Network Working Group
Request for Comments: 4726
Category: Informational

A. Farrel
Old Dog Consulting
J.-P. Vasseur
Cisco Systems, Inc.
A. Ayyangar
Nuova Systems
November 2006

A Framework for Inter-Domain Multiprotocol Label Switching Traffic Engineering

Status of This Memo

Copyright Notice

Abstract

This document provides a framework for establishing and controlling Multiprotocol Label Switching (MPLS) and Generalized MPLS (GMPLS) Traffic Engineered (TE) Label Switched Paths (LSPs) in multi-domain networks.

For the purposes of this document, 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).

Table of Contents

1. Introduction

The Traffic Engineering Working Group has developed requirements for inter-area and inter-AS Multiprotocol Label Switching (MPLS) Traffic Engineering in [RFC4105] and [RFC4216].

Various proposals have subsequently been made to address some or all of these requirements through extensions to the Resource Reservation Protocol Traffic Engineering extensions (RSVP-TE) and to the Interior Gateway Protocols (IGPs) (i.e., Intermediate System to Intermediate System (IS-IS) and OSPF).

This document introduces the techniques for establishing Traffic Engineered (TE) Label Switched Paths (LSPs) across multiple domains. In this context and within the remainder of this document, we consider all source-based and constraint-based routed LSPs and refer to them interchangeably as "TE LSPs" or "LSPs".

The functional components of these techniques are separated into the mechanisms for discovering reachability and TE information, for computing the paths of LSPs, and for signaling the LSPs. Note that the aim of this document is not to detail each of those techniques, which are covered in separate documents referenced from the sections of this document that introduce the techniques, but rather to propose a framework for inter-domain MPLS Traffic Engineering.

Note that in the remainder of this document, the term "MPLS Traffic Engineering" is used equally to apply to MPLS and Generalized MPLS (GMPLS) traffic. Specific issues pertaining to the use of GMPLS in inter-domain environments (for example, policy implications of the use of the Link Management Protocol [RFC4204] on inter-domain links) are covered in separate documents such as [GMPLS-AS].

For the purposes of this document, 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. Wholly or partially overlapping domains (e.g., path computation sub-domains of areas or ASes) are not within the scope of this document.

1.1. Nested Domains

Nested domains are outside the scope of this document. It may be that some domains that are nested administratively or for the purposes of address space management can be considered as adjacent domains for the purposes of this document; however, the fact that the domains are nested is then immaterial. In the context of MPLS TE, domain A is considered to be nested within domain B if domain A is

wholly contained in domain B, and domain B is fully or partially aware of the TE characteristics and topology of domain A.

2. Signaling Options

Three distinct options for signaling TE LSPs across multiple domains are identified. The choice of which options to use may be influenced by the path computation technique used (see section 3), although some path computation techniques may apply to multiple signaling options. The choice may further depend on the application to which the TE LSPs are put and the nature, topology, and switching capabilities of the network.

A comparison of the usages of the different signaling options is beyond the scope of this document and should be the subject of a separate applicability statement.

2.1. LSP Nesting

Hierarchical LSPs form a fundamental part of MPLS [RFC3031] and are discussed in further detail in [RFC4206]. Hierarchical LSPs may optionally be advertised as TE links. Note that a hierarchical LSP that spans multiple domains cannot be advertised in this way because there is no concept of TE information that spans domains.

Hierarchical LSPs can be used in support of inter-domain TE LSPs. In particular, a hierarchical LSP may be used to achieve connectivity between any pair of Label Switching Routers (LSRs) within a domain. The ingress and egress of the hierarchical LSP could be the edge nodes of the domain in which case connectivity is achieved across the entire domain, or they could be any other pair of LSRs in the domain.

The technique of carrying one TE LSP within another is termed LSP nesting. A hierarchical LSP may provide a TE LSP tunnel to transport (i.e., nest) multiple TE LSPs along a common part of their paths. Alternatively, a TE LSP may carry (i.e., nest) a single LSP in a one-to-one mapping.

The signaling trigger for the establishment of a hierarchical LSP may be the receipt of a signaling request for the TE LSP that it will carry, or may be a management action to "pre-engineer" a domain to be crossed by TE LSPs that would be used as hierarchical LSPs by the traffic that has to traverse the domain. Furthermore, the mapping (inheritance rules) between attributes of the nested and the hierarchical LSPs (including bandwidth) may be statically pre-configured or, for on-demand hierarchical LSPs, may be dynamic

according to the properties of the nested LSPs. Even in the dynamic case, inheritance from the properties of the nested LSP(s) can be complemented by local or domain-wide policy rules.

Note that a hierarchical LSP may be constructed to span multiple domains or parts of domains. However, such an LSP cannot be advertised as a TE link that spans domains. The end points of a hierarchical LSP are not necessarily on domain boundaries, so nesting is not limited to domain boundaries.

Note also that the Interior/Exterior Gateway Protocol (IGP/EGP) routing topology is maintained unaffected by the LSP connectivity and TE links introduced by hierarchical LSPs even if they are advertised as TE links. That is, the routing protocols do not exchange messages over the hierarchical LSPs, and LSPs are not used to create routing adjacencies between routers.

During the operation of establishing a nested LSP that uses a hierarchical LSP, the SENDER_TEMPLATE and SESSION objects remain unchanged along the entire length of the nested LSP, as do all other objects that have end-to-end significance.

2.2. Contiguous LSP

A single contiguous LSP is established from ingress to egress in a single signaling exchange. No further LSPs are required to be established to support this LSP so that hierarchical or stitched LSPs are not needed.

A contiguous LSP uses the same Session/LSP ID along the whole of its path (that is, at each LSR). The notions of "splicing" together different LSPs or of "shuffling" Session or LSP identifiers are not considered.

2.3. LSP Stitching

LSP Stitching is described in [STITCH]. In the LSP stitching model, separate LSPs (referred to as a TE LSP segments) are established and are "stitched" together in the data plane so that a single end-to-end Label Switched Path is achieved. The distinction is that the component LSP segments are signaled as distinct TE LSPs in the control plane. Each signaled TE LSP segment has a different source and destination.

LSP stitching can be used in support of inter-domain TE LSPs. In particular, an LSP segment may be used to achieve connectivity between any pair of LSRs within a domain. The ingress and egress of the LSP segment could be the edge nodes of the domain in which case

connectivity is achieved across the entire domain, or they could be any other pair of LSRs in the domain.

The signaling trigger for the establishment of a TE LSP segment may be the establishment of the previous TE LSP segment, the receipt of a setup request for TE LSP that it plans to stitch to a local TE LSP segment, or a management action.

LSP segments may be managed and advertised as TE links.

2.4. Hybrid Methods

There is nothing to prevent the mixture of signaling methods described above when establishing a single, end-to-end, inter-domain TE LSP. It may be desirable in this case for the choice of the various methods to be reported along the path, perhaps through the Record Route Object (RRO).

If there is a desire to restrict which methods are used, this must be signaled as described in the next section.

2.5. Control of Downstream Choice of Signaling Method

Notwithstanding the previous section, an ingress LSR may wish to restrict the signaling methods applied to a particular LSP at domain boundaries across the network. Such control, where it is required, may be achieved by the definition of appropriate new flags in the SESSION-ATTRIBUTE object or the Attributes Flags TLV of the LSP_ATTRIBUTES object [RFC4420]. Before defining a mechanism to provide this level of control, the functional requirement to control the way in which the network delivers a service must be established. Also, due consideration must be given to the impact on interoperability since new mechanisms must be backwards compatible, and care must be taken to avoid allowing standards-conformant implementations that each supports a different functional subset in such a way that they are not capable of establishing LSPs.

3. Path Computation Techniques

The discussion of path computation techniques within this document is limited significantly to the determination of where computation may take place and what components of the full path may be determined.

The techniques used are closely tied to the signaling methodologies described in the previous section in that certain computation techniques may require the use of particular signaling approaches and vice versa.

Any discussion of the appropriateness of a particular path computation technique in any given circumstance is beyond the scope of this document and should be described in a separate applicability statement.

Path computation algorithms are firmly out of the scope of this document.

3.1. Management Configuration

Path computation may be performed by offline tools or by a network planner. The resultant path may be supplied to the ingress LSR as part of the TE LSP or service request, and encoded by the ingress LSR as an Explicit Route Object (ERO) on the Path message that is sent out.

There is no reason why the path provided by the operator should not span multiple domains if the relevant information is available to the planner or the offline tool. The definition of what information is needed to perform this operation and how that information is gathered, is outside the scope of this document.

3.2. Head-End Computation

The head-end, or ingress, LSR may assume responsibility for path computation when the operator supplies part or none of the explicit path. The operator must, in any case, supply at least the destination address (egress) of the LSP.

3.2.1. Multi-Domain Visibility Computation

If the ingress has sufficient visibility of the topology and TE information for all of the domains across which it will route the LSP to its destination, then it may compute and provide the entire path. The quality of this path (that is, its optimality as discussed in section 3.5) can be better if the ingress has full visibility into all relevant domains rather than just sufficient visibility to provide some path to the destination.

Extreme caution must be exercised in consideration of the distribution of the requisite TE information. See section 4.

3.2.2. Partial Visibility Computation

It may be that the ingress does not have full visibility of the topology of all domains, but does have information about the connectedness of the domains and the TE resource availability across the domains. In this case, the ingress is not able to provide a

fully specified strict explicit path from ingress to egress. However, for example, the ingress might supply an explicit path that comprises:

- explicit hops from ingress to the local domain boundary
- loose hops representing the domain entry points across the network
- a loose hop identifying the egress.

Alternatively, the explicit path might be expressed as:

- explicit hops from ingress to the local domain boundary
- strict hops giving abstract nodes representing each domain in turn
- a loose hop identifying the egress.

These two explicit path formats could be mixed according to the information available resulting in different combinations of loose hops and abstract nodes.

This form of explicit path relies on some further computation technique being applied at the domain boundaries. See section 3.3.

As with the multi-domain visibility option, extreme caution must be exercised in consideration of the distribution of the requisite TE information. See section 4.

3.2.3. Local Domain Visibility Computation

A final possibility for ingress-based computation is that the ingress LSR has visibility only within its own domain, and connectivity information only as far as determining one or more domain exit points that may be suitable for carrying the LSP to its egress.

In this case, the ingress builds an explicit path that comprises just:

- explicit hops from ingress to the local domain boundary
- a loose hop identifying the egress.

3.3. Domain Boundary Computation

If the partial explicit path methods described in sections 3.2.2 or 3.2.3 are applied, then the LSR at each domain boundary is responsible for ensuring that there is sufficient path information added to the Path message to carry it at least to the next domain boundary (that is, out of the new domain).

If the LSR at the domain boundary has full visibility to the egress then it can supply the entire explicit path. Note, however, that the ERO processing rules of [RFC3209] state that it should only update the ERO as far as the next specified hop (that is, the next domain boundary if one was supplied in the original ERO) and, of course, must not insert ERO subobjects immediately before a strict hop.

If the LSR at the domain boundary has only partial visibility (using the definitions of section 3.2.2), it will fill in the path as far as the next domain boundary, and will supply further domain/domain boundary information if not already present in the ERO.

If the LSR at the domain boundary has only local visibility into the immediate domain, it will simply add information to the ERO to carry the Path message as far as the next domain boundary.

Domain boundary path computations are performed independently from each other. Domain boundary LSRs may have different computation capabilities, run different path computation algorithms, apply different sets of constraints and optimization criteria, and so forth, which might result in path segment quality that is unpredictable to and out of the control of the ingress LSR. A solution to this issue lies in enhancing the information signaled during LSP setup to include a larger set of constraints and to include the paths of related LSPs (such as diverse protected LSPs) as described in [GMPLS-E2E].

It is also the case that paths generated on domain boundaries may produce loops. Specifically, the paths computed may loop back into a domain that has already been crossed by the LSP. This may or may not be a problem, and might even be desirable, but could also give rise to real loops. This can be avoided by using the recorded route (RRO) to provide exclusions within the path computation algorithm, but in the case of lack of trust between domains it may be necessary for the RRO to indicate the previously visited domains. Even this solution is not available where the RRO is not available on a Path message. Note that when an RRO is used to provide exclusions, and a loop-free path is found to be not available by the computation at a downstream border node, crankback [CRANKBACK] may enable an upstream border node to select an alternate path.

3.4. Path Computation Element

The computation techniques in sections 3.2 and 3.3 rely on topology and TE information being distributed to the ingress LSR and those LSRs at domain boundaries. These LSRs are responsible for computing paths. Note that there may be scaling concerns with distributing the required information; see section 4.

An alternative technique places the responsibility for path computation with a Path Computation Element (PCE) [RFC4655]. There may be either a centralized PCE, or multiple PCEs (each having local visibility and collaborating in a distributed fashion to compute an end-to-end path) across the entire network and even within any one domain. The PCE may collect topology and TE information from the same sources as would be used by LSRs in the previous paragraph, or though other means.

Each LSR called upon to perform path computation (and even the offline management tools described in section 3.1) may abdicate the task to a PCE of its choice. The selection of PCE(s) may be driven by static configuration or the dynamic discovery.

3.4.1. Multi-Domain Visibility Computation

A PCE may have full visibility, perhaps through connectivity to multiple domains. In this case, it is able to supply a full explicit path as in section 3.2.1.

3.4.2. Path Computation Use of PCE When Preserving Confidentiality

Note that although a centralized PCE or multiple collaborative PCEs may have full visibility into one or more domains, it may be desirable (e.g., to preserve topology confidentiality) that the full path not be provided to the ingress LSR. Instead, a partial path is supplied (as in section 3.2.2 or 3.2.3), and the LSRs at each domain boundary are required to make further requests for each successive segment of the path.

In this way, an end-to-end path may be computed using the full network capabilities, but confidentiality between domains may be preserved. Optionally, the PCE(s) may compute the entire path at the first request and hold it in storage for subsequent requests, or it may recompute each leg of the path on each request or at regular intervals until requested by the LSRs establishing the LSP.

It may be the case that the centralized PCE or the collaboration between PCEs may define a trust relationship greater than that normally operational between domains.

3.4.3. Per-Domain Computation Elements

A third way that PCEs may be used is simply to have one (or more) per domain. Each LSR within a domain that wishes to derive a path across the domain may consult its local PCE.

This mechanism could be used for all path computations within the domain, or specifically limited to computations for LSPs that will leave the domain where external connectivity information can then be restricted to just the PCE.

3.5. Optimal Path Computation

There are many definitions of an optimal path depending on the constraints applied to the path computation. In a multi-domain environment, the definitions are multiplied so that an optimal route might be defined as the route that would be computed in the absence of domain boundaries. Alternatively, another constraint might be applied to the path computation to reduce or limit the number of domains crossed by the LSP.

It is easy to construct examples that show that partitioning a network into domains, and the resulting loss or aggregation of routing information may lead to the computation of routes that are other than optimal. It is impossible to guarantee optimal routing in the presence of aggregation / abstraction / summarization of routing information.

It is beyond the scope of this document to define what is an optimum path for an inter-domain TE LSP. This debate is abdicated in favor of requirements documents and applicability statements for specific deployment scenarios. Note, however, that the meaning of certain computation metrics may differ between domains (see section 5.6).

4. Distributing Reachability and TE Information

Traffic Engineering information is collected into a TE Database (TED) on which path computation algorithms operate either directly or by first constructing a network graph.

The path computation techniques described in the previous section make certain demands upon the distribution of reachability information and the TE capabilities of nodes and links within domains as well as the TE connectivity across domains.

Currently, TE information is distributed within domains by additions to IGPs [RFC3630], [RFC3784].

In cases where two domains are interconnected by one or more links (that is, the domain boundary falls on a link rather than on a node), there should be a mechanism to distribute the TE information associated with the inter-domain links to the corresponding domains. This would facilitate better path computation and reduce TE-related crankbacks on these links.

Where a domain is a subset of an IGP area, filtering of TE information may be applied at the domain boundary. This filtering may be one way or two way.

Where information needs to reach a PCE that spans multiple domains, the PCE may snoop on the IGP traffic in each domain, or play an active part as an IGP-capable node in each domain. The PCE might also receive TED updates from a proxy within the domain.

It is possible that an LSR that performs path computation (for example, an ingress LSR) obtains the topology and TE information of not just its own domain, but other domains as well. This information may be subject to filtering applied by the advertising domain (for example, the information may be limited to Forwarding Adjacencies (FAs) across other domains, or the information may be aggregated or abstracted).

Before starting work on any protocols or protocol extensions to enable cross-domain reachability and TE advertisement in support of inter-domain TE, the requirements and benefits must be clearly established. This has not been done to date. Where any cross-domain reachability and TE information needs to be advertised, consideration must be given to TE extensions to existing protocols such as BGP, and how the information advertised may be fed to the IGPs. It must be noted that any extensions that cause a significant increase in the amount of processing (such as aggregation computation) at domain boundaries, or a significant increase in the amount of information flooded (such as detailed TE information) need to be treated with extreme caution and compared carefully with the scaling requirements expressed in [RFC4105] and [RFC4216].

5. Comments on Advanced Functions

This section provides some non-definitive comments on the constraints placed on advanced MPLS TE functions by inter-domain MPLS. It does not attempt to state the implications of using one inter-domain technique or another. Such material is deferred to appropriate applicability statements where statements about the capabilities of existing or future signaling, routing, and computation techniques to deliver the functions listed should be made.

5.1. LSP Re-Optimization

Re-optimization is the process of moving a TE LSP from one path to another, more preferable path (where no attempt is made in this document to define "preferable" as no attempt was made to define "optimal"). Make-before-break techniques are usually applied to ensure that traffic is disrupted as little as possible. The Shared

Explicit style is usually used to avoid double booking of network resources.

Re-optimization may be available within a single domain. Alternatively, re-optimization may involve a change in route across several domains or might involve a choice of different transit domains.

Re-optimization requires that all or part of the path of the LSP be re-computed. The techniques used may be selected as described in section 3, and this will influence whether the whole or part of the path is re-optimized.

The trigger for path computation and re-optimization may be an operator request, a timer, information about a change in availability of network resources, or a change in operational parameters (for example, bandwidth) of an LSP. This trigger must be applied to the point in the network that requests re-computation and controls re-optimization and may require additional signaling.

Note also that where multiple mutually-diverse paths are applied end-to-end (i.e., not simply within protection domains; see section 5.5) the point of calculation for re-optimization (whether it is PCE, ingress, or domain entry point) needs to know all such paths before attempting re-optimization of any one path. Mutual diversity here means that a set of computed paths has no commonality. Such diversity might be link, node, Shared Risk Link Group (SRLG), or even domain disjointedness according to circumstances and the service being delivered.

It may be the case that re-optimization is best achieved by recomputing the paths of multiple LSPs at once. Indeed, this can be shown to be most efficient when the paths of all LSPs are known, not simply those LSPs that originate at a particular ingress. While this problem is inherited from single domain re-optimization and is out of scope within this document, it should be noted that the problem grows in complexity when LSPs wholly within one domain affect the re-optimization path calculations performed in another domain.

5.2. LSP Setup Failure

When an inter-domain LSP setup fails in some domain other than the first, various options are available for reporting and retrying the LSP.

In the first instance, a retry may be attempted within the domain that contains the failure. That retry may be attempted by nodes wholly within the domain, or the failure may be referred back to the LSR at the domain boundary.

If the failure cannot be bypassed within the domain where the failure occurred (perhaps there is no suitable alternate route, perhaps rerouting is not allowed by domain policy, or perhaps the Path message specifically bans such action), the error must be reported back to the previous or head-end domain.

Subsequent repair attempts may be made by domains further upstream, but will only be properly effective if sufficient information about the failure and other failed repair attempts is also passed back upstream [CRANKBACK]. Note that there is a tension between this requirement and that of topology confidentiality although crankback aggregation may be applicable at domain boundaries.

Further attempts to signal the failed LSP may apply the information about the failures as constraints to path computation, or may signal them as specific path exclusions [EXCLUDE].

When requested by signaling, the failure may also be systematically reported to the head-end LSR.

5.3. LSP Repair

An LSP that fails after it has been established may be repaired dynamically by re-routing. The behavior in this case is either like that for re-optimization, or for handling setup failures (see previous two sections). Fast Reroute may also be used (see below).

5.4. Fast Reroute

MPLS Traffic Engineering Fast Reroute ([RFC4090]) defines local protection schemes intended to provide fast recovery (in 10s of msecs) of fast-reroutable packet-based TE LSPs upon link/SRLG/Node failure. A backup TE LSP is configured and signaled at each hop, and activated upon detecting or being informed of a network element failure. The node immediately upstream of the failure (called the PLR, or Point of Local Repair) reroutes the set of protected TE LSPs onto the appropriate backup tunnel(s) and around the failed resource.

In the context of inter-domain TE, there are several different failure scenarios that must be analyzed. Provision of suitable solutions may be further complicated by the fact that [RFC4090] specifies two distinct modes of operation referred to as the "one to one mode" and the "facility back-up mode".

The failure scenarios specific to inter-domain TE are as follows:

- Failure of a domain edge node that is present in both domains. There are two sub-cases:
  - The Point of Local Repair (PLR) and the Merge Point (MP) are in the same domain.
  - The PLR and the MP are in different domains.
- Failure of a domain edge node that is only present in one of the domains.
- Failure of an inter-domain link.

Although it may be possible to apply the same techniques for Fast Reroute (FRR) to the different methods of signaling inter-domain LSPs described in section 2, the results of protection may be different when it is the boundary nodes that need to be protected, and when they are the ingress and egress of a hierarchical LSP or stitched LSP segment. In particular, the choice of PLR and MP may be different, and the length of the protection path may be greater. These uses of FRR techniques should be explained further in applicability statements or, in the case of a change in base behavior, in implementation guidelines specific to the signaling techniques.

Note that after local repair has been performed, it may be desirable to re-optimize the LSP (see section 5.1). If the point of re-optimization (for example, the ingress LSR) lies in a different domain to the failure, it may rely on the delivery of a PathErr or Notify message to inform it of the local repair event.

It is important to note that Fast Reroute techniques are only applicable to packet switching networks because other network technologies cannot apply label stacking within the same switching type. Segment protection [GMPLS-SEG] provides a suitable alternative that is applicable to packet and non-packet networks.

5.5. Comments on Path Diversity

Diverse paths may be required in support of load sharing and/or protection. Such diverse paths may be required to be node diverse, link diverse, fully path diverse (that is, link and node diverse), or SRLG diverse.

Diverse path computation is a classic problem familiar to all graph theory majors. The problem is compounded when there are areas of "private knowledge" such as when domains do not share topology

information. The problem can be resolved more efficiently (e.g., avoiding the "trap problem") when mutually resource disjoint paths can be computed "simultaneously" on the fullest set of information.

That being said, various techniques (out of the scope of this document) exist to ensure end-to-end path diversity across multiple domains.

Many network technologies utilize "protection domains" because they fit well with the capabilities of the technology. As a result, many domains are operated as protection domains. In this model, protection paths converge at domain boundaries.

Note that the question of SRLG identification is not yet fully answered. There are two classes of SRLG:

- those that indicate resources that are all contained within one domain

- those that span domains.

The former might be identified using a combination of a globally scoped domain ID, and an SRLG ID that is administered by the domain. The latter requires a global scope to the SRLG ID. Both schemes, therefore, require external administration. The former is able to leverage existing domain ID administration (for example, area and AS numbers), but the latter would require a new administrative policy.

5.6. Domain-Specific Constraints

While the meaning of certain constraints, like bandwidth, can be assumed to be constant across different domains, other TE constraints (such as resource affinity, color, metric, priority, etc.) may have different meanings in different domains and this may impact the ability to support Diffserv-aware MPLS, or to manage preemption.

In order to achieve consistent meaning and LSP establishment, this fact must be considered when performing constraint-based path computation or when signaling across domain boundaries.

A mapping function can be derived for most constraints based on policy agreements between the domain administrators. The details of such a mapping function are outside the scope of this document, but it is important to note that the default behavior must either be that a constant mapping is applied or that any requirement to apply these constraints across a domain boundary must fail in the absence of explicit mapping rules.

5.7. Policy Control

Domain boundaries are natural points for policy control. There is little to add on this subject except to note that a TE LSP that cannot be established on a path through one domain because of a policy applied at the domain boundary may be satisfactorily established using a path that avoids the demurring domain. In any case, when a TE LSP signaling attempt is rejected due to non-compliance with some policy constraint, this should be reflected to the ingress LSR.

5.8. Inter-Domain Operations and Management (OAM)

Some elements of OAM may be intentionally confined within a domain. Others (such as end-to-end liveness and connectivity testing) clearly need to span the entire multi-domain TE LSP. Where issues of topology confidentiality are strong, collaboration between PCEs or domain boundary nodes might be required in order to provide end-to-end OAM, and a significant issue to be resolved is to ensure that the end-points support the various OAM capabilities.

The different signaling mechanisms described above may need refinements to [RFC4379], [BFD-MPLS], etc., to gain full end-to-end visibility. These protocols should, however, be considered in the light of topology confidentiality requirements.

Route recording is a commonly used feature of signaling that provides OAM information about the path of an established LSP. When an LSP traverses a domain boundary, the border node may remove or aggregate some of the recorded information for topology confidentiality or other policy reasons.

5.9. Point-to-Multipoint

Inter-domain point-to-multipoint (P2MP) requirements are explicitly out of the scope of this document. They may be covered by other documents dependent on the details of MPLS TE P2MP solutions.

5.10. Applicability to Non-Packet Technologies

Non-packet switching technologies may present particular issues for inter-domain LSPs. While packet switching networks may utilize control planes built on MPLS or GMPLS technology, non-packet networks are limited to GMPLS.

On the other hand, some problems such as Fast Reroute on domain boundaries (see section 5.4) may be handled by the GMPLS technique of segment protection [GMPLS-SEG] that is applicable to both packet and non-packet switching technologies.

The specific architectural considerations and requirements for inter-domain LSP setup in non-packet networks are covered in a separate document [GMPLS-AS].

6. Security Considerations

Requirements for security within domains are unchanged from [RFC3209] and [RFC3473], and from [RFC3630] and [RFC3784]. That is, all security procedures for existing protocols in the MPLS context continue to apply for the intra-domain cases.

Inter-domain security may be considered as a more important and more sensitive issue than intra-domain security since in inter-domain traffic engineering control and information may be passed across administrative boundaries. The most obvious and most sensitive case is inter-AS TE.

All of the intra-domain security measures for the signaling and routing protocols are equally applicable in the inter-domain case. There is, however, a greater likelihood of them being applied in the inter-domain case.

Security for inter-domain MPLS TE is the subject of a separate document that analyzes the security deployment models and risks. This separate document must be completed before inter-domain MPLS TE solution documents can be advanced.

Similarly, the PCE procedures [RFC4655] are subject to security measures for the exchange computation information between PCEs and for LSRs that request path computations from a PCE. The requirements for this security (set out in [RFC4657]) apply whether the LSR and PCE (or the cooperating PCEs) are in the same domain or lie across domain boundaries.

It should be noted, however, that techniques used for (for example) authentication require coordination of secrets, keys, or passwords between sender and receiver. Where sender and receiver lie within a single administrative domain, this process may be simple. But where sender and receiver lie in different administrative domains, cross-domain coordination between network administrators will be required in order to provide adequate security. At this stage, it is not proposed that this coordination be provided through an automatic process or through the use of a protocol. Human-to-human

coordination is more likely to provide the required level of confidence in the inter-domain security.

One new security concept is introduced by inter-domain MPLS TE. This is the preservation of confidentiality of topology information. That is, one domain may wish to keep secret the way that its network is constructed and the availability (or otherwise) of end-to-end network resources. This issue is discussed in sections 3.4.2, 5.2, and 5.8 of this document. When there is a requirement to preserve inter-domain topology confidentiality, policy filters must be applied at the domain boundaries to avoid distributing such information. This is the responsibility of the domain that distributes information, and it may be adequately addressed by aggregation of information as described in the referenced sections.

Applicability statements for particular combinations of signaling, routing, and path computation techniques to provide inter-domain MPLS TE solutions are expected to contain detailed security sections.

7. Acknowledgements

The authors would like to extend their warmest thanks to Kireeti Kompella for convincing them to expend effort on this document.

Grateful thanks to Dimitri Papadimitriou, Tomohiro Otani, and Igor Bryskin for their review and suggestions on the text.

Thanks to Jari Arkko, Gonzalo Camarillo, Brian Carpenter, Lisa Dusseault, Sam Hartman, Russ Housley, and Dan Romascanu for final review of the text.

8. Normative References

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

[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.

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

[RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, September 2003.

[RFC3784] Smit, H. and T. Li, "Intermediate System to Intermediate System (IS-IS) Extensions for Traffic Engineering (TE)", RFC 3784, June 2004.

9. Informative References

[BFD-MPLS] Aggarwal, R., Kompella, K., Nadeau, T., and G. Swallow, "BFD For MPLS LSPs", Work in Progress, June 2006.

[CRANKBACK] Farrel, A., et al., "Crankback Signaling Extensions for MPLS Signaling", Work in Progress, May 2005.

[EXCLUDE] Lee, CY., Farrel, A., and DeCnodder, "Exclude Routes - Extension to RSVP-TE", Work in Progress, August 2005.

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

[GMPLS-AS] Otani, T., Kumaki, K., Okamoto, S., and W. Imajuku, "GMPLS Inter-domain Traffic Engineering Requirements", Work in Progress, August 2006.

[GMPLS-E2E] Lang, J.P., Rekhter, Y., and D. Papadimitriou, Editors, "RSVP-TE Extensions in support of End-to-End Generalized Multi-Protocol Label Switching (GMPLS)-based Recovery", Work in Progress, April 2005.

[GMPLS-SEG] Berger, L., Bryskin, I., Papadimitriou, D., and A. Farrel, "GMPLS Based Segment Recovery", Work in Progress, May 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.

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

[RFC4204] Lang, J., "Link Management Protocol (LMP)", RFC 4204, October 2005.

[RFC4216] Zhang, R. and J.-P. Vasseur, "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.

[RFC4420] Farrel, A., Papadimitriou, D., Vasseur, J.-P., and A. Ayyangar, "Encoding of Attributes for Multiprotocol Label Switching (MPLS) Label Switched Path (LSP) Establishment Using Resource ReserVation Protocol-Traffic Engineering (RSVP-TE)", RFC 4420, February 2006.

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

[RFC4657] Ash, J. and J. Le Roux, "Path Computation Element (PCE) Communication Protocol Generic Requirements", RFC 4657, September 2006.

[STITCH] Ayyangar, A. and J.-P. Vasseur, "LSP Stitching with Generalized MPLS TE", Work in Progress, September 2005.

Authors' Addresses

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

Jean-Philippe Vasseur
Cisco Systems, Inc
1414 Massachusetts Avenue
Boxborough, MA 01719
USA
EMail: jpv@cisco.com

Arthi Ayyangar
Nuova Systems
EMail: arthi@nuovasystems.com

Full Copyright Statement

Intellectual Property

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.