Network Working Group
Request for Comments: 5298
Category: Informational

T. Takeda, Ed.
NTT
A. Farrel, Ed.
Old Dog Consulting
Y. Ikejiri
NTT Communications
JP. Vasseur
Cisco Systems, Inc.

August 2008

Analysis of Inter-Domain Label Switched Path (LSP) Recovery

Status of This Memo

Abstract

Protection and recovery are important features of service offerings in Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS) networks. Increasingly, MPLS and GMPLS networks are being extended from single domain scope to multi-domain environments.

Various schemes and processes have been developed to establish Label Switched Paths (LSPs) in multi-domain environments. These are discussed in RFC 4726: "A Framework for Inter-Domain Multiprotocol Label Switching Traffic Engineering".

This document analyzes the application of these techniques to protection and recovery in multi-domain networks. The main focus for this document is on establishing end-to-end diverse Traffic Engineering (TE) LSPs in multi-domain networks.

Table of Contents

1. Introduction

Protection and recovery in Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS) networks are described in [RFC4428]. These are important features for service delivery in MPLS and GMPLS networks.

MPLS and GMPLS networks were originally limited to single domain environments. Increasingly, multi-domain MPLS and GMPLS networks are being considered, where a domain is considered to be any collection of network elements within a common sphere of address management or path computational responsibility. Examples of such domains include Interior Gateway Protocol (IGP) areas and Autonomous Systems (ASes).

[RFC4726] provides a framework for inter-domain MPLS and GMPLS traffic engineering. It introduces and discusses the various schemes and processes to establish Label Switched Paths (LSPs) in multi-domain environments.

However, protection and recovery introduce additional complexities to LSP establishment. Protection LSPs are generally required to be path diverse from the working LSPs that they protect. Achieving this is particularly challenging in multi-domain environments because no single path computation or planning point is capable of determining path diversity for both paths from one end to the other.

This document analyzes various schemes to realize MPLS and GMPLS LSP recovery in multi-domain networks. The main focus for this document is on establishing end-to-end diverse Traffic Engineering (TE) LSPs in multi-domain networks.

1.1. Terminology

The reader is assumed to be familiar with the terminology for LSP recovery set out in [RFC4427], and with the terms introduced in [RFC4726] that provides a framework for inter-domain Label Switched Path (LSP) setup. Key terminology may also be found in [RFC4216] that sets out requirements for inter-AS MPLS traffic engineering.

The following key terms from those sources are used within this document.

- Domain: See [RFC4726]. A domain is considered to be any collection of network elements within a common sphere of address management or path computational responsibility. Note that nested domains continue to be out of scope. Section 1.2 provides additional details.

- Working LSP: See [RFC4427]. The working LSP transports normal user traffic. Note that the term LSP and TE LSP will be used interchangeably.

- Recovery LSP: See [RFC4427]. The recovery LSP transports normal user traffic when the working LSP fails. The recovery LSP may also carry user traffic even when the working LSP is operating normally and transporting the user traffic (e.g., 1+1 protection). The recovery LSP is sometimes referred to as a protecting LSP.

- Diversity: See [RFC4726]. Diversity means the relationship of multiple LSPs, where those LSPs do not share some specific type of resource (e.g., link, node, or shared risk link group (SRLG)). Diversity is also referred to as disjointness.

  Diverse LSPs may be used for various purposes, such as load-balancing and recovery. In this document, the recovery aspect of diversity, specifically the end-to-end diversity of two traffic engineering (TE) LSPs, is the focus. The two diverse LSPs are referred to as the working LSP and recovery LSP.

- Confidentiality: See [RFC4216]. Confidentiality refers to the protection of information about the topology and resources of one domain from visibility by people or applications outside that domain.

1.2. Domain

In order to fully understand the issues addressed in this document, it is necessary to carefully define and scope the term "domain".

As defined in [RFC4726], a domain is considered to be any collection of network elements within a common sphere of address management or path computational responsibility. Examples of such domains include IGP areas and Autonomous Systems. Networks accessed over the GMPLS User-to-Network Interface (UNI) [RFC4208], and Layer One Virtual Private Networks (L1VPNs) [RFC4847] are special cases of multi-domain networks.

Example motivations for using more than one domain include administrative separation, scalability, and the construction of domains for the purpose of providing protection. These latter "protection domains" offer edge-to-edge protection facilities for spans or segments of end-to-end LSPs.

As described in [RFC4726], there could be TE parameters (such as color and priority) whose meanings are specific to each domain. In such scenarios, in order to set up inter-domain LSPs, mapping functions may be needed to transform the TE parameters based on policy agreements between domain administrators.

1.3. Document Scope

This document analyzes various schemes to realize MPLS and GMPLS LSP recovery in multi-domain networks. It is based on the existing framework for multi-domain LSP setup [RFC4726]. Note that this document does not prevent the development of additional techniques where appropriate (i.e., additional to the ones described in this document). In other words, this document shows how the existing techniques can be applied.

There are various recovery techniques for LSPs. For TE LSPs, the major techniques are end-to-end recovery [RFC4872], local protection such as Fast Reroute (FRR) [RFC4090] (in packet switching environments), and segment recovery [RFC4873] (in GMPLS).

The main focus of this document is the analysis of diverse TE LSP setup schemes that can be used in the context of end-to-end recovery. The methodology is to show different combinations of functional elements such as path computation and signaling techniques.

[RFC4105] and [RFC4216] describe requirements for diverse LSPs. There are various types of diversity, and this document focuses on node, link, and shared risk link group (SRLG) diversity.

Recovery LSPs may be used for 1+1 protection, 1:1 protection, or shared mesh restoration. However, the requirements for path diversity, the ways to compute diverse paths, and the signaling of these TE LSPs are common across all uses. These topics are the main scope of this document.

Note that diverse LSPs may be used for various purposes in addition to recovery. An example is for load-balancing, so as to limit the traffic disruption to a portion of the traffic flow in case of a single node failure. This document does not preclude use of diverse LSP setup schemes for other purposes.

The following are beyond the scope of this document.

- Analysis of recovery techniques other than the use of link, node, or SRLG diverse LSPs (see Section 1.4).

- Details of maintenance of diverse LSPs, such as re-optimization and Operations and Maintenance (OAM).

- Comparative evaluation of LSP setup schemes.

1.4. Note on Other Recovery Techniques

Fast Reroute and segment recovery in multi-domain networks are described in Section 5.4 of [RFC4726], and a more detailed analysis is provided in Section 5 of [RFC5151]. This document does not cover any additional analysis for Fast Reroute and segment recovery in multi-domain networks.

The recovery type of an LSP or service may change at a domain boundary. That is, the recovery type could remain the same within one domain, but might be different in the next domain or on the connections between domains. This may be due to the capabilities of each domain, administrative policies, or to topology limitations. An example is where protection at the domain boundary is provided by link protection on the inter-domain links, but where protection within each domain is achieved through segment recovery. This mixture of protection techniques is beyond the scope of this document.

Domain diversity (that is, the selection of paths that have only the ingress and egress domains in common) may be considered as one form of diversity in multi-domain networks, but this is beyond the scope of this document (see Section 2.2).

1.5. Signaling Options

There are three signaling options for establishing inter-domain TE LSPs: nesting, contiguous LSPs, and stitching [RFC4726]. The description in this document of diverse LSP setup is agnostic in relation to the signaling option used, unless otherwise specified.

Note that signaling option considerations for Fast Reroute and segment recovery are described in [RFC5151].

2. Diversity in Multi-Domain Networks

This section describes some assumptions about achieving path diversity in multi-domain networks.

2.1. Multi-Domain Network Topology

Figures 1 and 2 show examples of multi-domain network topologies. In Figure 1, domains are connected in a linear topology. There are multiple paths between nodes A and L, but all of them cross domain#1-domain#2-domain#3 in that order.

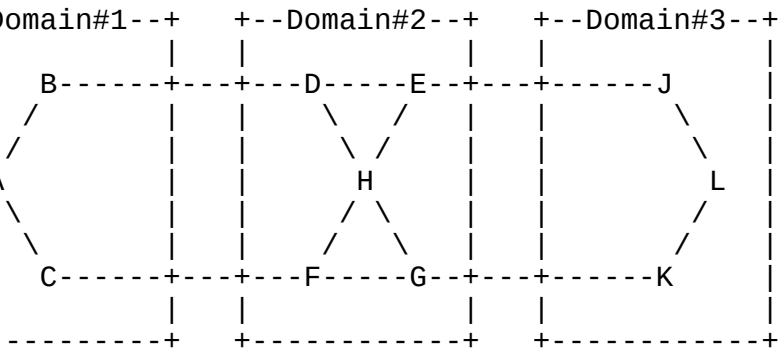

Figure 1: Linear Domain Connectivity

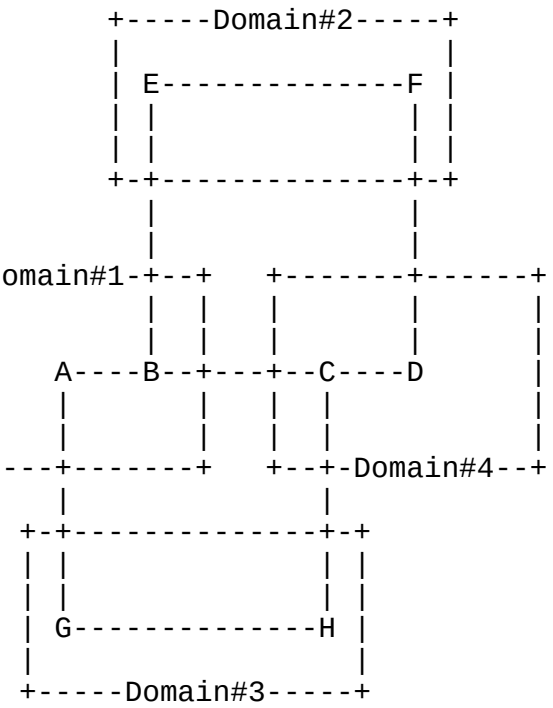

Figure 2: Meshed Domain Connectivity

In Figure 2, domains are connected in a mesh topology. There are multiple paths between nodes A and D, and these paths do not cross the same domains. If inter-domain connectivity forms a mesh, domain-level routing is required (even for the unprotected case). This is tightly coupled with the next-hop domain resolution/discovery mechanisms used in IP networks.

In this document, we assume that domain-level routing is given via configuration, policy, or some external mechanism, and that this is not part of the path computation process described later in this document.

Domain-level routing may also allow "domain re-entry" where a path re-enters a domain that it has previously exited (e.g., domain#X-domain#Y-domain#X). This requires specific considerations when confidentiality (described in Section 3.2) is required, and is beyond the scope of this document.

Furthermore, the working LSP and the recovery LSP may or may not be routed along the same set of domains in the same order. In this document, we assume that the working LSP and recovery LSP follow the same set of domains in the same order (via configuration, policy or some external mechanism). That is, we assume that the domain mesh topology is reduced to a linear domain topology for each pair of working and recovery LSPs.

In summary,

- There is no assumption about the multi-domain network topology. For example, there could be more than two domain boundary nodes or inter-domain links (links connecting a pair of domain boundary nodes belonging to different domains).

- It is assumed that in a multi-domain topology, the sequence of domains that the working LSP and the recovery LSP follow must be the same and is pre-configured.

- Domain re-entry is out of scope and is not considered further.

2.2. Note on Domain Diversity

As described in Section 1.4, domain diversity means the selection of paths that have only the ingress and egress domains in common. This may provide enhanced resilience against failures, and is a way to ensure path diversity for most of the path of the LSP.

In Section 2.1, we assumed that the working LSP and the recovery LSP follow the same set of domains in the same order. Under this assumption, domain diversity cannot be achieved. However, by relaxing this assumption, domain diversity could be achieved, e.g., by either of the following schemes.

- Consider domain diversity as a special case of SRLG diversity (i.e., assign an SRLG ID to each domain).

- Configure domain-level routing to provide domain-diverse paths (e.g., by means of AS_PATH in BGP).

Further details of the operation of domain diversity are beyond the scope of this document.

3. Factors to Consider

3.1. Scalability versus Optimality

As described in [RFC4726], scalability and optimality are two conflicting objectives. Note that the meaning of optimality differs depending on each network operation. Some examples of optimality in the context of diverse LSPs are:

- Minimizing the sum of their cost while maintaining diversity.

- Restricting the difference of their costs (for example, so as to minimize the delay difference after switch-over) while maintaining diversity.

By restricting TE information distribution to only within each domain (and not across domain boundaries) as required by [RFC4105] and [RFC4216], it may not be possible to compute an optimal path. As such, it might not be possible to compute diverse paths, even if they exist. However, if we assume domain-level routing is given (as discussed in Section 2), it would be possible to compute diverse paths using specific computation schemes, if such paths exist. This is discussed further in Section 4.

3.2. Key Concepts

Three key concepts to classify various diverse LSP computation and setup schemes are presented below.

o With or without confidentiality

- Without confidentiality

  It is possible to specify a path across multiple domains in signaling (by means of the Resource Reservation Protocol-TE (RSVP-TE) Explicit Route Object (ERO)), and to obtain record of the inter-domain path used (by means of the RSVP-TE Record Route Object (RRO)). In this case, it is clear that one domain has control over the path followed in another domain, and that the path actually used in one domain is visible from within another domain.

  Examples of this configuration are multi-area networks, and some forms of multi-AS networks (especially within the same provider). In these cases, there is no requirement for confidentiality.

- With confidentiality

  Where confidentiality of domain topology and operational policy is required, it is not possible to specify or obtain a full path across other domains. Partial paths may be specified and reported using domain identifiers or the addresses of domain border routers in the EROs and RROs.

  Examples of this configuration are some forms of multi-AS networks (especially inter-provider networks), GMPLS-UNI networks, and L1VPNs.

o Multi-domain path computation, per-domain path computation, and inter-domain collaborative path computation

- Multi-domain path computation

  If a single network element can see the topology of all domains along the path, it is able to compute a full end-to-end path. Clearly, this is only possible where confidentiality is not required.

  Such a network element might be the head-end Label Switching Router (LSR), a Path Computation Element (PCE) [RFC4655], or a Network Management System (NMS). This mode of path computation is discussed in Sections 4 and 5.

- Per-domain path computation

  The path of an LSP may be computed domain-by-domain as LSP signaling progresses through the network. This scheme requires ERO expansion in each domain to construct the next segment of the path toward the destination. The establishment of unprotected LSPs in this way is covered in [RFC5152].

- Inter-domain collaborative path computation

  In this scheme, path computation is typically done before signaling and uses communication between cooperating PCEs. An example of such a scheme is Backward Recursive Path Computation (BRPC) [BRPC].

  It is possible to combine multiple path computation techniques (including using a different technique for the working and recovery LSPs), but details are beyond the scope of this document.

o Sequential path computation or simultaneous path computation

- Sequential path computation

  The path of the working LSP is computed without considering the recovery LSP, and then the path of the recovery LSP is computed. This scheme is applicable when the recovery LSP is added later as a change to the service grade, but may also be used when both the working and recovery LSPs are established from the start.

  Using this approach, it may not be possible to find diverse paths for the LSPs in "trap" topologies. Furthermore, such sequential path computation approaches reduce the likelihood of selecting an "optimal" solution with regards to a specific objective function.

- Simultaneous path computation

  The path of the working LSP and the path of the recovery LSP are computed simultaneously. In this scheme, it is possible to avoid trap conditions and it may be more possible to achieve an optimal result.

Note that LSP setup, with or without confidentiality, depends on per-domain configuration. The choice of per-domain path computation or inter-domain collaborative path computation, and the choice between sequential path computation or simultaneous path computation can be determined for each individual pair of working/recovery LSPs.

The analysis of various diverse LSP setup schemes is described in Sections 4 and 5, based on the concepts set out above.

Some other considerations, such as network topology-specific considerations, addressing considerations, and SRLG diversity are described in Sections 6, 7, and 8.

4. Diverse LSP Setup Schemes without Confidentiality

This section examines schemes for diverse LSP setup based on different path computation techniques assuming that there is no requirement for domain confidentiality. Section 5 makes a similar examination of schemes where domain confidentiality is required.

4.1. Management Configuration

[RFC4726] describes this path computation technique where the full explicit paths for the working and recovery LSPs are specified by a management application at the head-end, and no further computation or signaling considerations are needed.

4.2. Head-End Path Computation (with Multi-Domain Visibility)

Section 3.2.1 [RFC4726] describes this path computation technique. The full explicit paths for the working and recovery LSPs are computed at the head-end either by the head-end itself or by a PCE. In either case, the computing entity has full TE visibility across multiple domains and no further computation or signaling considerations are needed.

4.3. Per-Domain Path Computation

Sections 3.2.2, 3.2.3, and 3.3 of [RFC4726] describe this path computation technique. More detailed procedures are described in [RFC5152].

In this scheme, the explicit paths of the working and recovery LSPs are specified as the complete strict paths through the source domain followed by either of the following:

- The complete list of boundary LSRs or domain identifiers (e.g., AS numbers) along the paths.

- The LSP destination.

Thus, in order to navigate each domain, the path must be expanded at each domain boundary, i.e., per-domain. This path computation is performed by the boundary LSR or by a PCE on its behalf.

There are two schemes for establishing diverse LSPs using per-domain computation. These are described Sections 4.3.1 and 4.3.2.

4.3.1. Sequential Path Computation

As previously noted, the main issue with sequential path computation is that diverse paths cannot be guaranteed. Where a per-domain path computation scheme is applied, the computation of second path needs to be aware of the path used by the first path in order that path diversity can be attempted.

The RSVP-TE EXCLUDE_ROUTE Object (XRO) [RFC4874] can be used when the second path is signaled to report the details of the first path. It should be noted that the PRIMARY_PATH_ROUTE Object defined in [RFC4872] for end-to-end protection is not intended as a path exclusion mechanism and should not be used for this purpose.

The process for sequential path computation is as follows:

- The working LSP is established using per-domain path computation as described in [RFC5152]. The path of the working LSP is available at the head-end through the RSVP-TE RRO since domain confidentiality is not required.

- The path of the recovery LSP across the first domain is computed excluding the resources used by the working LSP within that domain. If a PCE is used, the resources to be avoided can be passed to the PCE using the Exclude Route Object (XRO) extensions to the PCE Protocol [PCEP-XRO], [PCEP].

- The recovery LSP is now signaled across the first domain as usual, but the path of the working LSP is also conveyed in an RSVP-TE XRO. The XRO lists nodes, links and SRLGs that must be avoided by the LSP being signaled, and its contents are copied from the RRO of the working LSP.

- At each subsequent domain boundary, a segment of the path of the recovery LSP can be computed across the new domain excluding the resources indicated in the RSVP-TE XRO.

This scheme cannot guarantee to establish diverse LSPs (even if they could exist) because the first (working) LSP is established without consideration of the need for a diverse recovery LSP. It is possible to modify the path of the working LSP using the crankback techniques [RFC4920] if the setup of the recovery LSP is blocked or if some resources are shared.

Note that, even if a solution is found, the degree of optimality of the solution (i.e., of the set of diverse TE LSPs) might not be maximal.

4.3.2. Simultaneous Path Computation

Simultaneous path computation gives a better likelihood of finding a pair of diverse paths as the diversity requirement forms part of the computation process. In per-domain path computation mechanisms, there are several aspects to consider.

Simultaneous path computation means that the paths of the working and recovery LSPs are computed at the same time. Since we are considering per-domain path computation, these two paths must be computed at the boundary of each domain.

The process for simultaneous path computation is as follows:

- The ingress LSR (or a PCE) computes a pair of diverse paths across the first domain. If a PCE is used, PCEP offers the ability to request disjoint paths.

- The working LSP is signaled across the first domain as usual, but must carry with it the requirement for a disjoint recovery LSP and the information about the path already computed for the recovery LSP across the first domain. In particular, the domain boundary node used by the recovery LSP must be reported.

- Each domain boundary router, in turn, computes a pair of disjoint paths across the next domain. The working LSP is signaled as usual, and the computed path of the recovery LSP is collected in the signaling messages.

- When the working LSP has been set up, the full path of the recovery LSP is returned to the head-end LSR in the signaling messages for the working LSP. This allows the head-end LSR to signal the recovery LSP using a full path without the need for further path computation.

Note that the signaling protocol mechanisms do not currently exist to collect the path of the recovery LSP during the signaling of the working LSP. Definition of protocol mechanisms are beyond the scope of this document, but it is believed that such mechanisms would be simple to define and implement.

Note also that the mechanism described is still not able to guarantee the selection of diverse paths even where they exist since, when domains are multiply interconnected, the determination of diverse

end-to-end paths may depend on the correct selection of inter-domain links. Crankback [RFC4920] may also be used in combination with this scheme to improve the chances of success.

Note that even if a solution is found, the degree of optimality of the solution (i.e., set of diverse TE LSPs) might not be maximal.

4.4. Inter-Domain Collaborative Path Computation

Collaborative path computation requires the cooperation between PCEs that are responsible for different domains. This approach is described in Section 3.4 of [RFC4726]. Backward recursive path computation (BRPC) [BRPC] provides a collaborative path computation technique where the paths of LSPs are fully determined by communication between PCEs before the LSPs are established. Two ways to use BRPC for diverse LSPs are described in the following sections.

4.4.1. Sequential Path Computation

In sequential path computation, the path of the working LSP is computed using BRPC as described in [BRPC]. The path of the recovery LSP is then computed also using BRPC with the addition that the path of the working LSP is explicitly excluded using the XRO route exclusion techniques described in [PCEP-XRO].

In this case, the working LSP could be set up before or after the path of the recovery LSP is computed. In the latter case, the actual path of the working LSP as reported in the RSVP-TE RRO should be used when computing the path of the recovery LSP.

This scheme cannot guarantee to establish diverse LSPs (even if they exist) because the working LSP may block the recovery LSP. In such a scenario, re-optimization of the working and recovery LSPs may be used to achieve fully diverse paths.

4.4.2. Simultaneous Path Computation

In simultaneous path computation, the PCEs collaborate to compute the paths of both the working and the recovery LSPs at the same time. Since both LSPs are computed in a single pass, the LSPs can be signaled simultaneously of sequentially according to the preference of the head-end LSR.

Collaborative simultaneous path computation is achieved using the Synchronization Vector (SVEC) object in the PCE Protocol [PCEP]. This object allows two computation requests to be associated and made dependent. The coordination of multiple computation requests within the BRPC mechanism is not described in [BRPC]. It is believed that it is possible to specify procedures for such coordination, but the development of new procedures is outside the scope of this document.

This scheme can guarantee to establish diverse LSPs where they are possible, assuming that domain-level routing is pre-determined as described in Section 2. Furthermore, the computed set of TE LSPs can be guaranteed to be optimal with respect to some objective functions.

5. Diverse LSP Setup Schemes with Confidentiality

In the context of this section, the term confidentiality applies to the protection of information about the topology and resources present within one domain from visibility by people or applications outside that domain. This includes, but is not limited to, recording of LSP routes, and the advertisements of TE information. The confidentiality does not apply to the protection of user traffic.

Diverse LSP setup schemes with confidentiality are similar to ones without confidentiality. However, several additional mechanisms are needed to preserve confidentiality. Examples of such mechanisms are:

- Path key: A path key is used in place of a segment of the path of an LSP when the LSP is signaled, when the path of the LSP is reported by signaling, or when the LSP’s path is generated by a PCE. This allows the exact path of the LSP to remain confidential through the substitution of "confidential path segments" (CPSs) by these path keys.

  [PCE-PATH-KEY] describes how path keys can be used by PCEs to preserve path confidentiality, and [RSVP-PATH-KEY] explains how path keys are used in signaling. Note that if path keys are signaled in RSVP-TE EROs they must be expanded so that the signaling can proceed. This expansion normally takes place when the first node in the CPS is reached. The expansion of the path key would normally be carried out by the PCE that generated the key, and for that reason, the path key contains an identifier of the PCE (the PCE-ID).

- LSP segment: LSP segments can be pre-established across domains according to some management policy. The LSP segments can be used to support end-to-end LSPs as hierarchical LSPs [RFC4206] or as LSP stitching segments [RFC5150].

The end-to-end LSPs are signaled indicating just the series of domains or domain border routers. Upon entry to each domain, an existing trans-domain LSP is selected and used to carry the end-to-end LSP across the domain.

Note that although the LSP segments are described as being pre-established, they could be set up on demand on receipt of the request for the end-to-end LSP at the domain border.

It is also worth noting that in schemes that result in a single contiguous end-to-end LSP (without LSP tunneling or stitching), the same concept of LSP segments may apply provided that ERO expansion is applied at domain boundaries and that the path of the LSP is not reported in the RSVP-TE RRO.

These techniques may be applied directly or may require protocol extensions depending on the specific diverse LSP setup schemes described below. Note that in the schemes below, the paths of the working and recovery LSPs are not impacted by the confidentiality requirements.

5.1. Management Configuration

Although management systems may exist that can determine end-to-end paths even in the presence of domain confidentiality, the full paths cannot be provided to the head-end LSR for LSP signaling as this would break the confidentiality requirements.

Thus, for LSPs that are configured through management applications, the end-to-end path must either be constructed using LSP segments that cross the domains, or communicated to the head-end LSR with the use of path keys.

5.2. Head-End Path Computation (with Multi-Domain Visibility)

It is not possible for the head-end LSR to compute the full end-to-end path of an inter-domain LSP when domain confidentiality is in use because the LSR will not have any TE information about the other domains.

5.3. Per-Domain Path Computation

Per-domain path computation for working and recovery LSPs is practical with domain confidentiality. As when there are no confidentiality restrictions, we can separate the cases of sequential and simultaneous path computation.

5.3.1. Sequential Path Computation

In sequential path computation, we can assume that the working LSP has its path computed and is set up using the normal per-domain technique as described in [RFC5152]. However, because of confidentiality issues, the full path of the working LSP is not returned in the signaling messages and is not available to the head-end LSR.

To compute a disjoint path for the recovery LSP, each domain border node needs to know the path of the working LSP within the domain to which it provides entry. This is easy for the ingress LSR as it has access to the RSVP-TE RRO within first domain. In subsequent domains, the process requires that the path of the working LSP is somehow made available to the domain border router as the recovery LSP is signaled. Note that the working and recovery LSPs do not use the same border routers if the LSPs are node or SRLG diverse.

There are several possible mechanisms to achieve this.

- Path keys could be used in the RRO for the working LSP. As the signaling messages are propagated back towards the head-end LSR, each domain border router substitutes a path key for the segment of the working LSP's path within its domain. Thus, the RRO received at the head-end LSR consists of the path within the initial domain followed by a series of path keys.

  When the recovery LSP is signaled, the path keys can be included in the ERO as exclusions. Each domain border router in turn can expand the path key for its domain and know which resources must be avoided. PCEP provides a protocol that can be used to request the expansion of the path key from the domain border router used by the working LSP, or from some third party such as a PCE.

- Instead of using path keys, each confidential path segment in the RRO of the working LSP could be encrypted by the domain border routers. These encrypted segments would appear as exclusions in the ERO for the recovery LSP and could be decrypted by the domain border routers.

  No mechanism currently exists in RSVP-TE for this function, which would probably assume a domain-wide encryption key.

- The identity of the working LSP could be included in the XRO of the recovery LSP to indicate that a disjoint path must be found.

  This option could require a simple extension to the current XRO specification [RFC4874] to allow LSP identifiers to be included.

It could also use the Association Object [RFC4872] to identify the working LSP.

This scheme would also need a way for a domain border router to find the path of an LSP within its domain. An efficient way to achieve this would be to also include the domain border router used by the working LSP in the signaling for the recovery LSP, but other schemes based on management applications or stateful PCEs might also be developed.

Clearly, the details of this alternative have not been specified.

5.3.2. Simultaneous Path Computation

In per-domain simultaneous path computation the path of the recovery LSP is computed at the same time as the working LSP (i.e., as the working LSP is signaled). The computed path of the recovery LSP is propagated to the head-end LSR as part of the signaling process for the working LSP, but confidentiality must be maintained, so the full path cannot be returned. There are two options as follows.

- LSP segment: As the signaling of the working LSP progresses and the path of the recovery LSP is computed domain-by-domain, trans-domain LSPs can be set up for use by the recovery LSP. When the recovery LSP is signaled, it will pick up these LSP segments and use them for tunneling or stitching.

  This mechanism needs coordination through the management plane between domain border routers so that a router on the working path can request the establishment of an LSP segment for use by the protection path. This could be achieved through the TE MIB modules [RFC3812], [RFC4802].

  Furthermore, when the recovery LSP is signaled it needs to be sure to pick up the correct LSP segment. Therefore, some form of LSP segment identifier needs to be reported in the signaling of the working LSP and propagated in the signaling of the recovery LSP. Mechanisms for this do not currently exist, but would be relatively simple to construct.

  Alternatively, the LSP segments could be marked as providing protection for the working LSP. In this case, the recovery LSP can be signaled with the identifier of the working LSP using the Association Object [RFC4872] enabling the correct LSP segments to be selected.

- Path key: The path of the recovery LSP can be determined and returned to the head-end LSR just described in Section 4.4.2, but each CPS is replaced by a path key. As the recovery path is signaled each path key is expanded, domain-by-domain to achieve the correct path. This requires that the entity that computes the path of the recovery LSP (domain border LSR or PCE) is stateful.

5.4 Inter-Domain Collaborative Path Computation

Cooperative collaboration between PCEs is also applicable when domain confidentiality is required.

5.4.1. Sequential Path Computation

In sequential cooperative path computation, the path of the working LSP is determined first using a mechanism such as BRPC. Since domain confidentiality is in use, the path returned may contain path keys.

When the path of the recovery LSP is computed (which may be before or after the working LSP is signaled) the path exclusions supplied to the PCE and exchanged between PCEs must use path keys as described in [PCEP-XRO].

5.4.2. Simultaneous Path Computation

As described in Section 4.4.2, diverse path computation can be requested using the PCEP SVEC Object [PCEP], and BRPC could be extended for inter-domain diverse path computation. However, to conform to domain confidentiality requirements, path keys must be used in the paths returned by the PCEs and signaled by RSVP-TE.

Note that the LSP segment approach may not be applicable here because a path cannot be determined until BRPC procedures are completed.

6. Network Topology Specific Considerations

In some specific network topologies the schemes for setting up diverse LSPs could be significantly simplified.

For example, consider the L1VPN or GMPLS UNI case. This may be viewed as a linear sequence of three domains where the first and last domains contain only a single node each. In such a scenario, no BRPC procedures are needed, because there is no need for path computation in the first and last domains even if the source and destination nodes are multi-homed.

7. Addressing Considerations

All of the schemes described in this document are applicable when a single address space is used across all domains.

There may also be cases where private address spaces are used within some of the domains. This problem is similar to the use of domain confidentiality since the ERO and RRO are meaningless outside a domain even if they are available, and the problem can be solved using the same techniques.

8. Note on SRLG Diversity

The schemes described in this document are applicable when the nodes and links in different domains belong to different SRLGs, which is normally the case.

However, it is possible that nodes or links in different domains belong to the same SRLG. That is, an SRLG may span domain boundaries. In such cases, in order to establish SRLG diverse LSPs, several considerations are needed:

- Record of the SRLGs used by the working LSP.

- Indication of a set of SRLGs to exclude in the computation of the recovery LSP’s path.

In this case, there is a conflict between any requirement for domain confidentiality, and the requirement for SRLG diversity. One of the requirements must be compromised.

Furthermore, SRLG IDs may be assigned independently in each domain, and might not have global meaning. In such a scenario, some mapping functions are necessary, similar to the mapping of other TE parameters mentioned in Section 1.2.

9. Security Considerations

The core protocols used to achieve the procedures described in this document are RSVP-TE and PCEP. These protocols include policy and authentication capabilities as described in [RFC3209] and [PCEP]. Furthermore, these protocols may be operated using more advanced security features such as IPsec [RFC4301] and TLS [RFC4346].

Security may be regarded as particularly important in inter-domain deployments and serious consideration should be given to applying the available security techniques, as described in the documents referenced above and as set out in [RFC4726].

Additional discussion of security considerations for MPLG/GMPLS networks can be found in [SECURITY-FW].

This document does not of itself require additional security measures and does not modify the trust model implicit in the protocols used. Note, however, that domain confidentiality (that is the confidentiality of the topology and path information from within any one domain) is an important consideration in this document, and a significant number of the mechanisms described in this document are designed to preserve domain confidentiality.

10. References

10.1. Normative References

[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.

[RFC4216] Zhang, R., Ed., and J.-P. Vasseur, Ed., "MPLS Inter-Autonomous System (AS) Traffic Engineering (TE) Requirements", RFC 4216, November 2005.

[RFC4427] Mannie, E., Ed., and D. Papadimitriou, Ed., "Recovery (Protection and Restoration) Terminology for Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4427, March 2006.

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

10.2. Informative References

[RFC3812] Srinivasan, C., Viswanathan, A., and T. Nadeau, "Multiprotocol Label Switching (MPLS) Traffic Engineering (TE) Management Information Base (MIB)", RFC 3812, June 2004.

[RFC4090] Pan, P., Ed., Swallow, G., Ed., and A. Atlas, Ed., "Fast Reroute Extensions to RSVP-TE for LSP Tunnels", RFC 4090, May 2005.

[RFC4105] Le Roux, J.-L., Ed., Vasseur, J.-P., Ed., and J. Boyle, Ed., "Requirements for Inter-Area MPLS Traffic Engineering", RFC 4105, June 2005.

[RFC4206] Kompella, K. and Y. Rekhter, "Label Switched Paths (LSP) Hierarchy with Generalized Multi-Protocol Label Switching (GMPLS) Traffic Engineering (TE)", RFC 4206, October 2005.

[RFC4208] Swallow, G., Drake, J., Ishimatsu, H., and Y. Rekhter, "Generalized Multiprotocol Label Switching (GMPLS) User-Network Interface (UNI): Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Support for the Overlay Model", RFC 4208, October 2005.

[RFC4301] Kent, S. and K. Seo, "Security Architecture for the Internet Protocol", RFC 4301, December 2005.

[RFC4346] Dierks, T. and E. Rescorla, "The Transport Layer Security (TLS) Protocol Version 1.1", RFC 4346, April 2006.

[RFC4428] Papadimitriou, D., Ed., and E. Mannie, Ed., "Analysis of Generalized Multi-Protocol Label Switching (GMPLS)-based Recovery Mechanisms (including Protection and Restoration)", RFC 4428, March 2006.

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

[RFC4802] Nadeau, T., Ed., and A. Farrel, Ed., "Generalized Multiprotocol Label Switching (GMPLS) Traffic Engineering Management Information Base", RFC 4802, February 2007.

[RFC4847] Takeda, T., Ed., "Framework and Requirements for Layer 1 Virtual Private Networks", RFC 4847, April 2007.

[RFC4872] Lang, J., Ed., Rekhter, Y., Ed., and D. Papadimitriou, Ed., "RSVP-TE Extensions in Support of End-to-End Generalized Multi-Protocol Label Switching (GMPLS) Recovery", RFC 4872, May 2007.

[RFC4873] Berger, L., Bryskin, I., Papadimitriou, D., and A. Farrel, "GMPLS Segment Recovery", RFC 4873, May 2007.

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

[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.

[RFC5150] Ayyangar, A., Kompella, K., Vasseur, JP., and A. Farrel, "Label Switched Path Stitching with Generalized Multiprotocol Label Switching Traffic Engineering (GMPLS TE)", RFC 5150, February 2008.

[RFC5151] Farrel, A., Ed., Ayyangar, A., and JP. Vasseur, "Inter-Domain MPLS and GMPLS Traffic Engineering -- Resource Reservation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 5151, February 2008.

[RFC5152] Vasseur, JP., Ed., Ayyangar, A., Ed., and R. Zhang, "A Per-Domain Path Computation Method for Establishing Inter-Domain Traffic Engineering (TE) Label Switched Paths (LSPs)", RFC 5152, February 2008.

[BRPC] Vasseur, JP., Ed., Zhang, R., Bitar, N., and JL. Le Roux, "A Backward Recursive PCE-Based Computation (BRPC) Procedure to Compute Shortest Inter-Domain Traffic Engineering Label Switched Paths", Work in Progress, April 2008.

[PCE-PATH-KEY] Bradford, R., Vasseur, JP., and A. Farrel, "Preserving Topology Confidentiality in Inter-Domain Path Computation Using a Key-Based Mechanism", Work in Progress, May 2008.

[PCEP] Vasseur, JP., Ed., and JL. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol (PCEP)", Work in Progress, March 2008.

[PCEP-XRO] Oki, E., Takeda, T., and A. Farrel, "Extensions to the Path Computation Element Communication Protocol (PCEP) for Route Exclusions", Work in Progress, July 2008.

[RSVP-PATH-KEY] Bradford, R., Vasseur, JP., and A. Farrel, "RSVP Extensions for Path Key Support", Work in Progress, May 2008.

[SECURITY-FW] Fang, L., Ed., " Security Framework for MPLS and GMPLS Networks", Work in Progress, July 2008.

11. Acknowledgments

The authors would like to thank Eiji Oki, Ichiro Inoue, Kazuhiro Fujihara, Dimitri Papadimitriou, and Meral Shirazipour for valuable comments. Deborah Brungard provided useful advice about the text.

Authors' Addresses

Tomonori Takeda
NTT Network Service Systems Laboratories, NTT Corporation
3-9-11, Midori-Cho
Musashino-Shi, Tokyo 180-8585 Japan
EMail : takeda.tomonori@lab.ntt.co.jp

Yuichi Ikejiri
NTT Communications Corporation
Tokyo Opera City Tower 3-20-2 Nishi Shinjuku, Shinjuku-ku
Tokyo 163-1421, Japan
EMail: y.ikejiri@ntt.com

Adrian Farrel
Old Dog Consulting
EMail: adrian@olddog.co.uk

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

Full Copyright Statement

Intellectual Property