message Processing

Notify messages are introduced in [RFC3473]. They may be sent direct
rather than hop-by-hop, and so may speed the propagation of error
information. If a domain border router is interested in seeing
such messages (for example, to enable it to provide protection
switching), it is RECOMMENDED that the domain border router update
the Notify Request objects in the Path and Resv messages to show its
own address following the procedures of [RFC3473].

Note that the replacement of a Notify Recipient in the Notify Request
object means that some Notify messages (for example, those intended
for delivery to the ingress LSR) may need to be examined, processed
and forwarded at domain borders. This is an obvious trade-off issue
as the ability to handle notifiable events locally (i.e. within the
domain) may or may not outweigh the cost of processing and forwarding
Notify messages beyond the domain. Observe that the cost increases
linearly with the number of domains in use.

Also note that, as described in section 8, a domain administrator may
wish to filter or modify Notify messages that are generated within a
domain in order to preserve security or confidentiality of network
information. This is most easily achieved if the Notify messages are
sent via the domain borders.

4. **RSVP-TE Signaling Extensions**

   The following RSVP-TE signaling extensions are defined to enable
   inter-domain LSP setup.

4.1. **Control of Choice of Signaling Method**

   In many network environments there may be a network-wide policy that
   determines which one of the three inter-domain LSP techniques is
   used. In these cases, no protocol extensions are required.

   However, in environments that support more than one technique, an
   ingress node may wish to constrain the choice made by domain border
   nodes for each inter-domain TE LSP that it originates.

   [RFC4420] defines the LSP_Attributes object that can be used to
   signal required attributes of an LSP. The Attributes Flags TLV
   includes Boolean flags that define individual attributes.

   This document defines a new bit in the TLV that can be set by the
   ingress node of an inter-domain TE LSP to restrict the intermediate
   nodes to using contiguous signaling.

   Contiguous LSP bit (bit number assignment in Section 9.1).

   This flag is set by the ingress node that originates a Path message
   to set up an inter-domain TE LSP if it requires that the contiguous
   LSP technique is used. This flag bit is only to be used in the
   Attributes Flags TLV.

   When a domain border LSR receives a Path message containing this bit
   set (one), the node MUST NOT perform stitching or nesting in support
   of the inter-domain TE LSP being set up. When this bit is clear
   (zero), a domain border LSR MAY perform stitching or nesting
   according to local policy.

   This bit MUST NOT be modified by any transit node.

   An intermediate node that supports the LSP_Attributes object and the
   Attributes Flags TLV, and also recognizes the "Contiguous LSP" bit,
   but cannot support contiguous TE LSPs, MUST send a Path Error message
   with an error code "Routing Problem"/"Contiguous LSP type not
   supported" if it receives a Path message with this bit set.

   If an intermediate node receiving a Path message with the "Contiguous
   LSP" bit set in the Flags field of the LSP_Attributes, recognizes the
   object, the TLV, and the bit and also supports the desired contiguous
   LSP behavior, then it MUST signal a contiguous LSP. If the node is a
   domain border node, or if the node expands a loose hop in the ERO, it

MUST include an RRO Attributes subobject in the RRO of the
corresponding Resv message (if such an object is present) with the
"Contiguous LSP" bit set to report its behavior.

Domain border LSRs MUST support and act on the setting of the
"Contiguous LSP" flag.

However, if the intermediate node supports the LSP_Attributes object
but does not recognize the Attributes Flags TLV, or supports the TLV
but does not recognize this "Contiguous LSP" bit, then it MUST
forward the object unmodified.

The choice of action by an ingress node that receives a PathErr when
requesting the use of a contiguous LSP is out of the scope of this
document, but may include the computation of an alternate path.

5. Protection and Recovery of Inter-Domain TE LSPs

The procedures described in Sections 3 and 4 MUST be applied to all
inter-domain TE LSPs, including bypass tunnels, detour LSPs
[RFC4090], and segment recovery LSPs [RFC4873]. This means that these
LSPs will also be subjected to ERO processing, policies, path
computation, etc.

Note also that the paths for these backup LSPs needs to be either
pre-configured, computed and signaled with the protected LSP, or
computed on-demand at the PLR. Just as with any inter-domain TE LSP,
the ERO may comprise of strict or loose hops, and will depend on the
TE visibility of the computation point into the subsequent domain.

If loose hops are required created in the path of the backup LSP, ERO
expansion will be required at some point along the path: probably at
a domain border node. In order that the backup path remains disjoint
from the protected LSP(s) the node performing the ERO expansion must
be provided with the path of the protected LSPs between the PLR and
the MP. This information can be gathered from the RROs of the
protected LSPs and is signaled in the DETOUR object for Fast Reroute
[RFC4090], and using route exclusion [RFC4874] for other protection
schemes.

5.1. Fast Recovery Support Using MPLS-TE Fast Reroute (FRR)

[RFC4090] describes two methods for local protection for a
packet TE LSP in case of link, SRLG, or node failure. This section
describes how these mechanisms work with the proposed signaling
solutions for inter-domain TE LSP setup.

5.1.1. Failure Within a Domain (Link or Node Failure)

The mode of operation of MPLS-TE Fast Reroute to protect a
contiguous, stitched or nested TE LSP within a domain is identical to
the existing procedures described in [RFC4090]. Note that, in the
case of nesting or stitching, the end-to-end LSP is automatically
protected by the protection operation performed on the H-LSP or
stitching segment LSP.

No protocol extensions are required.

5.1.2. Failure of a Link at a Domain Borders

This cases arises where two domains are connected by a TE link. In
this case each domain has its own domain border node, and these two
nodes are connected by the TE link. An example of this case is where
the ASBRs of two ASes are connected by a TE link.

A contiguous LSP can be backed up using any PLR and MP, but if the
LSP uses stitching or nesting in either of the connected domains, the
PLR and MP MUST be domain border nodes for those domains. It will be
usual to attempt to use the local (connected by the failed link)
domain border nodes as the PLR and MP.

To protect an inter-domain link with MPLS-TE Fast Reroute, a set of
backup tunnels must be configured or dynamically computed between the
PLR and MP such that they are diversely routed from the protected
inter-domain link and the protected inter-domain LSPs.

Each protected inter-domain LSP using the protected inter-domain TE
link must be assigned to an NHOP bypass tunnel that is diverse from
the protected LSP. Such an NHOP bypass tunnel can be selected by
analyzing the RROs in the Resv messages of the available bypass
tunnels and the protected TE LSP. It may be helpful to this process
if the extensions defined in [RFC4561] are used to clearly
distinguish nodes and links in the RROs.

5.1.3. Failure of a Border Node

Two border node failure cases exist. If the domain border falls on a
link as described in the previous section, the border node at either
end of the link may fail. Alternatively, if the border falls on a
border node (as is the case with IGP areas) that single border node
may fail.

It can be seen that if stitching or nesting are used, the failed node
will be the start or end (or both) or a stitching segment LSP or
H-LSP in which case, protection must be provided to the far end of
stitching segment or H-LSP. Thus, where one of these two techniques
is in use, the PLR will be the upstream domain entry point in the

case of the failure of the domain exit point, and the MP will be the
downstream domain exit point in the case of the failure of the
domain entry point. Where the domain border falls at a single domain
border node, both cases will apply.

If the contiguous LSP mechanism is in use, normal selection of the
PLR and MP can be applied and any node within the domains may be used
to fill these roles.

As before, selection of a suitable backup tunnel (in this case an
NNHOP backup) must consider the paths of the backed up LSPs and the
available NNHOP tunnels by examination of their RROs.

Note that where the PLR is not immediately upstream of the failed
node, error propagation time may be delayed unless some mechanism
such as [BFD-MPLS] is implemented, or unless direct reporting, such
as through the GMPLS Notify message [RFC3473] is employed.

5.2. Protection and Recovery of GMPLS LSPs

[RFC4873] describes GMPLS based segment recovery. This allows
protection against a span failure, a node failure, or failure over
any particular portion of a network used by an LSP.

The domain border failure cases described in Section 5.1 may also
occur in GMPLS networks (including packet networks) and can be
protected against using segment protection without any additional
protocol extensions.

Note that if loose hops are used in the construction of the working
and protection paths signaled for segment protection then care is
required to keep these paths disjoint. If the paths are signaled
incrementally then route exclusion [RFC4874] may be used to ensure
that the paths are disjoint. Otherwise a coordinated path computation
technique such as that offered by cooperating Path Computation
Elements [RFC4655] can provide suitable paths.

6. Re-Optimization of Inter-Domain TE LSPs

Re-optimization of a TE LSP is the process of moving the LSP from the
current path to a more prefered path. This involves the determination
of the preferred path and make-before-break signaling procedures
[RFC3209] to minimize traffic disruption.

Re-optimization of an inter-domain TE LSP may require a new path in
more than one domain.

The nature of the inter-domain LSP setup mechanism defines how
re-optimization can be applied. If the LSP is contiguous then the
signaling of the make-before-break process MUST be initiated by the

   ingress node as defined in [RFC3209]. But if the re-optimization is
   limited to a change in path within one domain (that is, if there is
   no change to the domain border nodes) and nesting or stitching are in
   use, the H-LSP or stitching segment may be independently re-optimized
   within the domain without impacting the end-to-end LSP.

   In all cases, however, the ingress LSR may wish to exert control and
   coordination over the re-optimization process. For example, a transit
   domain may be aware of the potential for reoptimization, but not
   bother because it is not worried by the level of service being
   provided across the domain. But the cummulative effect on the
   end-to-end LSP may cause the head-end to worry and trigger an
   end-to-end reoptimization request (of course, the transit domain may
   choose to ignore the request).

   Another benefit to end-to-end reoptimization over per-domain
   reoptimization for non-contiguous inter-domain LSPs is that
   per-domain re-optimization is restricted to preserve the domain entry
   and exit points (since to do otherwise would break the LSP!). But
   end-to-end reoptimization is more flexible and can select new domain
   border LSRs.

   There may be different cost benefit analysis considerations to choose
   between end-to-end reoptimization and per-domain reoptimization. The
   greater the number of hops involved in the reoptimization, the higher
   the risk of traffic disruption. The shorter the segment reoptimized,
   the lower the chance of making a substantial improvement on the
   qulaity of the end-to-end LSP. Administrative policies should be
   applied in this area with care.

   [RFC4736] describes mechanisms that allow:

   - The ingress node to request each node with a loose next hop to
     re-evaluate the current path in order to search for a more optimal
     path.

   - A node with a loose next hop to inform the ingress node that a
     better path exists.

   These mechanisms SHOULD be used for re-optimization of a contiguous
   inter-domain TE LSP.

   Note that end-to-end reoptimization may involve a non-local
   modification and that might select new entry / exit points. In this
   case, we can observe that local reoptimization is more easily and
   flexibly achieved using nesting or stitching. Further, the "locality
   principle" (i.e., the idea of keeping information only where it is
   needed) is best achieved using stitching or nesting. That said, a
   contiguous LSP can easily be modified to take advantage of local

reoptimizations (as defined in [RFC4736]) even if this would require

   the dissemination of information and the invocation of signaling
   outside the local domain.

7. Backward Compatibility

   The procedures in this document are backward compatible with existing
   deployments.

   - Ingress LSRs are not required to support the extensions in this
     document to provision intra-domain LSPs. The default behavior by
     transit LSRs that receive a Path message that does not have the
     "Contiguous LSP" bit set in the Attributes Flags TLV of the
     LSP_Attribtes object or does not even have the object present is
     to allow all modes of inter-domain TE LSP, so back-level ingress
     LSRs are able to initiate inter-domain LSPs.

   - Transit, non-border LSRs are not required to perform any special
     processing and will pass the LSP_Attributes object onwards
     unmodified according to the rules of [RFC2205]. Thus back-level
     transit LSRs are fully supported.

   - Domain border LSRs will need to be upgraded before inter-domain
     TE LSPs are allowed. This is because of the need to establish
     policy, administrative, and security controls before permitting
     inter-domain LSPs to be signaled across a domain border. Thus
     legacy domain border LSRs do not need to be considered.

   The RRO additions in this document are fully backward compatible.

8. Security Considerations

   RSVP does not currently provide for automated key management.
   [RFC4107] states a requirement for mandatory automated key management
   under certain situations. There is work starting in the IETF to
   define improved authentication including automated key management for
   RSVP. Implementations and deployments of RSVP should pay attention to
   any capabilities and requirements that are outputs from this ongoing
   work.

   A separate document is being prepared to examine the security aspects
   of RSVP-TE signaling with special reference to multi-domain
   scenarios [MPLS-GMPLS-SEC]. [RFC4726] provides an overview of the
   requirements for security in an MPLS-TE or GMPLS multi-domain
   environment.

   Before electing to utilise inter-domain signaling for MPLS-TE, the
   administrators of neighboring domains MUST satisfy themselves of the
   existence of a suitable trust relationship between the domains. In
   the absence of such a relationship, the administrators SHOULD decide
   not to deploy inter-domain signaling, and SHOULD disable RSVP-TE on

   any inter-domain interfaces.

   When signaling an inter-domain RSVP-TE LSP, an operator MAY make use
   of the security features already defined for RSVP-TE [RFC3209]. This
   may require some coordination between the domains to share the keys
   (see [RFC2747] and [RFC3097]), and care is required to ensure that
   the keys are changed sufficiently frequently. Note that this may
   involve additional synchronization, should the domain border nodes
   be protected with FRR, since the MP and PLR should also share the
   key.

   For an inter-domain TE LSP, especially when it traverses different
   administrative or trust domains, the following mechanisms SHOULD be
   provided to an operator (also see [RFC4216]):

     1) A way to enforce policies and filters at the domain borders
        to process the incoming inter-domain TE LSP setup requests
        (Path messages) based on certain agreed trust and service
        levels/contracts between domains. Various LSP attributes
        such as bandwidth, priority, etc. could be part of such a
        contract.

     2) A way for the operator to rate-limit LSP setup requests
        or error notifications from a particular domain.

     3) A mechanism to allow policy-based outbound RSVP message
        processing at the domain border node, which may involve
        filtering or modification of certain addresses in RSVP
        objects and messages.

   Additionally, an operator may wish to reduce the signaling
   interactions between domains to improve security. For example, the
   operator might not trust the neighboring domain to supply correct or
   trustable restart information [RSVP-RESTART] and might ensure that
   the availablity of restart function is not configured in the Hello
   message exchange across the domain border. Thus, suitable
   configuration MUST be provided in an RSVP-TE implementation to
   enable the operator to control optional protocol features that may be
   considered a security risk.

   Some examples of the policies described above are as follows:

     A) An operator may choose to implement some kind of ERO filtering
        policy on the domain border node to disallow or ignore hops
        within the domain from being identified in the ERO of an
        incoming Path message. That is, the policy is that a node
        outside the domain cannot specify the path of the LSP inside the
        domain. The domain border LSR can make implement this policy in
        one of two ways:

      - It can reject the Path message.

      - It can ignore the hops in the ERO that lie within the domain.

  B) In order to preserve confidentiality of network topology, an
     operator may choose to disallow recording of hops within the
     domain in the RRO or may choose to filter out certain recorded
     RRO addresses at the domain border node.

  C) An operator may require the border node to modify the addresses
     of certain messages like PathErr or Notify originated from hops
     within the domain.

  D) In the event of a path computation failure, an operator may
     require the border node to silently discard the Path message
     instead of returning a PathErr. This is because a Path message
     could be interpretted as a network probe, and a PathErr provides
     information about the network capabilities and policies.

Note that the detailed specification of such policies and their
implementation are outside the scope of this document.

OAM mechanisms including [BFD-MPLS] and [RFC4379] are commonly used
to verify he connectivity of end-to-end LSPs and to trace their
paths. Where the LSPs are inter-domain LSPs, such OAM techniques MAY
require to be intercepted or modified at domain borders, or to be
passed transparently across domains. Further discussion of this topic
can be found in [INTERAS-PING] and [MPLS-GMPLS-SEC].

9. IANA Considerations

IANA is requested to make the codepoint allocations described in the
following sections.

9.1. Attribute Flags for LSP_Attributes Object

A new bit is to be allocated from the "Attributes Flags" sub-registry
of the "RSVP TE Parameters" registry.

Bit	Name	Path message	Resv message	RRO sub-object
XX	Contiguous LSP	Yes	No	Yes

The value XX is to be defined by IANA. A value of 4 is suggested.

9.2. New Error Codes

New RSVP error codes/values are required to be allocated from the
"Error Codes and Globally-Defined Error Value Sub-Codes" sub-registry
of the "RSVP Parameters" registry.

For the existing error code "Policy control failure" (value 2), two
new error values are suggested as follows. The values are suggested
and are for IANA determination.

    103 = Inter-domain policy failure
    104 = Inter-domain explicit route rejected

For the existing error code "Routing Problem" (value 24), two new
error values are suggested as follows. The values are suggested and
are for IANA determination.

    21 = Contiguous LSP type not supported
    22 = ERO conflicts with inter-domain signaling method

10. Acknowledgements

The authors would like to acknowledge the input and helpful comments
from Kireeti Kompella on various aspects discussed in the document.
Deborah Brungard and Dimitri Papdimitriou provided thorough reviews.

Thanks to Sam Hartman for detailed discussions of the security
considerations.

11. References

11.1. Normative References

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

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

[RFC3209] Awduche, D., et al, "Extensions to RSVP for LSP Tunnels",
          RFC 3209, December 2001.

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

[RFC4206] Kompella K., Rekhter Y., "LSP Hierarchy with Generalized
          MPLS TE", RFC 4206, October 2005.

   [RFC4420] Farrel, A., et al, "Encoding of Attributes for
             Multiprotocol Label Switching (MPLS) Label Switched Path
             (LSP) Establishment Using RSVP-TE", RFC 4420, February
             2006.

   [LSP-STITCHING] Ayyangar, A., Kompella, K., and Vasseur, JP., "Label
             Switched Path Stitching with Generalized MPLS Traffic
             Engineering", draft-ietf-ccamp-lsp-stitching, (work in
             progress).

11.2. Informative References

   [RFC2747] Baker, F., Lindell, B., and Talwar, B., "RSVP Cryptographic
             Authentication", RFC 2747, January 2000.

   [RFC3097] Braden, R., and Zhang, L., "RSVP Cryptographic
             Authentication -- Updated Message Type Value", RFC 3097,
             April 2001.

   [RFC4090] Ping Pan, et al, "Fast Reroute Extensions to RSVP-TE
             for LSP Tunnels", RFC 4090, May 2005.

   [RFC4105] LeRoux, JL., Vasseur, JP., Boyle, J., et al, "Requirements
             for Inter-Area MPLS Traffic Engineering", RFC 4105, June
             2005.

   [RFC4107] Bellovin, S. and Housley, R., "Guidelines for Cryptographic
             Key Management", BCP 107, RFC 4107, June 2005.

   [RFC4208] G. Swallow et al, "GMPLS UNI: RSVP-TE Support for the
             Overlay Model", RFC 4208, October 2005.

   [RFC4216] Zhang, R., et al, "MPLS Inter-Autonomous System (AS)
             Traffic Engineering (TE) Requirements", RFC 4216, November
             2005.

   [RFC4379] Kompella, K. and G. Swallow, "Detecting Multi-Protocol
             Label Switched (MPLS) Data Plane Failures", RFC 4379,
             February 2006.

   [RFC4561] Vasseur, JP., Ali, Z., and Sivabalan, S., "Definition of a
             Record Route Object (RRO) Node-Id Sub-Object", RFC 4561,
             June 2006.

   [RFC4655] Ash, G., Farrel, A., and Vasseur, JP., "Path Computation
             Element (PCE)-Based Architecture", RFC 4655, August 2006.

   [RFC4726] Farrel, A., Vasseur, J.P., and Ayyangar, A., "A
             Framework for Inter-Domain Multiprotocol Label Switching
             Traffic Engineering", RFC 4726, November 2006.

   [RFC4736] Vasseur, JP., et al, "Reoptimization of Multiprotocol Label
            Switching (MPLS) Traffic Engineering (TE) Loosely Routed
            Label Switched Path (LSP)", RFC 4736, November 2006.

   [RFC4873] Berger, L., et al, "GMPLS Based Segment Recovery",
            RFC 4873, May 2007.

   [RFC4874] Lee, CY., Farrel, A., and De Cnodder, S., "Exclude Routes -
            Extension to Resource ReserVation Protocol-Traffic
            Engineering (RSVP-TE)", RFC 4874, April 2007.

   [RFC4920] Farrel, A., et al, "Crankback Signaling Extensions for
            MPLS and GMPLS RSVP-TE", RFC 4920, July 2007.

   [BFD-MPLS] Aggarwal, R., et al, "BFD For MPLS LSPs",
            draft-ietf-bfd-mpls, (work in progress).

   [INTERAS-PING] Nadeau, T., and Swallow, G., "Detecting MPLS Data
            Plane Failures in Inter-AS and inter-provider Scenarios",
            draft-nadeau-mpls-interas-lspping, work in progress.

   [INTER-DOMAIN-PD-PATH-COMP] Vasseur, JP., Ayyangar, A., and
            Zhang, R., "A Per-domain path computation method for
            computing Inter-domain Traffic Engineering (TE) Label
            Switched Paths (LSPs)", draft-ietf-ccamp-inter-domain-pd-
            path-comp, (work in progress).

   [MPLS-GMPLS-SEC] Luyuan Fang, et al., "Security Framework for MPLS
            and GMPLS Networks", draft-ietf-mpls-mpls-and-gmpls-
            security-framework, work in progress.

   [RSVP-RESTART] Satyanarayana, A., and Rahman, R., "Extensions to
            GMPLS RSVP Graceful Restart", draft-ietf-ccamp-rsvp-
            restart-ext, work in progress.

12. Authors' Addresses

   Adrian Farrel
   Old Dog Consulting

   E-mail: adrian@olddog.co.uk

   Arthi Ayyangar
   Juniper Networks, Inc.
   1194 N.Mathilda Ave
   Sunnyvale, CA 94089
   USA

   E-mail: arthi@juniper.net

    Jean Philippe Vasseur
    Cisco Systems, Inc.
    300 Beaver Brook Road
    Boxborough , MA - 01719
    USA

    Email: jpv@cisco.com

Intellectual Property Statement

   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.

Disclaimer of Validity

   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, THE IETF TRUST 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.

Copyright Statement

   Copyright (C) The IETF Trust (2007).

   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.